Proceedings of the Maritime Situational Awareness Workshop 2019

Anne-Laure Jousselme, Paolo Braca, Leonardo Millefiori,
Francesca de Rosa, Maximilian Zocholl, Murat Uney,
Clément Iphar, Domenico Gaglione, Giovanni Soldi,
Nicola Forti, Enrica D'Afflisio

About CMRE

The Centre for Maritime Research and Experimentation (CMRE) is a world-class NATO scientific research and experimentation facility located in La Spezia, Italy.

The CMRE was established by the North Atlantic Council on 1 July 2012 as part of the NATO Science & Technology Organization. The CMRE and its predecessors have served NATO for over 50 years as the SACLANT Anti-Submarine Warfare Centre, SACLANT Undersea Research Centre, NATO Undersea Research Centre (NURC) and now as part of the Science & Technology Organization.

CMRE conducts state-of-the-art scientific research and experimentation ranging from concept development to prototype demonstration in an operational environment and has produced leaders in ocean science, modelling and simulation, acoustics and other disciplines, as well as producing critical results and understanding that have been built into the operational concepts of NATO and the nations.

CMRE conducts hands-on scientific and engineering research for the direct benefit of its NATO Customers. It operates two research vessels that enable science and technology solutions to be explored and exploited at sea. The largest of these vessels, the NRV Alliance, is a global class vessel that is acoustically extremely quiet.

CMRE is a leading example of enabling nations to work more effectively and efficiently together by prioritizing national needs, focusing on research and technology challenges, both in and out of the maritime environment, through the collective Power of its world-class scientists, engineers, and specialized laboratories in collaboration with the many partners in and out of the scientific domain.

Proceedings of the Maritime Situational Awareness Workshop 2019

Anne-Laure Jousselme, Paolo Braca, Leonardo Millefiori, Francesca de Rosa, Maximilian Zocholl, Murat Uney, Clément Iphar, Domenico Gaglione, Giovanni Soldi, Nicola Forti, Enrica D'Afflisio

Intentionally blank page

CMRE–CP-2020-001

Proceedings of the Maritime Situational Awareness Workshop 2019

Anne-Laure Jousselme, Paolo Braca, Leonardo Millefiori, Francesca de Rosa, Maximilian Zocholl, Murat Uney, Clément Iphar, Domenico Gaglione, Giovanni Soldi, Nicola Forti, Enrica D'Afflisio

Executive Summary: Between the 08 - 10 October 2019, the NATO Science and Technology Organization Centre for Maritime Research and Experimentation (CMRE) hosted the Maritime Situational Awareness Workshop 2019 (MSAW'19) at the Villa Marigola in Lerici, La Spezia (Italy). Under the theme *Science and technology meet operational needs*, the MSAW'19 thus aimed at encouraging engagement with operational experts and scientists from national governments, military, academia and industry to discuss their respective challenges regarding Maritime Situational Awareness (MSA). The objective of MSAW'19 was to foster the cross-fertilization of ideas from scientific and military domains, toward the design and implementation of future solutions tailored to MSA operational needs.

The MSAW'19 brought together about 170 participants from 23 countries including 18 NATO Nations: 14% from academia, 39% from applied research institutes, 31% from industry and 16% from the operational community. The programme mixed highly technical scientific contributions with implementation perspectives from industry. The workshop featured 42 oral presentations, 12 posters and 6 technical booths.

MSAW'19 kicked off with a welcome and opening session chaired by Dr Catherine Warner, Director of CMRE. Four keynote talks were provided by distinguished experts of radar signal processing, information fusion, target tracking and data fusion. Four projects funded by the European Commission (EC) were presented: RANGER (Radars Long Distance Maritime Surveillance and SAR Operations), co-sponsor of this workshop, MARISA (Maritime Integrated Surveillance Awareness), COMPASS-2020 (Coordination of Maritime Assets for Persistent and Systematic Surveillance), and ARCSAR (Arctic and North Atlantic Security and Preparedness Network). Mr Fabio Marziani of the NATO Communications and Information Agency (NCIA) chaired a session on NCIA's 2019 Defence Innovation Challenge that focused on challenges of the High North including improvements to MSA capability. The workshop closed with presentations by Dr Daniel Hutt and Robert Been on CMRE research activities, followed by a panel chaired by Dr Sandro Carniel (Head of Research, CMRE) where operational experts shared their views on the most promising technologies to support MSA, the topics where more research effort should be allocated, and the main challenges for MSA.

The MSAW'19 was co-sponsored by the EC project RANGER and NATO Allied Command Transformation as part of the CMRE Data Knowledge Operational Effectiveness programme. This document contains the proceedings of the MSAW'19 papers accepted for presentation.

Presentations are available at:
https//www.cmre.nato.int/msaw-2019-home/presentations.

Intentionally blank page

Proceedings of the Maritime Situational Awareness Workshop 2019

Anne-Laure Jousselme, Paolo Braca, Leonardo Millefiori, Francesca de Rosa, Maximilian Zocholl, Murat Uney, Clément Iphar, Domenico Gaglione, Giovanni Soldi, Nicola Forti, Enrica D'Afflisio

Abstract: The Maritime Situational Awareness Workshop (MSAW) was held 08 - 10 October 2019 in Lerici, La Spezia (Italy). CMRE organized this workshop with the objective of fostering the cross-fertilization of ideas from scientific and military domains, toward the design and implementation of future solutions tailored to Maritime Situational Awareness (MSA) operational needs. Centred on the theme of *Science and technology meet operational need*, the call for papers was distributed amongst the international MSA community during the early months of 2019. In response, CMRE received 48 submissions and several contributions from the European Commission (EC) funded MSA projects. The workshop brought together about 170 participants from 23 countries including scientists, engineers, researchers from scientific communities with national and international authorities, end users and operators, and industrial representatives. Overall the audience was composed of 14% academia, 39% applied research institutes, 31% industry and 16% operational community. This document contains the proceedings of the MSAW'19 papers accepted for presentation.

The MSAW'19 was co-sponsored by the EC project RANGER (Radars Long Distance Maritime Surveillance and SAR Operations) and NATO Allied Command Transformation as part of the CMRE Data Knowledge Operational Effectiveness programme.

Intentionally blank page

Contents

Intentionally blank page

1
Preamble

1.1 Organizing committee

GENERAL CHAIRS	Paolo Braca, Anne-Laure Jousselme
GENERAL CO-CHAIRS	Leonardo Millefiori, Francesca de Rosa
TECHNICAL PROGRAMME	Anne-Laure Jousselme, Leonardo Millefiori, Domenico Gaglione
LOCAL ORGANIZATION	Francesca de Rosa, Ilaria Saudella, Felicita Di Stefano
REGISTRATIONS	Maximilian Zocholl
PUBLICATION	Murat Uney, Clément Iphar, Maximilian Zocholl
PUBLICITY	Enrica D'Afflisio, Giovanni Soldi, Domenico Gaglione, Nicola Forti
WEBSITE	Xavier Berdaguer, Murat Uney

1.2   International board

Dr. Alfonso FARINA
Consultant - Leonardo S.p.A.

Prof. Dr. Wolfgang KOCH
Fraunhofer FKIE

Dr. Stefano CORALUPPI
Systems & Technology Research

Prof. Paulo Cesar COSTA
George Mason University

Dr. Michael WUNDER
Fraunhofer FKIE

Prof. Dimitris ZISSIS
University of Aegean

Mr. Lazaros KARAGIANNIDIS
Institute of Communications and Computer Systems

Mr. Evangelos SDONGOS
Institute of Communications and Computer Systems

Mr. Carmelo MASTROENI
Leonardo S.p.A.

Mr. Steven HORN
Defense Research and Development Canada

Cdr. Tim UNREIN
NATO MARCOM

Prof. James LLINAS
State University of New York at Buffalo

Prof. Peter WILLETT
University of Connecticut

Prof. Fulvio GINI
University of Pisa

Prof. Cyril RAY
Naval Academy Research Institute

Prof. James SCROFANI
U.S. Naval Postgraduate School

Mr. Dimitris KATSAROS
EXUS

Mr. Tuomas TAMMILEHTO
Laurea University

Mrs. Karna BRYAN
U.S. Department of Transportation

Mrs. Valeria FONTANA
Leonardo S.p.A.

Mr. Jonathan LOCKE
BAE Systems Submarines

Cdr. ITN Sandro FERRO
NATO MARCOM

1.3 Technical board

Dr Nicola FORTI
NATO STO CMRE

Dr Murat UNEY
NATO STO CMRE

Cdr Amleto GABELLONE
NATO STO CMRE

Dr Raffaele GRASSO
NATO STO CMRE

Cpt(N) Paolo LOMBARDI
NATO STO CMRE

Mrs Enrica D'AFFLISIO
NATO STO CMRE

Dr Domenico GAGLIONE
NATO STO CMRE

Dr Giovanni SOLDI
NATO STO CMRE

Dr Nadia BEN ABDALLAH
NATO STO CMRE

Mr Maximilian ZOCHOLL
NATO STO CMRE

Mr Gianfranco ARCIERI
NATO STO CMRE

Dr Clément IPHAR
NATO STO CMRE

1.4 Acknowledgements

We would like to express our gratitude to the keynote speakers, Dr Alfonso Farina, Mr Sergio Gallone, Mr Massimo Comparini, Dr James Llinas, Dr Felix Goavers, and Mr John Waterston for their invaluable contribution to the workshop, which stimulated discussions and throughout the event.

We would like to thank the panellists Cpt Jehan-Christophe Charles, Cdr Jorge Martinez, Lt Cdr Ivo Musulin and Mr Guy Thomas for sharing their operational perspectives about Maritime Situational Awareness during the closing session.

We would like to thank the European Union H2020 representatives of RANGER, MARISA, COMPASS2020 and ARCSAR for their presentations.

We would like to thank Mr Fabio Marzani from the NATO Communication Intelligence Agency for introducing the 2019 NITEC challenge, as well as Mr Dagfinn Vatne from the Norwegian Defence Research Establishment (FFI) for presenting the winning project.

We would like to thank CMRE Director, Dr Catherine Warner, for opening the event, Dr Sandro Carniel, CMRE Head of Research for chairing the panel session, Dr Daniel Hutt, CMRE Acting Chief of Staff, Dr Alain Maguer, CMRE Head of Engineering and Information Technology Department and Mr Robert Been, CMRE Head of Software Development, Data Management and Scientific Computer Information Systems, for their participation and interventions, as well as all CMRE scientists for their contribution.

We are grateful to all authors and presenters for the quality of the work submitted and the associated presentations, which provided the core scientific content of the workshop.

We would like to thank the chairs who kindly agreed to facilitate the sessions and who kept the (busy) agenda on schedule, as well as all the attendees for their insightful questions within the sessions.

Finally, we would like to thank our sponsors, Allied Command Transformation Innovation Branch and the EU H2020 project RANGER for providing the necessary financial support to ensure the success of this event.

The MSAW'19 Organizing Committee

1.5 Editorial notes

Maritime Situational Awareness (MSA) supports effective and efficient decision making and enables maritime operations to pre-emptively identify emerging safety, security or environmental issues so that a timely intervention is possible. MSA highly depends on the ability of sensing, collecting and processing technologies to handle the Big Data challenges brought by the ever increasing volume, velocity and variety of data, which often lack veracity. Understanding the maritime situation enables decision makers and emergency responders to focus on relevant events, to prevent malevolent acts, to minimize the impact of a possible threat, and/or to intervene in a timely manner. To reach a common and comprehensive understanding of the maritime operational environment, accurate, timely and standardized information need to be shared among nations, partners and civilian agencies, providing the required information superiority to successfully conduct maritime operations.

From coastal radar to satellite imagery, developments in Intelligence Surveillance and Reconnaissance (ISR) technology promise complementary and possibly persistent surveillance capabilities to build the maritime picture. This includes sensor technology deployed on-board Maritime Unmanned Systems (MUS), such as Unmanned Surface Vehicles (USV) and Unmanned Underwater Vehicles (UUV). These have recently demonstrated high performance in terms of scalability, adaptability, robustness, persistence and reliability. Deployed within a monitoring region, MUS can cooperatively form heterogeneous intelligent networks to detect, localize, and classify targets. Opening up new operating scenarios, MUS also introduce new scientific and technological challenges on autonomy, distributed and collective intelligence and sensing, data fusion, detection and tracking.

At the front-end of sensor systems (radar, sonar, camera, etc.) is signal processing, which uses mathematical models to investigate the principles of detection, localization, tracking and classification of targets. By associating and combining measurements of different sensors, data fusion techniques help to build a coherent and clear picture of a region of interest wherein multiple objects appear, move and disappear. At the same time, in the era of Big Data, processing infrastructures provide an unprecedented capability to gather, store and process massive amounts of data in real-time. In the maritime domain, sources are heterogeneous and provide data in many formats (structured or unstructured); moreover, data are often intermittent, sparse and noisy. Computational scalability and data models allow cueing together all the available data, discovering (otherwise hard to find) patterns.

While humans can leverage from their experience to distinguish between normal and anomalous patterns, such interpretation and contextualization capabilities are everything but immediate for machines. Artificial Intelligence (AI) aims at mimicking some human cognitive abilities to automate routine tasks in an efficient way. Machine learning focuses on the learning process from large amounts of data, to further perform classification or prediction tasks. Supervised machine learning, such as deep learning techniques, together with unsupervised approaches, such as clustering, have recently demonstrated that significant improvements are possible in the maritime domain (e.g., ship classification, maritime traffic characterisation and pattern recognition). Other AI approaches can support understanding by bringing transparency, interpretation, or explanation to reasoning and decision making, addressing cognitive tasks such as vessels behaviour analysis or intent assessment. In addition to automated reasoning and uncertainty handling, knowledge representation can

provide harmonized and rationalized terminology of the maritime domain, required for information sharing environments or integrating solutions toward technical, procedural and human maritime interoperability. To improve human-machine teaming, knowledge acquisition techniques can bring the human "in the loop" in the early stages of the system design. Capturing knowledge and know-how in a structured and reproducible way helps in the understanding of aspects, such as procedures of maritime operations, information needs, and cognitive bias.

In this context, the NATO Science and Technology Organization Centre for Maritime Research and Experimentation (CMRE) organized the Maritime Situational Awareness Workshop 2019 (MSAW'19) meant to be a forum for discussion, exchanging ideas and perspectives, to present and discuss advanced technologies, innovative concepts, and emerging scientific challenges with respect to current and future MSA operational needs.

This workshop is in-line with a series of events on the topic that have been organized by CMRE over the last ten years, including the *NATO Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness*, 15-17 September 2009, the *Symposium on Port and Regional Maritime Security*, 21-23 May 2012, the *SCI-275 Maritime Situational Awareness enabled by Space-based systems Workshop*, 24-26 February 2015, the *Table Top eXercise (TTX) on Maritime Situational Awareness*, 15-18 May 2017 and the *Maritime Big Data Workshop*, 9-10 May 2018.

The MSAW'19 brought together about 170 participants from 23 countries, 18 NATO nations including scientists, engineers, researchers from scientific communities with national and international authorities, end users, operators, and industrial representatives. Overall, the participants of the workshop were composed of 14% academia, 39% applied research institutes, 31% industry and 16% operational community. The programme mixed highly technical scientific contributions with implementation perspectives from industry. The workshop featured 42 oral presentations, 12 posters and 6 technical booths.

Thanks to the numerous contributions received, we have been able to offer a rich, unique and interesting programme covering diverse aspects of MSA through technical sessions on signal processing, artificial intelligence, deep learning and classification, big data architectures and systems, maritime anomaly detection and behavioural analysis, ethics sensing and surveillance technologies, maritime unmanned systems and persistent surveillance.

Four keynote talks were provided by distinguished experts of radar signal processing, information fusion, target tracking and data fusion. Dr Alfonso Farina and Sergio Gallone from Leonardo S.p.A., Italy, together with Massimo Comparini, e-GEOS Chief Executive Officer, Telespazio Head of Line of Business Geoinformation, Italy, shared the floor for opening the workshop with the presentation entitled *Maritime surveillance: radar technologies and scenario characteristics*. Dr James Llinas, Emeritus Professor at the State University of New York at Buffalo, US, presented remotely in the afternoon of the first day, *Re-examining fusion-sensemaking-decision-making interdependencies again*. Dr Felix Govaers, Deputy Head of Department Sensor Data and Information Fusion at Fraunhofer institute FKIE, in Bonn, Germany, presented the keynote talk on the second day, about *Push and pull in digitalization: Technology drivers for sensor data fusion*. Finally, Mr John Waterston, Program Manager in the Strategic Technology Office at the Defense Advance Research Program Agency, US, concluded with the last keynote talk, presenting *Ocean of Things*.

Four projects funded by the European Commission (EC) were presented: RANGER (Radars Long Distance Maritime Surveillance and SAR Operations), co-sponsor of this workshop, MARISA (Maritime Integrated Surveillance Awareness), COMPASS2020 (Coordination of Maritime Assets for Persistent and Systematic Surveillance), and ARCSAR (Arctic and North Atlantic Security and Preparedness Network). Mr Fabio Marziani of the NATO Communications and Information Agency (NCIA) chaired a session on NCIA's 2019 Defence Innovation Challenge that focused on challenges of the High North including improvements to MSA capability. The winning entry was Dual-use of AIS data, combining AIS tracking with social network analysis for increased maritime network awareness submitted by the US Navy Post Graduate School and Norway's defence research agency FFI. A brief on the project was presented by Dagfinn Vatne of FFI. The concept presented was to combine historical and live AIS data with social network analysis (SNA), in order to identify and geo-locate suspicious actors. Other NCIA Defence Innovation Challenge winners were also present at the workshop with several posters on their work.

The workshop closed with presentations by Dr Daniel Hutt and Robert Been on CMRE research activities. A expert panel was animated by Dr Sandro Carniel (Head of Research, CMRE) with Cdr Jorge Martinez (Spanish Navy, NATO Combined Joint Operations from the Sea Centre of Excellence); Cpt Jehan-Christophe Charles (Ret. French Navy); Lt Cdr Ivo Musulin (HRV Navy, NATO Shipping Centre, MARCOM); and Guy Thomas, DCSA, MBA (Advisor for Maritime Situational Awareness, Multinational Maritime Security Centre of Excellence). Experts provided their views on the most promising technologies to support MSA, the topics where more research effort should be allocated, and the main challenges for MSA.

This workshop was co-sponsored by the EC H2020 project RANGER (Radars Long Distance Maritime Surveillance and SAR Operations), and NATO Allied Command Transformation, as part of the CMRE Data Knowledge Operational Effectiveness programme.

1.6   Keynote talks

Maritime surveillance: radar technologies and scenario characteristics

by Dr Alfonso Farina, Sergio Gallone, Massimo C. Comparini

Since the beginning of the radar history Maritime Situational Awareness (MSA) has been identified as the first operational application for both commercial and defense areas. After the WW2 the first operational standards were defined for on board radar equipment, followed by the regulations for fixed radar installations. Vessel Traffic System (VTS) radars were the extension of the on board equipment, since they are the same apparatus equipped with larger antennas, to obtain longer detection ranges and better azimuth accuracy with medium-high altitude installations. In the last decade solid state VTS radars were developed replacing the use of magnetrons, obtaining lower costs for maintenance and logistics, and introducing the Doppler processing in the VTS radars. Modern radars still need to be improved, in term of performances, considering the complex environmental and operative scenario and the more stringent operational needs. Examples are the followings:

- *Sea clutter ("sea spikes") is a relevant restraint, and its characterization is still under study.*

- *The large variance of the target size (from sea tank to rubber or wooden boats) require high radar resolution, both in range and azimuth, and high dynamic amplitude range.*

- *The operative centers allow interoperability and remote relocation, thus the radar network works at different level of data information displays and different requirements for the data fusion*

In the presentation some of the topics listed above are discussed together with several examples of live data.

***

Dr Alfonso Farina received the doctor degree in Electronic Engineering from the University of Rome (IT) in 1973. In 1974, he joined Selenia, then Selex ES, where he became Director of the Analysis of Integrated Systems Unit and subsequently Director of Engineering of the Large Business Systems Division. In 2012, he was Senior VP and Chief Technology Officer of the company, reporting directly to the President. From 2013 to 2014, he was senior advisor to the CTO. He retired in October 2014. From 1979 to 1985, he was also professor of "Radar Techniques" at the University of Naples (IT). He is the author of more than 800 peer-reviewed technical publications and of books and monographs (published worldwide), some of them also translated in to Russian and Chinese. Some of the most significant awards he's received include: (2004) Leader of the team that won the First Prize of the first edition of the Finmeccanica Group (45,000 employees) Award for Innovation Technology, out of more than 330 submitted projects by the Companies of Finmeccanica Group; (2005) International Fellow of the Royal Academy of Engineering, U.K., and the fellowship was presented to him by HRH Prince Philip, the Duke of Edinburgh; (2010) IEEE Dennis J. Picard Medal for Radar Technologies and Applications for "Continuous,

Innovative, Theoretical, and Practical Contributions to Radar Systems and Adaptive Signal Processing Techniques"; (2012) Oscar Masi award for the AULOS® "green" radar by the Italian Industrial Research Association (AIRI); (2014) IET Achievement Medal for "Outstanding contributions to radar system design, signal, data and image processing, and data fusion". He is a Visiting Professor at University College London (UCL), Dept. Electronic and Electrical Engineering, CTIF (Center for TeleInFrastructures) Industry Advisory Chair, and a Distinguished Lecturer (DL) of IEEE AESS.

Sergio Gallone graduated in Electronic Engineering with maximum grade, in the University of Roma "La Sapienza" on 1984. He joined Selenia S.p.A on 1985 and was involved as Project Manager/Project Leader of many innovative radar projects, for both military and civilian applications, in naval and ground domain, from UHF to X-bands, for the detection/recognition of aircraft, ships and personnel/ground vehicles. He obtained 2 patents on sea clutter reduction methods. He is now in the central CTO of Leonardo S.p.A.

Massimo Claudio Comparini has a Master Degree in Electrical Engineering, Remote Sensing and Radar Systems, University of Rome "La Sapienza" (Italy), and a Master Degree in Strategy (MBA), Graduate School of Business, Stanford University, CA (USA).

He began his career in the space industry in 1983 at Selenia Spazio (later Alenia Spazio), holding positions of increasing responsibility, up to that of Chief Technology Officer. After the establishment of the joint venture Thales Alenia Space, he took up the role of Deputy Chief Technical Officer, CTO of TAS Italia and Vice President for R&D, Technology, Product Policy and IPR. In 2013 he was appointed Chief Technical Officer at Telespazio, a joint venture between Finmeccanica and Thales.

He is also the Chairman of Space Innovation in the Italian Technology Platform SPIN IT. Author of over 80 papers, he published works in the field of space and microwave technologies, innovation and technology management. He holds seats in several steering committees and boards of scientific associations and International conferences. He has had several Academic chairs in the Engineering and Management Departments of Italian Major Universities. Comparini is Chairman of the Space Innovation technology platform in Italy, and is a National expert in the space configuration of the European Horizon 2020 programme. In 2016, he has been appointed as CEO of e-GEOS, an ASI/Telespazio company, which has the mission to commercialize world-wide COSMO-SkyMed radar Constellation data, applications and services and to provide multimission Geospatial integrated solutions, based on innovative satellite data processing chain.

Reexamining Fusion–Sense Making–Decision Making Interdependencies Again

by Dr James Llinas

Many decision-making process models show some level of interdependency between the decision-making (DM) process and its subfunctions and the information available to that DM process. One example is the well-known "OODA: Observe-Orient-Decide-Act" model of Boyd and another well-known example is Klein's Naturalistic or Recognition-Primed Decision model. In the Information and Data Fusion (IDF) community, the point of view for process/system development is often centered about forming a least-uncertain estimate of a situational state as derived from multiple sources and sensors. However, there has been, unfortunately, relatively little interaction between the IDF, Sensemaking, and DM communities (e.g., the International Information Fusion Conferences of the last 5 years collectively have fewer than a dozen papers on DM), and the critical examination of inter-process interdependencies needed to co-design these processes for optimum performance.

This presentation suggests that IDF, Sensemaking and DM processes cannot be optimally designed without consideration of such interdependencies, and a variety of details of such interdependencies of these processes are discussed. Among the issues commented on are: little accounting for the "Dual-process Models" of DM (factors of which have in fact been addressed by Nobel Laureates), integration of Information Foraging operations and the link to Sensemaking processes supportive of DM, important temporal effects in inter-process design, metrics involved in measuring and evaluating process interdependencies, new factors on the input side such as unstructured and linguistic data, among others. The spirit of the paper is as a research challenge for the IDF, cognitive and DM communities and an appeal for joint efforts to evolve optimal designs of these important interdependent processes.

Dr James Llinas is an Emeritus Professor in the Departments of Industrial and Systems Engineering and dual-appointed in the Department of Mechanical and Aerospace Engineering at the State University of New York at Buffalo, in Buffalo, New York, USA. He is the Founder and Executive Director of the Center for Multisource Information Fusion, the only systems-oriented academic research center for Information Fusion in the United States; the Center has been in existence for some 20+ years and has conducted extensive and distinctive research for a wide variety of governmental and civilian clients. Dr. Llinas brings over 35 years of experience in multisource information processing and data fusion technology to his research, teaching, and business development activities. He is an internationally-recognized expert in sensor, data, and information fusion, co-authored the first integrated book on Multisensor Data Fusion, and has lectured internationally for over 30 years on this topic. He is a recipient of the Joe Mignona Award given to one person annually by the US Data Fusion Group for lifetime achievements in Data Fusion. In 1998, as the International Society for Information Fusion was being developed, he was voted as the First President of the Society. He has been a consultant and advisor to U.S., Canadian, Australian, Norwegian, Swedish, and Spanish Ministries of Defense. In addition to co-authoring the first integrated text on data fusion, Dr. Llinas is also the co-editor of the "Handbook of Multisensor Data Fusion" published in 2001, a second edition in 2009, of a new Handbook of Distributed Multisensor Data Fusion published in 2012, as well as a recent text on Context-Enhanced Information Fusion published in 2016.

Push and Pull in Digitalization: Technology Drivers for Sensor Data Fusion

by Dr Felix Goavers

Comprehensive networking, ubiquitous sensor technology, robots, drones, artificial intelligence, machine learning, ... and many more of these buzzwords! Digitalization has long since dominated almost all areas of modern life. Defence & Security is no less affected by the "digital revolution" in all its branches. On the contrary – many fundamental technologies have their roots in this area.

For a wide range of applications, we consider trends and practical examples to shed some light on the impact of comprehensive digitalization on sensor data fusion, i.e. on surveillance, resources management and supporting smart decisions. Fusion technology is basic for acing successfully and responsibly in the increasingly complex technosphere that is surrounding us.

Cognitivity of communication and sensor systems, automation of platforms up to levels of autonomy, manned-unmanned teaming, and handling of big or sparse data are technologies enabled by sensor data fusion and intelligent resources management and will be discussed by considering examples. Some more visionary ideas of fusion algorithms that are exploiting quantum physics inspired mathematics will conclude the talk.

Dr Felix Goavers received his Diploma in Mathematics and his Ph.D. with the title "Advanced data fusion in distributed sensor applications" in Computer Science, both at the University of Bonn, Germany. Since 2009 he works at Fraunhofer FKIE in the department for Sensor Data Fusion and Information Processing where he now leads the research group "Distributed Systems". The research of Felix Govaers is focused on data fusion for state estimation in sensor networks. This includes track-extraction, processing of delayed measurements as well as the Distributed Kalman filter and trackto-track fusion. He is also interested in advances in state estimation such as particle flow and homotopy filters and the random finite set theory approaches.

Felix Govaers is an active member of the ISIF community since 2008, he has been organizing the ISIF co–sponsored SDF Workshop in Germany for many years as the Technical Program Chair. For the upcoming FUSION in 2016 he supports the organization team as the Publication Chair. Since 2014 he also serves as an Associate Editor for the IEEE Transactions on Aerospace and Electronic Systems journal.

Ocean of Things

by Mr John Waterston

DARPA's Ocean of Things (OoT) program enables persistent maritime situational awareness over large ocean areas by deploying thousands of low-cost, intelligent floats that drift as a distributed sensor network. Each float manages a suite of commercially available sensors to collect environmental data such as sea surface temperature, sea state, and location as well as activity data about vessels and marine mammals moving across the ocean. The floats periodically transmit processed data, or immediately report events based on internal prioritization schemes. Messages travel via commercial satellite to a government cloud for storage and real-time analysis. Cloud-based data analytics feature machine learning aimed at discovering emergent features and behaviors from sparse data. The multiple performers manufacturing floats and developing software are being led by a government management team to employ commercial design methodology and agile best practices. At-sea float deployments are planned in two phases over 2019 (1-month) and 2020 (3-month). Program benefits include ocean environmental products derived from high-density, in-situ measurements and analytical applications, which can simultaneously provide users a range of outputs to include ocean circulation prediction, vessel and marine mammal tracking, and dynamic ocean resource management.

Mr John Waterston joined DARPA in January 2017 as a Program Manager in the Strategic Technology Office (STO). Mr. Waterston's interests are in maritime autonomy, undersea warfare, and distributed sensing. He seeks disruptive concepts to transform the future defense enterprise.

Mr. Waterston currently manages a diverse portfolio of advanced projects. POSYDON has developed an acoustic positioning, navigation, and timing system to enable future undersea operations. TEAC investigates fundamental techniques to cohere acoustic energy transmissions in the ocean environment. Ocean of Things seeks to transform environmental and operational awareness of the ocean using low-cost, intelligent floats.

Prior to arriving at DARPA, Mr. Waterston was a Senior Program Manager within SRI International's Advanced Technology and Systems Division. There he led defense research programs in various technical areas to include cyber warfare, arctic sensing, and unmanned vehicles. Outside of industry, Mr. Waterston serves as a Commander in the U.S. Naval Reserves, where he applies his submarine operational experience in support of warfare systems requirements development and platform integration.

A native of Dallas, TX., Mr. Waterston graduated from the United States Naval Academy with a bachelor's degree in Electrical Engineering and entered the Submarine Warfare Community upon commissioning. He graduated from Stanford University with a Master of Science in Electrical Engineering.

2

Radar systems

Contents

RANGER: Radars and Early Warning Technologies for Long Distance Maritime Surveillance

Lazaros Karagiannidis[1], Dimitrios Dres[2], Eftychios Protopapadakis[2,] Frédéric Lamole[3], François Jacquin[3], Gilles Rigal[4,] Eleftherios Ouzounoglou[1], Dimitris Katsaros[5], Alexandros Karalis[5], Luigi Pierno[6], Carmelo Mastroeni[6], Marco Evangelista[6], Valeria Fontana[6], Domenico Gaglione[7], Giovanni Soldi[7], Paolo Braca[7], Sari Sarlio-Siintola[8], Evangelos Sdongos[1,] Angelos Amditis[1]

[1]Institute of Communication and Computer Systems (ICCS), Iroon Polytechniou Str. 9, GR-15773, Athens, Greece

[2]Telesto Technologies Ltd 62 Imitou Str, 15561 Cholargos, Greece

[3] Embedded Systems Dpt / HF division 5 , rue Brindejonc des Moulinais, 31506 Toulouse Cedex 5

[4]Embedded Systems Dpt / HF division Les Hauts de la Duranne, 370, rue René Descartes , 13857 Aix-en-Provence Cedex 3

[5]EXUS Software LTD, 07932357, Old Broad Sreett 25 Tower 42, London EC2N 1PB, United Kingdom

[6]Leonardo S.p.A. (LDO), Via Tiburtina km 12,400 - Roma - 00131 - Italy

[7]NATO STO Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy

[8]Laurea University of Applied Sciences Ratatie 22, 02200 VANTAA, Finland

ABSTRACT

In this paper, we present a high-level view of RANGER, a novel platform that combines innovative radar technologies with cutting edge technological solutions for early warning, in view of delivering a surveillance platform offering detection, recognition, identification and tracking of suspicious vessels, i.e. capabilities that exceed those of current radar systems. The RANGER platform consists of two radar technologies, a novel Over-The-Horizon (OTH) Radar combined with a Multiple- input and Multiple-output Radar (MIMO) implemented exploiting the latest photonics advancements, a Uniform Communication Gateway (UCG) for seamless, secure and interoperable integration of radars and legacy systems, an Early Warning System (EWS) exploiting data fusion and deep and adaptable machine learning schemes able to automatically detect radar targets and produce early warnings, an Advanced User Interface (AUI) for visualization of the maritime operational picture, and a Common Information Sharing Environment (CISE) gateway adapter for sharing information with the CISE network. The RANGER system aims at significantly progressing the accuracy and long distance detection, identification and recognition rates as well as improving the provision of early warnings and alerts related to maritime operations, thus drastically improving the response and intervention capacity of related European services.

Keywords: Over-The-Horizon (OTH) Radar; Multiple-Input Multiple-Output (MIMO) radar; machine learning; maritime surveillance; early warning system, ethics compliance

1. INTRODUCTION

Maritime surveillance and, more generally, the monitoring of the maritime traffic in the Exclusive Economic Zone (EEZ) are today of critical importance, especially for the Mediterranean Sea that concentrates 25% of the global traffic and is subject to critical situations and irregular activities. With the vital importance of maritime trade to the integrity of the global economy, maritime security remains of great interest. Widespread criminal trafficking, piracy, and terrorism threaten infrastructure, vessels and supply chains within this vast environment. As a consequence, navies and coast guards are adapting themselves to confront asymmetric and criminal threats in the support of their pivotal role in helping to ensure maritime security, which proves to be an increasingly challenging task.

Traditional watch-keeping methods including visual observation; the use of audible warnings such as foghorns, bridge-to-bridge Very High Frequency (VHF) radio communications; and radar (including the Automatic Radar Plotting Aid) have all improved over the years but still suffer from lack of accuracy and time delays due to limited detection range.

A brief look at previous or existing systems shows that most mature civilian systems have been developed and used for oceanographic applications (e.g. current studies). Some experimentations for ship detection have however been performed without leading though to a general operational use for maritime traffic surveillance. Evaluation of HFSW radar system [1] [2] for this purpose has shown that it has a limited range when applied to small ship detection (100 to 150 km) and their robustness to sea and ionospheric clutter and jamming are not demonstrated although essential for small ship detection.

The majority of the modern HFSW radars are actually derived from an architecture originally designed by D. Barrick [3] at the end of the 70s. The main improvements performed in the last years have been made to waveform (updated to well-known FMCW modulation) or antenna shapes. Recently, Onera, with the collaboration of Pierre & Marie Currie University, studied on improving the excitation of the surface wave and reducing the self-perturbation generated by sky wave propagation [4] but still at an low Technology Readiness Level (TRL). Therefore, the aforementioned systems suffer from a number of which make them unsuitable for effective maritime traffic monitoring with current challenges (ie small boats), thus dictating disruptive technologies and approaches.

It is therefore imperative to exploit current and emerging technological advancements in view of providing the appropriate tools that, when combined with existing and legacy systems will be able to lead to significant reduction of the loss of lives at sea, a better control of irregular maritime traffic, while safeguarding internal security by preventing cross-border crime such as trafficking and smuggling.

In this context, the European project RANGER [5] innovates by combining novel and ground-breaking radar technologies with innovative supporting technological solutions for early warning, in view of delivering a surveillance platform that will offer detection, recognition, and identification as well as tracking of suspicious vessels capabilities far beyond existing legacy radar systems, seamlessly fitting and contributing to the Common Information Sharing Environment (CISE) framework through the provision of on-demand CISE compliant services.

In this paper, we present a high-level view of the RANGER system, concept, overall architecture and main components. The remainder of the paper is structured as follows: Sec. 2 presents an overview of the RANGER system, followed by a description of the RANGER radars (Sec. 3) and its Core Platform modules; Early Warning System, Uniform Communication gateway, Advanced User Interface, CISE translation gateway (Sec. 4). Sec. 5 briefly presents the RANGER Ethics compliance framework, and Sec. 6 concludes the paper.

2. RANGER SYSTEM OVERVIEW

The RANGER system architecture follows a modular approach supporting seamless integration of radar solutions, interoperability with legacy systems, enhanced situation awareness and usability. The RANGER system is composed of the following elements:

The *Over-the-Horizon* (*OTH) radar* is a HFSWR system that has the ability to detect targets at very large ranges, beyond the horizon (several hundreds of kilometers), vessels above 10 m length.

The *Photonics Enhanced Multiple-Input Multiple-Output* (*PE-MIMO) radar* is an innovative system that supports the OTH radar for detecting targets (<10m) in close range over several kilometers with high resolution.

The *Early Warning System (EWS)* constitutes the back-end of the RANGER system, and is responsible for early detection of events, data storage and provision of warnings and alerts. It is the core element, which is closely interdependent with the data fusion and machine learning module of the RANGER platform.

The *Data Fusion* module performs data fusion from different sensors to obtain a set of tracks (routes) of targets.

The *Machine Learning* module employs machine learning algorithms in order to derive conclusions about the characteristics of the detected/tracked vessels.

The *Uniform Communication Gateway (UCG)* acts as the interoperability layer of the RANGER system and enables the integration of legacy surveillance radars, Automatic Identification System (AIS) data and complementary data sources related to maritime environment exploiting the Inter VTS Exchange Format Service (IVEF)[6] for vessel track data.

The *CISE translation gateway* enables the integration of the CISE compliant RANGER services (OTH radar track service, the PE-MIMO radar track service and the RANGER EWS service) with the CISE network.

The *Advanced User Interface (AUI)* is the standalone user interface and front-end of the RANGER system. It displays the different tracks from RANGER data sources (OTH, PE-MIMO, AIS, Legacy System) as well as alerts and notifications from EWS, and will give the possibility to an operator to see information on a selected track.

The functional architecture of the RANGER system is depicted in Figure 1, and illustrates the various functional elements of the RANGER system: the sensing elements (OTH radar, PE-MIMO radar, AIS, Legacy Systems), the core platform (EWS, Data Fusion, Machine Learning), the RANGER gateways (UCG and CISE translation gateway), and the front-end and user interface (AUI).

Security as a non-functional requirement of the RANGER system design spans through all the layers of the RANGER architecture and covers physical infrastructure protection, perimeter and network security as well as user and data security.
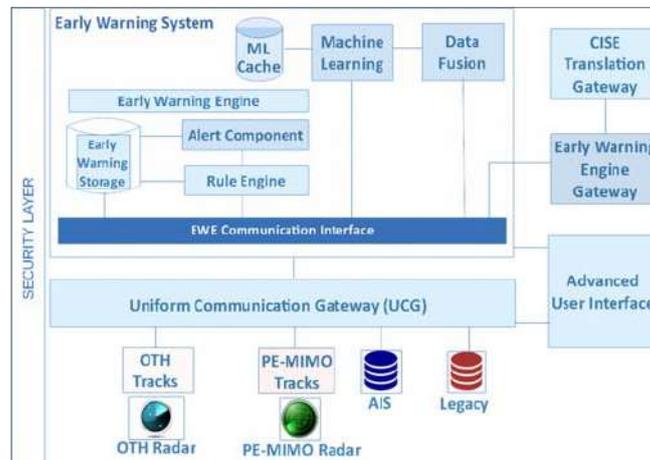
Figure 1: RANGER Functional Architecture

3. RANGER RADARS

3.1 Over-The-Horizon (OTH) Radar

The OTH radar system architecture, proposes new waveforms, processing methods and antenna concepts to meet the challenges that are imposed by High Frequency Surface Wave Radars (HFSWR). It is built on a bi-static architecture as illustrated in Figure 2.
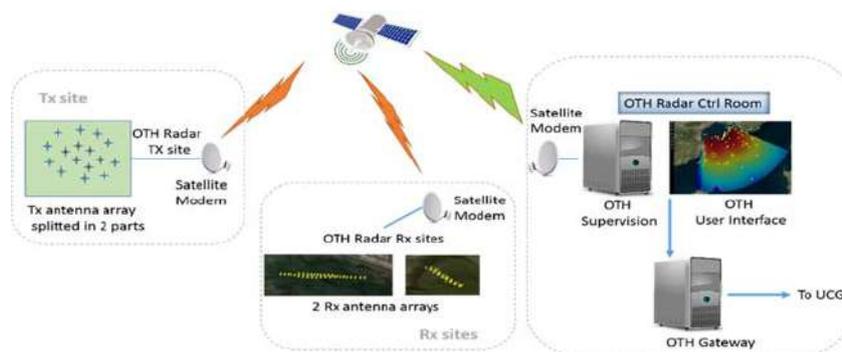
Figure 2: HFSWR Bi-static Architecture

The system consists of a Transmission Site (divided into two frequencies) and a Reception Site (with 2 antenna arrays allowing the multiplicative principle for both frequencies) located close to the shoreline.

Evaluation of the performance is realized during two pilot demonstrations in two different countries, each one dedicated to specific criteria measurement:

- In the Gulf of Lion, south of France, the radar is optimized to detect small dimension boats (and/or of weak Surface Radar Equivalent) within long range. The performance is optimized for maximum efficiency; this is accomplished through numerous transmitting antennas, and a stronger transmission power level.
- In Greek archipelago, the radar is demonstrating the ability to be integrated as a deployable system on remote areas and capable of taking into account the « mask » provided by the islands. The performance is optimized for detecting and tracking vessels masked by the islands on quite shorter distances.

In both cases, it is possible to derive specifications in range, size of boat, alarm rate, and other parameters. In order to address these objectives, diverse factors are taken into account:

- Use of a secondary transmission frequency which allows simultaneously:
  - To move the interference masks due to the reflectivity of the sea and the ionosphere, producing pertinent information for detection
  - To provide complementary information on the Surface Radar Equivalent of small sized boats.
- An antenna network development based on the multiplicative principle allowing to improve performances:
  - By reducing ionospheric disturbances
  - By increasing resolution
  - By handling both frequencies
- An adjustment in the digitalization means in order to simplify the developments, the validation, and the system's deployability.

With these new waveforms and new technical concepts, the coverage exceeds 150 Nm with limited power (500 W) and provides detection capability for vessels over 10m in length. Figure 3 depicts a typical coverage map of the OTH radar that visualizes different tracks of vessels.

The OTH detection capabilities during the 2nd French Pilot, with only a sum of 300 W, as amplifiers power output splitted on 10 transmitting antennas, are up to 230 km (125 NM) for the smaller vessels (between 10 m and 30 m length) and above 320 km (170 NM) for the bigger vessels (above 30 m length) or beyond for some cases, with a detection ratio of 90%.

Figure 3: HFSWR coverage map

3.2 Photonics Enhanced - Multiple-input Multiple-output (PE-MIMO) Radar

A short range PE-MIMO radar providing very high resolution in the range (<10m) and angle (<2deg), detection of very high number of targets and furthermore cost effectiveness is being developed in RANGER, complementary to the long range OTH radar. The PE-MIMO is based on Frequency modulate Continuous Wave radar in S-band, compliant with the IALA V 128 recommendations[7] for coastal radar.

This radar performs an outstanding bearing resolution exploiting a combination of "Multi-Static" MIMO design approach. This is ensured by using all the antennas to simultaneously transmit orthogonal signals (in the present application this is achieved by Time Domain Multiplexing of Mt transmitters) and to detect the backscatters by all the receive antennas at the same time. Under this condition, a virtual antenna array of size $N = Mt \times Mr$ (where Mr is the number of receive antennas simultaneously used) is formed, which provides N times more positioning data compared to single transmit and receive antenna radar technology counterparts.

A high level Radar Architecture is displayed in Figure 4.
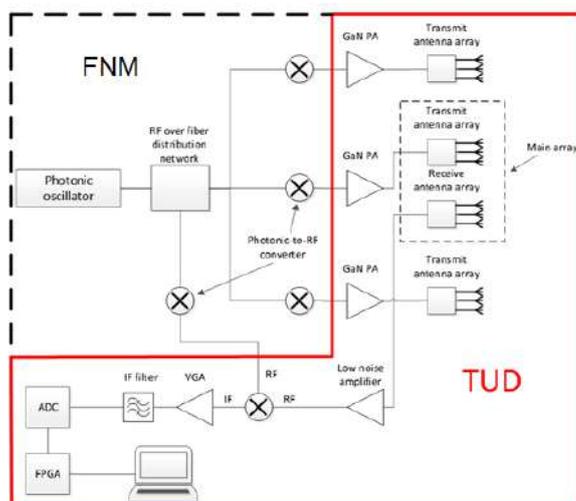
Figure 4: PE-MIMO Radar Architecture

The implementation of the PE-MIMO radar exploits a set of innovative system solutions, with special focus in the use of Photonic Technologies. Photonic Technologies are fundamentally exploited to realize Optoelectronics Oscillators (OEO) that perform critical system functions in a FMCW radar such as generation of low phase noise stable Local Oscillator and Clock for Direct Digital Synthesizer (DDS). Fibre optics microwave photonics (also named Radio over Fiber - RoF) links perform efficient signal distributions of RF chirp to transmitting and to receiving antennas that in the case of MIMO can be at a distance up to kilometers far from each other.

Moving from electrical to Microwave Photonics technologies, transmission of microwave signals enables us to experience a set of basic advantages related to:

- Wide Fiber Optic Bandwidth (> 100 THz) and Channel Capacity

- Low Fiber Optic Attenuation (< 0.2 dB/km)

- Reduced Weight & Size (0.7 kg/km for fibre with 567 kg/km for coax. cables)

- No EM Interference or Susceptibility

Outcomes deriving from cost reduction by architecture simplification or performance increase not achievable in electronic domain have also to be considered. This is the case, e.g., for RF carrier generation performed by using optical cavities (Photonic Oscillators) rather than electric resonators (typically multiplied quartz). The improved

performances, in term of strong Phase Noise reduction, of Photonic Oscillators with respect to conventional electric approaches are due to the high Q cavities realized by using very long and very low loss optical fibers.

4. RANGER CORE PLATFORM

4.1 Early Warning System

The EWS is the back-end of the RANGER system, and is constituted by a Data Fusion (DF) engine, a Machine Learning (ML) module and an Early Warning Engine. It is responsible to store the incoming data, to fuse the detected events from different sources (OTH and PE-MIMO radars, legacy systems), to correlate these events with the machine learning data and to provide early warnings and alerts.

4.2 Data fusion

The main role of the data fusion module is to take all the available measurements at a particular time step, say t, that could be measurements from different sensors (OTH radar, PE-MIMO radar, AIS data, measurements from existing legacy systems), and fuse them in order to obtain a set of tracks (routes) related to the targets present in the maritime scene. Two sequential tasks are performed to extract the tracks: data association and hybrid tracking.

The data association is required because of the uncertain origin of each measurement: indeed, in a region with multiple targets, as the one that RANGER is expected to deal with, it is not known if a measurement has actually originated from a target and, even if it was, which target generated it. The data association module takes the available measurements at each time step, t, and associates them to each of the already existing tracks (measurement to track association, M2TA). The measurements which are not associated to any of the existing tracks, can be either false alarms, due to sea-clutter or interference, or generated by recently detected targets and therefore used to initialize new tracks. RANGER uses a probabilistic data association approach, therefore allowing a level of uncertainty.

The hybrid tracker, given the result of the data association and some prior knowledge about existence and location of the targets at the previous time step, t-1, updates their status by means of filtering techniques based on Kalman filtering or more general particle filtering. In literature, several algorithms and methodologies have been presented for multi-sensor multi-target tracking for maritime surveillance, e.g. [8]. Recently, a Bayesian approach method for multi-sensor multi-target tracking based on the belief propagation (BP) scheme has been proposed [9]. It combines good performance with an attractive scaling of computational complexity. It allows estimating adaptively unknown environmental or model parameters, as, for example, the probability of detection of the targets [10]. Extensions to these works will be considered for the Data Fusion module.

The Key Perfomance Indicators used to evaluate the performance of the Data Fusion module are listed in the following table:

Table 1: KPI of Data Fusion Module

Name	Description	Value
Time-on-Target (ToT)	Fraction of time (expressed in percentage) for which a target is successfully tracked. For example: • 0 % : the target is not tracked at all. • 100 % : the target is tracked for its entire lifetime.	> 60 %
Track Fragmentation (TF)	Number of fused tracks associated to a true trajectory.	≤ 6

For the 1st RANGER French Pilot conducted off the shore of Marseille on October 16th 2018, an average ToT of 60% and a TF value of 3.2 were obtained during real-time operations. However, by processing offline all the measurements (comprising the ones not processed due to network/communication issues), an average ToT of 67.2 % and an average TF value of 2.2 were obtained.

4.3 Machine learning

The Machine Learning (ML) module operates using data from multiple data sources, as they are provided by the DF module. As the name implies, ML module employs various machine learning methods to derive conclusions about the characteristics of the detected/tracked vessels. Vessel characteristics are, first and foremost, related to vessel type/size,

as well as more sophisticated aspects (e.g. with regard to vessel behavioral patterns, common routes). ML has three main subsystems:

- Outlier detector, which detects differentiation from the norm by comparing the last 5-minute values for a ship, against all the others.
- Length estimator, which categorize the ship's length in one out of 5 predefined length categories.
- Statistical anomaly capturing, which is a generic system allowing for specific condition violation. The conditions are provided by the local authorities and require the track history of a ship to be calculated.

Outlier detection is a combination of stacked autoencoders (SAs)[11] and density-based clustering, using OPTICS algorithm. The SAs are responsible for the non-linear dimensionality reduction. The OPTICS algorithm uses as input the SAs outputs and compares ships' patterns [12]. An illustration on how the outlier detection scheme works, can be found in Figure 5. Length estimation constitutes a traditional multi-class classification problem. We have set a number of classes and the length range for each of these classes. Given a set of n past fussed data files, we use four values, i.e. longitude, latitude, speed and course as input to the classifier. The output is the corresponding class. The class (length) is estimated using five categories: i) <10 meters, ii) 10-15 meters, iii) 15-20 meters, iv) 20-25 meters and v) >25 meters. A set of seven different types of classifiers were evaluated for the suitability on length category estimation, namely: Classification trees, discriminant analysis, k nearest neighbors, naïve Bayes, support vector machines and their ensembles[13]. Despite the relatively small length sizes, i.e. four out of five categories are below 25 meters, an ensemble of three classifiers, using five-minute track history, provides an average of 60% in F1 score. Classification tree ensemble provided the best performance compared to other classifiers, and is used by ML module.
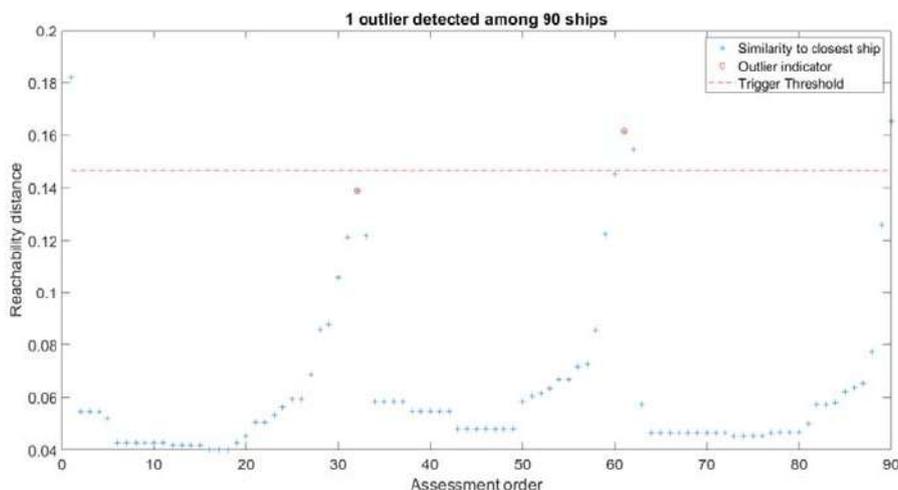
Figure 5: Outlier detection illustration: for the specific moment one ship, out of 90, has significant differences to its behavior compared to the next closest ship.

Last, but not least, variations compared to user-imposed standards are examined by the statistical anomaly capturing module. Any outcomes are forwarded to rule engine, and support the alert activation by the Early Warning Engine.

4.4 Early Warning Engine

The Early Warning Engine (EWE) is the core component of the EWS and its main functionalities are the generation of alerts (e.g.: Navigation in restricted area, High absolute speed, Unknown origins and others) and warnings based on predefined though configurable rules and to operate as the communication layer of the EWS. The maritime rules that indicate anomalies or abnormal behavior on see, have been identified with the assistance of the RANGER platform end users (operators) by providing a set of parameters and conditions that are fully modifiable through the Rule Engine. The Rule Engine processes data coming from the Data Fusion and Machine Learning modules and is capable

of applying rules simultaneously on various data streams, regardless of the communication protocols. The definition of time windows for the incoming data, the filtering of data and finally the queries between incoming data streams are features of the Rule Engine that enable the complex processing of the incoming events. The output of the rule application is fed to the Alert Component which stores the created alerts internally and shares them either with the AUI, in a standalone mode, or with the CISE translation gateway (through the CISE translation gateway).

4.6 Uniform Communication Gateway

The Uniform Communication Gateway (UCG) is the main gateway link between RANGER project's Radar sensors, Legacy systems and AIS Systems with the Early Warning Engine (EWE) and the Advanced User Interface (AUI). It is responsible to collect and merge heterogeneous data into a common and comprehensive picture to be shared among all relevant parties of the RANGER platform.

In other words UCG is the interoperability layer of the RANGER platforms facilitating the seamless integration of various maritime related data sources. This includes also satellite data from Copernicus Marine Environment Monitoring Service[14] such as significant weights height and other forecast data that are of interest in maritime surveillance and search and rescue operations. The UCG implementation is based on the Inter VTS Exchange Format (IVEF) service. IVEF is an ongoing standardization effort by a large group of industrial members of IALA which intends to provide a common framework for the exchange of information between shore-based e-Navigation systems, such as VTS systems, eNavigation stakeholders and relevant external parties[6]. Several IVEF servers and clients can be implemented, depending on the data sources and the ICT implementation. In RANGER the UCG translated and merges the data, while IVEF servers and clients have been implemented for sharing data with the EWE and the AUI. Furthermore, a standalone UCG Web Interface and Dashboard is implemented for the control and management of the UCG, the Copernicus satellite data, as well as display of health status information of the active processes, the collection of real time statics of message data flows, data traffic load and volume. Figure 6 illustrates the functional architecture of the UCG (left side) and the dashboard with the real time data traffic statistics.

Figure 6: UCG Functional Architecture and Dashboard

4.7 Advanced User Interface (AUI)

The RANGER AUI is a component specifically designed to provide multiple categories of users with the functionalities required to exploit the results of both the OTH and MIMO radars, data fusion results and legacy system information, by using the IVEF protocol and data format, which facilitates its deployment and extension for other sensors, without requiring extensive training. This component is a map-centric Graphical User Interface (GUI) component displaying multi-scale maps, S57 nautical charts, as well as topological information, radar cells and targets, with level of confidence clues. Moreover, the AUI provides the capability to display the alarms defined by the Rule Engine (EWE). Figure 7 illustrates a screenshot of the AUI where different tracks of the radars (OTH, PE-MIMO, legacy, fused) and AIS are displayed in different colors, including information of tracks and data sources (left side), as well as the alert panel which includes information about the early warnings and alerts.

Figure 7: AUI Visual Interface and Alert Display

4.8 CISE Translation Gateway

Currently national authorities in Europe, responsible for different aspects of surveillance, (e.g. border control, safety and security, fisheries control, customs, defense) collect data separately and often do not share them. A Common Information Sharing Environment (CISE) is being developed jointly by the European Commission and EU/EEA members with the support of relevant agencies[15]. This environment will integrate existing surveillance systems and networks making different systems interoperable and giving to all the authorities concerned, the access to the information they need for their missions at sea. CISE is a key innovation of the European maritime governance and a key element of the European Digital Agenda; it represents a pillar of the European Action Plan for the European Maritime Security Strategy[16] and supports the development of the Blue Economy of the European Union.

The objective of CISE is to improve maritime awareness by enhancing the maritime public authorities' abilities to monitor, detect, identify and track in order to better understand occurrences at sea, finding reasoned grounds for reaction measures on the basis of combining new information with existing knowledge. In this context, it is very important to ensure that RANGER platform is compliant with the CISE framework, in order to share the data acquired by OTH and PE-MIMO radar with all the authorities authorized to access them.

The CISE translation Gateway includes all necessary interfaces between RANGER services and the CISE framework: a set of services provided by RANGER are made compliant with CISE, including the individual sensor data services (OTH and PE-MIMO services), the Fused Data service and the Alerts services generated by the Early Warning Engine (EWE). The specific gateway is an external aggregator of interfaces to the RANGER platform to facilitate interconnection with CISE nodes. The CISE translation Gateway allows the RANGER internal data model to "speak the same language" of the CISE Data Model, adapting the internal transport protocol to the CISE Service Model.
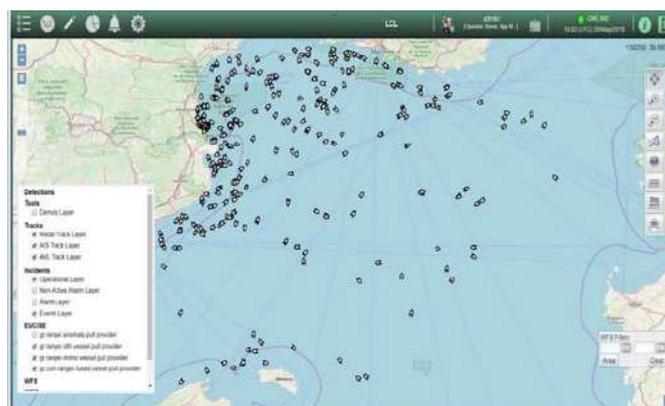
Figure 8: RANGER vessels on CISE Italian Node LCL

Figure 8 shows the Light Client (LCL is a CISE advanced service) of the CISE Italian Node displaying on map the vessels received from the RANGER services as demonstrated during the second French Pilot in May 2019.

Based on the implementation of the RANGER translation gateway, RANGER is not only compatible with CISE, but it has been also fully integrated with the operational CISE network. In particular the RANGER platform is now configured in the CISE network as a participant of the Greek CISE node and its services are available to all the nodes and legacy systems connected to CISE.

5. RANGER COMPLIANCE FRAMEWORK AND ETHICS

RANGER needs to take into consideration legal, ethical and societal dimensions, since only ethically sustainable solutions are viable in the long run. Ethics as part of this compliance framework can serve as a practical angle for engaging the ways in which security is produced. However, ethics should be seen as a way of putting critiques to work, not just as a legitimizing function of "ethics approval"[17] .With such critical ethics approach we are seeking to maximize the benefits of RANGER and to prevent and minimize its ethical risks. Both desk-top studies and workshops have been used in the data collection and analysis of this ethics work, including also Societal Impact Assessment (SIA).

Fundamental European values and norms expressed in international and European law as well as various ethical questions raised by the end-users and other stakeholders establish the value base of the RANGER ethical requirements and the RANGER Code of Conduct. The principles touch the totality of ethical and societal considerations of the RANGER solutions: the technology, how this technology will be used, as well as the whole RANGER business model/procurement as part of the European Maritime Surveillance ecosystem.

The biggest ethical challenge concerns the use of RANGER in the border control. This is related to the tensions between humanitarianism and security. On the other hand, this is not only a challenge for RANGER, but for the whole EU maritime surveillance policy and practices which - according to several scholars - are more focused on safety, technology and security businesses than on human rights and saving lives.

RANGER's impact on the wildlife and humans is a third ethical issue which emerged especially during the initial societal impact assessment workshops organized as part of this deliverable work. Regardless of whether the impact on wildlife and humans is real or only a fear, it is ethically and societally important issue to take into account. How to tackle these challenges concerns both the design of the RANGER technology, the location and installation of the radars, as well as the use of the technology in various maritime surveillance activities.

Data and information poses another set of ethical questions. Privacy and protecting personal data is always a concern, although, the current radar technology cannot capture sensitive or personal information. However, since the RANGER data combined with other data can violate privacy and personal data protection, the adaptation the Privacy by Design/Default, as well as proper data security architecture and data management are all essential. Also relevant are the right way of utilizing the data, and the avoidance of leakage and misuse of that data.

6. CONCLUSION

In this paper, we presented a high-level view of RANGER, a novel platform that combines innovative radar technologies with cutting edge technological solutions for early warning, to deliver a surveillance platform offering detection, recognition, identification and tracking of suspicious vessels and contributing to improved Search and Rescue operations. The RANGER system as a whole has been already fully implemented, integrated, tested and validated in 2 pilot demonstrations in France, while another two pilot demonstrations in Greece are planned. The preliminary results such as the detection capabilities of the OTH radar, the time-on-target of the data fusion module, and the vessel length classification performance indicator of the Machine Learning module, as well as the RANGER system as a whole are very promising. Some of the most challenging aspects in maritime surveillance have been addressed; continuous long range detection and tracking, short range detection and tracking of small vessels, improved situational awareness and early warnings provision through novel data fusion and machine learning techniques, interoperability, seamless integration with legacy systems and compatibility with CISE, advanced visualization of the

maritime picture for improved operations and decision making. All these results which will be complemented until the realization of the last two pilot demonstrations and the final system validation create new opportunities for EU maritime surveillance systems and operators. The RANGER system is currently being developed in the context of the European project RANGER and is validated in 4 pilot demonstrations in Mediterranean Sea (France and Greece), in real operational environment.

ACKNOWLEDGMENTS

REFERENCES

[1] H. J. Roarty, E. Rivera Lemus, E. Handel, S. M. Glenn, D. E. Barrick, and J. Isaacson, "Performance Evaluation of SeaSonde High-Frequency Radar for Vessel Detection," *Marine Technology Society Journal*, vol. 45, no. 3, pp. 14–24, May 2011.

[2] C. Hernandez, "Using HFSW radar into 21-century mexico's navy operations," *AMSC5*.

[3] L. Petrillo, F. Jangal, M. Darces, J.-L. Montmagnon, and M. Hélier, "Towards a Better Excitation of the Surface Wave," *Progress In Electromagnetics Research M*, vol. 13, pp. 17–28, 2010.

[4] N. Bourey, F. Jangal, M. Darces, and M. Hélier, "Enhancing field strength in HF propagation by using a transition between a metamaterial and the sea," in *2013 7th European Conference on Antennas and Propagation (EuCAP)*, 2013, pp. 2680–2684.

[5] "Ranger Project," *Ranger*. [Online]. Available: http://ranger-project.eu/. [Accessed: 05-Oct-2017].

[6] "Inter-VTS Exchange Format Service V-145," *IALA AISM*. [Online]. Available: https://www.iala-aism.org/product/inter-vts-exchange-format-service-145/.

[7] "V-128 Operational and Technical Performance Requirement for VTS Equipment.pdf." .

[8] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, "Maritime Surveillance Using Multiple High-Frequency Surface-Wave Radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 5056–5071, Aug. 2014.

[9] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A Scalable Algorithm for Tracking an Unknown Number of Targets Using Multiple Sensors," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3478–3493, Jul. 2017.

[10] F. Meyer, P. Braca, F. Hlawatsch, M. Micheli, and K. D. LePage, "Scalable Adaptive Multitarget Tracking Using Multiple Sensors," in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.

[11] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Computational Intelligence and Neuroscience*, 2018. [Online]. Available: https://www.hindawi.com/journals/cin/2018/7068349/. [Accessed: 10-Jun-2019].

[12] E. Protopapadakis, A. Voulodimos, A. Doulamis, N. Doulamis, D. Dres, and M. Bimpas, "Stacked Autoencoders for Outlier Detection in Over-The-Horizon Radar Signals," *Computational Intelligent and Neuroscience*, 2017.

[13] E. Protopapadakis, A. Voulodimos, A. Doulamis, S. Camarinopoulos, N. Doulamis, and G. Miaoulis, "Dance Pose Identification from Motion Capture Data: A Comparison of Classifiers," *Technologies*, vol. 6, no. 1, p. 31, Mar. 2018.

[14] "Home," *Copernicus*. [Online]. Available: http://marine.copernicus.eu/.

[15] "Integrated maritime surveillance," *Maritime Affairs - European Commission*, 28-Sep-2016. [Online]. Available: https://ec.europa.eu/maritimeaffairs/policy/integrated_maritime_surveillance_en.

[16] "Maritime Security Strategy," *Maritime Affairs - European Commission*, 28-Sep-2016. [Online]. Available: https://ec.europa.eu/maritimeaffairs/policy/maritime-security_en.

[17] M. Leese, K. Lidén, and B. Nikolova, "Putting critique to work: Ethics in EU security research," *Security Dialogue*, vol. 50, no. 1, p. 096701061880955.

Fifteen years of HF-Surface-Wave radar for maritime surveillance

Widera, J.*[a], Helzel, T*[a]., Kniephoff, M. *[a], Petersen, L. *[a]

[a]Helzel Messtechnik GmbH, Carl-Benz-Strasse 9, 24568 Kaltenkirchen, Germany

Abstract

The Helzel Over the Horizon radar (OTHR) is an FMCW HF radar for monitoring targets within the exclusive economic zone (EEZ). It has many advantages compared to classical pulsed HF radars like a much lower rf-noise emission, no blind range and less in-band noise. It is originating from the WERA (WEllen RAdar) HF radar for oceanographic applications. We describe the development from the HF radar WERA to the maritime surveillance radar OTHR and basic principles and technologies for this type of HF radar to use it as an instrument in a network of multi sensor trackers. It is an example of technology transfer and ongoing knowledge exchange between scientists, engineers and maritime stakeholders.

Keywords: OTHR, HF, surface wave radar, surveillance, ship detection, WERA

INTRODUCTION

Many different surveillance instruments are monitoring the maritime exclusive economic zone (EEZ). These have all their advantages and limitations. Two main instruments are shore-based radars and satellites. Microwave radars are constantly scanning the area around them with a low data latency but with a limited line of sight and should be installed on a high position. Satellites, on the other hand, are monitoring very large areas but with a high data delay and gaps in time and coverage.

An instrument that is monitoring large areas with a low data latency is the High Frequency radar (HF radar). First HF radars for ship detection work similar to microwave radars in a pulsed mode and their lower frequency allows a longer range but the lower frequency goes hand-in-hand with bandwidth- and noise-problems. Operators of microwave radars get much easier bandwidths than operators of HF radars because the scale for the same bandwidth between both techniques is 1:1000. This means HF operators often need several percent of their frequency as bandwidth while microwave operators are often below one percent. Another disadvantage of pulsed microwave radars is the blind range for the first kilometers resulting from the switching between sending and receiving.

For observation of currents and waves oceanographers are using HF radars with frequency modulated continuous waves (FMCW) instead of pulsed signals for observation of currents and waves to overcome these problems. These specific HF radars are sensitive at small bandwidths and do not have a blind range, albeit there are oceanographic HF radars using FMiCW that still have this blind range. In 1955, Crombie found specific signals from the ocean's surface when radiating HF waves [1] which were related to the ocean currents. Later, Hasselmann linked these and other specific signals and to ocean waves [2] in 1971. These specific signals are very hard to detect and the hardware has to have a very low noise floor and a long coherent measurement time to extract these hidden information from the ocean. This is possible with FMCW radars that can send and receive at the same time with a low output power that reduces unnecessary noise. Nowadays, HF radars like WERA are the state of the art for observing currents and waves in large areas (120 degrees of azimuth, more than 200km range). The technical development over the last decades lead to a high reliability of this technology [3].

During one of these technical development steps, where Gurgel tried to remove the noise from ships [4], first experiments with the FMCW HF radar WERA for ship detection were made. Over the years the techniques for detecting

ships improved. In 2011 Helzel manufactured HF radar for ship detection hardware for OEM partners and in 2017 launched it's own radar for ship detection, named "Over the Horizon Radar". This process will be shown in the next section. In a second section we will describe basic principles for ship detection.

HISTORY OF OTHR

OTHR is a successor of WERA that is used for observing oceanographic parameters. WERA had been developed at the University of Hamburg in the 90's and was later industrialized by Helzel Messtechnik GmbH. Experiments to reduce the noise created by ships led to the idea to spin-off an explicit HF radar for monitoring vessel movements that was later named Over The Horizon Radar (OTHR). Improvements at this radar were made as well as a multi sensor tracker developed that form the basics of the nowadays OTHR.

Table 1: Summary of the development of WERA/OTHR

Date	Development step	Published
1995	First experiments with new low noise radar concept	Yes [5]
1999	First industrialized WERA systems	Yes [6]
2004	First ship detection experiments with WERA by University of Hamburg	Yes [4]
2006	Raytheon used WERA at Germany Bight for ship tracking tests	Classified
2007	Start of Ship detection and tracking software development at HELZEL in co-operation with EADS and University of Hamburg and TU Hamburg-Harburg	Yes [7,8]
2008	French researchers used WERA for ship detection evaluations	Yes [9]
2009	NATO Research center (NURC) starts evaluation with WERA and develops multi sensor tracker	Yes[10, 11]
2010	WERA for ship tracking is registered as Dual-use instrument at German Export Control Authority (BAFA)	-
2011	First OEM contract with another radar manufacturer to use the HF radar core for their product.	No
2011	First export of HF radar system for ship tracking	No
2011	NATO Research (CMRE, successor of NURC) started WERA ship tracking evaluation in German Bight	Yes [12,13]
2013-2014	Ship detection and tracking experiments with 5.3 MHz WERA in the polar region by the Norwegian Navy (FFI)	Classified
2016	Three HF radar systems for ship tracking operational	Classified
2017	Spin-off of OTHR as a specialized HF radar	No
2018	Multi sensor tracker experiments in the Gulf of Guinea	Yes [14]

A group around Essen and Gurgel at the University of Hamburg developed the HF-WERA system in the 90's. They started in the 80's with a commercial HF radar from NOAA[16] for measurements of ocean currents in the German Bight and steadily improved the radar's performance[17]. At some point, the existing hardware wasn't able to fulfill the requirements of these scientists which lead to a complete redesign of the HF radar: A saw-tooth-like frequency modulated continuous wave (FMCW) replaced the pulsed signal to remove the blind range in the first few kilometers, improved the range-resolution by a factor of ten and increased the sensitivity of the signal to noise ratio (SNR). The antenna-design changed from four antenna square-array to a linear array of 16 antennas allowing using beam-forming algorithms that lead to a significant improvement of azimuthal resolution [5,6] and improved results in heavy ship traffic [18]. For a detailed review about current status of HF radars for oceanographic applications see [19].

This experimental radar was then industrialized by Helzel Messtechnik GmbH in 1999.

Experiments with WERA to reduce noise during wave measurements lead to the byproduct of ship detection. Ships can heavily influence the performance for getting wave and current information. They are creating a static signal on top of a Gaussian noise. These static signals – or ships – can be identified by the Constant False Alarm Rate (CFAR) algorithm [20]. Identifying, allows removing ships from the spectra for better results of wave parameters [4]. This method was applied at 2004 at the University of Hamburg and can be seen as the first ship tracking experiment with a WERA HF radar and the beginning of OTHR.

Between 2005 and 2008 a network of different stakeholders established a monitoring system in the German Bight. This North-Sea Monitoring Project (OMS) was supported by the North-German state Schleswig-Holstein. During this project, several HF radars have been used to monitor this area.

In a cooperation with EADS and ATLAS the ship detection algorithm had been improved and a probabilistic tracker was implemented [7,8] around 2007. To improve the false detection rate, physical parameters, like the range power dependency have been implemented and neighboring grid cells taken into account to create an adaptive power threshold that had a much better performance than the first simple implementation.
These improved detects were used in the first ship tracking algorithm again. This new algorithm was using a Kalman like Alpha-Beta tracker to generate tracking states which consists of position and velocity for possible targets. First comparisons with AIS showed a difference between AIS and radar positions of less than 1 km for 77% for all targets in AIS range.
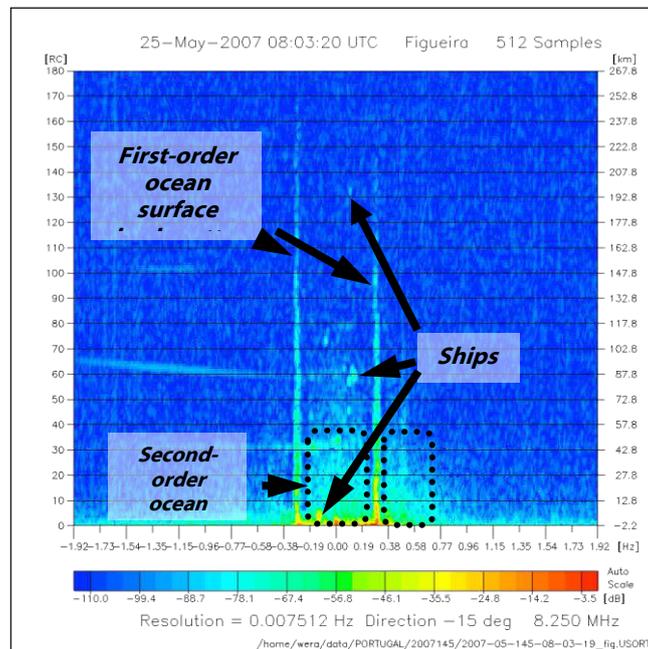
Fig. 1: Typical Range-Doppler map

From 2008 on, French researches started to use WERA for ship detection [9]. They applied the Morphological Component Analysis (MCA) techniques to detect ships. MCA is a technique that allows to extract targets from a spectrum, depending on their geometrical appearance. This algorithm works directly on the Range-Doppler-Map (fig. 1). The research showed that this algorithm has an improved performance compared with CFAR when the background has a lot of noise.

In 2009 the NATO Undersea Research Center (NURC) started as to investigate WERA radars. Their first project was to analyze the dependency between range and signal amplitude of targets, where they found a good fit to a Raleigh distribution and explained discrepancies between distribution and real data with noise sources 10]. In a second project a

first kind of multi sensor tracker was established by combining the results of two overlapping WERA systems[11] and comparing it with AIS data.

In the same year, WERA was used especially for Ship Tracking the first time [21].

The Center for Maritime Research and Experimentation (CMRE), successor of NURC, continued to work with data from the German Bight and with two WERA stations at the Ligurian Coast (West Italy) [12,13]. They analyzed the performance of data fusion and two different tracker algorithms for multi target tracking.
Performance of data fusion was measured by comparing single WERA with the combination of two WERA. Here they found a significant improvement of about $\sqrt{2}$ for the accuracy, False Alarm Rate and Time on Target. The two different tracker algorithms were implemented and compared to the CFAR detection algorithm. The tracker outscored the CFAR detection method in nearly all scenarios and especially the False Alarm Rate by a factor of ten.

In 2017, the gained experiences and scientific research in ship tracking with WERA that took place for over 13 years lead to the decision to spin-off a specialized HF radar from the WERA. This was called Over The Horizon Radar (OTHR). It is a specialized version of an HF radar that is focusing on the techniques to detect ships with specific hardware and software components.

Field experiments from an OTHR-like system show the automatic combination of two HF radars with satellite- and land-derived AIS data [14]. With this technique the software combines HF radar with AIS and marks the ships with the information from AIS[15]. When the ship leaves the coverage of AIS it will be still identifiable with HF radar information. This comes in handy for automated monitoring of the EEZ.

BASIC PRINCIPLES

PHYSICAL Aspects:

Several physical aspects influence the performance of the OTHR. The signal generator creates the signal which is amplified and sent to the transmit antennas. This transmitted signal hits a target, reflects and travels to the receive antennas and then to the receivers where it will be sampled for analyzation.

A simple method to increase the detection range can be achieved by increasing the output power but this has a few drawbacks. The OTHR generates the main signal and amplifies this with the low-noise power amplifier. This amplification results in a power output up of 0.1kW to 1 kW which is quite small compared to pulsed radars that are using several kW peak power. Increasing the output power increases the range but the effect is quite costly: The output power scales to the fourth power of the range [fig. 2]. Another costly factor is the distance between the transmit and receive antennas. Since the OTHR is sending and receiving at the same time, the distance between sending and receiving
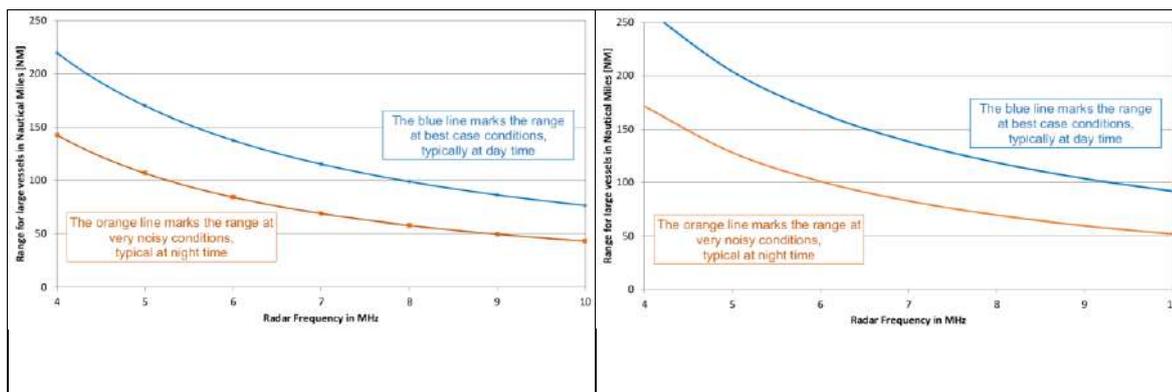
Fig. 2: Typical ranges for detection large ships for 100 (left) and 1000 Watt (right).

antennas has to be increased when power is increasing, otherwise the receiver will drive into saturation by the direct signals from the transmit antennas to the receive antennas.

The design of the transmit and receive antennas has a significant influence on the performance of the radar. A well-tuned antenna-array transmits and receives the electromagnetic waves very good. The receive array consists of a linear array of 16 monopole antennas with a length less than $\lambda/4$ and a distance of $\lambda/2$ that are individually tuned to fit to the environment. Current development allows the use of active broadband antennas that are much smaller in size and do not have to be tuned. The transmit array consists of four $\lambda/4$ monopole antennas that are ordered rectangulary which allows to transmit most of the energy to the ocean and provides a null in the direction towards the receive array.

This specific set-up of receive antennas allows the usage of beamforming algorithms that is the most accurate range-azimuth-resolver algorithm for ship detection at the moment. Another popular set-up of receive antenna array is using much fewer antennas and is using the direction finding method instead of the beamforming method to resolve the location of possible targets. The drawback of this direction finding method compared to beamforming method is a decreased performance in observing oceanographic parameters in the presence of ships [18].

The transmit antenna is radiating a vertical polarized electromagnetic wave. Electromagnetic waves normally propagate in a straight line resulting in shadow zone behind the horizon (space and sky waves) but due to the salt the conductivity in the ocean is much higher than in the air resulting in a wave that is following the curvature of the earth (ground wave) and seeing behind or Over The Horizon.

The frequency of the electromagnetic wave also influences the performance – longer waves have a longer range while shorter waves have a shorter range. HF is defined for wavelengths between 10 (30 MHz) and 100 meters (3 MHz), which results in theoretical ranges between 30 and 500km.

Another frequency related performance impact is the size and orientation of the target. Since the electromagnetic waves of an OTHR are vertically polarized the target has to have some vertical structure that is conductive. The beam then does only see the normal projection (silhouette) of the ship, the radar cross section. This radar cross section heavily depends on the orientation and vertical size of the ship regarding to the radar beam. Apart of simulations there is no deterministic way to guess the radar cross section of a ship. This means that two identical ships in the same distance but with different courses can result in only one detection as well as different sea states change the effective vertical size.

Another important parameter is the radar bandwidth. It is the bandwidth in the frequency allocation, that directly affects the range resolution as well as the noise. Increasing the bandwidth allows the radar to decrease the range-cell depth. The noise automatically decreases for a range cell when its size decreases.

NOISE, Jamming and countermeasures

The OTHR's performance is mainly limited by noise. Sources of noise are the the system's noise figure and external noise. The system's noise figure is already below the environmental noise by design so we only care about external noise. Various techniques exist to identify, remove and avoid same external noise.

The OTHR can detect and differentiate between targets and noise simultaneously [22]. This is possible due to the specific frequency modulation the radar is using. This Radio Frequency Interference (RFI) reduction is the basic technique to remove external noise.

Additional, the OTHR can identify and avoid bands with strong external noise by means of the listen-before-talk method [23]: The radar is scanning the environment before a measurement in the allocated frequency band. Then, the radar may select a reduced bandwidth in this band that has the lowest noise level. This method is especially useful, when noise sources are changing slowly in the order of several minutes.

If the noise source is changing rapidly, eg. due to electronic countermeasure, both mentioned methods to reduce the noise aren't sufficient anymore. Then, frequency hopping can be activated. In this case, the radar constantly scans the environment and changes it's frequency bandwidth on demand.

External noise sources can be located by using the azimuthal resolution of the radar. If two radars can identify the source, then the location can be derived by triangulation.

MULTI Sensor Tracking

Multi sensor tracking is the combination of several sensors of same or different type. This combination leads to significant improvements in the accuracy and the reliability of the tracker.

From a statistical point of view, the detection of a position from one single source for one target follows a Gaussian distribution. Combining two or more of these detects will improve the accuracy square-root like [24].

It is always preferable not to rely on one system or technique because if one sensor fails the remaining ones still operate – so, from a technical point of view, multiple sensors create redundancy.

Different types of sensors can create an additional level of protection against jamming. If the sensor network is using different frequencies and modulations techniques it is harder to interfere them all simultaneously. If combining different techniques, e.g. radar with optical satellite measurements the level of protection is growing constantly.

Fig. 3: Example of a multi sensor tracker result with OTHR and AIS data. Purple lines are OTHR results and blue lines from AIS. With courtesy of Vlatacom, Serbia [14].

Nowadays, it is standard to combine different radar sources with AIS information in a multi sensor tracker to monitor the exclusive economic zone (fig. 3) [14]. Here, techniques like the suspicious target detection and the seamless combination of AIS and HF data helps operators to classify ships [15].

ENHANCED Antenna Design

The main characteristics of a radar, like the pattern and sensitivity, can be influenced by the antenna layout. The number of antennas with constant spacing defines the effective area and the width of the main lobe, that means the more the better. For an antenna array length of some 100m for 16 antennas, space is often the limiting factor. Modern techniques like MIMO can reduce the required array length by keeping a small footprint. Active antennas are less sensitive than passive antennas but smaller in size which reduces the visual impact (fig. 4) and got other advantages as described below.

Fig. 4: Two different kinds of antennas. Left: active receive antenna with low visual impact, right: passive transmit antenna.

Multiple Input Multiple Output (MIMO) is a common technique in wireless communication systems where multiple transmit antennas and multiple receive antennas are combined. Adding a second transmit array in a

defined distance extends the receive array virtually by this distance. This allows to set up a real receive array of eight antennas that results in eight virtual antennas, summing up to 16 antennas [25] (In practice you build up nines antennas to have one overlapping antenna between real and imaginary antennas).

Another way to enhance the radar's flexibility is to use Dual Frequency, where two different transmit arrays with two different frequencies and one receive array with active broadband antennas can be used to use two frequencies alternating [26]. This method will result in two different radar systems in one: One system with the lower frequency has a better range then the system with the lower frequency, while the system with the higher frequency can probability detect smaller ships than the system with the lower frequency.

Summary

The development of FMCW HF radar for oceanographic observations lead to the usage of this technology for the maritime surveillance.

Many improvements have been made and these from the oceanographic system (WERA) influenced the ship detection (OTHR) and vice versa. It is a good example for knowledge transfer between scientists and engineers.

OTHR is best used as a part of a multi sensor tracker but can also be used solo since it covers the range of the EEZ in nearly real-time without a blind range.

REFERENCES
[1] Crombie D. D. "Doppler spectrum of sea echo at 13.56 Mc/s.", Nature, 175, 681-682, (1955).
[2] Hasselmann, K. "Determination of ocean wave spectra from Doppler radio return from the sea surface, ", Nature PhysicalScience 229(1971)
[3] Helzel, T., Holton, L., Petersen, L. and Rautenbach C. "Monitoring of the Coastal Zones and the EEZ by Ocean Radar A review of the potential of this technology. ", ICOE2019 5th International Conference in Ocean Engineering, At Le Meridien Village Hall Lane, Pointe aux Piments, Mauritius, (2019)
[4] Gurgel, K.-W. and Schlick, T., " HF radar Wave Measurements in the Presence of Ship Echoes - Problems and Solutions. ",Oceans 2005 Europe, (2005).
[5] Gurgel, K-W., and G. Antonischki. "Remote sensing of surface currents and waves by the HF radar WERA." Seventh

IEE Conference on Electronic Engineering in Oceanography, 211-217 (1997).

[6] Gurgel, K.-W., Antonischki, G., Essen, H. H. and Schlick, T. (1999). "Wellen Radar (WERA): a new ground-wave HF radar for ocean remote sensing." Coastal engineering, 37(3-4), 219-234 (1999).

[7] Dzvonkovskaya and H. Rohling, "Target Detection with Adaptive Power Regression Thresholding for HF radar", CIE International Conference on Radar ICR-2006, 183-186, DOI: 10.1109/ICR.2006.343550 (2006).

[8] Dzvonkovskaya A., Gurgel K.-W., Rohling H., and Schlick T., "Low Power High Frequency Surface Wave Radar Application for Ship Detection and Tracking", Proceedings of the International Conference RADAR-2008, 654-659, (2008).

[9] Grosdidier S., Baussard A., "Ship detection based on morphological component analysis of high-frequency surface wave radar images", IET Radar, Sonar and Navigation, (2012).

[10] Maresca, S., Greco, M., Gini, F., Grasso, R., Coraluppi, S., and Thomas, N., "The HF surface wave radar WERA. part I: Statistical analysis of recorded data", 2010 IEEE Radar Conference (pp. 826-831), IEEE (2010).

[11] S. Maresca, J. Horstmann, R. Grasso, M. Coffin, K. Gurgel and T. Schlick, "Performance assessment of HF radar ship detection," 2011 12th International Radar Symposium (IRS), Leipzig, pp. 131-136. (2011).

[12] Maresca S., Braca P., Horstmann, J., Grasso R., "Maritime surveillance using multiple high-frequency surface-wave radars", IEEE Transactions on Geoscience and Remote Sensing, Vol. 52 (8), 5056 – 5071, (2014).

[13] Braca, P., Maresca, S., Grasso, R., Bryan, K., & Horstmann, J., "Maritime surveillance with multiple over-the-horizon HFSW radars: An overview of recent experimentation.", IEEE Aerospace and Electronic Systems Magazine, 30(12), 4-18 (2015).

[14] Nikolic, D., Stojkovic, N., Popovic, Z., Tosic, N., Lekic, N., Stankovic, Z., & Doncov, N. "Maritime Over the Horizon Sensor Integration: HFSWR Data Fusion Algorithm. ", Remote Sensing, 11(7), 852 (2019).

[15] Zimmermann, R., Gern, T., Sandler, M., "MULTI SENSOR TRACKING OF SMALL TARGETS", WSS2008 – the 1st International Waterside Security Conference and Exhibition, Copenhagen, (2008).

[16] Essen, H.-H., Mittelstaedt, L., Schirmer, F. "On near-shore surface current measurements by means of radar", Deutsche Hydrographische Zeitschrift 36, pp. 1...14, (1981).

[17] Essen H.-H., Gurgel K.-W. and Schirmer F., "Horizontal and temporal variability of surface currents in the Lübecker Bucht as measured by radar", Deutsche Hydrographische Zeitschrift, 41, 57-74, (1991).

[18] Gurgel, K.-W., G. Antonischki and T. Schlick, "A comparison of surface current fields derived by beam forming and direction finding techniques as applied by the HF radar WERA," IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development, Singapore, pp. 1805-1807 vol.4., (1997).

[19] Rubio, A., Mader, J., Corgnati, L., Mantovani, C., Griffa, A., Novellino, A., Quentin, C., Wyatt, L., Schulz-Stellenfleth, J., Horstmann, J., Lorente, P., Zambianchi, E., Hartnett, M., Fernandes, C., Zervakis, V., Gorringe, P., Melet and A., Puillat, I., "HF radar Activity in European Coastal Seas: Next Steps toward a Pan-European HF radar Network. ", Frontiers in Marine Science. 4. 8. 10.3389/fmars.2017.00008, (2017).

[20] Finn, H. M. and Johnson, R. S., "Adaptive detection mode with threshold control as a function of spacially sampled clutter-level estimates; ", RCA Rev., vol. 29, pp. 141-464, (1968).

[21] Gurgel, K.-W., Dzvonkovskaya, A. and Schlick, T. , "HF radar WERA Application for Ship Detection and Tracking 1. ", (2010).

[22] Gurgel, K.-W., "Remarks on Signal Processing in HF Radars Using FMCW Modulation. ", (2009).

[23] Gurgel, K.-W., Barbin, Y., "Suppressing radio frequency Interference in HF radars". Sea Technology. 49. 39-42., (2008).

[24] Braca, P., Marano, A., Matta, V., and Willett P., "Multitarget-multisensor ML and PHD: Some asymptotics, " in Proc. 15th Int. Conf. Inf. FUSION, Singapore, pp. 2347–2353, (2012).

[25] Dzvonkovskaya, A., Merz, C., Liu, Y., Weisberg, R., Helzel, T. and Petersen, L., "Initial surface current measurements on the West Florida shelf using WERA HF ocean radar with multiple input multiple output (MIMO) synthetic aperture. ", 2014 Oceans - St. John's, OCEANS 2014, (2014).

[26] Helzel T., Petersen, L., Widera, J., Bennis, A.-C., Benoit, L. and Barbin, Y. "Dual frequency ocean radar concept to measure ocean currents and waves at the Raz Blanchard." Proceeding of OCEANS'17 MTS/IEEE, Aberdeen, UK, 19-22th June 2017, (2017).

Beyond the horizon: High Frequency Surface Wave Radar

Frédéric Lamole[1,] François Jacquin[1], Gilles Rigal[2]
[1] Embedded Systems Dpt / HF division 5, rue Brindejonc des Moulinais, 31506 Toulouse Cedex 5
[2] Embedded Systems Dpt / HF division   Les Hauts de la Duranne, 370, rue René Descartes, 13857 Aix-en-Provence Cedex 3

ABSTRACT

In this paper, we present a high-level view of the High Frequency Surface Wave Radar developed for the Radars for long distance maritime surveillance and SAR operations (RANGER) project ([1]). The over-the-horizon radar proposes a new system architecture, with new waveforms, new processing methods, and new antenna concepts to provide solutions to meet HFSWR (High Frequency Surface Wave Radar) challenges as increasing the long range distance detection for small vessels with improved accuracy.

Keywords: Over-the-horizon (OTH) Radar; High Frequency (HF) Surface Wave Radar (SWR); maritime surveillance; Radar Cross Section (RCS)

1. INTRODUCTION

The potential offered by HF surface wave radars for maritime surveillance and in particular for the entire Exclusive Economic Zone (EEZ), is a sea area for a state which has rights from its baseline out to 200 nautical miles from the coast defined by the United Nations Convention on the Law of the Sea in 1982[2], has recently been rediscovered, with the overhaul of national strategies related to security and surveillance of the maritime environment. Thus, over the past decade or so, the evolution of the strategic context has led to international instability, primarily affecting fragile states, promoting illicit and violent activities, generating potential risks for societies as: security threats, caused by trafficking (narcotics, psychotropic substances, weapons), illegal transport of migrants, maritime terrorism and piracy, have replaced the military threat. At the same time, the economic activities carried out by sea have increased very significantly, proportionally much higher than the increase in world economic development [3]. This development requires a stronger presence of coastal States to ensure and strengthen safety and security at sea.

One of the main challenges of HF radar industry is to provide a technology able to manage physics specific constraints of HF band. Conception of an efficient HF radar solution requires management of the following physical aspects:

- Propagation loss: It is affected by the sea conductivity (function of salinity and temperature) and land-sea transitions (island on the covered area).
- Sea clutter: it is a noise composed of all parasite echoes generated by sea surface irregularities.
- Ionospheric clutter: it is a noise composed of self-generated interferences via sky wave propagation.
- HF noise: It is a versatile environmental noise in HF band (atmospheric noise, galactic noise, man-made noise). It is seasonal and also has high day/night variation. The HF noise level variation at a given time around the globe is of around 20dB.
- Spectral congestion in HF band.

DIGINEXT's achievements within the RANGER framework take into account its background in the HF radar as well as its existing sites. Nevertheless, the new radar developed (HFSWR type) needed new technical concepts to be refined and to support itself on other technological bricks which go further than this first accomplishment in order to give answer to RANGER's broken objectives.

DIGINEXT's contributions in the field are numerous:

- Develop a system which allows small dimension boats to be detected (and/or of weak Surface Radar Equivalent) within long range. On the French sites, performances must be optimized for maximum efficiency; this is accomplished through numerous transmitting antennas, and a stronger transmission power level.
- Design and develop a deployable system on remote areas and capable of taking into account the interference caused by the islands. This system will introduce limited performances, coherent with the budget of this project, however it will allow evaluating the HFSWR systems' performances surrounded by the Greek archipelago's islands masking boats' radars.
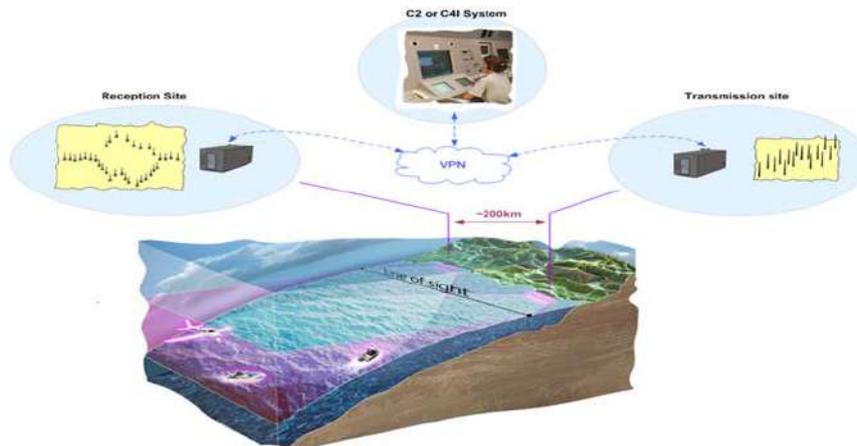
Figure 1: Global view of HF Surface Wave Radar

In order to answer to these objectives, diverse factors will be taken into account:
- Use of a secondary transmission frequency which allows simultaneously:
- To integrate the interference masks due to the reflectivity of the sea and the ionosphere, producing pertinent information for detection.
- To provide complementary information on the Radar cross section of small sized boats.
- An antenna array development allowing to improve performances.

2. SIMULATOR

Before its implementation, the capability of an HF radar needs to be evaluated to validate the potential of installation sites. The HF radar simulator is a software intended to allow the evaluation of the performances of one radar configuration in its environment of exploitation.

2.2 Simulator's specifications

The objective of this software is to allow the user to know in advance and in an intuitive way the parameters of coverage and the level of detection performances concerning a radar installation. The user can then study / modify the various accessible components of the radar in the simulator to obtain a configuration which responds to the established needs according to the following constraints:
- coverage area,
- expected capacity of detection in distance according to the ship sizes,
- available or necessary frequency bands for the operation of the system,
- radiated power available or required for the operation of the system,
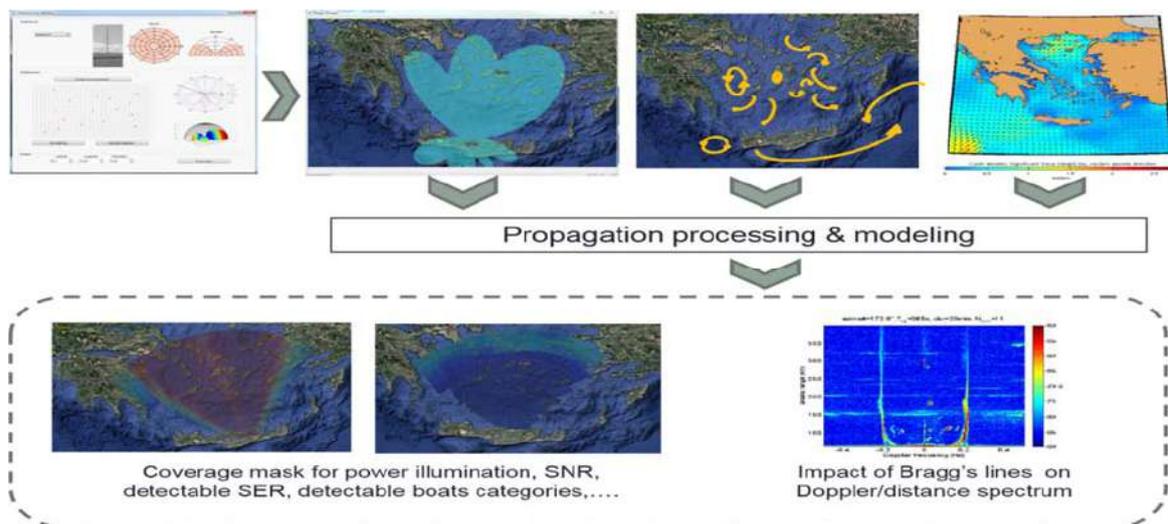
Figure 2: Simulator description

The simulation software (summarizes with the figure 2) produces a set of results that can be evaluated by the user. For every tested configuration the user knows the detection capabilities of the sensor in the area according to:

- the boat categories with the Radar cross section (RCS) and their particularities in HF band,
- the ground conductivity of the sites (transmitting and receiving)
- the sea state and its roughness: The classical model for the roughness of the sea is based on the studies of D. Barrick **Error! Reference source not found.**. This model takes into account the impact of the wind direction versus the propagation.
- the seasons and their specific electromagnetic noises (night and day) in HF band **Error! Reference source not found.**
- the impact of the geography on the covered zone with the natural or artificial obstacles (though the Millington model).

2.2 RCS study

The Radar Cross Section (RCS) is a key parameter of the rada**r** equation. It determines greatly the radar performances. The RCS of an object is specific to each radar frequency especially in the HF band. Knowing RCS of maritime target makes possible to:

- estimate performance for a particular radar configuration
- optimize radar setup for particular kind of target.

With High Frequency Surface Wave (HFSW) Radar it is very difficult to accurately measure RCS for different type of vessel. Moreover it is quite impossible to make measurements in laboratory (anechoic chamber) because of the required dimensions. It is why we choose to estimate RCS by simulation. RCS is calculated with the software "CST Micro Wave Studio EM solver". The transient solver was chosen for the RCS simulations. 3D-model of ship is taken from different kind of sources. Most of the models are downloaded from internet under free user license.
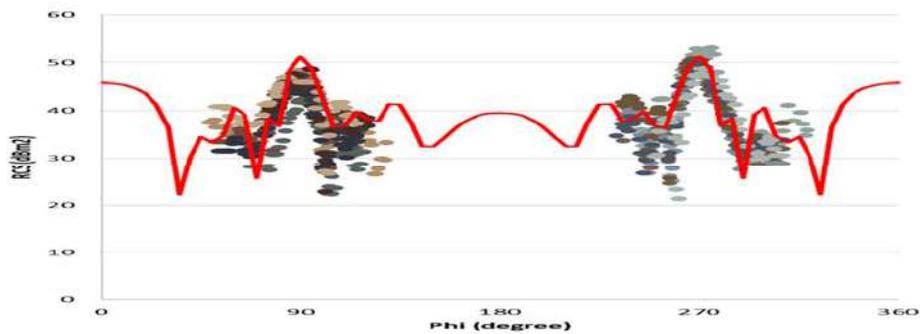
Figure 3: Simulated RCS (red) and received powers

The RCS simulations have been validated with different known trajectories based on their AIS (Automatic Identification System) data transmitted by the vessels. The figure 3 displays the adequacy between the received power by the HF Radar for a ferry (200 m length) and its RCS simulation at 4.5 MHz. It is to note that based on these simulation for several frequencies from 4 MHz to 10 MHz, the RCS vary significantly depending on the chosen one and on the radar angle of view or can be very flat depending on the vessel's dimensions. These results are to be merged with the surface wave's propagation losses for selecting the appropriate frequency.

The figure 4 shows the general view of the simulator interface with the different configuration tools or information zones which are available.
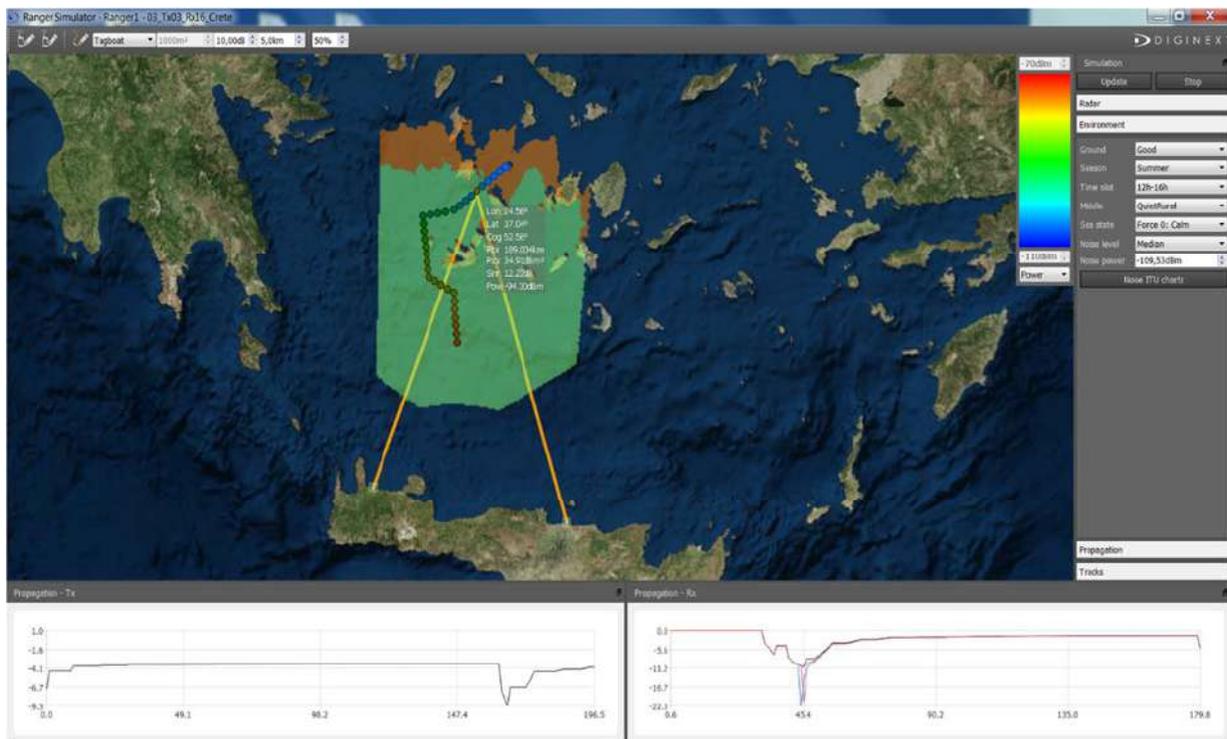
Figure 4: Simulator's general view

The figure 5 shows the Millington modeling (for the sea-land propagation losses) applied over the Aegean Sea for the 2 frequencies used for the RANGER's Pilots.
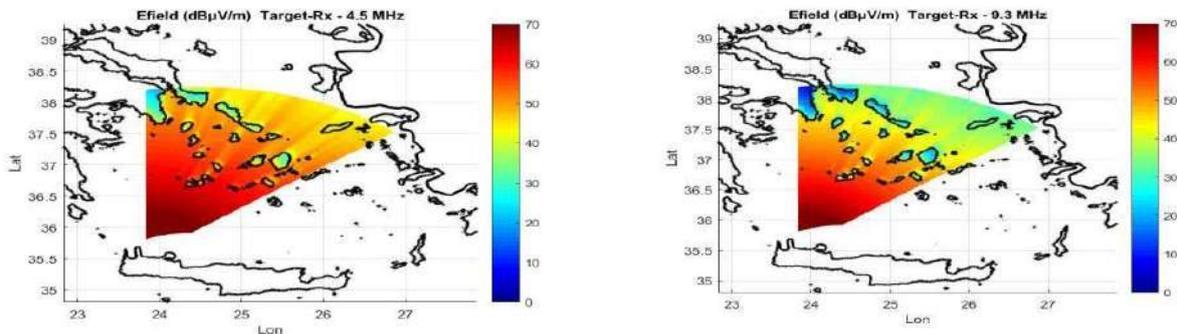
Figure 5: Island's impact

On the left picture is illustrated the E-field from the Rx point of view at 4.5 MHZ and, at 9.3 MHz on the right. For both cases, the impact of islands appears clearly just behind them, but in addition, there is an increase of loss for 9.3MHz frequency (around 10 dB).

3. ENVIRONMENTAL INTEGRATION

High accuracy for radars is synonym of high constraints on technical specifications. The first ones to come to mind are the size of the antenna networks and the power of the transmitters; however, the main one is the spectrum availability. In parallel, environmental rules, like urbanism and electromagnetic exposure, exist without any considerations about radar needs. As a consequence, the design of a new type of radar has to balance between technical constraints and regulations. The approach is structured in 3 domains.

3.1 Spectrum

The list order is pragmatic: Spectrum is the fundamental parameter for the best performance and it defines the radar's parameters to take into account to work safe and to be well integrated in the environment.

Radio waves have been in use for a long time, spectrum occupancy is high and the introduction of a new type of radar needs to take into account existing services and their international priority. The organizations in charge of the spectrum are:

- International Telecommunication Union –Recommendation (ITU-R)   for international level
- European Administrations for regional level (Europe)
- National administration

Practically, it is the national administration who wishes to implant a radar which manages the search for frequency allocation. The European Conference of Postal and Telecommunications Administrations (CEPT) coordinates the frequency assignment between them and can develop rules of sharing at European level. Work starts with its national frequency table, followed by an analysis of regional agreements. All the process is based on the methodology and international agreements of the shared frequencies defined in the ITU-R **Error! Reference source not found.**.The variety of the applications of radiolocation has forced ITU to define technical criteria to help cohabitation in the various wavebands (see **Error! Reference source not found.**).

3.2 Safety at work and around

A minimum requirement for the developed equipment is to respect the European directives on the subjects of conformity and security, which correspond to a civil use of the radars. However, the specificity of the use of RANGER radars could involve a deployment on military bases or platforms which have particular specifications (sensitivity of the weapons to the electromagnetic fields, for example). A thorough analysis must be carried out within the framework of the project to determine the standard levels to follow to achieve the most adapted one for the operational deployment of the radar.

Regarding the phase of prototyping, for the Exposure of workers to electromagnetic fields, the working conditions around HFSWR radar must respect directive **Error! Reference source not found.** on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (electromagnetic fields). System is designed to avoid direct contact current. For electric and magnetic fields, in operation, the safe guard distance for workers is less than 5 meters from each antenna. Therefore worker can access the antenna array keeping the above mentioned restriction.

For the Public exposure to electromagnetic fields, each EU Member State is responsible for providing adequate health protection for its citizens and for deciding on measures to be implemented. However, to support Member States and to ensure the highest level of health protection for European citizens, the EU asks governments to assess exposure to electromagnetic fields and to take appropriate action when reference levels are exceeded – as laid down in COUNCIL RECOMMENDATION of 12 July 1999 on the limitation of exposure of the general public to electromagnetic fields from 0 Hz to 300 GHz **Error! Reference source not found.**.

3.3 Environment

Installation of a new type of radar equipment need a large surface of flat ground being able to reach 1000m of length, and the proximity with the coast for HF radar. The rare places being able to correspond to the preceding criteria are often close to those protected to ensure the safeguard of biodiversity **Error! Reference source not found.**.The two French sites of the OTH radar are installed in activity in Natura 2000 area. They passed the authorization process by limiting the footprint of the antenna arrays, being careful of existing plants and taking into account reproduction period to plan the building of sites.

4. HFSW Radar

The general development of the HFSW radar, during the RANGER project, is the use of 2 simultaneous frequencies and the application of the multiplicative principle for the beamforming using 2 independent receiving antenna arrays.

4.1 French installation

The HF Radar is installed in a bi-static configuration and for the French case on three sites:
- The transmitting site is composed of 16 antennas but divided in 2 sub-arrays for using two frequencies:
  - 10 antennas for the first frequency
  - 6 antennas for the second frequency
- The receiving site is composed of 2 subarrays
  - One main array with up to 32 antennas
  - A second additional sub-arrays with 8 antennas for the multiplicative beamforming principle

The Rx antenna arrays are installed in a salt-march (Natura 2000 areas) in compliance environmental constraints.

Figure 6: Rx antennas in French's installation

- The Control Room site is linked to the other RANGER components such as the Uniform Communication Gateway (UCG) through the OTH (or HF) Gateway.

The figure 9 summarizes this HFSWR configuration for the 2nd RANGER French Pilot.
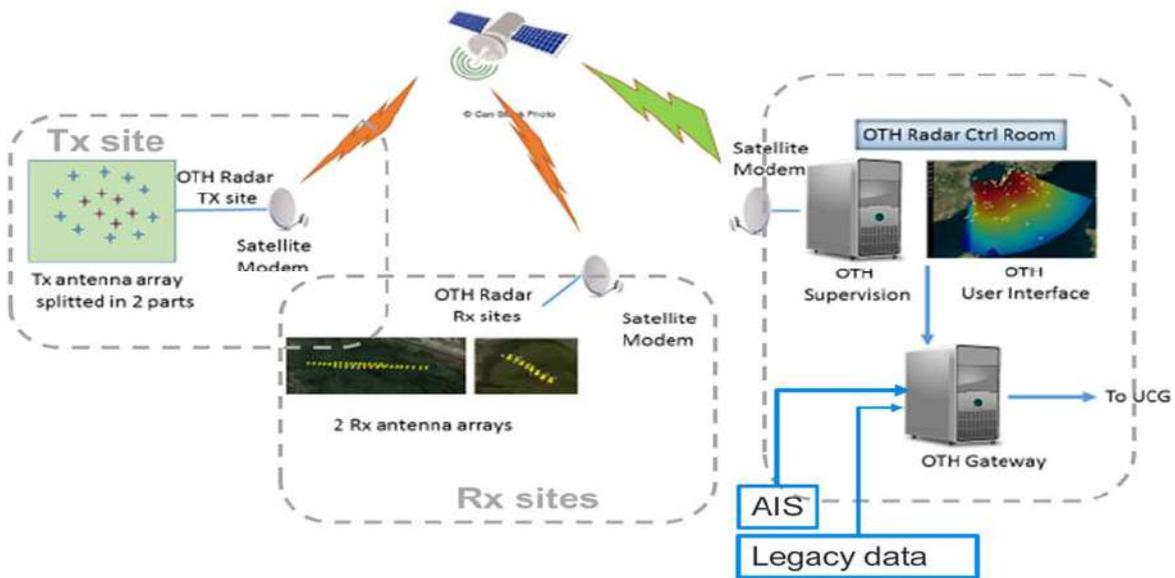
Figure 9: Schematic representation of the French HF Radar configuration

The ionospheric reflection is a self-interferer that can reduce the detection capability in specific areas of the Doppler/Distance map. The design of the antenna arrays (Tx and Rx), combined with a specific process for the antenna's weightings, allow to decrease the impact of the Ionospheric clutter. It is to note that the mitigation is applied on the both sites (Tx and Rx).
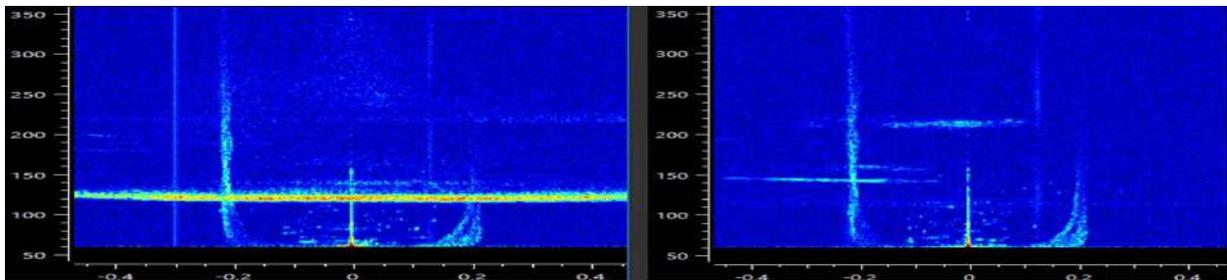
Figure 10: Ionospheric clutter: Before and after mitigation

The OTH detection capabilities during the RANGER's 2nd French Pilot (figure 11), with only a sum of 300W as amplifiers power output (per lobe) splitted on 10 transmitting antennas, are up to 230 km (125 NM) for the smaller vessels (between 10 m and 30 m length) and above 320 km (170 NM) for the bigger vessels (above 30 m length) or beyond for some cases, with a detection ratio of 90%.
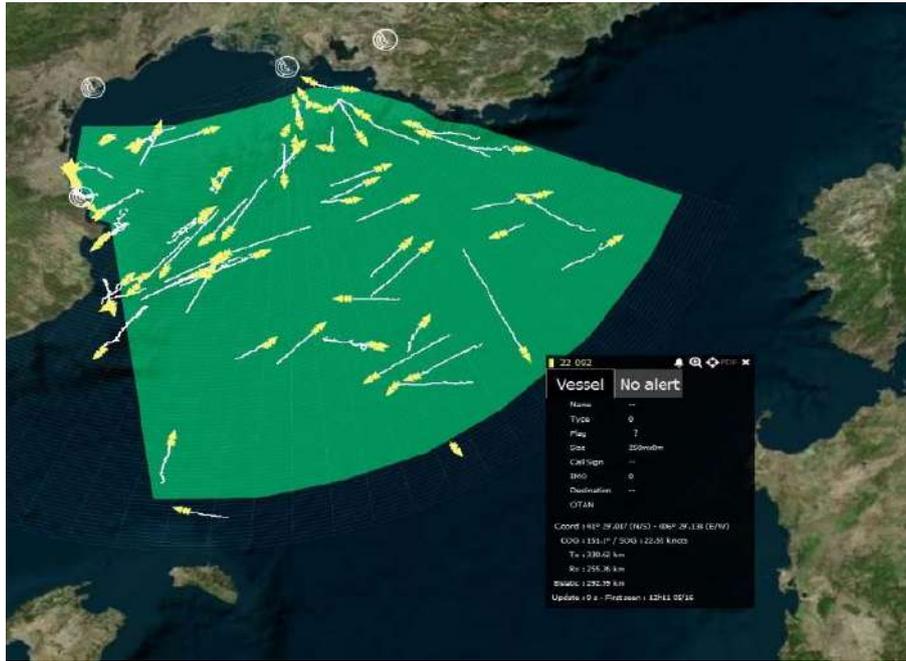
Figure 11: HFSWR Tracks with the 170 NM coverage map

4.2 Greek installation

The Greek HF Radar is also in a bistatic configuration. The crossing of both patterns (Tx and Rx) defines the coverage area of the Radar. The range of the coverage depends on the radar's link budget, that means the signal power amplification and Tx/Rx gains mainly but also the number of antennas on the both sites.

In the case of the Geek RANGER Pilot, the HF configuration is reduced compared to the French installation:

- o 3 antennas for the transmission site
- o 16 active antennas for the reception site

The optimization of antennas positioning as well as their weightings has been realized as in the French case.

Based on these developments, the coverage for these reduced installations is illustrated with the figure 12.
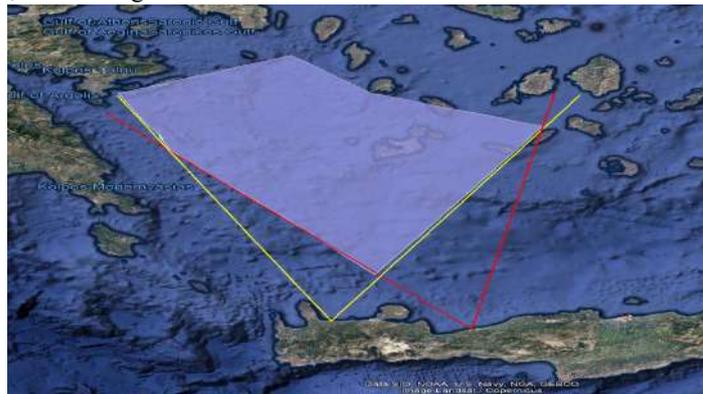
Figure 12: HFSWR coverage for the reduced installation and low power in Greek area

CONCLUSION

In this paper, we presented a high-level view of the Over-The-Horizon radar developed for the long range detection taking into account the interference from islands in the propagation wave. The system is currently being developed in the context of the European project RANGER and is going to be validated in several pilot scenarios in the Mediterranean Sea.

ACKNOWLEDGMENTS

REFERENCES

[1] "Ranger Project," Ranger. [Online]. Available: http://ranger-project.eu/. [Accessed: 05-Oct-2017

[2] "Part V – Exclusive Economic Zone, Article 56". Law of the Sea. United Nations. Retrieved 2011-08-28 https://www.un.org/depts/los/convention_agreements/texts/unclos/part5.htm

[3] United Nations conference on trade and development, review of maritime transport, 2016, https://unctad.org/en/PublicationsLibrary/rmt2016_en.pdf

[4] ITU-R Radio Regulations – ed 2016

[5] ITU-R M1638-1 Characteristics of and protection criteria for sharing studies for radiolocation (except ground based meteorological radars) and aeronautical radionavigation radars operating in the frequency bands between 5 250 and 5 850 MHz

[6] DIRECTIVE 2013/35/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 26 June 2013 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (electromagnetic fields) http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:179:0001:0021:EN:PDF

[7] COUNCIL RECOMMENDATION of 12 July 1999 on the limitation of exposure of the general public to electromagnetic fields (0 Hz to 300 GHz) - 1999/519/EC http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1999:199:0059:0070:EN:PDF

[8] EU-wide Natura 2000 network of protected areas(http://natura2000.eea.europa.eu/#)

[9] D.E.Barrick, Theory of ground-gave propagation across a Rough sea at decameter wavelengths , January 1970

[10] Recommendation ITU-R P372-11 Radio Noise (09/2013)

A MIMO architecture for SaR application

Francesco Prodi*[a], Luigi Pierno [a], Alfonso Farina (Consultant [a])
[a]Leonardo Electronics SpA, via Tiburtina km 12,4 00131 Roma, Italy

ABSTRACT

This paper propose a MIMO architecture proposed for a S-band VTS coastal radar aiming to detect, up to 10 km range, small RCS targets (<1m2) in sea state 0-5, due to rubber boats of illicit trafficking of humans and materials in SaR (Search and Rescue Maritime) Surveillance systems (*RANGER system*). The radar architecture is formed of a filled RX ULA (Uniform Linear Array) with Mr elements, and of a co-located TX ULA with Mt elements separated by $l=M_r*\lambda/2$. This basic configuration is equivalent to a uniform filled ULA with *MrMt* elements and MIMO orthogonality of TX-s waveforms is obtained by Time Domain Multiplexing of the transmitted large band FM chirp between different TX-s. The potential leakage between Doppler domain and azimuth domains put some constrains on MIMO processing. In particular it is shown how a pseudo-random law of TX array scanning may mitigate this leakage. It is then shown that a small boat may be well detected thank to fine range resolution (large bandwidth), narrow beam (large TX array aperture) and fine Doppler processing. Finally it is shown that a bi-static architecture may further improve detection performance due to both clutter cell and bi-static clutter Reflectivity reduction.

Keywords: MIMO, FMCW, TDM-MIMO, FMCW radar, SaR (Search and Rescue), STAP

1. MIMO ARCHITECTURE

MIMO architecture is composed of a filled *RX ULA* (*Uniform Linear Array*) with M_r elements, and of a co-located *TX ULA* with M_t elements separated by $M_r*\lambda/2$ (see Figure 1). Large band FM-CW pulses are sequentially transmitted with a given PRT (Pulse Repetition Time) cyclically on the TX ULA: that is TX1, TX2 up to TX M_t in the first cycle, again TX1, TX2 up to TX M_t in the second cycle, and so on for N_r cycles. This approach implements the MIMO architecture based on Time Domain Multiplexing TDM and the *Time On Target*, that is the processed frame length, is $T_o=M_t N_r*T$, where T is PRT. In standard [1] MIMO architectures, pulses could be made orthogonal to each other and transmitted simultaneously and $T_o= N_r*T$ can be made shorter, keeping the same energy on target wrt the TDM architecture. Let B the instantaneous Bandwidth of TX pulses and τ pulse-width. In order to maximize pulse transmitted energy $\tau\approx$PRT is assumed. Power transmitted PW by each TX is typically few Watts. Sequential transmission of TX-s limit duty-cycle of each TX to $100/M_t$.

Typically we assume the following ranges for system parameters: $M_t \in [5,20]$, $M_r\in [5,20]$ $N_r\in [4,32]$, T$\in [200,1000]$ µsec, and B$\in$ [50-200] MHz. The parameter set is obviously adapted on system specifications (i.e. radar range, target RCS, TX power..) and on hardware resources (typically M_t and M_r).

The general architecture receiver chain, typical of a FM-CW radar, for each of the M_t receivers, is reported in Figure 1. After deramping mixer the received chirp is baseband converted to a complex envelope which is indicatively the sum of beat frequencies corresponding to the range of scatterers. Range profile is then recovered, for each received pulse, by a FFT of slant range samples and also undesired harmonics of the scatterers are present in the spectrum. For each range sample the data structure is structured in the 3D complex hologram H(M_t, M_r, N_r) of Figure 1.
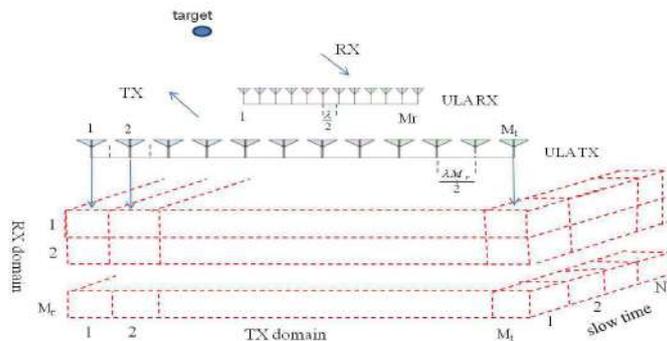
Figure 1 MIMO architecture with data structure

The four input domains together with conjugate domains and relative resolutions obtained after signal processing are reported in Table 1

input domain	output domain	processing goal
fast time	range	range resolution
RX	azimuth	azimuth resolution
TX	azimuth	azimuth resolution
slow time	Doppler	clutter cancellation

Table 1: processing domains

where *fast time* domain is the single sweep of length T sampled at range bin rate while *slow time* domain is the processing interval $T_0 = M_t N_r * T$, with sampling time $M_t T$ in the slow-time algorithm (Par. 2.1) and T in TX staggering algorithm (Par. 2.2).

Range compression along fast time domain and Doppler processing along slow time are conventionally implemented with weighted FFT. Compression of RX-TX allows synthesis of a narrow azimuth beam, and is peculiar of MIMO processing. Also in this case azimuth compression is based on 1D FFTs along RX and TX domains. the present MIMO architecture implies only one 'big' 1D FFT along $M_r M_t$ samples of the virtual ULA elements. In this perspective the peculiar sequence of the three domains - i.e. range-azimuth-Doppler is arbitrary.

We observe that two constrains, make signal processing more complicate than a three cascaded 1D FFTs. They are:

a. range migration compensation
b. leakage between slow time and spatial-TX domains due to TDM transmission operation

Constraint a) implies, for large arrays (indicatively $M_r > 5$) splitting azimuth beam-forming into two separate TX and RX compressions. It will be considered later.

Constraint b) lead to a deviation from the standard MIMO processing based of use on orthogonal waveforms (more generally: linearly independent) transmitted by different TX-s in the same frame time of length T. This fact is now explained is now described.

Referring to the standard (orthogonal waveform simultaneous Tx pulses) based MIMO architecture (i.e. /1/, Chapter 1) with co-located and parallel TX and RX ULAs, signal y_i, (i=1,2..M_r) received by the i-th receiver antenna, and scattered by a point target with complex Reflectivity γ, and azimuth ϑ is:

$$\mathbf{y}_i = \gamma \sum_{j=1}^{M_t} a_{ij}(\theta)\mathbf{x}_j \qquad \text{i=1,2..}M_r \qquad (1)$$

with $x_1, x_2 .... x_{Mt}$ orthogonal signal *contemporary* transmitted by antennas 1-2...M_t, a_{ij} phasors accounting of delay from j-th TX element and i-th RX element.

By projection of y_i (i=1,2..M_r) on the elements x_j (j=1,2.. M_r) we recover the baseband (M_r, M_t) matrix S^{ort} with elements:

$$s_{ij}^{ort} = \mathbf{y}_i \mathbf{x}_k^T = \gamma \sum_{j=1}^{M_t} a_{ij}(\theta)\mathbf{x}_i \mathbf{x}_k^T = \gamma a_{ik}(\theta) \qquad (2)$$

This projection may be implemented in each receiver at IF level by a bank of filters matched on $x_1, x_2 .... x_M$, or at baseband level.

In practice the not perfect orthogonality of signals y_i (i=1,2..M_r) lead leakage between target having different azimuth ϑ. Moreover the use on only one waveform, say x, transmitted cyclically in time division through all M_r antenna *TX-s*, simplify MIMO architecture and at the same time prevent any spatial leakage between TX locations. Naturally this simplification has a draw-back inherent to target radial migration from frame to frame. Supposing uniform radial speed eq. (1) is now:

$$\mathbf{y}_{ij} = \gamma a_{ij}(\vartheta) \exp(-\mathbf{j}2\pi f_d(j-1)T)\mathbf{x} \qquad (3)$$

with $f_d = 2v_r/\lambda$ is target Doppler frequency. Now, baseband signal matrix S has elements:

$$s_{ij} = \mathbf{y}_{ij}\mathbf{x}^T = \gamma a_{ij}(\vartheta)\exp(-\mathbf{j}2\pi f_d(j-1)T) \qquad (4)$$

Now, for filled uniform array the following relation holds:

$$a_{ij} = \exp[-j2\pi\sin(\vartheta)(\frac{i}{M_r}+j)] \qquad \text{i=0,}M_r\text{-1} \qquad \text{j=0,}M_t\text{ - 1} \qquad (5)$$

Combining eq. (4) with eq. (5), elements of s_{ij} may be expressed as:

$$s_{ij} = \gamma \exp[-\mathbf{j}2\pi(\frac{\sin(\vartheta)}{M_r}i + (\sin(\vartheta) + f_d T)j)] \qquad i=0,M_r\text{-}1; j=0,M_t\text{-}1 \qquad (6)$$

The phase term in eq.(6) is a bilinear form in i (RX index) and j (time index). We observe that, while coefficient of i depends only on ϑ -signal reception by different RX antennas is simultaneous-, the coefficient of j depends both on ϑ and f_d, showing leakage between azimuth and Doppler domains typical of *STAP* (*Space-Time Adaptive Processing*) context [2]. We try to break this leakage using two different techniques: *slow time processing* and TX staggering.

Slow time: TX array is scanned cyclically. So phase dependence versus time is:

$$\varphi_j = 2\pi \sin(\vartheta)\operatorname{mod}(j, M_t) + 2\pi f_d T \; j \qquad j=0,1.... \qquad (7)$$

with *mod* modulus operator. Sampling φ_j at $M_t T$, rather than T:

$$\varphi_k = 2\pi \sin(\vartheta) j_0 + 2\pi f_d T \; j_k \qquad j_k=j_0 + (k\text{-}1)^*M_t , \; k=1,2... \qquad (8)$$

phase variation received from a given TX antenna j_0 depends only on target Doppler. So spectral analysis is performed along slow time domain, while azimuth compression along TX-RX 2D domains (see **Figure 1**). The drawback of space and time domains de-leakage is to perform Doppler processing at reduced rate 1/(N*T) rather than 1/T.

TX staggering: if Tx-s antennas are activated with a pseudo-random permutation law, eq. (7) is modified into:

$$\varphi_j = 2\pi \sin(\vartheta)\,perm(\operatorname{mod}(j, M_t)) + 2\pi f_d T \; j \qquad j=1,2.... \qquad (9)$$

where *perm* is a random but known permutation law which is changed from on TX cycle to the next. *i.e*: for M_t=5, in place of the sequence (1,2,3,4,5, 1,2,3,4,5..) is transmitted the sequence *(4,1,3,2,5, 2,5,4,1,3...).* In this way the linear behaviour of target phase versus index j (time) can be somewhere separated from the non-linear behaviour of the permutation law. We note that the same scope could be obtained with a Non Uniform Linear Array. The drawback of this technique, however, is a limitation of the Doppler azimuth scene, as will be shown later.

2. SIGNAL PROCESSING

Slow time and TX staggering processing will be described respectively in 2.1 and 2.2, and represented in **Figure 2**.

2.1 Slow time processing

The different step of processing flow, reported in Figure 2a, are now described in details. Due to TX time switching Doppler processing is done before azimuth (MIMO) processing. Range migration compensation is not necessary for narrow TX ULAs, and allows azimuth compression by 1-D FFT processing along the virtual 1-D array.

step	domain
range compression & spurious mitigation	fast time
doppler processing & equalization	slow-time
range migration compensation	range-domain
azimuth compression (TX-RX)	TX-RX domain
CFAR detection	range-doppler map

a) slow time

step	domain
range compression & spurious mitigation	fast time
range migration compensation	range
doppler compression & TX beamforming	TX-slow time
RX beamforming	RX
CFAR detection	range-doppler map

b) TX staggering

Figure 2 processing flow: a): slow time , b) TX staggering

Range compression and spurious suppression

Spuriouses are originated as Inter-Modulation product of:

- the target beat frequencies with the strong direct path leakage between the Tx and the Rx
- beats among multiple real targets beats frequencies due to large RCS close targets

If not mitigated they may strongly limit radar dynamic along range domain. Spurious cancellation technique is proposed in [3].

Doppler processing

Doppler processing is performed along slow-time domain on the M_rM_t vectors of data hologram (Figure 1). Each N_r length complex vector is first processed with a double MTI (Moving Target Indicator) *(optional)* and a weighted 1-D FFT. Let h_{ijk} be the generic element of the transformed hologram, where i,j and k are respectively *RX, TX* and Doppler filter indexes. Doppler equalized elements p_{ijk} are obtained by:

$$p_{ijk} = h_{ijk} \exp(\mathbf{j}2\pi \frac{k\,j}{M_r N_d}) \qquad i=0,M_r\text{-}1; \qquad j=0,M_t\text{-}1 \qquad k=0,N_d\text{-}1 \qquad (10)$$

where N_d is number of Doppler filters ($N_d > N_r$ with zero-padded FFT-s). In eq. (10) $2\pi k/(M_r N_d)$ represents indeed the opposite of target phase change experienced between two adjacent TX-s and relative to the *k-th* filter. This phase compensation is possible only in the Fourier Transformed hologram *H*, where each Doppler filter corresponds to the peculiar Doppler term to compensate. This fact explain why in the diagram flow of Figure 2a Doppler processing is anterior to MIMO processing.

Range delay compensation

In a MIMO system 3D hologram samples are affected by different range delays because of target radial motion and of no co-located *TX and RX ULAs*. According to a consolidated rule, the larger not-compensated range delay should be less than a fraction ρ of range resolution (i.e. [2]). Target migration constrain is $\Delta_{tar}=2v_rT_o<\rho/2$, while array migration constrain is $\Delta_{TX}\approx l_{TX}\sin(\vartheta)$, where (see Figure 3) v_r is target radial velocity, l_{TX} is TX array length, ϑ is target azimuth with respect to TX-RX array bore-sight, and $T_o=M_t N_r{*}T$ is the processed frame length. As small boat have small velocities ($|v_r|<10$ m/sec is a valid hypothesis for small boats) range delay comp. implementation depends basically from TX array length, indicatively it is necessary for $l_{TX}> 1$ m. While intrinsically a simple operation (shift of complex envelop along range domain) it prevents azimuth beam-forming implementation by 1D FFTs.
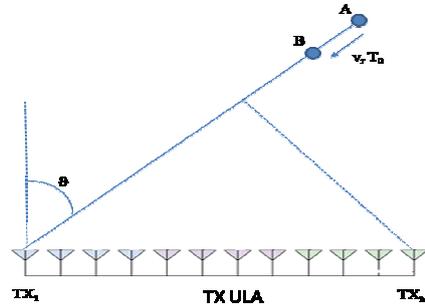
Figure 3 range delay due to target radial motion and antenna spatial distribution

Azimuth beam-forming will be described assuming or less delay compensation.

a) **No range delay compensation**.

If in the hologram P (see eq. (10)), we keep fixed a generic doppler filter k, we obtain a 2D hologram $P_k(i,j)$ $i=1.2..M_r$; $j=1.2..M_t$. Let $\boldsymbol{p}_k$ be the $M_r{*}M_t$ length vector obtained by joining the columns of P_k, starting from the 1st end ending with the Mth. $\boldsymbol{p}_k$ contains the elements of the virtual filled *ULA* with M_tM_r elements (see [1], par. 1.3). Azimuth compressed vector $\boldsymbol{v}_k$ is obtained, for each k, by:

$$\mathbf{v}_k = fft\,(\mathbf{w}^T\mathbf{p}_k, n_a) \qquad k=1,2..N_d \qquad (11)$$

with $\boldsymbol{w}$ weight vector and, using *MatLab* language, *fft($\boldsymbol{a},n_a$)* indicates *Fast Fourier Transform* applied to vector $\boldsymbol{a}$ and zero-padded to n_a ($n_a \geq M_tM_r$). Corresponding azimuth angles (not dependent on k) are given by:

$$\vartheta_j = \arccos(1 - \frac{2j}{n_a}) - \frac{\pi}{2} \qquad\qquad j=0,1,2\ldots n_a\text{-}1 \qquad\qquad (12)$$

b) **Range delay compensation**.

Referring to eq. (12) with formal substitution of n_b in place of n_a, the set $sin(\vartheta_j)$, $j=0,1,2\ldots n_b\text{-}1$, with is represented on a uniform grid of [-1,1]. It can be easily verified that if *TX* array is *range compensated* on ϑ_j, the maximum uncompensated delay over the *TX* array length l_{TX}, within the sector $(\vartheta_{j-1}, \vartheta_{j+1})$, is $2*l_{TX}/n_b$. We say that the TX array is range compensated on the generic angle *of* ϑ when the range profiles indexed by i (RX index), j (TX index), k (Doppler index) are range shifted by $(j\text{-}1)l \sin(\vartheta)$, $j=1,2\ldots M_r$ and $l=M_r*\lambda/2$ adjacent TX element separation (see Figure 3). Then the azimuth-doppler map is reconstructed in the current angular sector, using for instance DFT evaluated to a reduced number of bins of the frequency (in this case azimuth) domain. This two-step compensation-compression procedure is repeated for each angular sector. Minimal computation charge is achieved by minimization of n_b, given the constrain:

$$2\frac{l_{TX}}{n_b} < \frac{\rho}{2} \qquad\qquad (13)$$

For instance: given l_{TX}=20 m, r=1 m at least n_b=80 different azimuthal sectors are required.

CFAR processing

After azimuth beam-forming we have a 3D complex hologram V with elements v_{pkj} where $p=1,2\ldots N_s$ is range index, N_s number of range bin, $k=1,2\ldots N_d$ is Doppler index, and $j=1,2\ldots n_a$ is azimuth index. If we fix a generic azimuth cell j, $V_k(p,j)$ represents the range Doppler map for azimuth bin j. Target detection are found on the range Doppler map with a *CFAR* (Constant False Alarm Rate) detector on $/V_k/$. Eventually a morphological detector of clutter may inhibit Doppler range bin from CFAR application in the case of appreciable clutter, maybe residual from MTI processing.

2.2 TX staggering processing

TX staggering processing flow is represented in **Figure 2**b. The scope is to achieve doppler processing at rate 1/T rather than $1/(M_tT)$. After *Range compression and spurious mitigation*, range migration compensation has to be performed for each angular sector (see 2.1) before *Doppler processing & TX beamforming*, in order to equalize paths from target to the different antenna TX-s. Then *RX beamforming* allows elimination of TX beamforming aliasing due to large separation of TX elements. Finally CFAR detection is performed for each azimuthal cell on range-doppler maps. Here, we limit to analyze *Doppler processing & TX beamforming*, being the peculiar step of processing chain. Moreover the whole processing flow has not been detailed because *Doppler processing & TX beamforming* suffer of dynamic limitation.

Doppler processing &TX beamforming

As shown in Figure 4, the successive *RX-TX* (M_r, M_t) 2D hologram acquired at *TX* cycles *1,2..N_r*, are merged in the time domain to form a 2D (M_r, M_t*N_r). The k-th raw represents the signal *s* received by the *k-th* receiver during the whole frame time T_o. Signal scattered by a scenario of N_{rifl} reflectors, is:

$$s(j) = \sum_{p=1}^{N_{rifl}} A_p \exp(\mathbf{j}2\pi \frac{j l_p}{N_d}) \exp(\mathbf{j}2\pi \frac{perm(j)\ k_p}{M_t}) \qquad\qquad j=1,2\ldots N_{tot} \qquad\qquad (14)$$

where $N_{tot}=M_t*N_r$ and (A_p, l_p, k_p) respectively complex amplitude, doppler index and TX beam index of the p-th scatterer, with $l_p \in (1,N_d)$, $k_p \in (1,M_t)$, $p \in (1,N_{rifl})$. In eq. (14) the phase of each point reflector is sum of a pseudo random term (*array term*) and a linear term associated to the radial target motion (*doppler term*).

The following spectral techniques have been investigated in order to estimate target Reflectivity $f(l,k)$: matched filtering, CLEAN and direct spectral estimation.
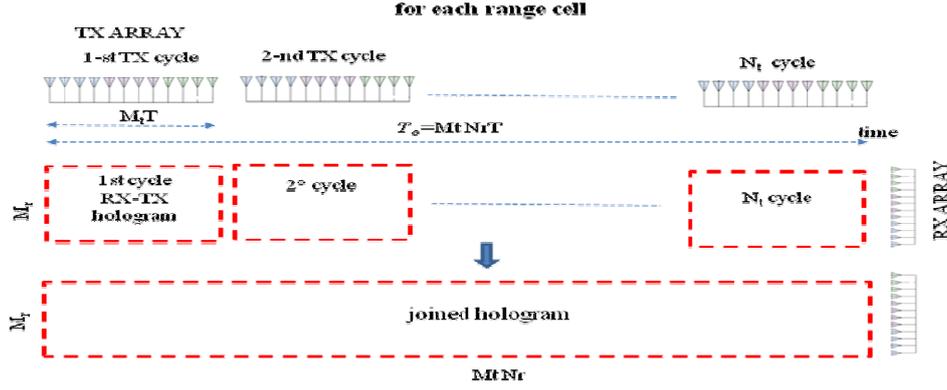
Figure 4 Combination of TX-RX hologram

a) matched filtering

Matched filter estimate $\hat{f}_{mf}$ is recovered from signal s as:

$$\hat{f}_{mf}(l,k) = \left| \sum_{j=1}^{Ntot} s(j) \exp(-j2\pi \frac{j\,l}{N_d}) \exp(-j2\pi \frac{perm(j)\,k}{M_t}) \right| \qquad l=1,2...N_d \qquad k=1,2.M_t. \qquad (15)$$

In the case of a single scatterer (A_1, l_1, k_1) samples

$$\hat{f}_{mf}(l,k) = \sum_{j=1}^{Ntot} \exp(-j2\pi \frac{j\,(l-l_1)}{N_d}) \exp(-j2\pi \frac{perm(j)\,(k-k_1)}{M_t}) \qquad (16)$$

then if $k=k_1$: $\left| \hat{f}_{mf}(l_1,k_1) \right| = N_{tot} A_1$, whereas if $k{\neq}k_1$ $\left| \hat{f}_{mf}(l_1,k) \right| \cong \sqrt{N_{tot}} A_1$ due to the de-coherent effect of pseudo-random $perm(j)\,(k-k_1)$ not perfectly compensated. Then the reconstructed doppler-azimuth dynamic is limited to $dyn = (N_{tot})_{dB}$. For instance $dyn=25\ dB$ for $N_{tot} = 640$.

b) CLEAN technique

Application of CLEAN technique [4] may improve the dynamic range of the doppler- azimuth 2D map obtained at step a) by *detecting & removing* in sequence the peaks, starting from the larger to the smaller. By simulation on a three point target reflectivity 50 dB dynamic all the three reflectors have been detected.

c) direct spectral estimation

Optimum Unbiased Estimation and MMSE (Minimum Mean Square Estimation) have been tested for the direct estimation of target parameters (A_p, l_p, k_p), $p=1,N_{rifl}$ from eq. (14). Our implementations represent a 2D extensions of estimate techniques applied respectively in [6] and [7] to a 1D case.

Unfortunately at this moment we don't experience an improvement with respect to matched filtering, that is image dynamic does not improve.

2.3 Comparison of the two algorithms

TX staggering technique is more attractive because it works at rate $f_s=1/T$ rather than f_s/M_t. Unfortunately it seem to suffer dynamic limitation and application of 2D Optimum Unbiased Estimation and MMSE don't seem to overcome this problem. The reason could be that 2D spectral techniques try to recover a 2D Reflectivity distribution from a 1D hologram (see eq. 14). Moreover, CLEAN technique show to be quite sensitive to peak parameter estimation. More effort will be done do overcome this problem increasing spatial diversity through consideration of *Not Uniform Arrays*.

Moreover, the necessity to indicate a robust solution for RANGER project lead to analyze the slow time f_s/M_t solution with more detail (see Sect. 3).

3. ANALYSIS OF PERFORMANCE

In 3.1 the main MIMO parameters, essentially M_r , T and N_r , will be specified from system requirements: range $R_0=12$ Km, azimuth resolution $\vartheta_b<0.2°$ at boresight, processing time $T_0>100$ msec, and a non ambiguous interval $d_v=12.5$ m/sec for target radial speed. In 3.2 we show the reconstructed azimuth-Doppler map relative to a four 0 dB m^2 point scatterers target, in the hypothesis of two different radar antenna heights.

3.1 MIMO parameter specification

A basic specification for Ranger project is a filled ULA array with $M_r=M_t=20$. Assuming λ=0.1 m (S band) array width is $l_{TX}=20$ m. In this way azimuth beam-forming gives an azimuth resolution $\vartheta_b=1/(M_r\,M_t)=2.5*10^{-3}$ rad $(≈0.15°)$ satisfying requirement ($\vartheta_b<0.2°$ at bore-sight). $d_v=12.5$ m/sec implies $T \leq M_t\lambda/(2\,d_v) = 2*10^{-4}$ (sec), and we fix $T=200$ μsec. We fix also $N_r≥32$, that is an integration time larger than $T_0>M_r*M_{t*}N_r T=128$ msec. Large coherent integration gain $g_{dB}=10*log(M_r*M_{t*}N_r)=41$ dB largely compensate for the 3dB loss suffered within de-chirping step by the signal scattered by a target located at maximum range $R_0=12$ Km, which overlap with the beat chirp for only half length. Moreover $N_r≥32$ implies a fine radial velocity resolution $\Delta_v \leq d_v/N_r = 0.4$ m/sec. Spurious suppression technique described in [2] considers the option to re-transmit four times with each TX, instead on one, using still T as PRT, such as to give a further improvement to spurious cancellation.

3.2 Azimuth Doppler map reconstruction

Azimuth-Doppler map reconstruction is simulated on I/Q 3D complex hologram G for a given range bin in output of compression & spurious mitigation (Figure 2). A simulation reconstructs the range-doppler map implementing *Doppler processing* and *azimuth beamforming*. For simplicity range migration effects has not been considered in the generation of G. Here the main assumptions of simulation: <u>Target model:</u> *0 dB m^2 SW 0;* <u>SNR</u> : *≥ -10 dB* on the single (range compressed) pulse at range $R_0=15$ Km on standard *RCS=0 dB m^2* target. <u>Sea Clutter model:</u> compound clutter with an amplitude Gamma distribution and Gaussian spectrum, see [8] and [9]. Correlation coefficient versus time is modelled as: $\rho(t)=exp[-2(\pi\,\sigma_f\,t)^2]\,exp(j2\pi\,\eta_f\,t)$ with $\sigma_f=2\sigma_v/\lambda$, η_f mean clutter frequency and σ_v is clutter velocity standard deviation. $\eta_f=0$ and $\sigma_v=2$ m/sec are assumed. We have assumed the improvement of GIT (Georgia Institute of Technology) Clutter Reflectivity model described in [9], Sea State=3 is assumed. <u>Antenna target height</u> $h_a=10$ m sufficient for a target distance of a rubber boat at 17Km , $h_a=50$ m (radar site elevated due to coast morphology constraint) <u>Chirp bandwidth</u> *B=100MHz* achieving a nominal range resolution $r_y=1.5$ m. <u>Vertical Polarization</u> is assumed conservatively, being sea clutter returns larger with respect to Horizontal polarization, especially at low grazing angles. <u>Target Reflectivity</u> Composed of 4 point reflectors, all with Reflectivity *0 dB m^2*, radial velocity respectively [2,0.1,3,5] m/sec and azimuth [-60.2 -10.2 4.3 40.7]. Azimuth-Doppler maps have been evaluated respectively for $h_a=10$ m (Figure 5a) and $h_a=50$ (Figure 5b) and respectively at 4 and 6 Km from radar. In the first case sea clutter level is lower with respect to azimuth side-lobe level, practically negligible. In the second case clutter is clearly visible around null Doppler and reduce detection performance for low radial velocity target: indicatively for $|v_r| < \sigma_f$ $(\eta_f=0)$.

(a)   *(b)*
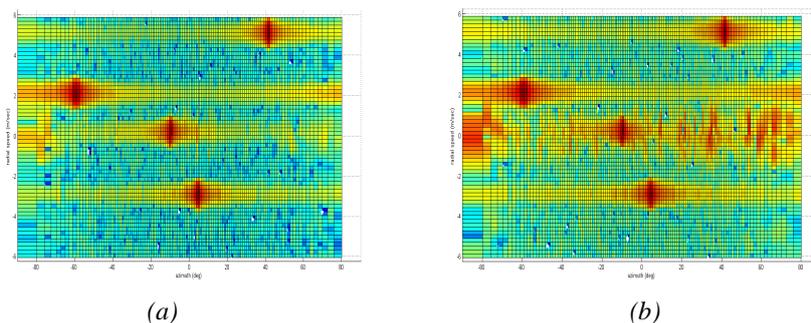
Figure 5 Azimuth Doppler maps reconstructed Reflectivity for $h_a=10$ m (a) and $h_a=50$ m (b)

T stagger from frame to frame, with 2 PRTs around nominal PRT, may in order resolve doppler ambiguity, by extending the non ambiguous interval of radial speed from [-6.25, 6.25] m/sec to [-12.5, 12.5] m/sec, including also fast boats. In presence of large clutter ($h_a≥50$ m and/or sea state>5) proper selection of T around nominal $T=200$ μsec, avoids Doppler superposition of folded target with sea clutter. Though super-clutter visibility appear realistic for the low speed point

target - and we expect a confirm by live data analysis - Doppler processing is still necessary within MIMO architecture based on *Time Domain Multiplexing* in order to separate time and space domains.

4. CONCLUSIONS

Two signal processing for a MIMO architectures based on *Time Domain Multiplexing* have been proposed in this paper: *slow time sampling* and *TX staggering*. In the first, each receiver process echoes received sequentially by the same TX element. In this way phase change do to target (doppler effect) is separated by target change due to phase change due to TX location switch. Doppler processing at a reduced rate PRF/M_t, and the necessity to detect unambiguously low speed targets ($vr<6$ m/sec) implies to assume a low PRT, in our case $T=200$ μsec, implying an range of 12 Km largely compatible with RANGER project requirements. Due to narrow range resolution cell, narrow beam-width and slow grazing angle, due to low antenna height, a *1 m2* target is well detected in super-clutter visibility at all ranges.

The MIMO architecture shows key advantages for the detection of rubber boats with small size and reduced RCS (0dB m2) in sea Clutter: narrow range resolution cell due to large instantaneous BW, narrow beam width due to the distributed Tx antenna, and consequently the collected clutter is very small, moreover high clutter resolution due to long slow time (TM_t of 4 ms) that is useful to distinguish the target from the Clutter.

It has been simulated that the target can be detected in super clutter visibility. The low Pd typical of a CFAR detector in case of super clutter visibility is not a limitation for the radar mission of Search & Rescue; in this scenario the available time for detection can be very long (a few seconds), consequently cumulative Pd can be improved wrt single detection Pd by non-coherent integration of CFAR detections over multiple consecutive frames (T0 =128ms).

In order to reduce PRT limitation, and extend application of the same MIMO architecture to longer ranges (i.e. 50 Km), we have proposed to stagger TX element selection. The pseudo-random law of phase change results to be only partially orthogonal to the nearly linear doppler term. The effect is a limitation of azimuth-doppler map dynamic, nor application of 2D MMSE and CLEAN techniques has shown to be effective. Alternatively, in order to extend the range to 50Km , in the slow time processing architecture presented in chapter 3, the $T>700\mu sec$ must be set, consequently a non-ambiguous Doppler velocity of $d_v<3.5$ *m/sec* occurs. A T stagger frame to frame technique can overcome the Doppler ambiguity, but still it is difficult to distinguish the target from the clutter in the Range dopper maps, due to the clutter folding.

To overcome this problem, it has been estimated that, as it is for the case of 12 km range (Figure 5) also in the case of 50 km range, the detection is feasible in super clutter visibility.

REFERENCES

[1] J. Li, P. Stoica: *MIMO radar signal processing*, 2009 John Wiley & Sons.

[2] R. Klemm: *Principles of Space-Time Adaptive Processing,* The Institution of Engineering and Technology, 2006

[3] R. Lalli, M. Pullo, A. Manuale, A. Farina, L. Pierno, F. Prodi: *Spurious mitigation technique for TDM-MIMO FMCW radar*, 1st Maritime Situational Awareness Workshop (MSAW), Lerici, 8-10 Oct. 2019

[4] Peter W. Moo: Range-Doppler Migration in Coherent MIMO Radar, Proceedings of the 8th European Radar Conference, Manchester, 12-14 Oct. 2011

[5] Krzysztof Kulpa: *The CLEAN Type Algorithms for Radar Signal Processing* MRRS-2008, Symposium Proceedings. Kiev, Ukraine, September 22-24, 2008.

[6] T. Felhauer: *Optimum unbiased estimation, a favourable alternative to matched filter, and its suboptimum implementation by a novel mismatch filter,* IEE Radar conference, Brighton (UK), 1992.

[7] M. H. Ackroyd and F. Ghani, *Optimum Mismatched Filters for Sidelobe Suppression*, IEEE Transactions on Aerospace and Electronic Systems, Vol AES-9, No. 2, March 1973, pp 214-218.

[8] K. J. Sangston, A. Farina: *Coherent Radar Detection In Compound-Gaussian Clutter: Clairvoyant Detectors*, IEEE A&E Systems Magazine, November 2016, Tutorial, Part I of II, pp. 42-63.

[9] S. Watts, K. Wards, R. Tough: *Modelling the shape parameter of sea clutter,* 2009 International Radar Conference: *Surveillance for a Safer World***,** 12-16 Oct. 2009, Bordeaux

[10] V. G. Hansen, R. Mital: *An Improved Empirical Model for Radar Sea Clutter Reflectivity:* NRL/MR/5310-12-9346, April 2012.

*main author: francesco.prodi@leonardocompany.com; phone 0039 3289657249

Passive radar on moving platforms for maritime and coastal surveillance

Philipp Wojaczek[a], Fabiola Colone[b], Diego Cristallini[a],
Pierfrancesco Lombardo[b], Daniel O'Hagan[a]

[a]Fraunhofer Institute for High Frequency Physics and Radar Techniques FHR
Fraunhoferstr. 20, 53343 Wachtberg, Germany;
[b]University of Rome "La Sapienza", Via Eudossiana 18, 00184 Roma, Italy

ABSTRACT

This contribution presents the application of passive radar for maritime and coastal surveillance. The focus is on the application of passive radar on moving platforms for target detection. Its potential on moving target detection and surveillance applications will be presented with results of real data analysis.

Keywords: Passive radar, PCL, clutter suppression, moving target indication, detection

1. INTRODUCTION

For several years the radar research society has been focusing its attention towards passive radar, especially towards passive radar on moving platforms. A passive radar, also known as passive coherent location (PCL), is a special bi- or multistatic radar system, which exploits already existing transmitter infrastructure in order to detect and localize non-cooperative, moving targets. The probably most appealing advantage is the fact, that the passive radar system does not emit any signals, as it uses the emissions of e.g. broadcast or communication transmitters as illuminating signal sources. This fact enables a covert and in principle non-detectable operation, e.g. for reconnaissance. Furthermore, due to the lack of the adversary's knowledge about the passive radar location, it is very difficult to jam a passive radar system. Mounting a passive radar on a moving platform offers the radar operator new opportunities and possibilities. A moving platform could be e.g. a small boat, a small aircraft, or even an unmanned aerial vehicle (UAV). Due to the flexibility of the moving platform the range of action is increased compared to a stationary passive radar system. Therefore, it can serve as a gap filler for terrain hidden to a stationary passive radar, and it is possible to detect moving targets (PCL-MTI) at an earlier stage.

In the context of maritime operations different scenarios can be imagined. An appealing scenario is for example the exploitation of a small boat or a UAV as carrier platform which moves along a coast line. Equipped with a passive radar system it allows to covertly observe movement distant to the coastline, e.g. of non-cooperative small targets approaching a harbor[1,2] in order to increase the safety of harbors and other safety-critical infrastructure. A further possibility is to do covertly ground imaging by passive synthetic-aperture radar (PCL-SAR).[3,4] These images can be analyzed e.g. in order to detect structures and man-made objects, and can be exploited for terrain reconnaissance. By that, PCL-SAR supports the military planning.

One of the most exploited transmitters of opportunity are Digital Video Broadcasting – Terrestrial (DVB-T) transmitters, due to the manifold advantages, such as generally high transmit power level (in the order of tens of kilowatts), constant bandwidth leading to constant range resolution, and almost world-wide availability. DVB-T is a terrestrial transmission, therefore the reception and surveillance is generally limited to on land and to coastal areas. However, for surveillance of coastal areas it is an important and meaningful tool; for surveillance of sea one can think of using satellite based transmissions, e.g. Digital Video Broadcasting – Satellite (DVB-T), or Global Navigation Satellite Systems (GNSS) such as the Global Positioning System (GPS), or Galileo.[5–7]

In this paper the potential of passive radar on moving platforms for moving target indication will be presented by evaluation of measurement data from a maritime moving platform exploiting the emissions from a DVB-T

---

transmitter network. The paper is organized as follows: an introduction into passive radar signal processing exploiting DVB-T transmissions is given in Sec. 2. In Sec. 3 a clutter suppression technique and its adaptation for passive radar on moving platforms is given. In Sec. 4 results from data evaluation from a measurement campaign are presented. Outlook and conclusions are given in Sec. 5.

2. PASSIVE RADAR SIGNAL PROCESSING USING DVB-T

DVB-T is a variant of Digital Video Broadcasting (DVB), which is used for the transmission of audio and video content using terrestrial transmitters.[8] The used frequency band is in the UHF band: 474–786 MHz. One DVB-T channel covers in total a bandwidth of approximately $B_W = 7.61$ MHz in the UHF band, centered at carrier frequency f_C. The DVB-T signal is transmitted as continuous waveform (CW) signal. It constitutes in time-domain of so called symbols, each of constant duration T_S. The constant duration enables the radar engineer to process each DVB-T symbol separately as a type of pulse-Doppler radar, where the pulse repetition interval (PRI) equals the symbol duration: PRI $= T_S$. In frequency domain the DVB-T signal utilizes the orthogonal frequency division multiplex (OFDM) modulation scheme, which means that each symbol comprises K separately modulated carriers, where the each carrier is separated by approximately 1.116 kHz. Each carrier carries either deterministic information or non-deterministic information, where the deterministic carriers enable synchronization and estimation of the transmission channel $H(f)$. The signal received at the passive radar surveillance channel can be written as a function of time t as

$$s_R(t) = a_T s_T(t) \exp(j2\pi f_{DT}(\alpha_T)t) + \sum_{q=1}^{N_R} \int_{\Phi_q} a_q(\alpha)s(t - \tau_q)\exp(j2\pi f_{Dq}(\alpha)t)d\alpha$$
$$+ \sum_{g=1}^{N_G} a_g(\alpha, \gamma)s(t - \tau_g)\exp(j2\pi f_{Dg}(\alpha, \gamma)t) + e(t) \qquad 0 \leq t \leq T_0$$
(1)

where T_0 is the observation time, $s_T(t)$ is the transmitted signal and a_T is its corresponding complex amplitude at the surveillance antenna, N_R is the range extent of the observed scene (expressed in range gates), Φ_q is the azimuthal angular sector of the illuminated scene, $a_q(\alpha)$ is the complex amplitude of the echo received with time delay $\tau_q = R_q/c_0$ from azimuth angle $\alpha \in \Phi_q$, and R_q is the bistatic range of the q-th range gate. $f_{DT} = \frac{v_T}{c_0}f_T \cos \alpha_T$ is the Doppler frequency for the motion component relative to the transmitter. Finally, the third contribution of (1) corresponds to contributions of N_G moving targets of delay τ_g and bistatic Doppler shift $f_{Dg}(\alpha, \gamma)$. The component $e(t)$ is white Gaussian noise (WGN).

For the digital signal processing, the passive radar receiver synchronizes on the received signal $s_R(t)$, and then reconstructs the transmitted signal $s_T(t)$.[9] After reconstruction, a copy $\hat{s}_T(t)$ of the transmitted signal is available at the receiver, which is used as a filter $h(t)$ for the correlation process on a batch-wise basis, that is for each DVB-T symbol $m = [0, \ldots, M - 1]$ individually: $x_C(l, m) = s_R(l, m) * h(l, m)$. $l = t f_S$ represents the samples acquired with sampling frequency f_S, and $x_C(l, m)$ is a so-called slow-time–fast-time matrix resulting from the correlation process. The matrix $x_C(t, T_S)$ can be Fourier transformed into a range-Doppler map $z(l, f_D)$, via: $z(l, f_D) = \sum_m^{M-1} x_C(l, m) \exp(-j2\pi v m/M)$, where f_D represents the Doppler frequency in a particular Doppler bin.

3. PASSIVE RADAR FOR MOVING TARGET INDICATION

Referring to (1), one can easily see, that the clutter returns have a Doppler frequency depending on the velocity of the moving receiver, which can be seen in the range-Doppler map as an interference at Doppler frequencies $f_D = \frac{v_{Rx}}{c_0}f_C \cos \alpha$, usually termed as "clutter ridge". This is especially a problem for MTI (as well for active and passive radar), as signal echos of slow moving targets can fall into the clutter ridge and thus being hidden by the clutter. In order to suppress the clutter, various techniques were created for active radar, e.g. displaced phase center antenna (DPCA).[10]

DPCA is a simple technique, which requires only two channels in the basic configuration. Provided ideal conditions, it allows to completely suppress the clutter. Two antenna elements are displaced mounted along-track,

where the preceding element can be labeled leading antenna (LA) and the rear element can be labeled trailing antenna (TA). As the receiver moves, both elements, LA and TA, receive signal echos according to (1). Due to the receiver's movement, the phase center of the TA will be at the location, where the phase center of the LA has been a time step before. The spatial overlap of both phase centers – the DPCA condition – is a crucial requirement, as then both LA and TA receive clutter echoes with the same Doppler shift and delay. It can be expressed as a time delay T_D depending on the antennas' displacement d and on the receiver's velocity v_{Rx}: $T_D = d/v_{Rx} = K \cdot \text{PRI}, \quad K \in \mathbb{N}$. A further crucial requirement is that the transmitted radar pulses are time-invariant. While in active radar this requirement is usually possible to fulfill, as the transmitted waveform is under control of the radar operator, for a passive radar this is not feasibly, due to the dependence on the non-cooperative time-varying waveform. To overcome this limitation, the reciprocal (or inverse) filter $h(t) = \text{IFFT}\{\hat{S}_T^{-1}(f)\}$, where $\hat{S}_T(f) = \text{FFT}\{\hat{s}_T(t)\}$, can be used in order to create a time-invariant impulse response for the clutter suppression process.[11] Applying $h(t)$ on the data from both receiving channels, two slow-time–fast-time maps $x_C^{(LA)}(l,m)$ and $x_C^{(TA)}(l,m)$ for LA and TA can be created. To suppress the clutter, both maps are subtracted with the respective time delay T_D and Fourier transformed to create a range-Doppler map $z_D(l, f_D)$:

$$z_D(l, f_D) = \sum_{m=0}^{M-1} (x_C^{(TA)}(l,m) - x_C^{(LA)}(l, m - T_D)) \exp(j2\pi m f_D/M) \tag{2}$$

4. RESULTS OF PCL-MTI

A measurement campaign with a moving passive radar has been undertaken in order to show and evaluate the performance of passive radar for moving target indication. As receiving system the Parasol system,[12] developed by Fraunhofer FHR, has been used. The passive radar was carried by a small boat (see Fig. 1) which moved along defined trajectories in the Oslo fjord in Norway (see Fig. 2 with a velocity $v_{Rx} \approx 8.5\,\text{m/s}$. Two DVB-T transmitters, Tx-1 and Tx-2, transmitting on center frequency $f_C = 650\,\text{MHz}$ were in the vicinity of the trial site. Tx-1 is marked with a red circle in Fig. 2, while the direction to Tx-2 being further distant to the trial site is indicated with a black arrow in Fig. 2. The black dashed line in Fig. 2 indicates frequently traveled waterways, convenient for the detection of non-cooperative targets or ships in order to show the potential of passive radar for moving target indication.

Two discone antenna elements were mounted along-track as linear array on the boat, with bore sight to starboard. As the discone antenna elements are omnidirectional in azimuth, radiation absorbing material was mounted on the array's left side to achieve a side-looking condition.

The processing described in Sec. 2 and Sec. 3 was applied on the received data. The result before clutter suppression is shown in Fig. 3. The interference from the direct signal (approx. at Doppler frequency $f_D \approx -11\,\text{Hz}$) has been suppressed using the ECA-CD filter,[13] as the direct signal considerably limits the dynamic range in the range-Doppler map due its high power level. One can clearly see the clutter returns up to a bistatic range of 6 km, which overlap the echoes of slow moving targets. The results after clutter suppression are shown in Fig. 4. The improvement is clearly visible, as most clutter echoes are suppressed, so that targets are detectable against the background: two slow moving targets (marked with red ellipses in Fig. 4) at bistatic range of $\approx 400\,\text{m}$ and $\approx 600\,\text{m}$ and Doppler $\approx -17\,\text{Hz}$ and $\approx -20\,\text{Hz}$ are clearly visible against the background. A third moving target (marked with a red ellipse) at bistatic range $\approx 2000\,\text{m}$ and Doppler $\approx -7\,\text{Hz}$ is as well better detectable. Further results are presented in Figs. 5 and 6. Both Figures present results after clutter suppression, and in both Figures responses from the same target, but from different illuminations, can be seen: In Fig. 5 one can see the return from the target (marked with a black ellipse) illuminated from Tx-1. The target is close to the receiver at Doppler $\approx 21\,\text{Hz}$ and bistatic range $\approx 100\,\text{m}$. The second transmitter Tx-2 being further distant to the receiver than compared to Tx-1 illuminated the same target as well, but due to its greater distance to the receiver, is received second. The target response is shown in Fig. 6 at a bistatic distance of $\approx 19.8\,\text{km}$ at Doppler $\approx 23\,\text{Hz}$. The echo of the target is marked with a white ellipse. The simultaneous illumination by multiple transmitters and reception is very appealing as it allows the localization of detected targets via ellipsoid intersection.

These results show the potential of passive radar on moving platforms for detection and localization of slow moving targets.

Figure 1. Picture of the moving receiver platform. The linear array is mounted on the boat's stern (recognizable due to the RAM to achieve a side-looking condition).
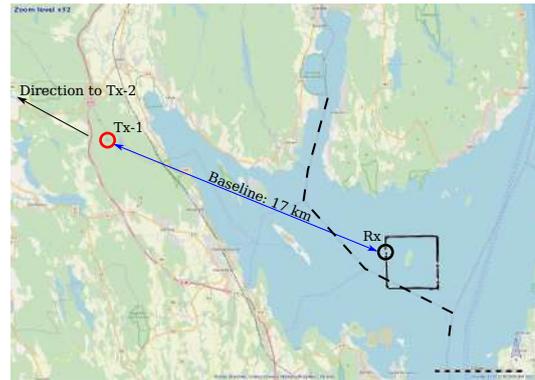
Figure 2. Map of the trial site (from OpenStreetMap). The receiver's trajectory is indicated in black. At the time of the evaluated data it was moving from North to South on the left side of the square (indicated with the black circle). The array steering direction was to starboard. The transmitter Tx-1 is indicated with the red circle, the baseline is indicated with the blue double-sided arrow. The direction to the second transmitter Tx-2 is highlighted with the black arrow.
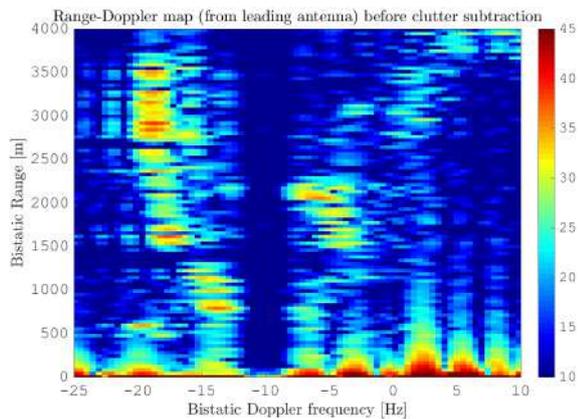
Figure 3. Range-Doppler map of evaluated data from the LA (before clutter suppression). The clutter responses are clearly to be seen and can overlap target echoes.
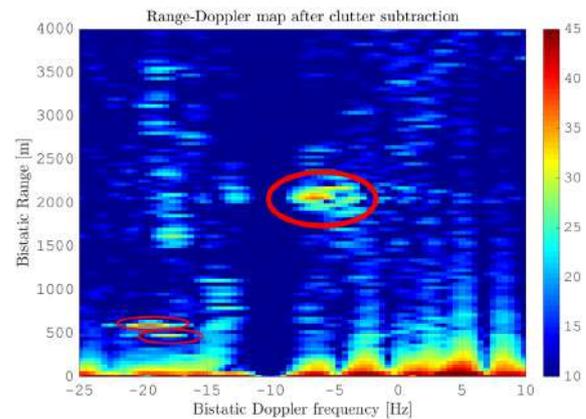
Figure 4. Range-Doppler map of evaluated data after clutter suppression. The overall clutter has been considerably reduced and allows to detect the three slow moving targets (marked with red ellipses).

5. OUTLOOK AND CONCLUSIONS

This contribution presented passive radar on moving platforms for the purpose of moving target detection using terrestrial broadcast transmitters as illuminators. The passive radar signal processing was introduced and results from evaluated measurement data were presented. The results show the feasibility of passive radar for moving target detection, even if the targets' echoes are buried in clutter. This shows the potential and the applicability for the surveillance of coast-line and maritime infrastructure, such as harbors and ports, and for reconnaissance of large areas. As next steps it is planned to increase the number of receiving elements to have more than two elements, and to adapt the algorithms in order to be able to do multi-array processing in order to further improve the clutter suppression and target detection.

6. ACKNOWLEDGMENTS
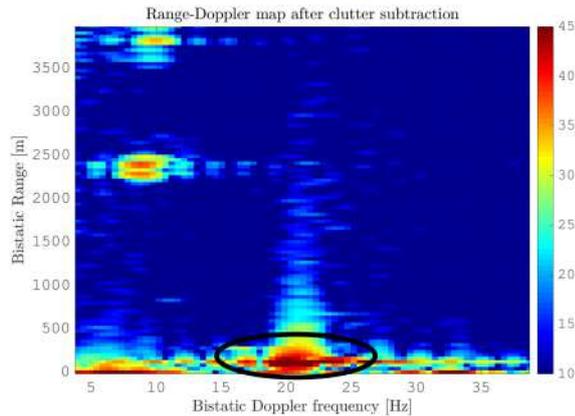
Figure 5. A moving target (marked with black ellipse) with high reflectivity being illuminated by the transmitter Tx-1 received first from the network.
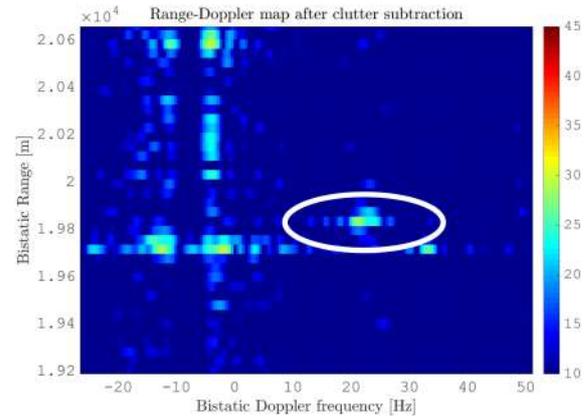
Figure 6. The echo from the same moving target (marked with white ellipse) as in Fig. 5 being illuminated by the transmitter Tx-2 received second from the network.

REFERENCES

[1] Bournaka, G., Baruzzi, A., Heckenbach, J., and Kuschel, H., "Experimental validation of beamforming techniques for localization of moving target in passive radar," in [*2015 IEEE Radar Conference (RadarCon)*], 1710–1713 (May 2015).

[2] Baruzzi, A., Bournaka, G., Heckenbach, J., and Kuschel, H., "Experimental validation of kalman filter based tracking in passive radar systems," in [*2015 IEEE Radar Conference (RadarCon)*], 1634–1637 (May 2015).

[3] Atkinson, G. M., Sayin, A., Cherniakov, M., Antoniou, M., Stove, A., and Underwood, C. I., "Passive SAR satellite system (PASSAT): Ground trials," in [*2018 International Conference on Radar (RADAR)*], 1–6 (Aug 2018).

[4] Wojaczek, P., Summers, A., Cristallini, D., Walterscheid, I., Lombardo, P., and Colone, F., "Results of airborne PCL under CCI conditions using DVB-T illuminators of opportunity," in [*2018 International Conference on Radar (RADAR)*], 1–6 (Aug 2018).

[5] Antoniou, M., Stove, A. G., Tzagkas, D., Cherniakov, M., and Ma, H., "Marine target localization with passive GNSS-based multistatic radar: Experimental results," in [*2018 International Conference on Radar (RADAR)*], 1–5 (Aug 2018).

[6] Santi, F., Pieralice, F., and Pastina, D., "Multistatic GNSS-based passive radar for maritime surveillance with long integration times: Experimental results," in [*2018 IEEE Radar Conference (RadarConf18)*], 1260–1265 (April 2018).

[7] Pisciottano, I., Cristallini, D., Schell, J., and Seidel, V., "Passive ISAR for maritime target imaging: Experimental results," in [*2018 19th International Radar Symposium (IRS)*], 1–10 (June 2018).

[8] ETSI, 650 Route des Lucioles F-06921 Sophia Antipolis Cedex - FRANC, *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television* (October 2015).

[9] Poullin, D., "Passive detection using digital broadcasters (DAB, DVB) with COFDM modulation," *IEE Proceedings - Radar, Sonar and Navigation* **152**, 143–152 (June 2005).

[10] Stimson, G. W., Griffiths, H. D., Baker, C. J., and Adamy, D., [*Introduction to airborne radar*], SciTech Publishing, Inc. (2014).

[11] Wojaczek, P., Colone, F., Cristallini, D., and Lombardo, P., "Reciprocal-filter-based STAP for passive radar on moving platforms," *IEEE Transactions on Aerospace and Electronic Systems* **55**, 967–988 (April 2019).

[12] Heckenbach, J., Kuschel, H., Schell, J., and Ummenhofer, M., "Passive radar based control of wind turbine collision warning for air traffic parasol," in [*2015 16th International Radar Symposium (IRS)*], 36–41 (June 2015).

[13] Schwark, C. and Cristallini, D., "Advanced multipath clutter cancellation in ofdm-based passive radar systems," in [*2016 IEEE Radar Conference (RadarConf)*], 1–4 (May 2016).

3
Artificial Intelligence for MSA

Contents

A Geospatial Complex Event Processing Engine for Abnormal Vessel Behavior Detection Suitable for Maritime Surveillance

Manolis Tsogas*[a], Polyzois Parthymos[a], Marios Moutzouris[a], Nektarios Patlakas[a], George Karagiannis[b], Antonis Kostaridis[a], Dimitris Diagourtas[a]

[a]Satways Ltd., 3 Christou Lada, Athens, Attiki, Greece 15233; [b]Hellenic Navy General Staff A4/IV, Hellenic Ministry of Defense, 227 Mesogeion, Attiki, Greece 11525

ABSTRACT

Modern maritime surveillance systems provide a huge amount of data for near real time processing due to the large number of ships sailing across the seas. Thus, it is necessary to be able to handle large amounts of information accurately and efficiently. This can be done by utilizing a Complex Event Processing (CEP) engine which can process streams of data available from different sensors. In this paper, we present an advanced computational rule-based engine (TRITON) that analyzes ship positions for the automatic detection of abnormal and/or specific vessel behavior following the CEP paradigm. The purpose of the system is to support maritime authorities' surveillance functions, by providing an enhanced situational picture in near real time. The system is able to analyze ship position reports using data from available tracking systems such as Maritime radars, Long Range Identification and Tracking (LRIT) systems, Terrestrial-AIS, Satellite-AIS, Vessel Monitoring System (VMS) and Earth Observation satellites, among others. Patterns, such as entering an area of interest, encounters at sea, approaches to shore, drifting and deviations from usual routes, are detected and operators can be automatically alerted in real time. The system reasons on over 20 different rules and is designed in such a flexible way that can be expanded very easily according to user needs.

Keywords: Complex event processing, geospatial, maritime, abnormal behavior, patterns, anomaly detection

1. INTRODUCTION

Vessel Traffic Services (VTS) are used for maritime traffic monitoring and are operated by harbor, port or national authorities in order to ensure that ships moving in areas of interest are following established traffic and safety rules. Over the last years, instead of using human operators to detect abnormal vessel behavior, which is a difficult task considering the huge amount of vessel traffic, it is more efficient to use automatic detection systems that are less prone to errors. Maritime domains that can benefit from such systems include safety, when detection and early warnings for vessels posing a potential risk [1], arriving at specific areas or following unusual routes are of great importance. Also, another domain is security, when it is necessary to track vessels involved in illegal activities, smuggling prohibited substances, or breaking embargos. Border protection is also another interesting field, where vessels crossing sea borders transferring goods or persons illegally must be detected. Finally, detection of illegal fishing activities within regulated zones is something that requires attention considering the depleting stocks.

Information flow processing models have become an important approach in order to cope with time constraints in a wide range of environments [2]. Complex event processing was originally developed as a set of tools and methodologies for detecting situations of interest under high throughput data streams in applications such as network intrusion detection, industrial control systems [3] and real time analytics. Over the last years, huge volumes of data are continuously generated due to the increasing number of applications especially in the maritime environment [4]. Therefore, efficient methods are required to determine the event patterns of interest and manage highly dynamic events in real-time. A methodology for applying CEP in the maritime environment is described in [5] where the authors have created a basic rule-based system using an event model for handling incoming and outcoming information for detecting simple anomalies produced from ships.

*m.tsogas@satways.net; phone +30 210 6840036; fax +30 210 6840037; satways.net

In this paper we present a novel abnormal vessel behavior detection service, which is called Triton and is implemented using a Geospatial Complex Event (GCEP) processing architecture, in order to identify and analyze motion patterns of vessels that indicate an ongoing situation that needs attention. Geospatial complex event processing combines geospatial information and detections from different and distributed sources of information to infer events or patterns that describe the current situation. It offers flexibility in applying rules only in specific areas of interest or in specific vessels according to their type or identifier and it allows to combine rules in order to identify more complex motions. Thus, the user can focus exactly where it is necessary, less processing power is required, and a broader set of abnormal vessel behavior patterns is supported.

2. DETECTION OF ABNORMAL BEHAVIOR

2.1 Complex Event Processing

Complex event processing (CEP) is the method for processing large volumes of incoming events in real time and detecting when there is an abnormal behavior using specified rules. In a system implementing CEP, events are arriving in a streaming fashion and the goal is to detect certain patterns defined by domain experts. Typical examples of applications where CEP can be used are in the financial sector for fraud detection, algorithmic trading and risk management. Also, it can be used in network and application monitoring for the detection of intrusion attempts and monitoring of service-level agreement (SLA). In the business process management and automation, it can be used for relationship management, workflow applications, process tracking, operational intelligence and reporting exceptions. And finally, it is useful for sensor network applications such as RFID reading, scheduling and control of production lines, analysis and management of public transport, air traffic and maritime surveillance.

The open source Esper platform was used [6] to provide native complex event processing capabilities during the implementation of the TRITON service. Esper offers low latency and high throughput, compliance with standards and is light-weighted in terms of memory, CPU and IO-usage. It is highly scalable, memory efficient, with low latency and the most important is real-time streaming-capable so it can be used for online and real-time arriving data and high-variety data, as well as for historical event analysis. Esper's compiler and runtime are not limited to running on a single machine and can be part of a distributed stream processing framework. Both modules can run in any architecture and any container, as they have no dependencies on external services and do not particularly require any threading model or model of how time advances and do not require any external storage. Also, Esper offers a scripting language which is called Event Processing Language (EPL), that implements and extends the SQL-standard and enables rich expressions to be applied in events over time. EPL works well with event-time and watermark-based time management.
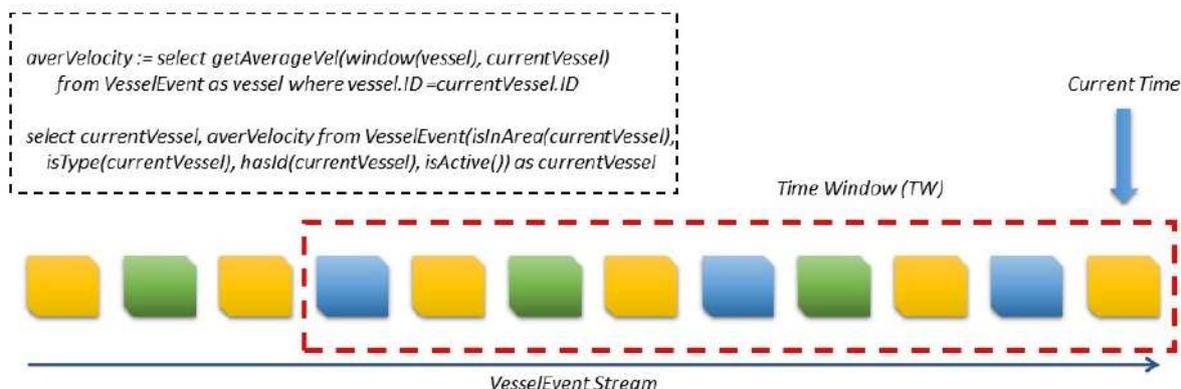
Figure 1. Rule example using EPL language

Figure 1 shows a rule example using the Event Processing Language to form a query in order to calculate the average velocity for every vessel. Every update in ship's state is represented by a new 'Vessel' event (ship ID is distinguished using different color in events). In order to find the average velocity a subquery (averVelocity) has been created, which

filters out only the vessels having the same id as the current one, using a time window in the past. A custom aggregation function (getAverageVel) has been created for estimating the average velocity using the filtered events. Finally, this subquery is part of a standard 'select' query that has been created for reading from the stream only events for vessels located in the user-specified area (isInArea) or having the requested id (hasId) and when this rule is scheduled to run (isActive).

2.2 Geospatial Complex Event Processing Service (TRITON)

TRITON is a service that processes data in real time providing enhanced situational awareness and fast response in critical situations. This is accomplished by producing alerts from various patterns which are manually or automatically enabled. The maritime domains in which this service can be applied include border control, counter smuggling, counter human trafficking, counter piracy, counter terrorism, naval and joint operations, traffic safety, crisis management, maritime search and rescue, pollution and safeguarding of the environment, protection against illegal fishing and port control.

The service can detect patterns and relationships between events using the EPL language. EPL provides out-of-the-box capabilities like pattern matching and detection, filtering, transformation, aggregation, event hierarchies, detecting relationships, thus it allows easier modelling of rules and detection of abnormal vessel behavior, when applied to incoming data streams of AIS, radar and fused tracks. Furthermore, a great number of custom aggregation functions have been implemented using the JTS Topology Suite which provide TRITON with geospatial capabilities and advanced features for handling complex scenarios. The user can select and apply multiple rules in specific vessels (selected using their type or their identification number), routes or sea areas or combine rules for more complex situations. The core of the service provides multiple algorithms such as Bayesian Inference Systems for utilizing prior knowledge.
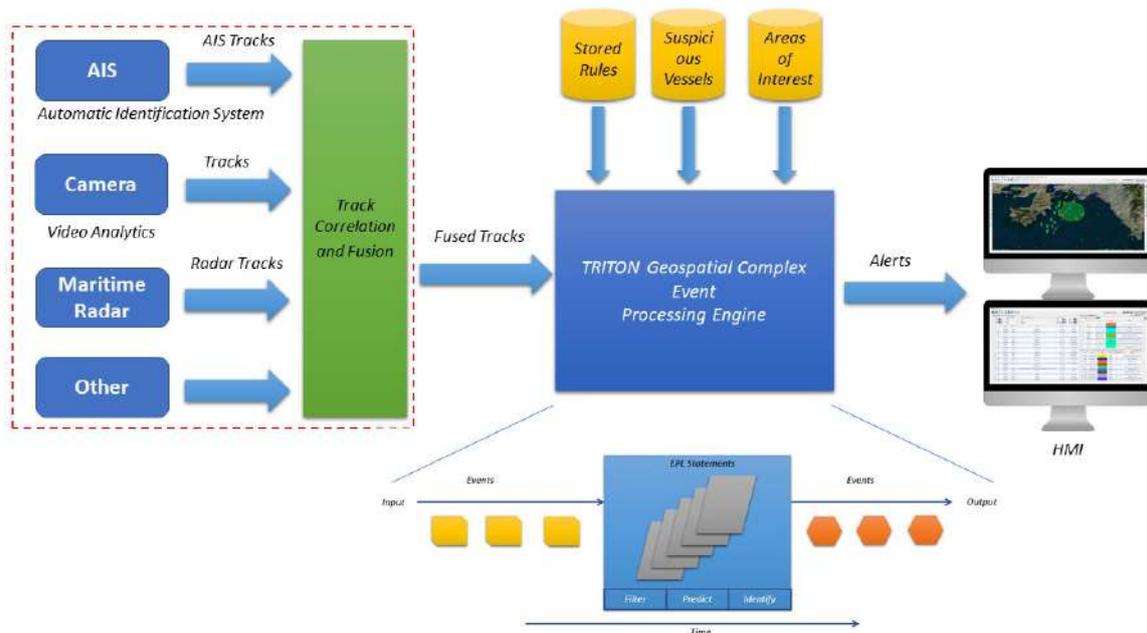
Figure 2. Architecture of Triton Service

The architecture of the engine is shown in figure 2. The input consists of vessel information (position, velocity, heading, identity, dimensions, etc.) received from various sources of information, like the Automatic Identification System (AIS) and maritime radars. If necessary, information about suspicious vessels can also be retrieved from the database, which also provides the coordinates of all the areas, borders and locations which are useful for identifying vessel behavior patterns in respect with their position in the sea environment. Due to the multiple sources of information it is necessary

to apply a fusion preprocessing step in order to correlate detections from different sensors and acquire a reliable and more accurate vessel state. Then, each vessel update is considered as an event, which feeds the input stream of the engine. The next step is to process the event stream by applying a set of rules, which define specific patterns to be detected, according to user preferences. The operator can enable and disable rules according to the current situation and apply them to specific zones in the monitored sea environment or exclude specific zones that are of no interest, through the user interface.

2.3 Abnormal Patterns

A list of patterns indicating abnormal or unwanted behavior from vessels can be found in [7]. TRITON provides a wide range of rules that can be used to detect those anomalies. Existing rules associated with geofence operations (Figure 3) include detection of vessels entering, exiting, crossing, moving away or approaching areas. Also, crossing, approaching or moving away from borders is supported in addition to approaching or moving away from single location points. These rules allow the user to completely monitor traffic near critical infrastructures like ports and drilling platforms, or detect vessels crossing country borders for disrupting irregular immigration. Finally, it is possible to detect a vessel which is approaching a specific area and has departed from a specific port or zone.
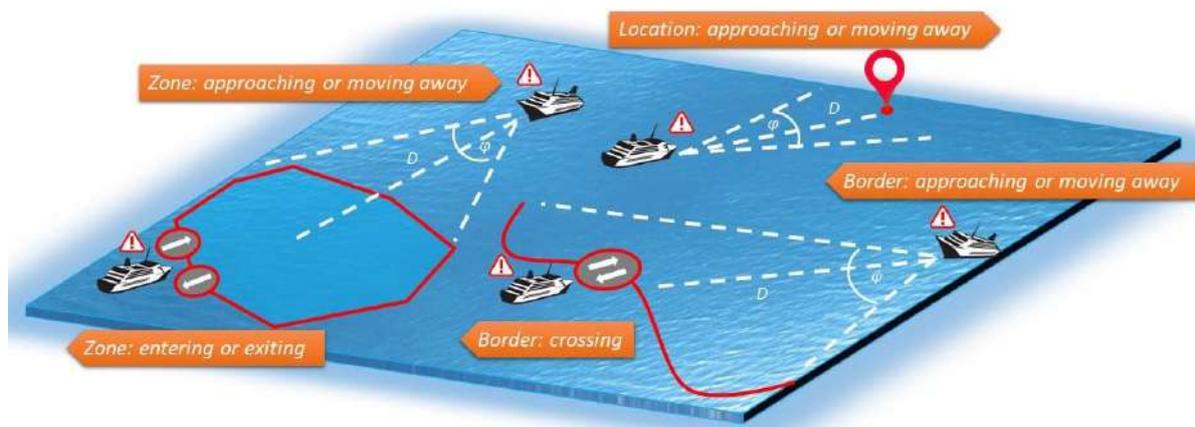
Figure 3. Geofence rules supported from Triton

Furthermore, TRITON provides a set of rules for recognizing specific patterns related with vessel interaction, such as a small boat approaching a larger vessel, two vessels sailing together in parallel, vessels rendezvousing and vessel continually crossing the route of another vessel. These rules are useful for identifying illegal activities such as smuggling scenarios. Also, rules related to ship safety, such as detection of an imminent collision and vessel domain violation are also provided, when it is necessary to know if two ships are too close together or following their current course will end up in a collision between them. There are also rules helping operators to identify possible spoofing or suspicious scenarios such as track splitting, inconsistent kinematic attributes, space-time incompatibilities, suspicious identity, AIS transmission stopped and contact lost. Finally, other rules provided by the TRITON include the detection of abnormal speed or course change, exceeding speed limit, vessel having an average velocity above or below user-defined threshold, detection of loitering ships and deviation from shipping lanes.

But the most important feature of TRITON is that it offers the ability to detect complex scenarios using combined rules. For example, a smuggling scenario would include the following actions. A cargo vessel transmitting AIS information is approaching a designated area near an island, which is isolated and where similar illegal actions are known to be taking place. Before the ship alters its course, it switches off its AIS transmitter (it is still detected by a radar station) and then it is making an abnormal course change and heading to the shore. When it is sufficiently near the coast it halts and is waiting for a small boat which is approaching it. The boat is rendezvousing with the cargo vessel, they stay side by side, completely stopped (transfer of illegal items may be taking place) and then the boat departs for the coast. The same

procedure may be repeated multiple times or several small boats to be used for transferring items to the shore. When their intended actions are completed, the cargo vessel departs and resumes its initial route, switching on its AIS transmitter. Using TRITON, one would be able to set multiple rules, which cover a specific test case like the one described, and to trigger a single alarm when this complex pattern is detected. In our case, the user should create a complex rule consisting from a geofence rule for approaching a specific area, a geofence rule for entering an area, an abnormal course deviation rule, a rule for AIS transmission stopped, a loitering rule and a small vessel approaching a bigger vessel rule.

3. RESULTS

3.1 Simulating data

TRITON service has been evaluated using large scale simulated data for reconstructing vessel traffic in Greek territorial waters. For this purpose, a simulator was developed as an extra service which have been adapted in the input of the Complex Event Processing Engine. The simulator can produce track updates based on predefined routes which are loaded from a database or created from the user of the service. Also, it is able to produce both AIS vessels (refresh rate is not constant and usually low) or vessels simulating radar detections (constant high refresh rate).

At the end, the simulator was able to produce traffic for all kinds of ships, such as passenger vessels travelling from major mainland ports (Piraeus) to Greek islands, cargo vessels and tankers crossing the Aegean and following the main trading shipping lanes, pleasure crafts and sailing boats making short trips from the Greek islands, fishing boats following random paths in specified areas and unknown contacts from radar stations. In order to increase the traffic load directed in the event processing service, the simulator was able to produce also stationary random vessels in all Greek ports. For the purpose of the evaluation we have used publicly available routes of Greek ferry boats around Greece and we have created additional routes for representing the trading shipping lanes, the paths of pleasure crafts and the random paths of fishing boats.
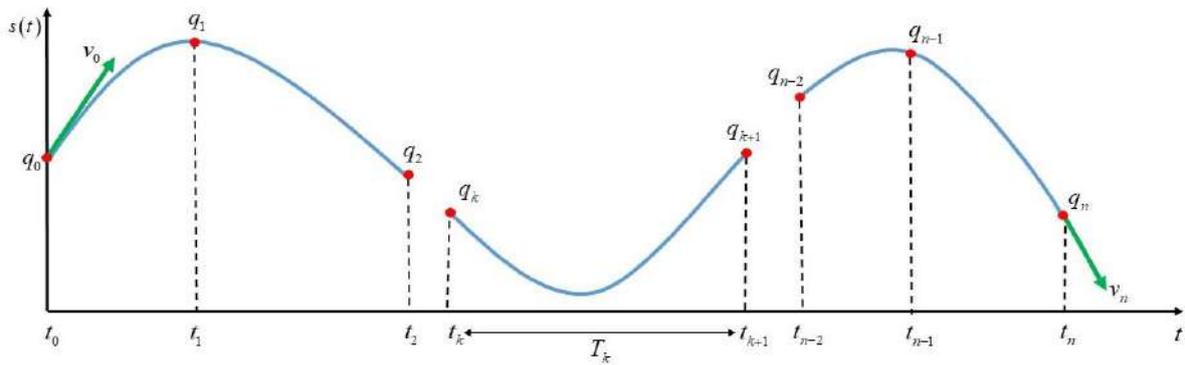
Figure 4: Definition of spline for interpolating ship trajectory

The core engine of the simulator is using an interpolation method based on cubic splines [8] in order to reconstruct the motion of a vessel following a specific path which is provided as a set of points in space (route). When $n+1$ points are given it is possible to use n polynomials of degree p, each one defining a segment of the trajectory. The overall function $s(t)$ defined in this manner is called spline of degree p. The desired degree of continuity between segments dictates the value of p. In our case, we want to ensure the continuity of velocities and accelerations at the time instants t_k, where the transition between two consecutive segments occurs. In that case it is sufficient to use polynomials of degree $p = 3$ (cubic spline).

$$s(t) = \{q_k(t), t \in [t_k, t_{k+1}], k = 0, \ldots, n-1\} \tag{1}$$

$$q_k(t) = a_{k0} + a_{k1}(t - t_k) + a_{k2}(t - t_k)^2 + a_{k3}(t - t_k)^3 \tag{2}$$

So, in order to estimate the overall trajectory, it is necessary to compute 4 coefficients for each polynomial, which means $4n$ coefficients for the whole trajectory. In order to solve this problem, the following conditions must be considered: $2n$ conditions for the interpolation of the given points, since each cubic function must cross the points at its extremities, $n-1$ conditions for the continuity of the velocities at the transition points and $n-1$ conditions for the continuity of the accelerations at the transition points. In this way, there are $4n-2$ conditions and therefore the remaining degrees of freedom are 2. So, two more conditions are necessary for computing the final spline, which were selected to be that the initial and final velocities should have fixed values selected by the user $\dot{s}(t_0) = v_0$ and $\dot{s}(t_n) = v_n$. Among other possible choices, one can assign the initial and final accelerations to fixed values, set the final velocity and acceleration equal to their initial values (cyclic conditions for periodic splines) or finally ensure the continuity of the jerk at every time instance t_k. Eventually, after gathering all conditions the following equation must be solved:

$$A(T) \cdot v = c(T, q, v_0, v_n) \tag{3}$$

where $T = [T_0, T_1, \ldots, T_{n-1}]^T$ are the time lengths of each segment, $q = [q_0, q_1, \ldots, q_n]^T$ are the points forming the selected route (x or y coordinates) and $v = [v_1, v_2, \ldots, v_{n-1}]^T$ are the velocities to be found. The $(n-1) \times (n-1)$ matrix $A(T)$ has a diagonal dominant structure and therefore it is always invertible if $T_k > 0$. Once, the inverse of A has been computed then velocities can be calculated from $v = A^{-1} \cdot c$ and therefore the problem is solved and an analytical form (position, velocity and acceleration in every desired time instance) of ship's trajectory can be found using spline polynomials (1), (2) and their derivatives.

It should be noted that since the initial selected ship route is defined using geographic coordinates, there is a transformation of each point (longitude, latitude) from the global coordinate system to the local tangential coordinate system. The interpolation process is repeated twice, one time for all x-coordinates and one time for all y-coordinates and in the end the output contains the two-dimensional positions (x, y) of the ship through time alongside with its velocity and acceleration in each position. Coordinates can be then transformed back to the global coordinate system. The output time is calculated using the following formula:

$$t_i = \begin{cases} i \cdot dt & vessel\_type = radar \\ t_{i-1} + dt_i & vessel\_type = ais \end{cases} \tag{4}$$

where dt is the refresh time in seconds. If the simulated vessel is a radar contact, then dt is constant and is set to 0.5 or 1 seconds. If the simulated vessel is received from an AIS station then dt_i is not constant, but it is approximated as a gaussian random variable with mean and standard deviation selected by the user.

3.2 System Architecture

Both services (TRITON and the simulator) are implemented as Java 8 standalone applications. They are using Jboss to access (read and write data) to the database and ActiveMQ for messaging purposes between them and also the graphical user interface. PostgreSQL is used as the main database, which is optimized for storing and querying data that represent objects defined using geospatial information.

The complete system is deployed inside a Virtual Machine (VM) running Debian as the operating system. The tightly isolated software container with an operating system and applications inside, ensures that development can be done in a controlled environment and then the whole system to be deployed in the final server without worrying about compatibility issues. Also, this way the computer resources assigned to TRITON VM can be modified according to the current needs. Furthermore, updates can be applied more easily without affecting the other applications running on the server, which may be VMs hosting other components of the system. The communication with other modules is accomplished using an adaptor, which translates data from the internal TRITON format to the one used from the overall external system. For example, TRITON has been deployed in various sites for the purpose of the Marisa Project [9]. In

this case, Kafka is the messaging server for communication purposes within the server, so an adaptor has been developed, which implements a bidirectional translation of information between Kafka and ActiveMQ. This way TRITON can be used in any system by spending minimal effort in developing the appropriate translation unit. In our case, for the purposes of this paper, data are generated by the simulator within the VM, but it is possible to provide both real data (using the appropriate adaptor) and simulated vessels at the same time, so it is possible to examine more complex scenarios in real conditions.
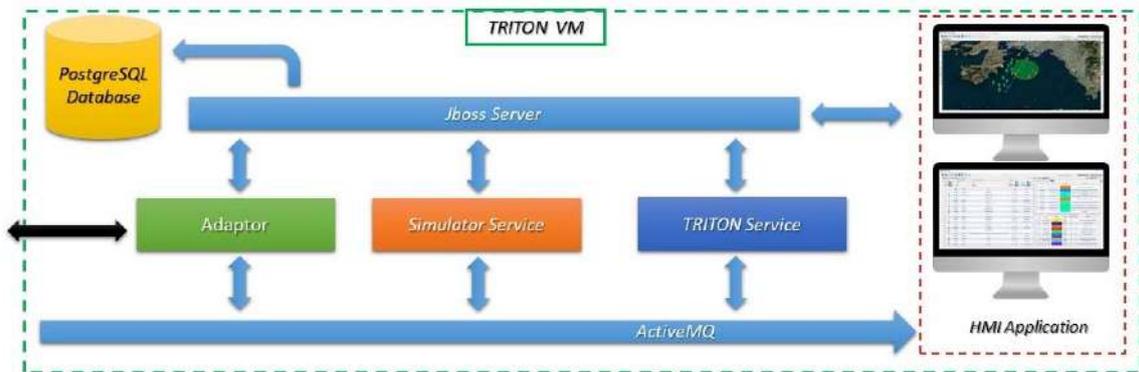
Figure 5: Internal architecture of TRITON Virtual Machine

3.3 Evaluation of service

In order to evaluate the service more accurately, the simulator was designed to generate vessels in random fashion according to the following rules. The total number of vessels that are created in every port, are uniformly distributed in the space $[0...15]$. Their initial position (longitude, latitude) is selected using a Gaussian distribution with mean value the longitude and latitude of the port respectively and standard deviation equal to 0.025 degrees in both axes. The type of the vessel is selected randomly according to a uniform distribution of all the available vessel types as defined in [10]. The following types -radar, mobile offshore drilling unit and nuclear ship- have also been included. Their position, heading and velocity are considered as Gaussian random variables with mean and standard deviation set by the user and are estimated accordingly in each update of the simulator. As far as it concerns the ferry routes, the generated vessels can only be passengers or high-speed vessels (if the total travelled distance is big enough). The selection of the type is also done randomly. The initial position of the vessel is selected by considering the starting point of each segment of the route as uniformly distributed and selecting a random one, from which each vessel begins its trip. Initial and final velocities are selected randomly within the range $[3...8]$ Knots. The AIS refresh rate is considered as a Gaussian random variable with mean and standard deviation set by the user. The same approach is also used for the routes occupied by pleasure crafts or sailing boats and the routes for the fishing vessels. As far as it concerns the radar tracks, the only change is that the refresh rate is constant (1 sec), their initial velocity is randomly selected in the range $[5...8]$ Knots and their final velocity randomly selected in the range $[8...12]$ Knots. Finally, in the four trading shipping lanes the generated vessels are either cargo ships or tankers and the generated vessels are 100 per direction. The rules applied are the same as in the passenger ships except for the initial velocity which is randomly selected in the range $[12...14]$ Knots and the final velocity which is randomly selected in the range $[16...19]$ Knots.

The total number of vessels generated by the simulator for each type of vessel is shown in figure 6. The total number is 2378 vessels of all types, while the real traffic is no more than 700 vessels around Greece. The passenger and high-speed vessels (294), cargo vessels and tankers (1130), pleasure crafts and sailing boats (80), fishing boats (50) and radar tracks (44) were the ships of interest (non-stationary) because most of the rules applied would track their behavior.

The evaluation of TRITON service took place using a laptop with 16GB of RAM and an i7-5500U CPU running at 2,4GHz with 4 logical dual core processors. Statistics about the CPU and Memory Heap utilization was performed using

the Java VisualVM. Also, the total number of threads running during the whole test, was recorded (64 threads allocated by the service).
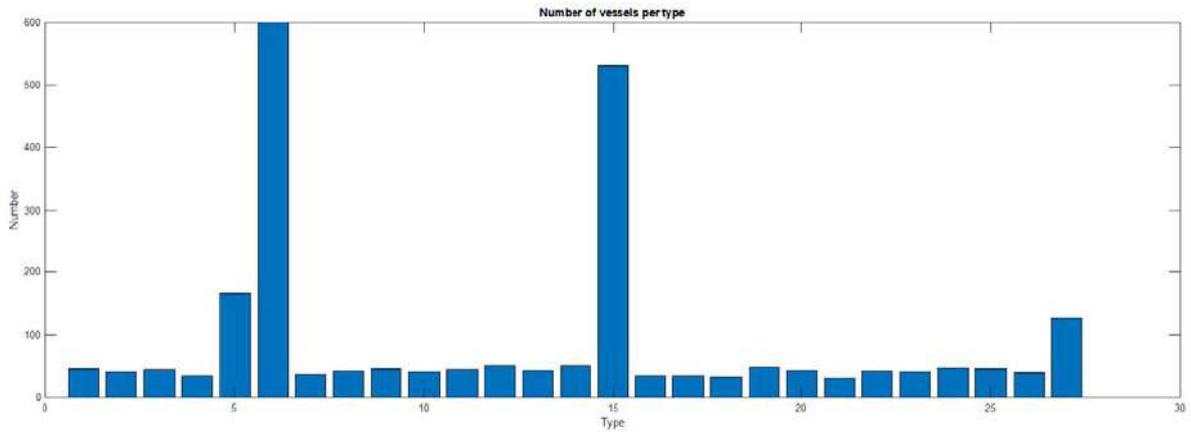
Figure 6: Total number of generated vessels per type. 1=pleasure craft, 2=not available, 3=radar, 4=vessel, 5=passenger, 6=tanker, 7=anti-pollution, 8=search & rescue, 9=tug, 10=port tender, 11=wig, 12=fishing, 13=ship No.18, 14=underwater operations, 15=cargo, 16=military, 17=sailing, 18=mobile offshore drilling, 19=pilot, 20=towing, 21=spare for assignment to local vessels, 22=law enforcement, 23=nuclear, 24=medical, 25=diving operations, 26=other, 27=high speed

The simulation lasted for 6 hours and 25 minutes. The CPU and memory utilization are shown in the next figures.
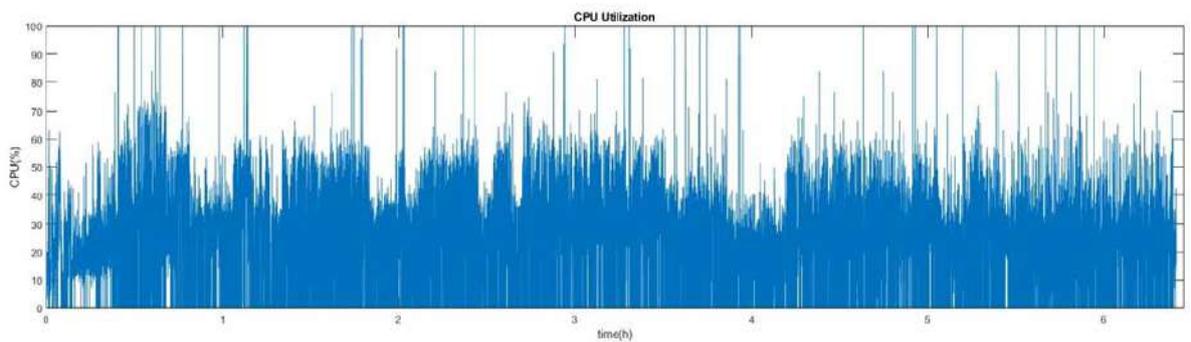
Figure 7: Usage of CPU
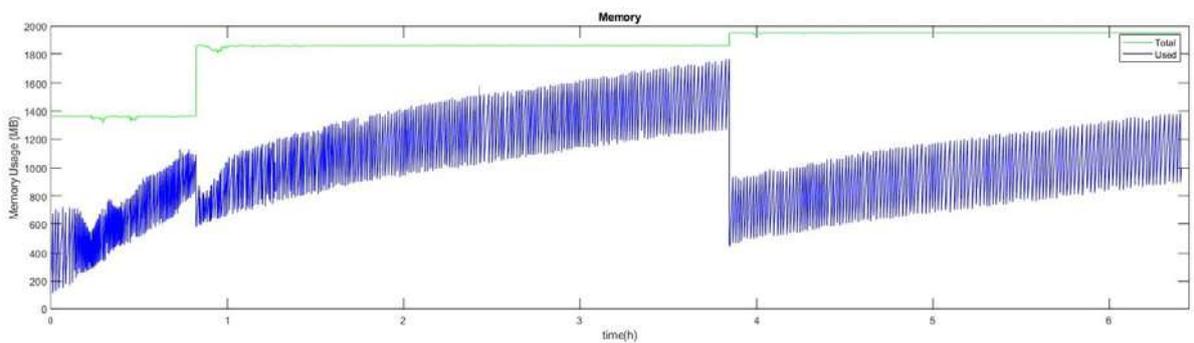
Figure 8: Total memory allocated for TRITON service (green) and memory used by the service (blue)

The average CPU usage was 22.4% and the standard deviation 18% during the whole time of the simulation. The average memory used by the service was 1GB and the maximum value was 1.7GB. When the peak value was reached, the garbage collector was called by the system and the memory usage dropped to 450MB and it continued to rise with the same rate as before, which means that the true need for memory is not so much big. From the figure is seen that if enough memory is pre-allocated at the beginning of the execution, then the rate of memory consumption is 1GB per 3 hours and then it resets when the garbage collector is called automatically.

Finally, in the next figure are displayed the generated events and alarms per minute. The total number of events created, was 1955548 and the total number of alarms 129768. The large number of alarms can be explained, because when configuring the rules, some of them were applied in the whole geographic area where the simulation took place and the snooze time (parameter for limiting alarm rate) was left to zero in order to stretch the limits of the engine.
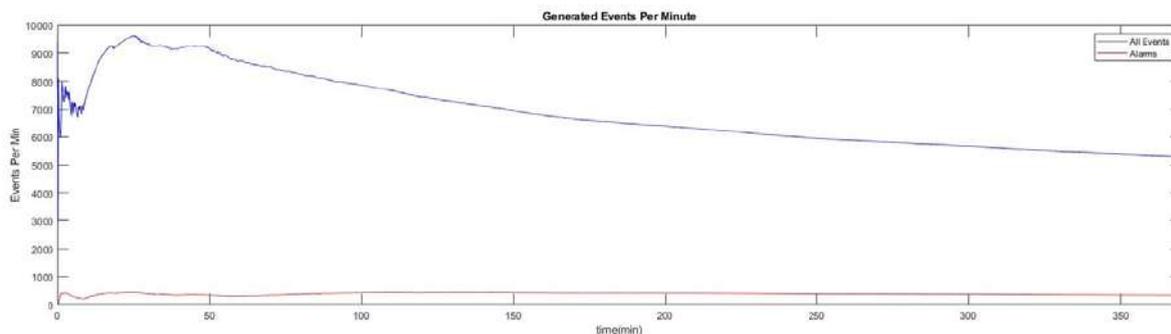
Figure 9: Rate of creation for all events (blue) and alarms (red)

In the above figure the maximum number of events per minute is 9639 and it happens at the beginning of the simulation when all vessels have been created. As time passes the rate is reduced, which is logical since a lot of vessels arrive at their destination and stop to transmit information. Knowing the total number of vessels, there is a 1:4 ratio of existing vessels to generated events.

TRITON service was able to handle the large amount of vessel traffic for the whole duration of the simulation. Even if the traffic (2700 vessels) was a lot higher than the average expected in Greece (~700 vessels), it was able to process events without any delays, keeping the CPU usage at low levels and efficiently handling the used memory.

4. CONCLUSIONS

The present work shows how the geospatial complex event processing paradigm can be used to build a robust, real time, with low latency, memory efficient and scalable service for abnormal vessel behavior detection. The service can recognize many patterns and it can be extended by adding new rules, when new anomalies should be detected. Thus, it is capable for supporting a wide range of maritime applications and to be a primary component of modern VTS platforms. TRITON has been tested with real and simulated data and has shown its efficiency in detecting anomalies. Future work includes the use of artificial intelligence for recognizing specific patterns using reinforcement and unsupervised learning. The work presented in this paper was partially funded by the ongoing EC H2020 Marisa (GA 740698) project.

REFERENCES

[1] Andersson, M. and Johansson, R., "Multiple sensor fusion for effective abnormal behaviour detection in counter-piracy operations," in Proceedings of International Waterside Security Conference, Carrara, IT (2010).
[2] Cugola, G., and Margara, A., "Processing flows of information: From data stream to complex event processing," ACM Computing Surveys 44(3), 15:1–15:62 (2012).
[3] T. Lu, X. Zha, and X. Zhao, "Multi-stage monitoring of abnormal situation based on complex event processing," Procedia - Procedia Comput. Sci. 96, 1361–1370 (2016).

[4] Tawsif, K., Hossen, J., Raja, J.E., Jesmeen, M. Z. and Arif, E., "A Review on Complex Event Processing Systems for Big Data," 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), 1-6 (2018).

[5] Terroso-Saenz, F., Valdes-Vela, M. and Skarmeta-Gomez, A. F., "A complex event processing approach to detect abnormal behaviours in the marine environment," Information Systems Frontiers 18(4), 765-780 (2016).

[6] EsperTech, "Esper Description," EsperTech, 2019, < http://www.espertech.com/esper/> (2019).

[7] Roy, J., Davenport, M., "Categorization of maritime anomalies for notification and alerting purpose," Nato Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness, 15-17 (2009)

[8] Biagiotti, L. and Melchiorri, C., Trajectory Planning for Automatic Machines and Robots, Springer-Verlag, Berlin Heidelberg, 166-172 (2008)

[9] Marisa, "Maritime Integated SurveillanceAwareness," Marisa Project, < https://www.marisaproject.eu/>.

[10] Navigation Center, "Ais Class A Ship Static and Voyage Related Data," U.S. Coast Guard, < https://www.navcen.uscg.gov/?pageName=AISMessagesAStatic>.

Detection of maritime anomalous behavior in a successful MARISA North Sea trial

Ali Mohamoud, Johan van de Pol, Eric den Breejen, Kim Veltman, Jos van der Velde, Tommaso Mannucci, Hanno Hildmann

TNO Defence, Security and Safety
Oude Waalsdorperweg 63, PO BOX 96864, 2509 JG, The Hague, Netherlands

Emails: {ali.mohamoud, johan.vandepol, eric.denbreejen, jos.vandervelde, kim.veltman, tommaso.mannucci, hanno.hildmann} @tno.nl

ABSTRACT

Maritime situational awareness (MSA) is of paramount importance for defense as well as civil authorities. We report on MARISA (MARitime Integrated Situational Awareness), a H2020 security project which was conducted to create improved situational awareness with a focus on delivering a complete and useful comprehension of the situation at sea. Specifically, behavioral analysis was used to narrow down the number of vessels in an environment and thereby help optimize decision making processes of the operational end-user. Vessel behavior was analyzed in real time with the actual sensor data of the Netherlands Coast Guard in relation with contextual information. In this paper we introduce the operational maritime scenario and the MARISA Toolkit application. Furthermore, the rule-based behavioral analysis (RBBA) service used for anomaly detection is presented and the implemented automated reasoning approach (based on Complex Event Processing (CEP) over input streams such as e.g., vessel motion, vessel metadata and contextual information) is discussed. The MARISA Toolkit configuration containing the RBBA service has been validated through a successful live trial in the North Sea, conducted under real operational conditions.

Keywords: maritime, surveillance, safety, anomaly detection, situation assessment, behavioral analysis, threat analysis

1. INTRODUCTION

Coast Guard organizations have nowadays detection systems at their disposal collecting large amounts of data that can be used for the broader maritime security and safety and in particular fisheries, illegal smuggling of goods, illegal trafficking, piracy and terrorism. It is of vital importance to correlate and fuse the large amount of heterogeneous data to extract relevant information for maritime surveillance practitioners creating enhanced maritime situational awareness for efficient decision making and response capabilities [2]. This work deals with a rule-based behavioral analysis (RBBA) service that analyses vessel behavior in relation with contextual information and detects anomalies. The technology makes use of automated reasoning based on Complex Event Processing (CEP) by processing input streams (e.g. vessel trajectories, vessel metadata and contextual information) using rules predefined by subject matter experts. The paramount goal of the behavioral analysis is to narrow down the number of vessels that are of interest in the huge amount of data and thereby help optimize the use of analytical capacity and the decision making processes of the operational end-user. The behavior analysis service – as part of the MARISA toolkit configuration for the North Sea – has been validated in a successful operationally relevant, live trial in the North Sea executed within the MARISA EU project. During the exercise, operationally known as *MARISA Alert*, the Netherlands Coastguard (NLCG) instructed three ships to sail anomalous patterns in maritime security and safety domain such as illegal diving activity near a ship wreck and transfer of contraband at sea. The behavior analysis service, connected to a live feed of the Coastal Surveillance System successfully captured, processed, analyzed & produced alerts in real time for both illegal diving and transfer of contraband use case without any false alarms. The NLCG operators underscored the added value and operational relevance of RBBA service's anomaly detection for early warning capabilities and support in rapid decision making.

2.   BACKGROUND AND CONTEXT OF THE MARISA PROJECT

Maritime Situational Awareness (MSA) is of paramount importance for defense as well as civil authorities. MSA is supported by the recognized maritime picture: acquired through correlation and fusion of data and information from land, sea, airborne, and satellite sensor systems, augmented with heterogeneous information from Geographical Information Systems (GIS) and vessel information databases. Detection and identification of all vessels, both cooperative and non-cooperative, is a key prerequisite for a successful and enhanced situational awareness of the maritime domain. The paramount purpose of MARISA project is to create improved situation awareness with a focus on delivering a complete and useful comprehension of the situation at sea. In doing so support for the practitioners can be ensured along the complete lifecycle of incidents at sea, from the observation of objects in the environment up to detection of anomalies and efficient planning of assets.

The RBBA service and its subsequent validation in the MARISA North Sea trial presented in this paper analyses in real time the motion patterns of vessels in relation to contextual data such as: vessel type and information, environment, and interaction with other objects and vessels. Subject matter expert can easily be integrated into the RBBA service approach as well as other relevant sources such as vessel databases, Electronic Navigational Charts (ENC), ship wreck data, historical data, risk maps and white-/blacklists.

2.1 Available data sources

The primary source of data for the MARISA North Sea Trial is the fused real-time Automatic Identification System (AIS) and Vessel Traffic System (VTS) radar tracks – also known as Inter VTS Exchange Format (IVEF) data. In addition, historic AIS data for the MARISA Alert exercise area is also used to create density maps. As additional contextual information the ENC chart for the Dutch Exclusive Economic Zone (EEZ) and information about ship wrecks in the North sea were used. In order to run the RBBA service in the MARISA toolkit on live IVEF data stream with vessel tracks, TNO developed and implemented an IVEF adaptor.

2.2 Approach and Methodology

The afore-mentioned RBBA service is part and parcel of the broader MARISA toolkit – the outcome of the H2020 MARISA project. The paramount purpose of MARISA is to create improved maritime situational awareness with a focus on delivering a complete and useful comprehension of the situation at sea. In doing so, support along the complete lifecycle of incidents at sea, from the observation of objects in the environment up to detection of anomalies and efficient planning of assets can be ensured for the practitioners. The MARISA project envisages to achieve the enhanced maritime situational awareness through a toolkit that provides a suite of services to correlate and fuse various heterogeneous data and information from different sources, including open sources. The MARISA toolkit builds upon the huge potential that arises from using the open access to "big data" for maritime surveillance: the availability of large to very large amounts of data, acquired from various sources ranging from sensors, satellites, open source, internal sources and extraction of valuable information from these amounts through advanced correlation, fusion and predictive analysis.

2.3 Validation

The MARISA project is a trial-based project where the developed toolkit can be verified and validated in a variety of operational trials. Effective verification and validation of the MARISA toolkit is ensured through active involvement of the user community; identification and definition of relevant scenarios, execution of real live exercises, setting up performance metrics and indicators, and validation of performance in the evaluation of operational trials. The work presented in this paper has firstly been verified in a simulation environment where realistic maritime traffic and anomalies have been generated and injected into the service. Validation of the work in a relevant environment has been done in MARISA North Sea trial hosted by the NLCG. Realistic relevant scenarios have been used with coastguard assets mimicking anomalous behavior for validation purposes. The validation of the RBBA service in detecting the anomalous behavior of vessels without any false alarms has been successfully carried out during the MARISA North Sea trial.

3.   RULE-BASED BEHAVIORAL ANALYSIS SERVICE

3.1 RBBA service concept

The RBBA service enables the analysis of vessel behavior to narrow down the potential number of ships of interest. Rule-based reasoning approaches appear especially well-suited for developing expert systems that reason about

situations of interest based on subject matter expert knowledge to assist decision makers [1]. The RBBA service analyses (in real time) the trajectories of vessels in relation to contextual data such as: vessel type and information, the environment, and interaction with other vessels. The service processes the input streams (e.g. vessel trajectories, vessel information and environment information) using predefined rules. The service's concept as depicted in Figure 1, uses Siddhi Complex Event Processing (CEP) for processing streams of vessel data in order to detect suspicious vessel patterns in space and time. The service contains two types of applications which form its building blocks: 1) predefined applications and 2) dynamic or configurable applications (also known as rules). The service deploys these applications for detecting abnormal vessel behavior and manages them through its application management component.
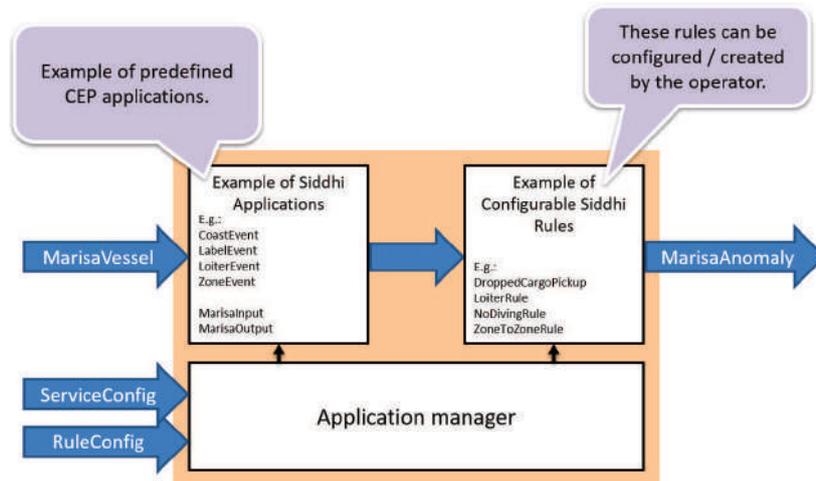
Figure 1. Overview over the RBBA service concept

The RBBA service reasons over user defined rules in order to detect maritime behavioral patterns that vary from vessel to vessel, vessel to area and vessel to object, and provides the operational user with an automatic, real-time alerts. A number of rules have been implemented in the RBBA service. The most prominent rules that have been implemented are:

- Loitering – triggers an alert when a vessel stops for a predefined time in an area of interest.
- No diving – alerts the user when a vessel loiters inside a no diving zone for a given time.
- Zone to zone – gives an alert when any vessel moves from one predefined zone to another predefined zone.
- Vessel to zone – gives an alert when any vessel moves to a predefined zone.
- Zone to vessel – gives an alert when any vessel moves from a predefined zone into the vicinity of a specific vessel.
- Dropped cargo – gives an alert when a small vessel loiters in an area within a specific time and distance after a large vessel has been there.
- Speed deviation – alerts the user when the speed of a vessel deviates from normal speeds in a given location
- Course deviation – alerts the user when the course of a vessel deviates from normal courses in a given location.

The RBBA service is designed in a modular way so that additional rules can easily be added and integrated into the service. Furthermore rules can be combined to address complex operational scenarios.

3.2 RBBA service configuration

The RBBA service can be set up in two ways; 1) service configuration – predefined applications have a few configuration options which can adjusted through the MARISA service configuration interface and 2) rule configuration – dynamic applications can be instantiated by the operator through the available MARISA rule configuration interface. The operator can configure and deploy those applications as rules in order to find the events and anomalies which are of interest. Each rule may be modified and deployed multiple times with different settings; The application manager handles all settings related tasks and handles the deployment of (new) Siddhi applications.

4. VALIDATION AND EVALUATION

4.1 The MARISA North Sea Trial

The MARISA North Sea trial is among the 5 operational trials planned to verify and validate the MARISA toolkit. The main objective of this trial is to validate a decision support tool fusing heterogeneous vessel information, assessing of situations, detecting anomalies and threats to support effective use of assets for risk mitigation and interdiction. The MARISA North Sea Trial involved; 1) the successful organization of the trial in terms of defining requirements from the user needs' point of view and creating operationally relevant scenarios; 2) the selection of suitable area of interest with relevant maritime traffic, where the live exercise can be executed and 3) the suitable configuration of the MARISA toolkit for the North Sea trial

4.2 Area of interest and scenarios

The trial took place in the North Sea (Figure 2), above the Netherlands Waddenzee and near the islands of Terschelling and Vlieland. It is a very busy area just South of the "Terschelling TSS" (Traffic Separation Scheme), within the Netherlands EEZ. The trial was organized in close collaboration between NLCG and TNO.
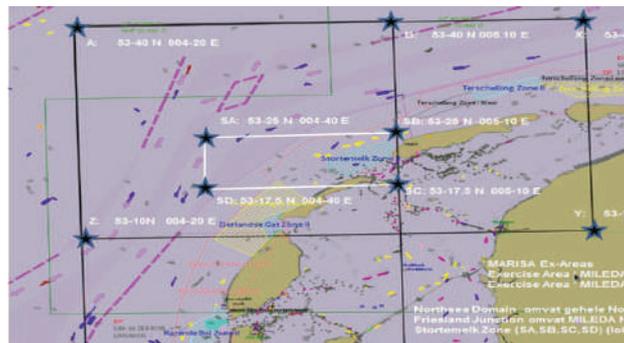
Figure 2. The MARISA North Sea trial execution area.

Resulting from the active engagement with the NLCG with regard to defining user needs and requirements analysis on the MARISA RBBA service, two operationally relevant scenarios for NLCG have been defined [4];

1) illegal diving activity – a scenario about a vessel moving from a fishing area, going to the vicinity of a ship wreck in the North Sea, loitering for a given time within a given distance from the wreck.
2) transfer of contraband at sea activity – a scenario about a vessel leaving a fishing area, crosses the traffic lane, passes a cargo vessel at a short distance and comes in the wake of the cargo vessel, slows down and even stops for a moment in the middle of the traffic lane to replicate the retrieval of 'cargo'.

4.3 MARISA Alert – the exercise

The MARISA North Sea Operational Trial has been code-named as MARISA Alert and was hosted by NLCG on 12th and 13th of September 2018 in Den Helder. MARISA Alert involved three ships of the NLCG: the watch ship "Guardian", patrol ship "Visarend" and support ship "Terschelling". The ships were instructed to sail 'anomalous patterns' (cf. Figure 3).
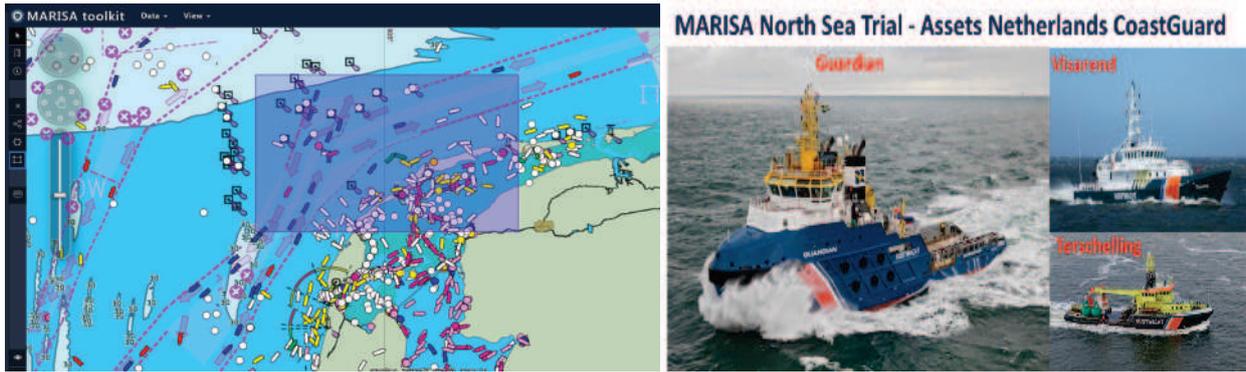
Figure 3. Execution of MARISA Alert. The MARISA exercise Area of Interest with live vessel data (left), NLCG vessels used during the MARISA alert (right).

4.4 MARISA toolkit configuration for the North Sea

For the North Sea trial, a specific MARISA toolkit configuration has been setup based on its operational relevance for the NLCG. The North Sea Trial demonstration was set up in the NLCG back-up operations facility in Den Helder where a live IVEF feed was injected to the Toolkit in real time. The IVEF stream is made available to the Toolkit by a MARISA IVEF adaptor developed by TNO, and is processed and analysed by behavioral and anomaly detection services and visualized on the MARISA Human Computer Interface (HCI). The RBBA service as part of the MARISA toolkit connects to MARISA IVEF adaptor to process the live vessel stream through the Big Data Infrastructure.

Figure 4. The MARISA toolkit demonstration in the NLCG backup facility.

4.5 The trial results and discussion

A dedicated MARISA toolkit configuration was deployed at the NLCG headquarters in Den Helder, Netherlands. There were two operators manning the toolkit during the trial – a MARISA consortium member and an operational end user – who used the MARISA toolkit in parallel with the operational backup system of the NLCG manned by two other NLCG operators for comparison and evaluation of the added value of MARISA RBBA service.

During the exercise, operationally code-named as MARISA Alert, the Netherlands Coastguard (NLCG) instructed three ships to sail anomalous patterns in maritime security and safety domain such as illegal diving activity near a ship wreck and transfer of contraband at sea. The RBBA service within the MARISA toolkit configuration was connected to a live IVEF feed. The service successfully captured and processed the incoming vessel stream, carried out vessel track analysis, generation of behavioral attributes based on user input through rule and service configuration, correlation with contextual information such as ship wreck database and ENC. The RBBA service was able to produce alerts in real time for both, the illegal diving activity as well as the transfer of contraband scenario without any false alarms. The NLCG operators underlined the added value and operational relevance of RBBA service's anomaly detection for early warning and rapid decision making capabilities. During the scenarios sailed by the instructed Coast Guard vessels the implemented rules in the RBBA service (created with expertise knowledge of the operators) detected the anomalous behaviors timely. The alerts generated were clearly understood by the operators. The presentation of the alerts on the screen of the HMI shows besides an icon on the location on the map additional information on the anomalous event.
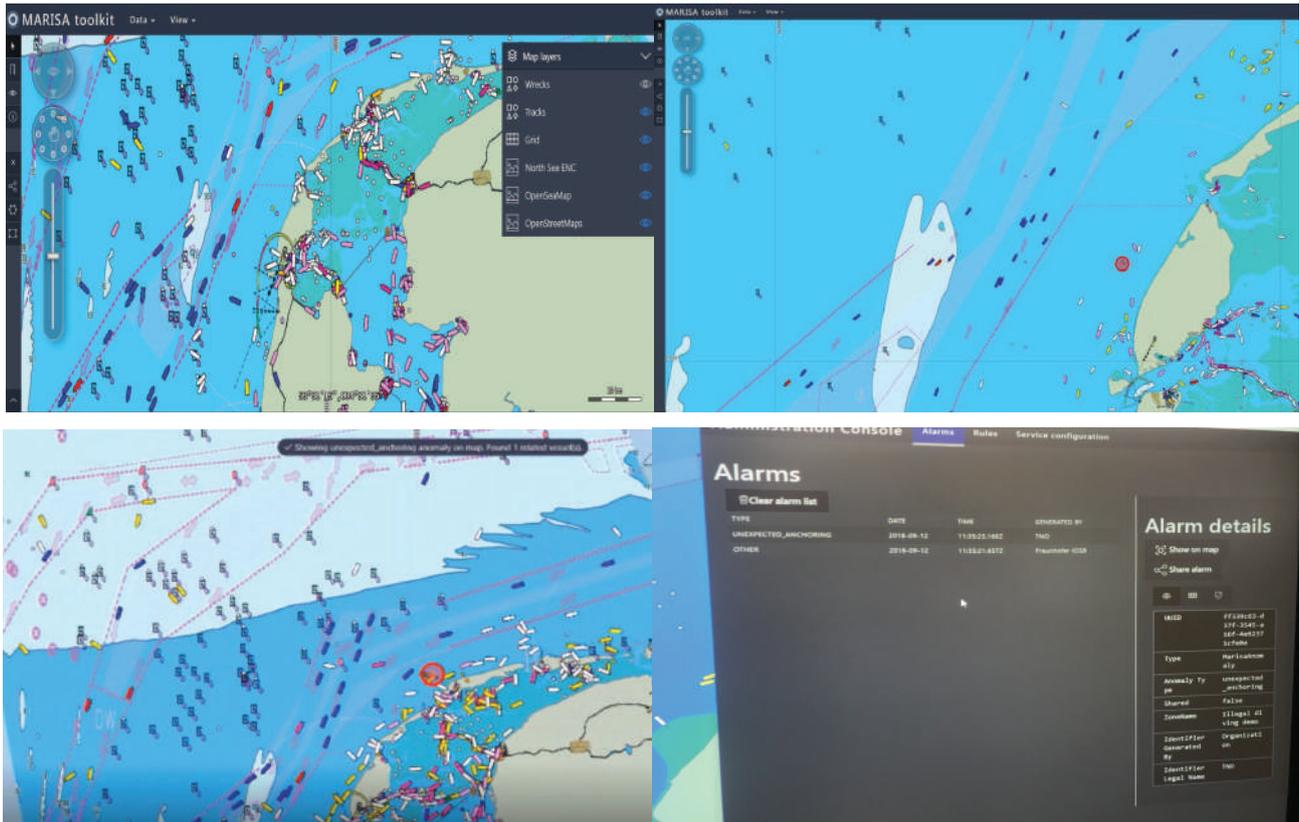
Figure 5. The MARISA toolkit HMI with RBBA service anomaly output.

REFERENCES

[1] Roy, R. "Automated Reasoning for Maritime Anomaly Detection", (2009), NATO Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness, 15-17 September 2009, La Spezia, Italy.

[2] Smith, A.J.E., Bergmans, J., "Overview of maritime situational awareness research at the Netherlands Organization for Applied Scientific Research TNO", (2009), NATO Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness, 15-17 September 2009, La Spezia, Italy.

[3] Dekker, R., Breejen, E., Mohamoud. A., "Maritime Situation Awareness Capabilities From Satellite And Terrestrial Sensor Systems", (2013), MAST (Maritime Systems and Technology) Europe Conference 4-6 June 2013, Gdansk, Poland.

[4] Smith, A.J.E., L. Anitori, A. Gabban, H. Greidanus, and J. Mallorqui, "Inference of vessel behaviour in the Dutch coastal North Sea", (2008), Fraunhofer Symposium Future Security, 10-11 September 2008, Karlsruhe, Germany.

Achieving maritime situational awareness using knowledge graphs: a study

Jacques Everwyn[a,b], Bruno Zanuttini[a], Abdel-Illah Mouaddib[a],
Sylvain Gatepaille[b] and Stephan Brunessaux[b]

[a]Normandie Univ., UNICAEN, ENSICAEN, CNRS; GREYC, 14000 Caen, France;
[b]Airbus Defence and Space, Elancourt, France

ABSTRACT

Currently, maritime surveillance operators have to monitor by hand the massive amount of data at their disposal to spot the events of interest, thus limiting their capabilities. Maritime data comes from various and heterogeneous sources that can be merged into a dynamic attributed knowledge graph which represents an evolving maritime situation. Using this graph, the automation of alert rising comes through a link prediction task: given some labels from expert knowledge, are there similar situations of interest elsewhere in the graph? In this article, we review link prediction techniques for situation awareness in a maritime context, and draw conclusions on how the addition of attributes in a dynamic graph model could improve results on this task.

Keywords: Dynamic knowledge graph, maritime situation, attributes, machine learning, link prediction

1. INTRODUCTION

The maritime domain is the theater of many unlawful activities that may go unnoticed: terrorism, piracy, smuggling, illegal immigration... For this reason, Maritime Situational Awareness (MSA) is of first importance to maritime security. It is defined by NATO as "The understanding of military and non-military events, activities and circumstances within and associated with the maritime environment that are relevant for current and future NATO operations and exercises where the Maritime Environment (ME) is the oceans, seas, bays, estuaries, waterways, coastal regions and ports".[1] MSA is often performed by surveillance operators who monitor the flow of data coming from maritime activities. This data is diverse, heterogeneous and coming from several sources: AIS (Automatic Identification System), radars, satellites, intelligence, websites... With more than 50.000 vessels sailing the oceans each day, there is a need for automation in the detection of illicit events.[2]

A maritime situation implies evolving entities: vessels, ports, countries... Such a situation can be represented by a *dynamic attributed* knowledge graph (DAKG), and understanding how its elements connect and jointly evolve gives valuable information pertaining to MSA. This task is here reduced to a link prediction problem. A *link*, or an *event*, is a relation between two entities at a given time point, for instance (Titanic ; :builtBy ; WhiteStarCompany ; 1909), and *attributed* means that entities have attributes whose values may change over time, e.g. (Titanic ; :passengers ; 2,344 ; April 10th 1912).

Generally, link prediction is performed by learning an embedding for each entity of the graph and predictions are made by ranking the events in the graph using these embeddings. This benefits to MSA in two ways:

- *data completion:* when monitoring an operational situation, the sensors and intelligence services do not always have all the needed information at their disposal. Using link prediction, missing data can be inferred to improve MSA;

- *automated alerts:* link prediction will discover events that a human operator would not have noticed in the massive dataset. Illegal activities could also be anticipated by making prediction in the future and evaluating the risk a ship represents based on its current and past behavior.

In this article, we review (1) two models on a dynamic (but not attributed) knowledge graph, (2) the literature on static/dynamic/attributed knowledge graphs, (3) how to apply DAKGs to MSA.

Further author information:
a: firstname.name@unicaen.fr
b: firstname.name@airbus.com

2. PREVIOUS WORK

The previous work related to this study can be broadly divided into four categories: maritime related work, static graphs, dynamic graphs and attributed graphs.

2.1 MSA

MSA often focuses on anomaly detection.[3] It can be tackled with clustering,[2] bayesian networks[4], self-organizing maps[5] and many others techniques[6]. Route estimation is also handled, e.g. with neural networks[7] or Extended Kalman filter.[8] These strategies perform well on trajectory analysis but do not take the whole context of a situation to detect an anomaly or a threat. To the best of our knowledge, this is the first attempt of using link prediction on DAKG to improve MSA.

2.2 Static Knowledge Graph

In a static setting, each node is represented by a single vector. This field is largely covered with a broad range of techniques. Translational models evaluate a fact by measuring the distance between the two entities, generally using the relation during the translation. TransE[9] is its most known representative. Semantic matching models are similarity-based and compare the latent semantics of entities and relations embeddings. RESCAL[10] was the first to do this and has been extended multiple times.[11,12] Neural network architectures have also been tried with NTN[13] or VGAE.[14] These models achieve great performances on static knowledge graphs but are not suited to deal with dynamic ones.

2.3 Dynamic Knowledge Graph

In a dynamic setting, each node is represented by a time series of vector modeling its evolution. This topic is emerging and has less contributions but advances have already been made. Leblay et al.[15] predict time validity for unannotated edges using side information in the learning process. Esteban et al.[16] update the knowledge graph using an event graph to add new information, and Trivedi et al.[17] extend the bilinear model (RESCAL) with a LSTM network in order to learn non-linearly evolving entities. Jiang et al.[18] incorporate the valid time of facts using a joint time-aware inference model based on Integer Linear Programming. Self-attention networks were tried by Sankar et al.[19]

Although these models are time-aware, they do not include attribute information in the relation prediction task and we will show that they are needed when dealing with MSA.

2.4 Attributes

KR-EAR[20] can predict discrete attribute values and find correlation between them. However, they are not included during the learning of relations and relations are not included in the learning of attributes.[21] propose a model that jointly learns KG^R and KG^A with a neural network and predicts continuous values with a regression task. However, neither model deals with temporal data.

Li et al.[22] propose a streaming model (SLIDE) on dynamic attributed networks using a sketching matrix that summarizes the currently observed links and node attributes. They review the challenges pertaining to such networks and real-world data, but they apply it on social networks (Epinions, DBLP, ACM) that have very different kinds of attributes and only a few widely separated timesteps (~20 timesteps from one month to one year each). All these models showed that the addition of attributes improves the results on link prediction.

Table 1 compares representative models from each category and shows that no model currently fits our needs perfectly (see Part 4 for detailed requirements).

	Static	Dynamic	Attributes	Near Real-Time
TransE[9]	X			
Know-Evolve[17]		X		
MT-KGNN[21]	X		Continuous	
SLIDE[22]		X	Discrete	~

Table 1: Application domain of models comparison

3. PROBLEM STATEMENT

In this section, we define a knowledge graph structure to represent a maritime situation.

3.1 Knowledge graph

DEFINITION 1 (FRAME). *A frame is a quadruple $F = \langle E, R, A, D \rangle$ where E, R, and A are finite sets of elements called* entities, relations, *and* attributes, *respectively, and $D : A \to S_a$ is a function assigning a* range $D(A)$ *to each attribute and S_a is a set of possible values for $a \in A$ (discrete or continuous).*

DEFINITION 2 (DYNAMIC ATTRIBUTED KNOWLEDGE GRAPH). *Let $F = \langle E, R, A, D \rangle$ be a frame. A* standard knowledge graph *on F is a couple $KG = \langle KG^R, KG^A \rangle$, where*

- *KG^R is a finite subset of $E \times R \times E \times \tau$, with τ the set of time points,*

- *KG^A is a finite subset of $E \times A \times D(A) \times \tau$ such that for all quadruples $(e, a, v, t) \in KG^A$, $v \in D(a)$ holds.*

For $KG = \langle KG^R, KG^A \rangle$, KG^R is called the relational *part of KG, and KG^A is called its* attributional *part.*

Intuitively, (e^s, r, e^o, t) is read "entity e^s is in relation r with entity e^o at time t", and (e, a, v, u) is read "entity e has value v for attribute a at time u". Given a knowledge graph KG, we always write KG^R (resp. KG^A) for its relational (resp. attributional) part. Figure 1 is an example of the previously defined KG.
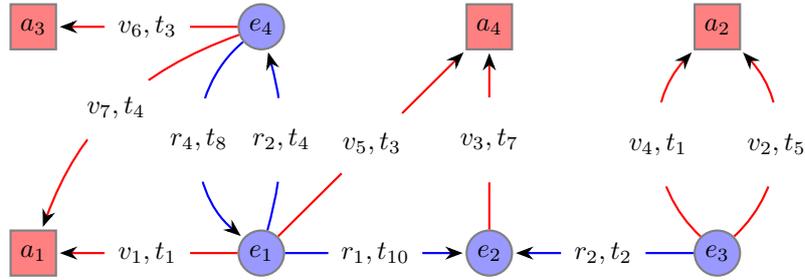
Figure 1: Example of knowledge graph KG on a frame $F = \langle E, R, A, D \rangle$. The nodes $e_i \in E$ are *entities* and the nodes $a_j \in A$ are *attributes*. The *relations* $r_n \in R$ are annotated on edges between two entities. The *values* v_k, annotated on the edges between two attributes, belong to the *domain* $D(A)$ of the attribute they are attached to, and the $t'_l s$ are the *timestamps* of the edges. Attributes of entities can change their value over time (e_3 and a_2) and two entities can have common attributes (e_4, e_1 and a_1) but not necessarily with the same value. The blue square nodes and edges are KG^R and the red edges with all the nodes are KG^A.

3.2 Prediction problem

We are interested in predicting the missing relations between entities and values of attributes in knowledge graphs. We focus on the case where for some timestamp, they can be predicted from the values of a subset of the relations and attributes at previous timestamps. In this article, we experiment attribute prediction (position) from past attributes.

4. APPLICATION TO MSA

"Real-world" datasets often have more constraints than the academic ones because of their specificities. Maritime datasets are no exception and the following challenges must be overcome.

- *Evolution of attributes*: a maritime situation is a fast evolving world with very little time between two events. In MSA, a good evolutionary model is needed for change detection and the granularity depends on the task. For a change in the position/course/speed of a vessel (dynamic attributes), the information must be given within minutes (e.g. rapid response needed in case of piracy). But to detect a change in a vessel particulars (identifier, name...), the granularity needed can be in hours or days.

- *Event and threat detection*: an event (or quadruple) represents a new relation between two entities or an abovementionned attribute evolution. Using knowledge graphs and machine learning, it could be possible to find events using latent features that cannot be perceived by a human or a rule. If event mining detects facts, threat detection is a task highly related to its context and definition. A nation will not consider a transhipping between two fishing vessels as a threat since they are more likely to exchange fish than warheads, but an NGO for ocean conservation can suspect illicit fishing of an endangered species. Performing this task still requires either expert knowledge or labeled events.

- *Streaming*: MSA requires a constant monitoring of maritime areas, meaning that the model must deal with a continuous flow of data.

- *Uncertainty*: maritime data often results from hard (sensors) and soft (websites, intelligence) data fusion. However, this data is not always 100% certain: an intelligence report may have a typo, sensors have a range and precision (e.g. +/- 500 meters), or collisions may happen when satellites receive signals.

- *Explainability*: link prediction models are often black boxes when it comes to the origin of the prediction. However, a surveillance operator needs to know why a prediction was made in order to understand it and justify any upcoming response to an event. Because operators still do not trust AI-based systems to take decisions, explainability is needed to take DAKG-based decisions for MSA.[30]

An illustration of all these concepts can be found in Figure 2.

5. EXPERIMENTS

In these experiments, the performed task is the prediction of the position of a vessel in the next time points. Obviously, there are many better fitted methods to do this (like regression or a Kalman filter), but the ultimate goal is to use the full capacities of the knowledge graph i.e. exploit all the relationships, events and attributes in the maritime surveillance ecosystem to perform better link predictions. Position prediction is just a reduction of this task to test knowledge graphs capabilities on MSA. As we could not find a method handling both time and attributes that can be tested on our data, the attribute *:location* is replaced by a relation *:isLocatedIn* between a vessel and an area.

5.1 Dataset

In the absence of publicly available maritime knowledge graph, we created our own in order to evaluate the models.

5.1.1 AIS data

The dataset used in our experiments is based on real maritime data: AIS messages transmitted by vessels. AIS is a short range (37-74km) ship-to-ship and ship-to-shore navigational data exchange system. It is currently the main source of information available in support of maritime surveillance. The satellite version of AIS (S-AIS) gives a broader range (~5000km) but the transmissions are less regular and more subject to signal collision.[2] AIS provides the following non-exhaustive list of information about ships: the unique identifier of the vessel (called MMSI), its longitude/latitude, its speed and course, the timestamp of the report, the type of ship, the destination.
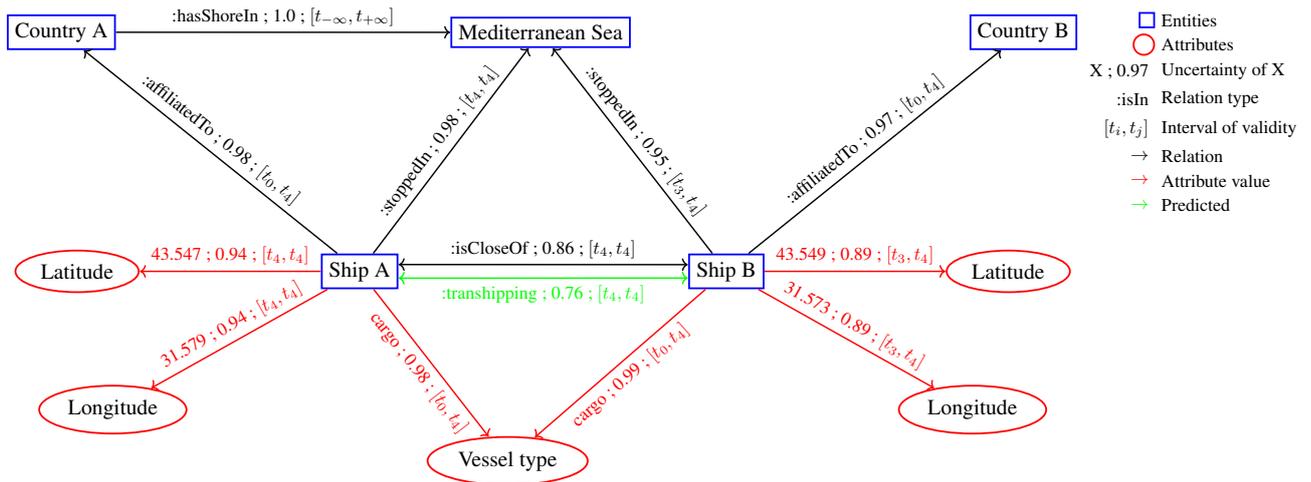
Figure 2: Example of KG^* at time t^* (best viewed in color). Events in red represent KG^A and the other events KG^R. Ship A and Ship B are vessels from different countries, moving in the Mediterranean Sea. They have a static attribute (vessel type) setting them as cargos and dynamic attributes (latitude and longitude) revealing their positons. Before t^*, the two vessels were moving in the Mediterranean Sea and had different positions, but now (t^*) they have stopped and are close to each other. A possible link prediction from $KG^{<t^*}$ is that the two vessels are performing transhipping. Note that the *:isStopped* relation can be deduced from the *speed* attribute going to zero, not represented here for the sake of clarity. If relations such as CountryA having a shore on the Mediterranean Sea are 100% sure, some are more uncertain: the position of Ship B is only 89% sure because the signal was picked up by a satellite in an area with high ship density. More, the uncertainty of predicted relations (*:transhipping*) depends on the uncertainty of the root events. Finally, to predict the transhipping action in time, the model must be updated with the root causes as soon as they are available, hence the need for streaming link prediction.

5.1.2 AIS to KG

A knowledge graph can be built using these AIS messages, where vessels are entities with attributes. Other entities can be added like nations (flag of the ship) or ports. However, the reviewed methods can only handle time, not attributes, hence the need to consider attributes as entities. In our work, the focus is on the evolution of the positions of vessels. Positions being continuous values, they need to be discretized to be casted as entities in the graph. Therefore, the studied area is converted into a grid made of 1km × 1km squares and each square is an entity (further referred to as "areas").

Moreover, AIS messages are on average received every three minutes so it can be a reasonable choice to separate each time point by three minutes, instead of having a time point every second as it happens in the data (different events can be attached to the same time point). Finally, as the chosen models can not always handle entities or relations not encountered during the training phase, the test set is filtered to remove any event involving an entity or relation not present in the train set. Note that only one relation type is considered here: "vessel :isLocatedIn area" ($|R| = 1$) and each event is represented by a quadruple $(e^s, r, e^o, t) \in KG^R$.

To summarize, we build the knowledge graph consisting of entities = {vessels, areas} and relation = {:isLocatedIn} over one month, we divide it into train/test sets and run the methods to predict the relation ":isLocatedIn" between vessels and areas.

The dataset covers the Gibraltar Strait from February 2nd, 2017 to March 2nd, 2017. It holds around 2.5k vessels and one million positions. The two evaluated methods are TransE[9] and Know-Evolve.[17] They use embeddings (vectors of real values) to represent entities and make link predictions.
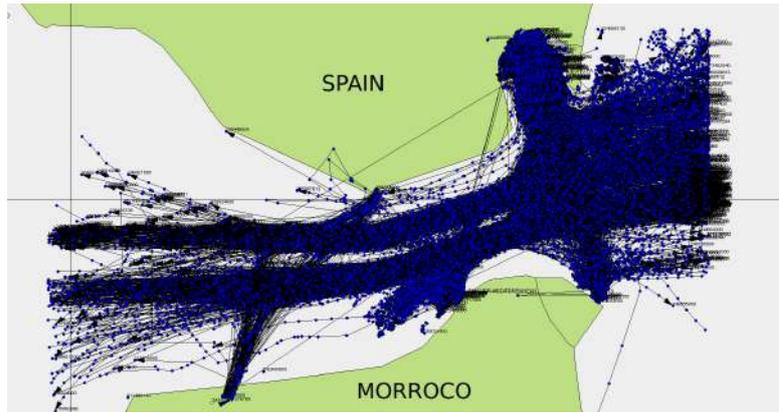
Figure 3: One week of maritime traffic in the Gibraltar Strait (best viewed in color)

5.2 Evaluation task

5.2.1 Link prediction

The evaluation is performed on the link prediction task: given a quadruple (e^s, r, e^o, t), e^o is replaced by every possible entity and the resulting quadruple is evaluated by the model. All the quadruples are then ranked in descending order of plausiblity and we record the Mean Average Rank (MAR) and the @Hits10 measure (one of the 10 best ranked quadruples is the true one). A lower rank means that the quadruple is classified better (the best rank being 1 and the worst the number of entities) and @Hits10 is expressed in percentage of correctly ranked quadruples i.e. higher is better. The filtering method of TransE[9] is applied, i.e. the quadruple is not ranked against corrupted (i.e. modified) quadruples that are true.

5.2.2 Sliding window evaluation

The performance is tested using the sliding window evaluation from Know-Evolve. We divide the test set into 8 different slides, each slide including one day of time (Know-Evolve uses 12 slides of two weeks each). This method is said to *"help to realize the effect of modeling temporal and evolutionary knowledge"*[17].

5.2.3 Static method on dynamic data

As it is a static method, the evaluation of TransE required some modifications of the dataset. All the timestamps t are removed and as a result, multiple occurences of the same triples (e^s, r, e^o) appear. Those are removed in order to have a unique representant for each triple and the dataset is then comprised of 102,470 (train) and 16,807 (test) events. The test set still only contains entities seen during training.
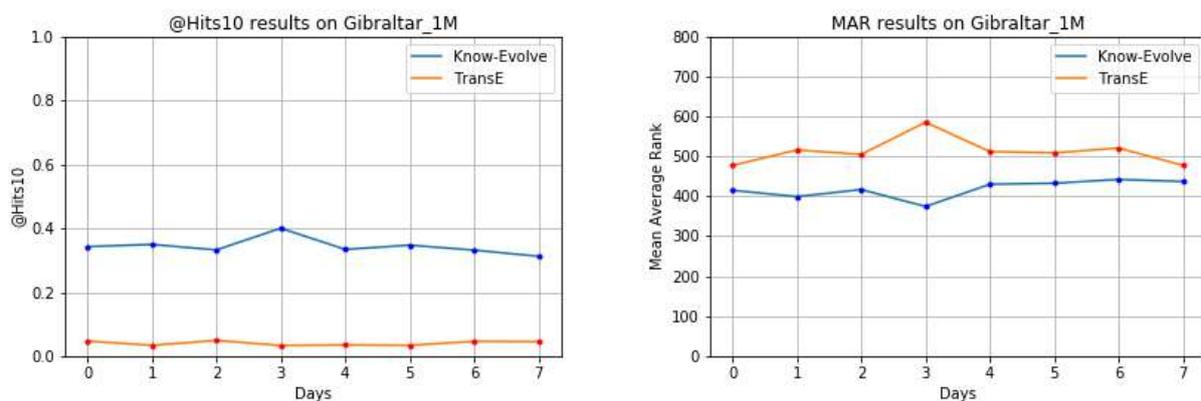
Figure 4: @Hits10 and MAR results of tested models

5.3 Results

5.3.1 Quantitative Analysis

Figure 4 show the results of the reviewed models over the Gibraltar1M dataset. Know-Evolve, being a temporal model, performs way better than TransE which struggle in @Hits10 prediction despite being not so far from Know-Evolve in Mean Rank. The reason is that TransE depend only on static entity embeddings to perform prediction. With an average @Hits10 of 34%, Know-Evolve captured the relationship between the vessels and the areas better than TransE but do not excels at the task.

5.3.2 Contextual analysis

This evaluation was made on a single task: predict in which area a vessel will be next. This task is made harder by the discretization of the positions: areas are independant and the graph does not tell if areas are close to each other or not. The only way to extract proximity is the analysis of a vessel's track (the succession of relations with area entities), meaning that two areas having a relation with a vessel in a short timespan may be close. More, a proximity relationship between two vessels in the same area could not be established because areas are too wide to consider two vessels as close (e.g. enough to perform an exchange of goods). At last, areas not seen in training cannot be predicted as next location due to the limitations of TransE. Know-Evolve somehow managed to find some connections between vessels and areas but the results are very unsatisfactory: a Mean Rank of 400 means that the correct area is on average ranked 400th, against MR = 20 on ICEWS.[17] Despite the difficulty induced by the discretization, position prediction is a simple task and the models performed poorly: they are not adequated to address this problem. The use of positions as continuous attributes could solve the abovementionned issues and improve the results on position prediction with knowledge graphs.

5.3.3 Experimental settings

We used the settings reported in the associated article[17] to run Know-Evolve. For TransE, we set batch size=200, learning rate = 0.001 and embedding dimension = 64.

6. CONCLUSION AND FUTURE WORK

In this article, we reviewed two link prediction techniques for a task: the evolution of the positions of vessels using a dynamic knowledge graph for Maritime Situational Awareness. We showed that relational data (KG^R) is not sufficient to modelize the movement of a vessel and that attributional information should be used (KG^A). We also exhibited the challenges that need to be overcome to apply DAKGs on MSA, and formalized the relation and attribute value prediction problem.

We foresee several tasks for future work: (1) make the prediction task more realistic by adding more entity and relation types in the dataset, such as ships going in and out of ports, or encounters between ships, (2) find a model that can handle both KG^R and KG^A for link and attribute prediction in a temporal setting, (3) perform threat and/or anomaly detection on DAKGs. These are the three requirements to fully evaluate the use of DAKGs on operational maritime data.

7. AKNOWLEDGEMENTS

REFERENCES

[1] NATO, "Mc msa draft definition," (2007).

[2] Le Guillarme, N. and Lerouvreur, X., "Unsupervised Extraction of Knowledge from S-AIS Data for Maritime Situational Awareness," *FUSION* , 8 (2013).

[3] Riveiro, M., Pallotta, G., and Vespe, M., "Maritime anomaly detection: A review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **8**, 1266 (Sept. 2018).

[4] Mascaro, S., Nicholso, A. E., and Korb, K. B., "Anomaly detection in vessel tracks using Bayesian networks," *International Journal of Approximate Reasoning* **55**, 84–98 (Jan. 2014).

[5] Riveiro, M., Johansson, F., Falkman, G., and Ziemka, T., "Supporting maritime situation awareness using self-organizing maps and gaussian mixture models," *Frontiers in Artificial Intelligence and Applications* **173**, 84 (2008).

[6] Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., and Huang, G.-B., "Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey," *arXiv:1606.00981 [cs]* (June 2016). arXiv: 1606.00981.

[7] Zissis, D., Xidias, E. K., and Lekkas, D., "Real-time vessel behavior prediction," *Evolving Systems* **7**, 29–40 (Mar. 2016).

[8] Perera, L. P. and Soares, C. G., "Ocean Vessel Trajectory Estimation and Prediction Based on Extended Kalman Filter," *2nd International Conference on Adaptive and Self-adaptive Systems and Applications* , 7 (2010).

[9] Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., and Yakhnenko, O., "Translating Embeddings for Modeling Multi-relational Data," 9 (2013).

[10] Nickel, M., Tresp, V., and Kriegel, H.-P., "A Three-Way Model for Collective Learning on Multi-Relational Data," 8 (2011).

[11] Nickel, M., Rosasco, L., and Poggio, T., "Holographic Embeddings of Knowledge Graphs," *AAAI* , 7 (2016).

[12] Liu, H., Wu, Y., and Yang, Y., "Analogical Inference for Multi-relational Embeddings," *ICML* , 11 (2017).

[13] Socher, R., Chen, D., Manning, C. D., and Ng, A., "Reasoning With Neural Tensor Networks for Knowledge Base Completion," *NIPS* , 10 (2013).

[14] Kipf, T. N. and Welling, M., "Variational Graph Auto-Encoders," *arXiv:1611.07308 [cs, stat]* (Nov. 2016). arXiv: 1611.07308.

[15] Leblay, J. and Chekol, M. W., "Deriving Validity Time in Knowledge Graph," in [*Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*], 1771–1776, ACM Press, Lyon, France (2018).

[16] Esteban, C., Tresp, V., Yang, Y., Baier, S., and Krompaß, D., "Predicting the Co-Evolution of Event and Knowledge Graphs," 8 (2016).

[17] Trivedi, R., Dai, H., Wang, Y., and Song, L., "Know-Evolve: Deep Temporal Reasoning for Dynamic Knowledge Graphs," *ICML* (May 2017).

[18] Jiang, T., Liu, T., Ge, T., Sha, L., Chang, B., Li, S., and Sui, Z., "Towards Time-Aware Knowledge Graph Completion," *International Conference on Computational Linguistics* , 10 (2016).

[19] Sankar, A., Wu, Y., Gou, L., Zhang, W., and Yang, H., "Dynamic Graph Representation Learning via Self-Attention Networks," *arXiv:1812.09430 [cs, stat]* (Dec. 2018). arXiv: 1812.09430.

[20] Lin, Y., Liu, Z., and Sun, M., "Knowledge Representation Learning with Entities, Attributes and Relations," *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI)* , 7 (2016).

[21] Tay, Y., Tuan, L. A., Phan, M. C., and Hui, S. C., "Multi-Task Neural Network for Non-discrete Attribute Prediction in Knowledge Graphs," 1029–1038, ACM Press (2017).

[22] Li, J., Cheng, K., Wu, L., and Liu, H., "Streaming Link Prediction on Dynamic Attributed Networks," in [*Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining - WSDM '18*], 369–377, ACM Press, Marina Del Rey, CA, USA (2018).

[23] Hoffart, J., Suchanek, F. M., Berberich, K., and Weikum, G., "YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia," *Artificial Intelligence* **194**, 28–61 (Jan. 2013).

[24] Boschee, E., Lautenschlager, J., O'Brien, S., Shell-man, S., Starz, J., and Ward, M., "Icews coded event data," (2015).

[25] Leetaru, K. and Schrodt, P. A., "GDELT: Global Data on Events, Location and Tone," *ISA Annual Convention* , 51 (2013).

[26] Roy, J., "Anomaly detection in the maritime domain," *Proceedings of SPIE - The International Society for Optical Engineering* (2008).

[27] Zhao, P., Aggarwal, C., and He, G., "Link prediction in graph streams," in [*2016 IEEE 32nd International Conference on Data Engineering (ICDE)*], 553–564, IEEE, Helsinki, Finland (May 2016).

[28] Rezende, D. J., Mohamed, S., and Wierstra, D., "Stochastic Backpropagation and Approximate Inference in Deep Generative Models," in [*International Conference on Machine Learning*], 1278–1286 (Jan. 2014).

[29] Chekol, M. W. and Pirro, G., "Marrying Uncertainty and Time in Knowledge Graphs," *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* , 7 (2017).

[30] Adadi, A. and Berrada, M., "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access* **6**, 52138–52160 (2018).

Maritime Situation Awareness through Data Analytics, Machine Learning and Risk Assessment Based on Ship Trajectories

Felix Opitz, Camilla Mohrdieck, Kaeye Dästner
Airbus Defence and Space GmbH, Wörthstr.85, 89077 Ulm, GERMANY

ABSTRACT

Modern surveillance networks are able to provide trajectories of many different types of vessels and aircraft worldwide. For example, the (Satellite-) Automatic Identification System (AIS) and Automatic Dependent Surveillance – Broadcast (ADS-B) are widely used in maritime and air surveillance. Both are cooperative technologies. Sensor networks based on ground installations or mounted on airborne and space-based platforms deliver object trajectories independent of any cooperation. These surveillance systems enable the extraction of mid- and long-term trajectories of objects. To generate situation awareness, the trajectories need to be placed into the right context and the intentions of the tracked objects need to be estimated. Activity-based intelligence and the determination of patterns of life support this goal. They are increasingly important components of modern surveillance systems, because they allow taking full advantage of large volumes of gathered data. Promising technologies that are at the basis of these efforts are Data Analytics and Machine Learning. In this paper, we present both methods in the context of use cases that address i) clustering techniques to identify areas of interest and patterns of life, ii) supervised machine learning for ship type and activity classification, and iii) the generation of thematic heat maps to enhance situational awareness and assist routing and mission planning. Furthermore, these new data analytic techniques have to be integrated in existing near real time surveillance systems. In our presentation, we therefore also briefly address examples of new system architectures as well as a new software tools that are adapted to achieve this.

Keywords: maritime situation awareness, modern surveillance systems, data analytics, machine learning, patterns of life, thematic heat maps

1. INTRODUCTION

1.1 Sources of Trajectories

An object trajectory is a sequence of chronologically ordered positions of the object that are collected by cooperative or uncooperative sources. Cooperative sources include, the Automatic Identification System (AIS) which is an automatic collision avoidance and tracking system used by ships and by vessel traffic services (VTS) as well as coastal authorities [1], [2]. The necessity to carry AIS equipment onboard a ship, depends on the ship's size and mission. Vessels fitted with AIS transceivers can be tracked by AIS base stations located along coast lines or by satellite based AIS receivers (S-AIS). AIS equipped ships have an assigned mobile service identity (MMSI) that allows their unique identification. The position information is based on GPS receivers complemented by additional electronic navigation sensors, such as a gyrocompass or rate of turn indicator. Analogously, aircraft equipped with Automatic Dependent Surveillance-Broadcast (ADS-B) transceivers can be tracked by associating the aircraft's unique ICAO number with its reported positions.

However, both trajectory sources depend on the object's willingness to be cooperative. If this is not given, or if the object does not fall under the carriage requirements (e.g. small ships), then the transmitted data might be incomplete, or completely missing or spoofed. In these situations, non-cooperative trajectory sources need to be used in order to complement or even generate an object track. Examples of uncooperative sources include costal surveillance radars, Ground Moving Target Indicating (GMTI) radar and image sequences from airborne and space borne imaging sensors (e.g. optical/ infrared cameras, synthetic aperture radar (SAR)). Partitioning all measured object positions resulting from heterogeneous sources into trajectories requires dedicated tracking systems [3], [4], [5].

1.2 Processing Architectures for Trajectories

The processing of continuous long-time trajectories implies a novel approach to surveillance. While real-time surveillance systems focus on the detection, identification and tracking of objects with minimum possible time delays after signal detection, their integration with data analysis and machine learning techniques requires a big data architecture – the so called λ-architecture. In this architecture, data processing is spread over three different layers [6], [7], [8]:

• Batch Layer

• Serving Layer

• Speed Layer

The batch layer continuously receives and stores the raw trajectories. This may occur on a distributed filesystem and/or database, e.g. Hadoop HDFS or Cassandra [9]. Here, the trajectories have to be cleaned, transformed and presented such that different applications can access relevant information quickly [10].

Subsequently, trajectories are processed with methods related to data analytics and supervised or unsupervised machine learning. E.g. trajectories can be used to generate pure statistics – like heat maps. Or their patterns of life can be analyzed and visited areas or covered routes can be extracted to study an object's normal behavior and to separate it from abnormal behavior. If trajectories are attributed with further information related to the objects' type or activity, they can be used to train classifiers or predictors. This is supported by well-established tools like Spark [11] and a broad spectrum of machine learning libraries [12], [13], which enables distributed processing of large data volumes. These processing methods are done within the serving layer and do not necessarily occur in real time.
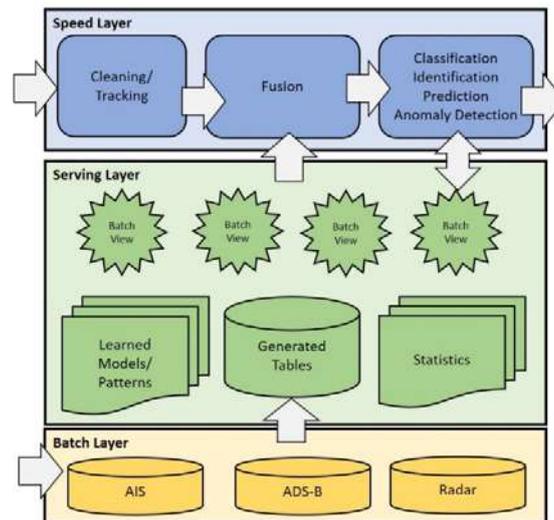
Figure 1. Lambda architectures

The real time aspects are considered in the third layer – the speed layer. Here, the learned patterns and classifiers/ predictors of the serving layer can be applied to the continual data stream to deliver real-time anomaly detection, predictive estimations, classification or even identification. For non-cooperative targets the speed layer possesses a tracking functionality, which can be used to aggregate single measurements to trajectories and to feed the batch layer.

Using this architecture allows a smooth integration of the big data and machine learning environment into already established and proven real time surveillance systems [8], [14]. Depending on the system's requirements, a real time system can be enriched with either batch and serving layer, or with an extraction of the serving layer containing trained models and batch views only. The latter approach is particularly suitable for systems used in the field that are often subject to hardware limitations.

2

2. UNSUPERVISED MACHINE LEARNING FOR TRAFFIC ASSESSMENT

2.1 Area of Interest Extraction by Point Clustering

There are several areas of interest a trajectory normally passes through. These could be e.g. sea ports, airports/heliports (for aircraft), specific areas associated with offshore drilling rigs, wind parks or fishing grounds. Plots in these areas are usually easy to extract due to their characteristics such as navigation status, low speed or high turn rate or simply by the fact that they are the origin or terminus of a trajectory. After extraction, these plots can be clustered. The convex hulls of these point clusters are candidates for areas of interest [15], [16], [17].
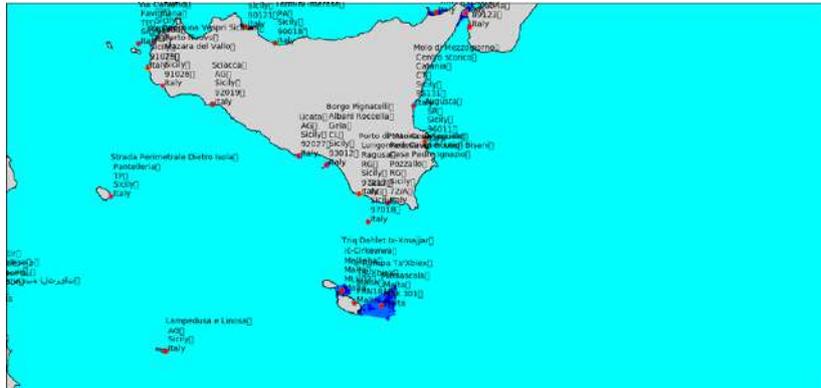
Figure 2. Sea ports (blue: convex cluster hulls, red: centers of clusters)

If the data used for clustering contains an identifier, e.g. MMSI, relations between objects and areas of interest can be extracted, such as the traveling behaviour between ports. This finally opens the door to data mining and relational graphs.
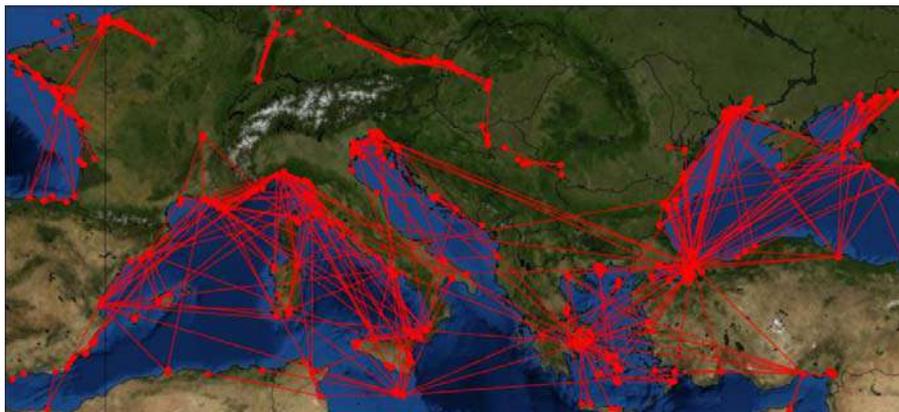
Figure 3. Trajectories between sea ports based on one-month AIS data

2.2 Pattern of Life Analysis

Besides the extraction of areas of interest, the clustering leads to the partition of trajectories into graphs, where the vertices (nodes) correspond to the areas of interest and the edges to the sub-trajectories with known starting and end node. This yields an inherent pattern of life analysis of the trajectories' objects. The patterns of life of objects are the basis for determining their usual or normal travelling behavior which constitutes the reference with respect to which any abnormal behavior is detected [7], [8], [18].

Figure 4. Ship routes between ports

3. SUPERVISED MACHINE LEARNING FOR CLASSIFICATION

Labelled data, e.g. AIS or ADS-B data, is typically used for supervised machine learning like classification topics. Each trajectory point has one or more labels, e.g. the ship type and navigation status for AIS and military flag and aircraft type for ADS-B [14], which can be learned to be associated with given features of the trajectory. These features are derived from the complete or windowed trajectories and are based on the object's positional and kinematical history. After training of a supervised machine learning algorithm in the serving layer, labels are predicted based on the positional or kinematical features of a trajectory. This is schematically depicted in   Figure **5** that shows a classification algorithm that is trained with given labels. The trained model is saved and serves as autonomous classifier in future applications where unlabelled data (e.g. radar data) is classified (i.e. the corresponding labels are predicted).
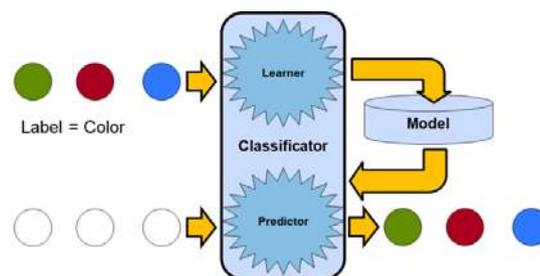
Figure 5. Principle use of a machine learning classifier

For classification tasks, a lot of classical machine learning algorithms are known, such as: Logistic Regression, Supported Vector Machine, Decision Trees, Ensembles of Decision Trees, e.g. Random Forest, Gradient Boost Tree and Multi-Layer Perceptron, which is the simplest form of a deep neural network [12], [13], [19]. The choice of a classification algorithm depends on the implementation constraints, e.g. integration environment, learning time, number of labels etc. as well as on the input data.

Figure **6** shows the result of a random forest classifier for five different ship types that was trained on AIS trajectories that were observed in the German Bight over a time period of 3 months [20]. From each trajectory roughly 20 discriminative features (static and dynamic ones) are manually extracted to generate feature vectors for the training procedure. After training, the classifier was used on a test set to predict the types of the test trajectories. The resulting confusion matrix displays the correctly (diagonal) and incorrectly (off-diagonal) predicted types (see table in Figure 8). Analysis of the classification results showed that the features that impacted the classification results most are the ship dimensions: length and width. Without them, the prediction performance drops by as much as 11%.
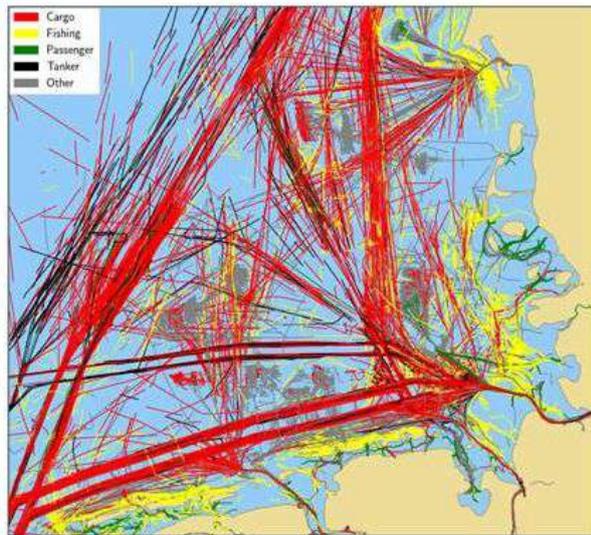
Figure 6. Ship classifier based on ship trajectories

A classifier which was trained in the serving layer can be delivered to the speed layer and can be applied to the real time data streams of the trajectories. Here, they can be used in two ways: first they can be used to detect spoofing or abnormal behavior of trajectories based on cooperative data sources such as the detection of illegal fishing. Second, they can be used to allow classification capabilities for unlabelled, often uncooperative data sources, like coastal radar networks or GMTI radar applications [8].

For example, for the scenario described in   Figure **6**, the training trajectories revealed that some ships actually report multiple different types in the course of their voyage. In these cases, the most often used type codes are "other", "not available", and "reserved for future use". However, the trained classifier predicted the type "fishing vessel" for many of these ships. Validation against open source ship libraries showed that the predicted type is correct in most cases.

4. HEAT MAP GENERATION AND VISUAL SITUATION ASSESSMENT

Heat maps offer an illustrative and intuitive way to improve situation awareness based on trajectories. Basically, they include the definition of a spatial grid and the evaluation of statistical properties for each grid cell. Examples are the number of trajectories crossing a grid cell, and the average speed or course inside a grid cell. Some statistics follow mono-modal distributions others are multi-modal. Comparing the actual situation with the heat map can enhance awareness of the situation. The generation of such statistics is supported in the serving layer with modern tools of distributed processing like Spark. With these tools it is possible to generate even global heat maps on a long-term basis and to adapt them to different time windows continuously. Additionally, the heat maps can be used to generate warnings or alarms.
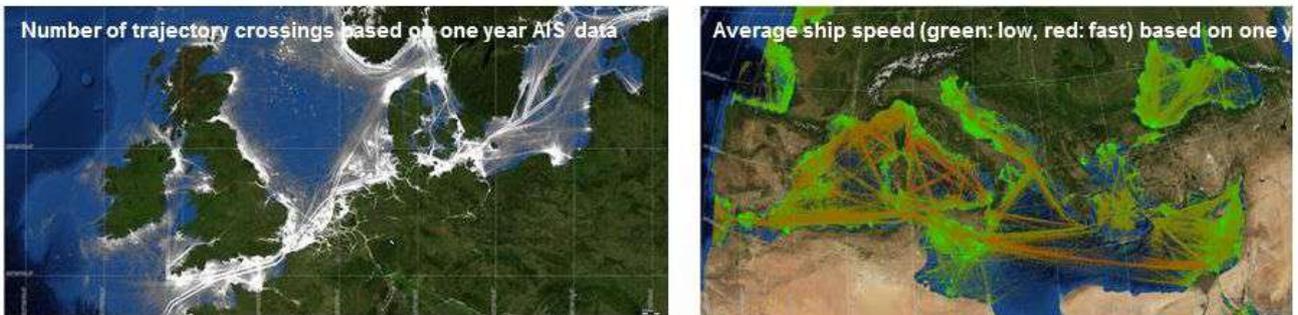
Figure 7. Examples of long-term heat maps

5

Another example of a thematic heat map is the time-and space resolved risk assessment presented in Figures 9-11. It has been developed in the course of a German-Canadian research project that was jointly conducted by Airbus Defence and Space, the Fraunhofer Institute for Communication, Information Processing and Ergonomics, FKIE, and Dalhousie University, Halifax [21].

The heat map assess the risk to a ship with a given ice class (or none at all) due to sea ice in a specific area and at a specific time with the aim of increasing the safety of the ship while operating in polar waters. The assessment is based on the internationally established Polar Operational Limitations Assessment Risk Indexing System (POLARIS) [22, 23] which was produced by the International Association of Classification Societies (IACS) to determine ship operational limits in ice. POLARIS uses a combination of the known ship ice classification limit and prevailing ice conditions in a given ice regime as published by the Canadian Ice Service to evaluate a risk score, the so called Risk Index Outcome (RIO), for each spatio-temporal grid cell. All numerical RIO values are associated with a recommendation of the type "operation permitted" or "operation permitted at limited speed" or "operation not permitted" to the ship operator. If a ship is escorted by an ice breaker, its RIO value is increased by 10, which increases the likelihood that the ship's intended operation is permitted. This is depicted in   Figure **8**.
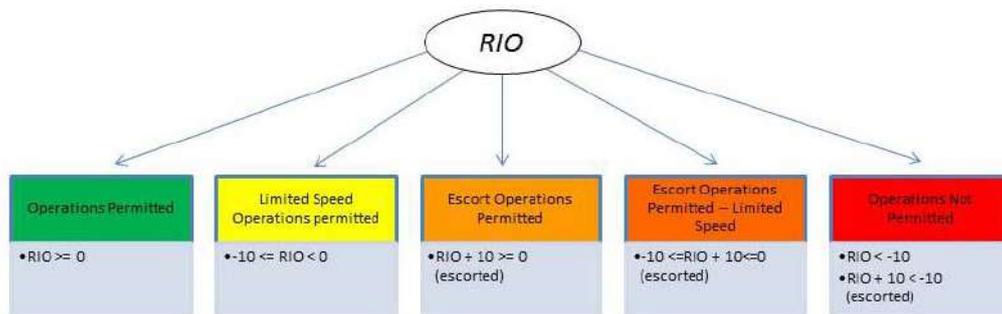
Figure 8. Operational limitations based on the POLARIS RIO value for operations in ice

Using the color code as given in   Figure **8**, we can use the POLARIS framework to compute heat maps of an area of interest for different time periods, for different ship ice class classifications and for different statistical aggregations of RIO values. For example, in   Figure **9** and   Figure **10** the area of interest corresponds to Canadian arctic waters including the area of the Northwest Passage and Hudson Bay. The area is tessellated into 1km x 1km grid cells and in both cases the final RIO value per grid cell is the average over the RIO values of that grid cell evaluated for weekly published ice charts of the years 2007 – 2014.

Figure **9** shows three heat maps all for the same week in August but for different ship ice classes: "NOT IS" means not ice-strengthened at all, "IA" is a commonly occurring Baltic ice class, whereas "PC 1" is the strongest polar ice class. It can be seen that the area is getting increasingly safer to navigation as the ice class of the ship increases.
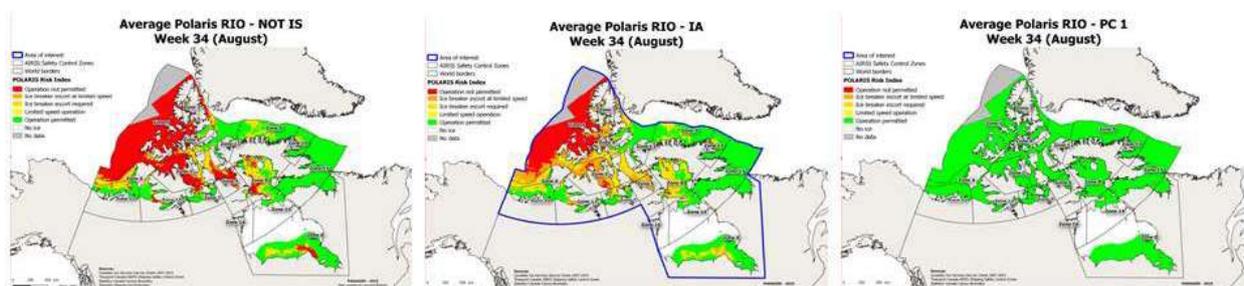
Figure 9. Operational limitations for 3 different ship ice classes during week 34 (August) in Canadian polar waters

In   Figure **10**, the heat maps correspond to the same time window (week 34) and to the same ship ice class (PC 7: lowest polar ice class) but display three different statistical aggregations of RIO values: the maximum (best case) RIO values per grid cell, the averaged RIO values per grid cell, and the minimum (worst case) RIO values per grid cell – always averaged over the RIO values of week 34 for the years 2007 – 2014.
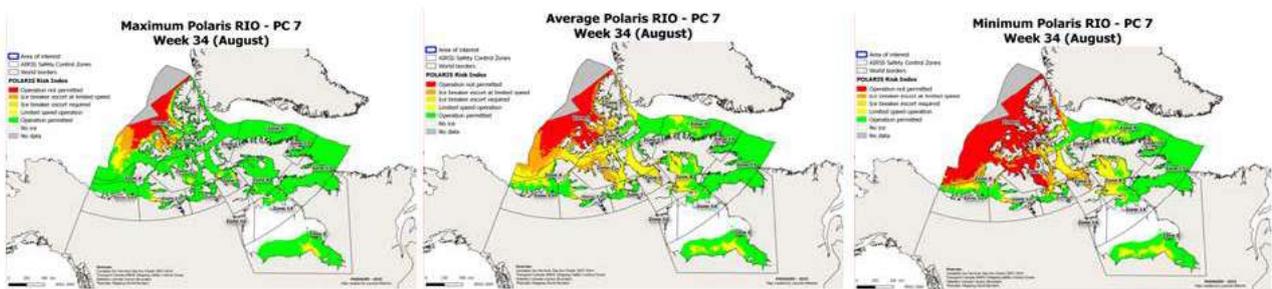
Figure 10. Operational limitations for 3 different statistical RIO evaluations during week 34 for ice class PC 7

The risk of a ship trajectory is computed by determining through which grid cells a trajectory passes. For easy visualization, the colours of the crossed grid cells can be projected onto corresponding track parts in order to create a coloured trajectory that immediately shows which operational limitations apply for which parts of the trajectory (see   Figure **11**).
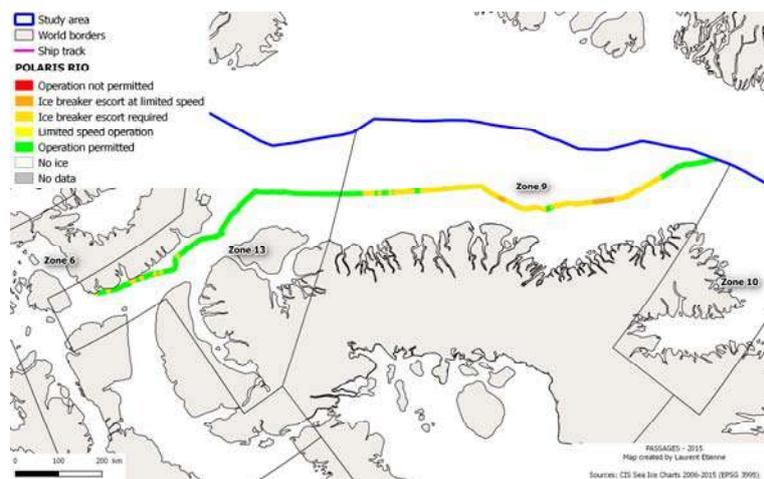
Figure 11. Ship trajectory risk (colored trajectory)

This type of risk evaluation can principally be used as a strategic planning tool as well as a tactical routing tool that guides the ship along the safest path while it is "en route". Furthermore, the assessment can be extended to also include the risk due to other factors such as bathymetry or weather. In both cases is the data-driven quantitative risk assessment a valuable tool to increase the awareness for maritime risks.

5.   CONCLUSION

Real time surveillance in defence, security and safety scenarios is a difficult task since the scenarios often take place in ambiguous and complex environments with asymmetric and manifold threats. Trying to address these challenges with data-driven methods results in big data problems, for which the user needs support to control large data volumes, high data turn-over rates and to be able to make the right decisions.

7

These needs are supported by innovative big data and machine learning methods that offer a chance to cope with existing and new challenges efficiently and effectively. Therefore, modern surveillance systems have to integrate big data concepts. Furthermore, globally and persistently available position data received by networked AIS and ADS-B transceivers and sensors deliver multitudes of trajectories. This paves the way for advanced data analytics and supervised and unsupervised machine learning. In this paper, we show how data-driven approaches can add to maritime situation awareness and decision support by applying these technologies to selected use cases.

REFERENCES

[1]   International Maritime Organization (IMO) http://www.imo.org
[2]   US Department of Homeland Security, Navigation Center, https://www.navcen.uscg.gov/
[3]   Blackman, S. and Popoli, R., [Design and Analysis of Modern Tracking Systems], Artech House, Boston, (1999)
[4]   Koch, W., [Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications], Springer Science & Business Media, (2013).
[5]   Bar-Shalom, Y., Rong Li, X., and Kirubarajan, T., [Estimation with applications to tracking and navigation: theory algorithms and software], John Wiley & Sons, (2004).
[6]   Nathan, M., and Warren, J., [Big Data: Principles and best practices of scalable realtime data systems], Manning Publications Co., (2015).
[7]   Opitz, F., Dästner, K., von Haßler zu Roseneckh-Köhler, B., and Schmid, E., "Data Analytics and Machine Learning based on Trajectories," NATO Specialist Meeting AI, Budapest, Hungary, (2018).
[8]   Dästner, K., et al., "Machine Learning Techniques for Enhancing Maritime Surveillance Based on GMTI Radar and AIS," International Radar Symposium (IRS), Bonn, Germany, (2018).
[9]   Apache Cassandra, "a no-sql database", cassandra.apache.org
[10] Syd Ali, B., Schuster, W., Ochieng, W., and Majumdar, A., "Analysis of anomalies in ADS-B and its GPS data," Springer, GPS Solutions, Volume 20, Issue 3, 429–438, (July 2016).
[11] Apache Spark, "a unified analytics engine for large-scale data processing", spark.apache.org
[12] scikit-learn, "Machine Learning in Python", scikit-learn.org
[13] Apache Spark MLlib, "a scalable machine learning library", spark.apache.org/mllib
[14] Dästner, K., Brunessaux, S., Schmid, E., von Haßler zu Roseneckh-Köhler, B., and Opitz, F., "Classification of Military Aircraft in Real-time Radar Systems based on Supervised Machine Learning with Labelled ADS-B Data," 12-th Symposium Sensor Data Fusion, Bonn, Germany, (October 2018).
[15] Zhongliang, F. et al., "A two-step clustering approach to extract locations from individual GPS trajectory data," ISPRS International Journal of Geo-Information 5.10, 166 (2016).
[16] Dästner, K. et al., "Exploratory data analysis for GMTI radar," Radar Symposium (IRS), 18th International. IEEE, Prague, Czech Republic, (2017).
[17] Le Guillarme, N., and Lerouvreur, X., "Unsupervised extraction of knowledge from S-AIS data for maritime situational awareness," Proceedings of the 16th International Conference on Information Fusion. IEEE, (2013).
[18] Pallotta, G., Vespe, M., and Karna, B., "Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anamoly Detection and Route Prediction," Entropy 15, 2218-2245, (2013).
[19] Géron, A., "Hands On Machine Learning with Scikit-Learn & TensorFlow,",O'Reilly, 5th Release, (2018).
[20] Kraus, P., Mohrdieck, C., and Schwenker, F., "Ship classification based on trajectory data with machine-learning methods," Proc. International Radar Symposium, Bonn, Germany, (2018).
[21] Battistello, G., Ulmke, M., and Mohrdieck, C.,"Enhanced maritime traffic picture for the Canadian Arctic," Proc.10th Future Security Research Conference, Berlin, Germany, (September 2015).
[22] Maritime Safety Committee (2014b), Technical background to POLARIS, International Association of Classification Societies (IACS).
[23] Stoddard, M. A., Etienne, L., Pelot, R., Fournier, M., and Beveridge, L.,"From sensing to Sense-Making Assessing and visualizing ship operational limitations in the Canadian Arctic using open-access ice data", Proc. ShipArc Conference, Malmö, Sweden, (August 2015).

IDCP as Kernel Element for Force Level Recognized Maritime Picture Generation

Jürgen Ziegler *

Industrieanlagen-Betriebsgesellschaft mbH (IABG), Einsteinstrasse 20, 85521 Ottobrunn, Germany, Dept. Intelligence, Surveillance & Reconnaissance

ABSTRACT

The Identification Data Combining Process (IDCP) is standardised in STANAG 4162. Edition 2 of the STANAG is promulgated and, e.g., implemented in NATO's ACCS (Air Command Control System). Currently Edition 3 together with Allied Identification Publication 01 is under ratification. The IDCP defines a standard for a real time process for identification and platform classification. The essential elements of the IDCP are a standardized process for the generation of ID results, standardized encoding and exchange of ID Source Information and the description of the generation and encoding of all operational and technical data that is necessary to use the process. IDCP can be regarded as hierarchical Bayesian Decision Network and provides a solid methodological foundation for identification. The IDCP can be integrated into a Coalition architecture in such a way that it guarantees the optimal force level use of identification information.

Keywords: Identification Data Combining Process, IDCP, STANAG 4162, force level function, Bayesian Network

1. INTRODUCTION

Interoperable Identification within multinational maritime forces is a key for the successful real-time processing of defence again agile threats like missiles. The paper describes why, where and how to use the IDCP according STANAG 4162 as an kernel element for the Force Level Recognized Maritime Picture Generation, i.e. how the IDCP can provide results which can directly be used by the subsequent processes, i.e. the Threat Evaluation and Weapon Assignment. Section 2 starts with a short explanation of the STANAG 4162 and the IDCP process. Section 3 describes the methodological foundation of the IDCP, i.e. its interpretation as hierarchical Bayesian Decision Network. Section 4 shows the possibilities to implement IDCP as force level function. Section 5 addresses the benefit using the IDCP.

2. IDCP – THE STANAG 4162

Before describing the application of the IDCP as Force level Function for identification, this chapter summarises the content of the STANAG 4162 and the essential properties of the IDCP.

2.1 Scope of STANAG 4162 and the Standard Related Documents

The current version of STANAG 4162 is the ratification draft of Edition 3 with its content elaborated in the Allied Identification Process Publication 01. Drafts of implementation guidance and user guidance (standard related documents SRD AIDPP-01.01 and SRD AIDPP-01.02. (See [1], [2], [3])) complement the STANAG 4162.

The AIDPP-01 provides a description of a standardized computer process for the automatic generation of an ID result, the IDCP. The IDCP is an automated data combining method which fuses identification data on detected objects from multiple and dissimilar sources and in all environments (land, air, maritime and space), providing ID Category recommendations to an operator responsible for identification. The operator can be located on a platform or at an operational site. The AIDPP-01 provides the required information to implement the IDCP such that it can be used for identification of objects in all environments using as many available sources as possible, and a maximum of interoperability with other host systems is guaranteed.

The AIDPP-01 provides a description and standardization of the data necessary for the IDCP. The AIDPP-01 defines the standards for the input to the IDCP, which is the so-called ID Source information. It prescribes how to convert and use all data, which are potentially usable for identification, so that IDCP can process the data. It is presented in the form of a systematic and hierarchical structuring of all possible ID Sources called the ID Source Taxonomy. For maritime

---

applications, an exhaustive taxonomy of ID Sources is defined for the identification of air, surface and – with limitations – sub-surface tracks. STANAG 4162 standardizes the conditioning data, which is necessary to adapt the IDCP to a diversity of missions in various environments and operational conditions.

The AIDPP-01 provides a description of the required processes for data generation to supply the IDCP with the necessary input. The document describes the process for the generation of ID Source information and the process for the generation of the configuration data. It describes how this data can be derived from operational and technical knowledge and from data, which is not IDCP-specific (such as Air Tasking Orders, Order of Battle and geographical information).

The IDCP supports the use of both locally generated and remote ID Source information. The AIDPP-01 therefore also describes the information exchange requirements resulting from the requirement to be able to use remote information. It defines how to map the information to be exchanged (the ID Source Information) to message formats of available message transfer systems. The paper about IDCP in completed with two additional supplementary documents: The standard related document (**SRD) AIDPP-01.01 "Implementation Guidance (IG) for IDCP"** presents information to generate an implementation of the IDCP, which is in accordance with the objective of the STANAG. The **SRD AIDPP-01.02 "User Guidance (UG) for IDCP"** presents the information to the user to understand the IDCP application and how to use an implementation of the IDCP in accordance with the objective of the STANAG.

2.2 IDCP Process

The IDCP Process comprises the following steps:
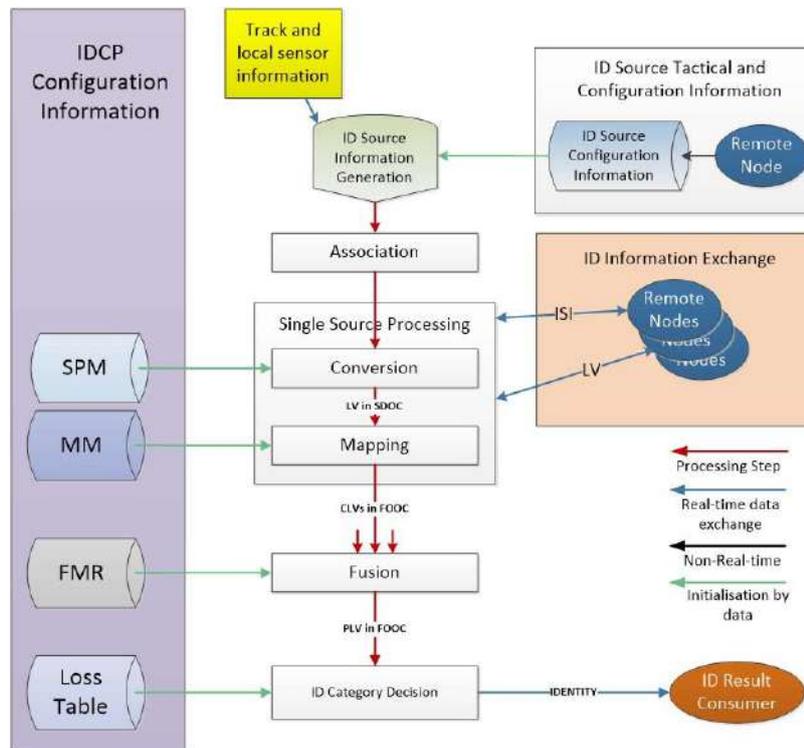
Figure 1: IDCP processing (according to [1])

2.2.1 The ID Source Information Generation

During this processing step, sensor information about a track is evaluated to generate information, which is relevant for identification. Usually this step is processed automatically. Host systems may also provide the functionality for operator interaction to check, change, delete and add ID Source Information. To be able to perform this step usually some tactical information is needed to configure the ID Source. Examples for ID Source information are Q&A-Mode 5 (it is checked whether a track answers to an IFF Mode 5 request); Movement and Procedural Routing (MPR)-Mission Plan (track

trajectory is compared with an air volume to extract information whether or not it complies with a mission plan); IDBO (track origin is compared with an area surrounding an airfield).

2.2.2 Transmission / Reception of ID Source Information

During this processing step, the different entities, which are able to generate ID information, shall provide this information to the IDCP. STANAG 4162 Edition 3 contains the standard and the information exchange requirements for this ID source data exchange.

2.2.3 Kernel Processing

The kernel processing comprise three parts: The first step, "**conversion**", translates the ID Source Information into a standardized form (a so-called Source Likelihood Vector) which is defined for every source type in STANAG 4162. This format contains the mathematical representation of the discriminatory ability and quality of an ID source.

The **mapping** step converts the ID Source-specific result into tactical information (the so-called Output Object Class), e.g. the information that an aircraft has provided a valid IFF Mode 5 response will be used to derive that the track is a friendly military track (with very high probability).

After the mapping all information from the various ID Sources is available in the same form of tactical information and can be used by the **fusion step** to generate the result, e.g. the probabilities that the track is one of the objects "Own Force Military", "Own Civil", "Non Aligned Military", "Non Aligned Civil", "Enemy Forces Military" or "Enemy Civil". This result is the so-called Posterior Likelihood Vector (PLV).

2.2.4 "ID Category Selection"

The PLV is not the adequate result to be presented to an ID operator within the situational picture. Therefore, this step generates an Identity (Recommendation) in light of a risk assessment. In an air defence environment this is usually defined by the STANAG 1241 [6], resulting in an identity selected from the list FRIEND, ASSUMED FRIEND, NEUTRAL, SUSPECT, HOSTILE, UNKNOWN. This result will be presented to the ID operator by using the appropriate symbology as defined in APP-6. Additionally the IDCP can also be used to generate a recommendation for the platform class of the track.

2.3 Necessary Data for the IDCP

There are several different types of information, which must be considered for the IDCP as described in STANAG 4162 Edition 3. The first two types represent the technical and operational knowledge about the ID source, which are used in a specified host system, and the tactical environment of the current mission. They give the opportunity to adapt the IDCP to the technical specifics of the host system and to the operational requirements and conditions of the mission.

The first type is the IDCP Conditioning Information. This information is mostly static during a mission. It can be used to adapt the IDCP to changes in the mission environment (e.g. if the alert state changes from Crisis to war) a different Loss Table will normally be used. It includes

a)   The knowledge about the discrimination capability and discrimination quality of the ID Sources. This information is represented by the so-called **Source Probability Matrix** (SPM). This information is used for the processing step "Conversion".

b)   The operational knowledge about the interpretation of the ID Source Information. This information is represented by the so-called **Mapping Matrix** (MM). This information is used for the processing step "Mapping".

c)   The a priori knowledge about the expectation concerning the distribution of the allegiance of the objects in a specific mission. This information is represented by the so-called **Force Mix Ratio** (FMR). This information is used in the "Fusion "step.

d)   The knowledge and rules concerning the appropriate selection of an ID Category. This information is represented by the so-called **Risk Matrix**. This information is used for the process step "ID Category Selection"

The second type of information is necessary for the **configuration of ID sources**. It consists of the tactical information that is necessary to generate the appropriate ID Source information. The format of this data is not part of the IDCP STANAG. The **Track information** is fundamental for the complete ID process. All ID Source information and all intermediate IDCP processing results as well as the final ID result are associated to an appropriate track. Additionally track information is also the basis for the generation of several types of ID Source information, e.g. for the ID Sources of the type Movement Planning and Routing. The **ID Source Information** contains the information usable for identification, which was generated by an ID Source, either locally or remotely. This data must be associated to a track for it to be useful.

The exchange (transmission / reception) of the ID Source information guarantees that the generation of the ID Category recommendation is based on all ID Source information, which is available within an Identification Network. Normally this data must be exchanged in (near) real-time. The **recommendation for the ID Result** is the final information generated by the IDCP process. Usually, the ID result is provided as an ID Category and / or a platform class. If approved, the ID result will be used for the subsequent tactical processing steps (e.g. the threat assessment) and will be forwarded to other participants of the tactical network.

3. IDCP AS A HIERARCHICAL BAYESIAN DECISION NETWORK

The IDCP can be represented as a Bayesian hierarchical decision network with discrete nodes (see [4] for a more general Bayesian Network for Identification). The possible results of the fusion step – the OOC – defines the states of the top node. This node determines the probabilities of the states of the ID Sources (the likelihood vectors for the different source types) by applying appropriate Conditional Probability Tables (CPTs). The CPTs relating the OOC and the likelihood vector for the different ID Source Types are given by the mapping matrices. The standards for the description of the ID Source Types define the states of the ID Source Type nodes. The probability distribution of the states of these nodes determine the probability of the states of the ID Source nodes (i.e. the nodes representing the ID Source Information), the CPTs are given by the Source Probability Matrices. The standards for encoding of the ID Source Information define the states of the ID Source nodes. Finally, the Risk Matrix is applied for the calculation of the risk of possible decision (ID Category selections) by using a value and decision node.
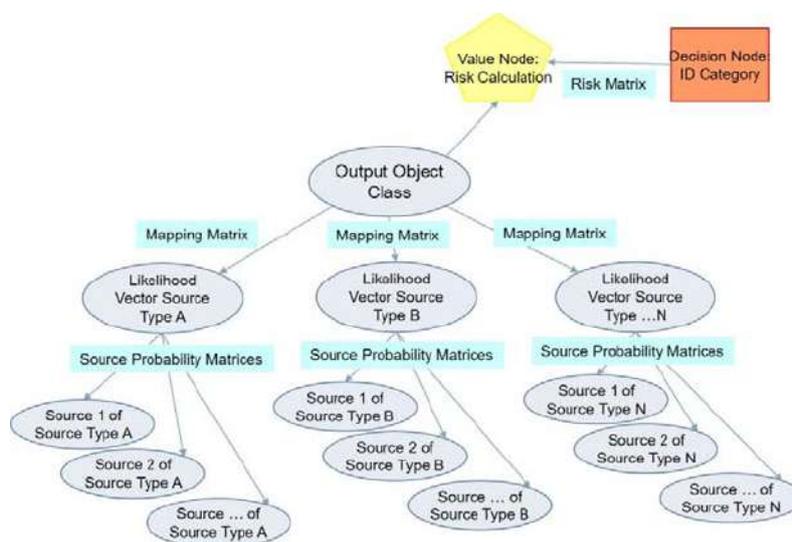
Figure 2: IDCP kernel processing as a hierarchical Bayesian Decision Network

4. IDCP IMPLEMENTATION AS FORCE LEVEL IDENTIFICATION FUNCTION

There is an high operational need to implement identification within a naval force in a way that the Identification is fully interoperable so that the ID results can be used within an automated processing. This is of special importance to enable naval forces to engage against threats like modern missiles, where the time to react is not sufficient for manual coordination of possibly contradictory ID results arising from incomplete availability of ID Source information and non-interoperable ID processing. STANAG 4162 defines the standardized data exchange of ID Source information. An IDCP implementation as force level function requires that all ID Source information is provided to the IDCP process(es) according to the STANAG which is (which are) used for the identification of tracks for specific environments. In [5] two basic architectural alternatives are described: The first is the implementation of IDCP as a function, which is not fully integrated in the Combat Direction System of the ships. In this case the IDCP is using the CDS specific ID Source information of the platforms, generates the standardized ID Source information, provides the information to the IDCP on other platforms. generates the ID Category Recommendation and provides it to the platform CDS. This is a solution for cases, where the integration of IDCP into an existing CDS might be too expensive. The second alternative is the implementation of an IDCP as CDS Identification component of the integration of newly implemented IDCP into the CDS. In this case, the IDCP within the CDS generates the ID Source information and provides it to the network and the CDS generates the ID result according to the STANAG for the application of the future processes.

Note that the experience with already available IDCP implementations shows that the implementation of the standardized fusion process is no great challenge. The effort to implement all ID Sources accurately is usually larger. Furthermore, there is a need for a tool (IDCP Conditioning Tool) for the generation of the necessary data for the IDCP by operational (educated) users, i.e. the source probability matrices, the mapping matrices and the risk matrix. This tool must additionally provide the functions to test the IDCP results applying a specific data set, e.g. for being sure that all mandatory ID rules are satisfied by the IDCP using this data set.

5. BENEFIT OF INTEROPERABLE IDCP IMPLEMENTATION

The IDCP is realized using a sound mathematical model, which enables the proper exploitation of all important technical and operational knowledge to its maximum extent. Therefore, the intermediate and the results are reliable, transparent and explainable. Under the assumption that the conditioning data set was generated correctly, the results will correspond to user expectations. The IDCP processing is based on an efficient algorithm based on Bayesian probability theory, which simplifies the implementation of an IDCP with full real-time capability. The IDCP can be implemented in a way that it is fully adaptable by definition and usage of appropriate Conditioning data sets. Therefore, it can be adapted without change of software, if there are changes in the technical equipment of host system e.g. due to future developments in technique, to different scenarios, and to different environmental conditions. The implementation and use of IDCP as force level identification process can guarantee that the recognized Picture of all force members coincides and that the recognized picture of them based on the use of all ID Source information available within the force.

6. CONCLUSION

This paper describes IDCP according to STANAG 4162 Edition 3, which is under ratification. IDCP is based on Bayesian probability theory and can be regarded as a hierarchical Bayesian Network. IDCP provides an approach for an implementation of an interoperable ID process, which can be used as force level identification process for naval forces. The successful application of the IDCP STANAG calls for its ratification and gradual implementation by a significant number of NATO members. On that condition, IDCP will improve the interoperability and quality of identification of NATO naval forces.

REFERENCES

[1] AIDPP-01 "Identification Data Combining Process"; Edition A Version 1; Study Draft 3; April 2016
[2] Standard Related Document "AIDPP-01.1 Implementation Guidance for Identification Data Combining Process"; Edition A Version 1; March 2017
[3] Standard Related Document "AIDPP-01.2 User Guidance for Identification Data Combining Process"; Edition A Version 1; March 2017
[4] Krüger, Max, Ziegler Jürgen and Heller Kathrin: Generic Bayesian Network for Identification and Assessment of Objects in Maritime Surveillance, International Conference on Information Fusion, 2012
[5] Ziegler, Jürgen: IDCP as a kernel element for the Force Track Evaluation and Engagement Coordination within the Framework Architecture 2030+, MTMD Forum in Halifax 2017
[6] STANAG 1241 MAROPS (Edition 5) – NATO STANDARD IDENTITY DESCRIPTION FOR TACTICAL USE, 2005

4

Radar signal processing

Contents

Synthesis of Constant Modulus Radar Signals in Spectrally Crowded Scenarios

Augusto Aubry[a], Antonio De Maio[a], Mark Govoni[b], and Luca Pallotta[c]

[a]Università di Napoli "Federico II", Department of Electrical Engineering and Technology Information (DIETI), via Claudio 21, I-80125 Napoli, Italy;
[b]Army Research Laboratory, Aberdeen Proving Ground, MD, USA.
[c]Università "Roma Tre" (previously with CNIT), Engineering Department, via Vito Volterra 62, I-00146 Roma, Italy.

ABSTRACT

This paper deals with the synthesis of waveforms optimizing radar performance while satisfying multiple spectral compatibility constraints. Precisely, on each shared bandwidth a precise and specific control on the injected interference energy is forced. Furthermore, constant modulus waveforms are considered at the design stage to comply with today amplifiers technology. To tackle the resulting NP-hard optimization problem, an iterative procedure based on the coordinate descent method is introduced. Hence, some case studies are reported to highlight the effectiveness of the technique.

Keywords: Waveform design, Spectral compatibility, Optimization theory

1. INTRODUCTION

Recent years have witnessed a growing interest in the Radio Frequency (RF) congestion problem being RF spectrum a commodity necessary for an ever-growing number of services and systems.[1] In particular, wider and wider amount of bandwidths are required by advanced wireless communication and remote sensing systems to accomplish their demanding tasks, e.g., high data-rates and reliable surveillance. As a consequence, a serious challenge is posed on the overall spectral usability by non-cooperative systems[2] and spectrum compatibility issue has attracted the interest of many scientists and engineers, currently becoming one amongst the hot topics in both regulation and research field. In particular, many papers in the open literature have dealt with the problem of designing radar signals with a suitable frequency allocation so as to induce acceptable interference levels on frequency-overlaid systems (see[3–10] and references therein).

In this paper, a new technique for constant modulus waveform synthesis specifically designed to face spectrally dense environments is proposed, where the Interference plus Noise Ratio (SINR) serves as the preferred performance metric. A local control on the interference energy radiated on each shared/reserved frequency bandwidth is performed. Along with a requirement on the maximum transmitted energy, a similarity constraint is enforced on the probing sequence in order to control significant waveform characteristics, e.g., Peak Sidelobe Level (PSL) and ISL. The optimization process is restricted to constant modulus waveforms for compatibility with current amplifier technology and is unprecedented in terms of synthesizing constant modulus codes subject to multiple spectral compatibility constraints. The resulting optimization problem belongs to the class of NP-hard problems and a new iterative algorithm based on the Coordinate Descent (CD) method is developed to synthesize optimized codes. At the analysis stage, some interesting case studies are illustrated to assess the capability of the devised algorithm to improve radar detection performance while guaranteeing spectral compatibility.

The paper is organized as follows. Section 2 introduces the system model and defines the waveform design problem. In Section 3, an efficient algorithm to handle the formulated optimization problem is devised. Section 4 is devoted to the performance analysis. Finally, Section 5 concludes the paper.

Further author information: (Send correspondence to A. Aubry)
A. Aubry: E-mail: augusto.aubry@unina.it

Notation

We adopt the notation of using boldface for vectors $\boldsymbol{a}$ (lower case), and matrices $\boldsymbol{A}$ (upper case). The n-th element of $\boldsymbol{a}$ and the (n,m)-th entry of $\boldsymbol{A}$ are denoted by a_n and $\boldsymbol{A}_{n,m}$, respectively. The transpose and the conjugate transpose operators are denoted by the symbols $(\cdot)^T$ and $(\cdot)^\dagger$ respectively. $\boldsymbol{I}$ and $\boldsymbol{0}$ denote respectively the identity matrix and the matrix with zero entries (their size is determined from the context). $\mathbb{R}^N$, $\mathbb{C}^N$, and $\mathbb{H}^N$ are respectively the sets of N-dimensional vectors of real numbers, of N-dimensional vectors of complex numbers, and of $N \times N$ Hermitian matrices. The curled inequality symbol $\succeq$ (and its strict form $\succ$) is used to denote generalized matrix inequality: for any $\boldsymbol{A} \in \mathbb{H}^N$, $\boldsymbol{A} \succeq \boldsymbol{0}$ means that $\boldsymbol{A}$ is a positive semi-definite matrix ($\boldsymbol{A} \succ \boldsymbol{0}$ for positive definiteness). For any complex number x, $|x|$ represents the modulus of x. Moreover, for any $\boldsymbol{x} \in \mathbb{C}^N$, $\|\boldsymbol{x}\|$ denotes the Euclidean norm and $\mathbf{diag}\,(\boldsymbol{x})$ indicates the diagonal matrix whose i-th diagonal element is the i-th entry of $\boldsymbol{x}$. Finally, $\mathbb{E}\,[\cdot]$ denotes statistical expectation and for any optimization problem $\mathcal{P}$, $v(\mathcal{P})$ is its optimal value.

2. SYSTEM MODEL

Let $\boldsymbol{c} = [c_1, \ldots, c_N]^T \in \mathbb{C}^N$ be the transmitted fast-time radar code vector with N the number of coded sub-pulses (code length). The waveform at the receiver end is down-converted to baseband, undergoes a sub-pulse matched filtering operation, and then is sampled. As a result, the N-dimensional column vector $\boldsymbol{v} = [v_1, \ldots, v_N]^T \in \mathbb{C}^N$ of the fast-time observations from the range-azimuth cell under test can be expressed as

$$\boldsymbol{v} = \alpha_T \boldsymbol{c} + \boldsymbol{n},$$

with α_T a complex parameter accounting for channel propagation and backscattering effects from the target within the range-azimuth bin of interest and $\boldsymbol{n}$ the N-dimensional column vector containing the filtered disturbance signal samples; $\boldsymbol{n}$ is modeled as a complex, zero-mean, circularly symmetric Gaussian random vector with covariance matrix $\mathbb{E}[\boldsymbol{n}\boldsymbol{n}^\dagger] = \boldsymbol{M}$.

As to the licensed emitters coexisting with the radar of interest, it is supposed that each of them is operating over a frequency band $\Omega_k = \left[f_1^k, f_2^k\right]$, $k = 1, \ldots, K$, where f_1^k and f_2^k denote the lower and upper normalized frequencies for the k-th system, respectively. To ensure spectral compatibility the radar has to shape properly its transmit waveform to manage the amount of interfering energy produced on the shared frequency bandwidths given by[7]

$$\int_{f_1^k}^{f_2^k} S_c\,(f)\,df \;\; = \;\; \boldsymbol{c}^\dagger \boldsymbol{R}_I^k \boldsymbol{c}, \tag{1}$$

where $S_c\,(f) = \left|\sum_{n=0}^{N-1} \boldsymbol{c}_n e^{-j2\pi f n}\right|^2$ is the Energy Spectral Density (ESD) of the fast-time code $\boldsymbol{c}$ and for $(m,l) \in \{1, \ldots, N\}^2$

$$\boldsymbol{R}_I^k(m,l) = (f_2^k - f_1^k) e^{j\pi(f_2^k + f_1^k)(m-l)} \mathrm{sinc}\left(\pi(f_2^k - f_1^k)(m-l)\right)$$

Thus, denoting by E_I^k, $k = 1, \ldots, K$, the acceptable level of disturbance on the k-th bandwidth, that is related to the quality of service required by the k-th telecommunication networks, the transmitted waveform has to comply with the constraints

$$\boldsymbol{c}^\dagger \boldsymbol{R}_I^k \boldsymbol{c} \leq E_I^k, \quad k = 1, \ldots, K. \tag{2}$$

Remarkably, unlike several approaches proposed in the open literature, e.g.,[11] where a weighted sum of the produced interference power is usually controlled, (2) provides a detailed control of the interference energy produced on each shared frequency bandwidth. Finally, the radar system may resort to a Radio Environmental Map (REM)[12] as well as suitable spectrum sensing modules[13] to get cognizance of the licensed emitters (e.g. their spatial location and bandwidth occupation).

2.1 Code Design Optimization Problem

In this subsection, a radar waveform design framework is introduced aimed at improving target detection probability while controlling both the amount of interfering energy produced in each shared bands and some hallmarks of the transmitted waveform. To this end, the SINR, defined as

$$\text{SINR}(\boldsymbol{c}) = |\alpha_T|^2 \boldsymbol{c}^\dagger \boldsymbol{R} \boldsymbol{c} \tag{3}$$

is considered as figure of merit, where $\boldsymbol{R} = \boldsymbol{M}^{-1}$ is the inverse of the interference covariance matrix. To control some attributes of the transmitted waveform, other than an energy requirement, namely $\|\boldsymbol{c}\|^2 \leq 1$, a similarity constraint is imposed on the probing signature, i.e., $\left\| \frac{\boldsymbol{c}}{\|\boldsymbol{c}\|} - \boldsymbol{c}_0 \right\|_\infty \leq \frac{\epsilon}{\sqrt{N}}$, where the parameter $0 \leq \epsilon \leq 2$ rules the size of the trust hypervolume, and $\boldsymbol{c}_0 = [c_{0,1}, \ldots, c_{0,N}]^T \in \mathbb{C}^N$ is a specific phase-only reference code, with $\|\boldsymbol{c}_0\|^2 = 1$. Indeed, an unconstrained optimization can lead to signals with high PSL/ISL values and more generally with an undesired ambiguity function behavior. These drawbacks can be partially circumvented forcing the solution to be similar to a known code $\boldsymbol{c}_0$ that shares some advantageous properties. Finally, to assure compatibility with current amplifier technology, the optimization process is restricted to constant modulus waveforms.

Summarizing, leveraging on the aforementioned guidelines, the waveform design problem of interest can be formulated as the following non-convex optimization problem

$$\mathcal{P} \begin{cases} \max_{\boldsymbol{c} \in \mathbb{C}^N} & \overline{\text{SINR}}(\boldsymbol{c}) \\ \text{s.t.} & \boldsymbol{c}^\dagger \boldsymbol{R}_I^k \boldsymbol{c} \leq E_I^k, k = 1, \ldots, K \\ & \|\boldsymbol{c}\|^2 \leq 1 \\ & |c_i| = |c_k|, (i, k) \in \{1, \ldots, N\}^2 \\ & \left\| \frac{\boldsymbol{c}}{\|\boldsymbol{c}\|} - \boldsymbol{c}_0 \right\|_\infty \leq \frac{\epsilon}{\sqrt{N}} \end{cases} \tag{4}$$

where

$$\overline{\text{SINR}}(\boldsymbol{c}) = \boldsymbol{c}^\dagger \boldsymbol{R} \boldsymbol{c}. \tag{5}$$

is the normalized SINR. Furthermore, re-parameterizing the optimization vector $\boldsymbol{c}$ as

$$\boldsymbol{c} = \sqrt{\frac{\rho}{N}} \left[e^{j(\psi_1 + \arg(c_{0,1}))}, \ldots, e^{j(\psi_N + \arg(c_{0,N}))} \right]^T \in \mathbb{C}^N,$$

with $\psi_i \in [-\pi, \pi]$, $i = 1, \ldots, N$, and $0 \leq \rho \leq 1$, the design problem boils down to

$$\bar{\mathcal{P}} \begin{cases} \max_{\boldsymbol{x}} & \overline{\text{SINR}}(\boldsymbol{x}) \\ \text{s.t.} & \bar{\boldsymbol{c}} = \sqrt{x_{N+1}} \left[e^{jx_1}, e^{jx_2}, \ldots, e^{jx_N} \right]^T \\ & \bar{\boldsymbol{c}}^\dagger \bar{\boldsymbol{R}}_I^k \bar{\boldsymbol{c}} \leq E_I^k, k = 1, \ldots, K \\ & 0 \leq x_{N+1} \leq 1 \\ & |x_i| \leq \delta, i = 1, \ldots, N \end{cases} \tag{6}$$

where $\boldsymbol{x} = [\psi_1, \psi_2, \ldots, \psi_N, \rho]^T \in \mathbb{R}^{N+1}$ is the new optimization vector (with x_i, $i = 1, \ldots, N$, the phase-code shift with respect to $c_{0,i}$, and x_{N+1} related to the actual code amplitude), $\delta = \arccos\left(1 - \frac{\epsilon^2}{2}\right)$,

$$\bar{\boldsymbol{R}}_I^k = \text{diag}(\boldsymbol{c}_0^*) \boldsymbol{R}_I^k \text{diag}(\boldsymbol{c}_0), \tag{7}$$

controls the k-th spectral constraint in terms of the new optimization variables, and, with a slight abuse of notation, $\overline{\text{SINR}}(\boldsymbol{x})$ is the function (5) evaluated at $\boldsymbol{c} = \sqrt{\frac{x_{N+1}}{N}} \left[e^{j(x_1 + \arg(c_{0,1}))}, \ldots, e^{j(x_N + \arg(c_{0,N}))} \right]^T$.

3. RADAR CODE OPTIMIZATION

In this section, an iterative algorithm that exploits the CD maximization paradigm[14,15] is developed to get optimized solutions with some quality guarantee to the NP-hard Problem $\mathcal{P}$. The idea is to alternate the maxization of the objective (5) among the entries of $\boldsymbol{x}$, namely optimizing one variable at time while keeping fixed the others. This is tantamount to solving appropriate univariate optimization problems in a loop.[14] With reference to $\bar{\mathcal{P}}$, at each iteration a specific entry of $\boldsymbol{x}$ is selected as variable to optimize. This leads to the following problem at step $n+1$

$$\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}} \begin{cases} \max_{x_d} & \widehat{\mathrm{SINR}}\left(x_d;\boldsymbol{x}^{(n)}_{-d}\right) \\ \text{s.t.} & f_k(x_d;\boldsymbol{x}^{(n)}_{-d}) \leq E_I^k, k=1,\ldots,K \\ & g(x_d) \leq 0 \end{cases} \tag{8}$$

where $\boldsymbol{x}^{(n)}$ is the optimized vector at step n, x_d is the variable to optimize at step $n+1$, and

$$\boldsymbol{x}^{(n)}_{-d} = [x_1^{(n)},\ldots,x_{d-1}^{(n)},x_{d+1}^{(n)},\ldots,x_{N+1}^{(n)}]^T \in \mathbb{R}^N.$$

Moreover,

$$\widehat{\mathrm{SINR}}\left(x_d;\boldsymbol{x}^{(n)}_{-d}\right) = \overline{\mathrm{SINR}}\left(x_1^{(n)},\ldots,x_{d-1}^{(n)},x_d,x_{d+1}^{(n)},\ldots,x_{N+1}^{(n)}\right),$$

is the normalized SINR function restricted to x_d only while the other parameters are those of the previous step; besides,

$$f_k(x_d;\boldsymbol{x}^{(n)}_{-d}) = \begin{cases} \left(\bar{z}_{k,d}^{(n)} + 2\Re\left\{z_{k,d}^{(n)}e^{-jx_d}\right\}\right)x_{N+1}^{(n)}, & d=1,\ldots,N \\ x_d\,p_k^{(n)}, & d=N+1 \end{cases}$$

is the restriction of the k-th spectral constraint, induced by (7), to x_d only keeping fixed the other variable (see[16] for the formal definition of the parameters $\bar{z}_{k,d}^{(n)}, z_{k,d}^{(n)}, p_k^{(n)}$ as well as technical details); finally,

$$g(x_d) = \begin{cases} |x_d| - \delta, & d=1,\ldots,N \\ \left|x_d - \frac{1}{2}\right| - \frac{1}{2}, & d=N+1 \end{cases}$$

Thus, denoting by $x_{d,n+1}^\star$ the optimal solution to $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$, the optimized radar code at step $n+1$ is $\boldsymbol{x}^{(n+1)} = [x_1^{(n)},\ldots,x_{d-1}^{(n)},x_{d,n+1}^\star,\ldots,x_{N+1}^{(n)}]^T$. As a result, starting from a feasible solution $\boldsymbol{x}^{(0)}$ a sequence $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(3)},$ ... of radar codes are obtained iteratively. A summary of the proposed approach can be found in **Algorithm 1**.

Remarkably, being $\boldsymbol{x}^{(n)}$ feasible to $\bar{\mathcal{P}}$ for any n,

$$\begin{aligned} \overline{\mathrm{SINR}}^{(n)} &= \widehat{\mathrm{SINR}}\left(x_d^{(n)};\boldsymbol{x}^{(n)}_{-d}\right) \\ &\leq \widehat{\mathrm{SINR}}\left(x_{d,n+1}^\star;\boldsymbol{x}^{(n)}_{-d}\right) \\ &= \overline{\mathrm{SINR}}^{(n+1)} \end{aligned}$$

implying that the objective function monotonically increases along with iterations; hence, due to the fact that the objective function is bounded (from above) the convergence of the sequence of objective values holds true.

In the next subsections, efficient algorithms to tackle $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$, $d=1,\ldots,N+1$, are developed. From an optimization theory point of view this is the main technical innovation of the paper.

Algorithm 1 Constant Modulus Code Design with Spectral Compatibility Requirements
---
Input: Initial feasible code $\boldsymbol{x}_0$ and the minimum required improvement $\bar{\epsilon}$;
Output: Optimal solution $\boldsymbol{x}^\star$;

1. **Initialization**.

   - Set $d := 1$ and $n := 0$ and compute the initial objective value $\overline{\mathrm{SINR}}^{(n)} = \overline{\mathrm{SINR}}(\boldsymbol{x}_0)$;

2. **Improvement**.

   - Solve $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ obtaining $x^\star_{d,n+1}$;
   - Set $n := n+1$,
     $$\boldsymbol{x}^{(n)} = \left[x_1^{(n-1)}, \ldots, x_{d-1}^{(n-1)}, x^\star_{d,n}, x_{d+1}^{(n-1)}, \ldots, x_{N+1}^{(n-1)}\right]^T,$$
     $\overline{\mathrm{SINR}}^{(n)} = \overline{\mathrm{SINR}}(\boldsymbol{x}^{(n)})$;

3. **Stopping Criterion**.

   - If $\mathrm{mod}(n, N+1) = 0$ and $\left|\overline{\mathrm{SINR}}^{(n)} - \overline{\mathrm{SINR}}^{(n-1)}\right| < \bar{\epsilon}$, stop. Otherwise, update d, i.e., $d = \mathrm{mod}(d - 1; N+1) + 1$, and go to the step 2;

4. **Output**.

   - Set $\boldsymbol{x}^\star = \boldsymbol{x}^{(n)}$.

3.1 Solution Teqcnique for $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$

Let us start with the study of Problem $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ as $d = N + 1$. In this case, it can shown that (8) reduces to a linear programming problem with a positive objective slope. Hence, the optimal solution is just the highest feasible value which is given by

$$x^\star_{N+1,n+1} = \min\left(\min_{k=1,\ldots,K}\left(\frac{E_I^k}{p_k^{(n)}}\right), 1\right)$$

with $p_k^{(n)}$ the parameter involved in the definition of $f_k(x_d; \boldsymbol{x}_{-d}^{(n)})$. With reference to $1 \leq d \leq N$, the following proposition paves the way for an efficient solution of $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$.

PROPOSITION 1. *Problem $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$, $d = 1, \ldots, N$, is equivalent to the following quadratic constrained fractional quadratic problem*

$$\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}} \begin{cases} \max\limits_{t} & \dfrac{a_d^{(n)}t^2 + b_d^{(n)}t - a_d^{(n)}}{1+t^2} \\ s.t. & a_{k,d}^{(n)}t^2 + b_{k,d}^{(n)}t + c_{k,d}^{(n)} \leq 0, \ k = 1, \ldots, K \\ & -\bar{\delta} \leq t \leq \bar{\delta} \end{cases} \tag{9}$$

where $\bar{\delta} = \tan\left(\frac{\delta}{2}\right)$ and $a_d^{(n)}$, $b_d^{(n)}$, $a_{k,d}^{(n)}$, $b_{k,d}^{(n)}$, $c_{k,d}^{(n)}$ ($k = 1, \ldots, K$) are real-valued coefficients depending on $\boldsymbol{x}_{-d}^{(n)}$. Precisely, given an optimal solution $t^\star$ to $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$, $x^\star_{d,n+1} = 2\arctan(t^\star)$ is an optimal solution to $\bar{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$. Moreover, if the feasible set of $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ is unbounded and $v(\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}) = a_d^{(n)}$, $x^\star_{d,n+1} = \pi$.

   Proof. See.[16] □

As shown in[16] the feasible set of $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ can be expressed as union of disjoint closed intervals, i.e.,

$$\mathcal{F} = \cup_{i=1}^{K_1} \mathcal{B}_i, \tag{10}$$

with $K_1 \leq K + 1$ and $\mathcal{B}_i$, $i = 1, \ldots, K_1$, a closed interval whose lower and upper boundaries depend on the specific parameter values of Problem $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$. Since, $\mathcal{F}$ can be evaluated with an overall computational

complexity $O(K \log_2(K))$ and the objective function of Problem $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ exhibits some monotonicities, it is shown in[16] that the optimal solution to $\hat{\mathcal{P}}_{d,\boldsymbol{x}^{(n)}}$ can be obtained in polynomial time. Finally, to initialize **Algorithm 1**, two heuristic procedures are considered. The former relies on the SemiDefinite Relaxation (SDR) and randomization paradigm, where the SDR of $\mathcal{P}$ is obtained removing both the similarity and the rank constraint. The latter considers a radar code design aimed at minimizing the energy transmitted on both licensed and jammed bandwidths, so as to reduce both produced and perceived interference, while fulfilling the constant modulus and similarity constraints. In particular, an iterative procedure based on alternating optimization is devised to handle the resulting optimization problem.

4. PERFORMANCE ANALYSIS

In this section the performance assessment of **Algorithm 1** is conducted. Specifically, achievable detection performance, spectral shape, and autocorrelation features, associated with the devised waveforms are analyzed. To this end, a radar whose baseband equivalent transmitted signal has a two-sided bandwidth of 810 kHz is considered. The disturbance covariance matrix is described as $\boldsymbol{M} = \sigma_0 \boldsymbol{I} + \sum_{k=1}^{K} \frac{\sigma_{I,k}}{\Delta f_k} \boldsymbol{R}_I^k + \sum_{k=1}^{K_J} \sigma_{J,k} \boldsymbol{R}_{J,k}$, where $\sigma_0 = 0$ dB is the thermal noise level; $K = 7$ is the number of licensed emitters; for $K = 1, \ldots, 7$, $\sigma_{I,k}$ accounts for the energy of the k-th coexisting telecommunication network operating on the normalized frequency band $[f_2^k, f_1^k]$ with $\Delta f_k = f_2^k - f_1^k$ the related bandwidth extent; $K_J = 2$ is the number of active and unlicensed narrowband interference sources; $\sigma_{J,k}$, $k = 1, \ldots, K_J$, accounts for the energy of the k-th interference source ($\sigma_{J,1\,dB} = 40$ dB, $\sigma_{J,2\,dB} = 50$ dB); $\boldsymbol{R}_{J,k}$, $k = 1, \ldots, K_J$, is the normalized disturbance covariance matrix of the k-th interference source, whose normalized carrier frequency and bandwidth are $f_{J,k}$ and $\Delta_{J,k}$ ($f_{J,1} = 0.7025$, $f_{J,2} = 0.7525$ and $\Delta_{J,k} = 510^{-3}$, $k = 1, 2$).

As to the overlaid telecommunication systems spectrally coexisting with the radar of interest, the following normalized baseband equivalent radar stop-bands are considered: $\Omega_1 = [0.06, 0.11]$, $\Omega_2 = [0.14, 0.26]$, $\Omega_3 = [0.27, 0.29]$, $\Omega_4 = [0.31, 0.37]$, $\Omega_5 = [0.39, 0.46]$, $\Omega_6 = [0.48, 0.51]$, $\Omega_7 = [0.82, 0.87]$. Hence, it is required that the radar probing waveform fulfills the spectral compatibility requirements corresponding to $E_{I\,dB}^k = 10 \log_{10}(E_I^k) = -30$ dB for $k = 3, 6$ and $E_{I\,dB}^k = -20$ dB for the other frequency bands. Finally, concerning the reference code $\boldsymbol{c}_0$, it is employed a unitary norm Linear Frequency Modulated (LFM) pulse with a duration of 200 μs and a chirp rate $K_s = (750 \times 10^3)/(200 \times 10^{-6})$ Hz/s which corresponds to $N = 162$ samples due to the considered sampling frequency.

In Figure 1 the probability of detection (P_d) versus $|\alpha_T|^2$ (assuming a false alarm probability of $P_{fa} = 10^{-4}$) is shown considering different similarity parameter values[*]. For comparison, it is also plotted the P_d associated with the upper bound to the normalized SINR provided by the SDR of $\mathcal{P}$. The results show that increasing the similarity parameter better and better performance are obtained as a result of the larger degrees of freedom available at the design stage. Interestingly, as ϵ is large enough, i.e. $\epsilon \geq 1.405$ for the case study under test, the upper bound performance is substantially achieved highlighting the effectiveness of the proposed algorithm.

---

[*]Note that for any ϵ, **Algorithm 1** is run with three different initializations: the first is obtained via SDR and randomization approach, the second minimizing the energy transmitted on the occupied bandwidths, and third corresponds to the optimized code at the previous ϵ value. Hence, among the three synthesized codes the one providing the higher $\overline{\text{SINR}}$ is picked up.
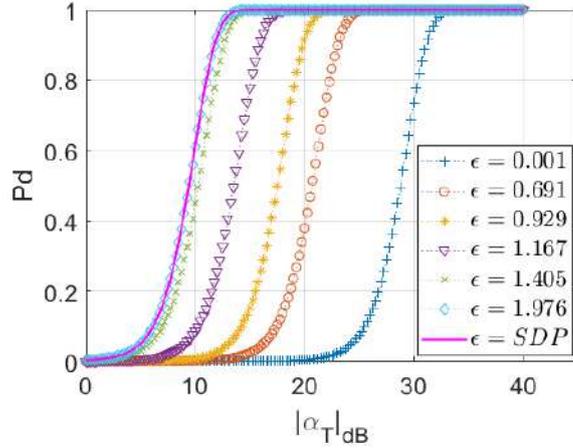
Figure 1. P_d versus $|\alpha_T|^2$ (in dB) of codes synthesized for different ϵ, assuming $P_{fa} = 10^{-4}$.

In Figure 2, the behavior of the codes synthesized through **Algorithm 1** (in terms of ESD and normalized ACF) is reported for three different values of epsilon, i.e., $\epsilon \in \{0.001, 1.405, 1.976\}$. In Figure 2(a), the stopbands (i.e., where the licensed systems are located) are shaded in light gray. Therein, the curves highlight the capability of the devised technique to control the amount of interference energy produced over the shared frequency bandwidths, as required by the enforced local spectral compatibility constraints. In this respect, note that the code designed for $\epsilon = 0.001$ is almost aligned with $\boldsymbol{c}_0$ and experiences just an energy modulation to satisfy the forced spectral constraints. Moreover, an improvement in the "useful" energy distribution is observed as the similarity parameter increases. As a result, interference rejection of unlicensed sources is achieved via a reduction of the radar emission in correspondence of the shared frequencies. Finally, inspection of Figure 2(b) suggests that better P_d values and subsequently, interference rejection are traded for ideal range resolutions and/or ISL/PSL.
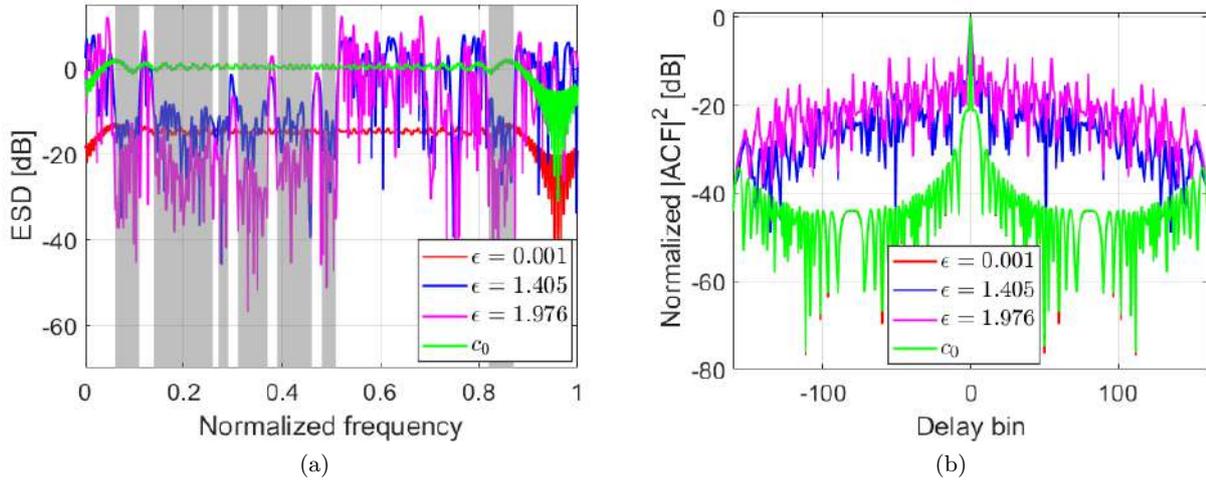
Figure 2. a) ESD; b) Squared modulus of the ACF. Green curve: reference code $\boldsymbol{c}_0$; red curve: $\epsilon = 0.001$; blue curve: $\epsilon = 1.405$; magenta curve: $\epsilon = 1.976$.

5. CONCLUSION

Radar waveform design intended for addressing the spectrum congestion problem while complying with the current amplifiers technology has been explored. Specifically, multiple spectral compatibility constraints have been enforced at the design stage to ensure a tight and local control on the interference energy induced on each shared/reserved bandwidth. Furthermore, to account for hardware limitations, the radar waveform is constrained to the class of constant modulus signals. Hence, a polynomial computational complexity solution technique, with ensured convergence properties, has been developed to synthesize optimized radar waveforms in terms of achieved SINR. The performance of the devised signals has been analyzed studying the trade-off among the achievable detection probability, spectral shape, and ACF features. As possible future research line, it might be worth to account for the presence of signal-dependent interference at the design stage.

REFERENCES

[1] He, H., Stoica, P., and Li, J., "Waveform design with stopband and correlation constraints for cognitive radar," in [*2010 2nd International Workshop on Cognitive Information Processing*], (14-16 Junel 2010).

[2] Griffiths, H., Cohen, L., Watts, S., Mokole, E., Baker, C., Wicks, M., and Blunt, S., "Radar spectrum engineering and management: Technical and regulatory issues," *Proceedings of the IEEE* **103**(1), 85–102 (2015).

[3] Wicks, M., "Spectrum crowding and cognitive radar," in [*2010 2nd International Workshop on Cognitive Information Processing*], 452–457 (14-16 June 2010).

[4] Gerlach, K., "Thinned spectrum ultrawideband waveforms using stepped-frequency polyphase codes," *IEEE Transactions on Aerospace and Electronic Systems* **34**, 1356–1361 (Oct 1998).

[5] Aubry, A., De Maio, A., Piezzo, M., and Farina, A., "Radar waveform design in a spectrally crowded environment via nonconvex quadratic optimization," *IEEE Transactions on Aerospace and Electronic Systems* **50**, 1138–1152 (April 2014).

[6] Aubry, A., De Maio, A., Piezzo, M., Naghsh, M. M., Soltanalian, M., and Stoica, P., "Cognitive radar waveform design for spectral coexistence in signal-dependent interference," in [*2014 IEEE Radar Conference*], 0474–0478 (May 2014).

[7] Aubry, A., Carotenuto, V., and De Maio, A., "Forcing multiple spectral compatibility constraints in radar waveforms," *IEEE Signal Processing Letters* **23**, 483–487 (April 2016).

[8] Aubry, A., Carotenuto, V., De Maio, A., Farina, A., and Pallotta, L., "Optimization theory-based radar waveform design for spectrally dense environments," *IEEE Aerospace and Electronic Systems Magazine* **31**, 14–25 (December 2016).

[9] Wu, L., Babu, P., and Palomar, D. P., "Transmit waveform/receive filter design for mimo radar with multiple waveform constraints," *IEEE Transactions on Signal Processing* **66**, 1526–1540 (March 2018).

[10] Cheng, Z., Liao, B., He, Z., and Li, J., "Spectrally compatible waveform design for mimo radar in the presence of multiple targets," *IEEE Transactions on Signal Processing* **66**, 3543–3555 (July 2018).

[11] Cui, G., Yu, X., Foglia, G., Huang, Y., and Li, J., "Quadratic optimization with similarity constraint for unimodular sequence synthesis," *IEEE Transactions on Signal Processing* **65**, 4756–4769 (September 2017).

[12] Zhao, Y., Gaeddert, J., Bae, K., and Reed, J., "Radio environment map-enabled situation-aware cognitive radio learning algorithms," in [*Proceedings of Software Defined Radio Technical Conference*], (November 2006).

[13] Aubry, A., Carotenuto, V., De Maio, A., and Govoni, M. A., "Multi-snapshot spectrum sensing for cognitive radar via block-sparsity exploitation," *IEEE Transactions on Signal Processing* **67**, 1396–1406 (March 2019).

[14] Wright, S. J., "Coordinate descent algorithms," *Math. Programm.* **151**(1), 3–34 (2015).

[15] Razaviyayn, M., Hong, M., and Luo, Z., "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization* **23**(2), 1126–1153 (2013).

[16] Aubry, A., De Maio, A., Martino, L., and Govoni, M. A., "Constant modulus radar waveform design for spectral compatibility," *submitted on IEEE Transactions on Signal Processing* (2018).

A Constrained Least Squares Approach for 2D PBR Localization

Augusto Aubry[a], Vincenzo Carotenuto[b], Antonio De Maio[a], and Luca Pallotta[c]

[a]Università di Napoli "Federico II", Department of Electrical Engineering and Technology Information (DIETI), via Claudio 21, I-80125 Napoli, Italy;
[b]CNIT udr Università "Federico II", via Claudio 21, I-80125 Napoli, Italy;
[c]Università "Roma Tre" (previously with CNIT), Engineering Department, via Vito Volterra 62, I-00146 Roma, Italy.

ABSTRACT

A new algorithm for Passive Bistatic Radar (PBR) localization exploiting multiple illuminators of opportunity is proposed. To capitalize a-priori information on the receiving antenna beampattern extent, specific constraints are forced to the target localization process. At the estimator design stage the problem is formulated according to the constrained Least Squares (LS) framework. The results highlight that interesting improvements with respect to some counterparts can be achieved.

Keywords: Passive Bistatic Radar (PBR), Multiple Transmitter of Opportunity, Elliptic Localization, Range Measurements

1. INTRODUCTION

Passive Bistatic Radar (PBR) is widely recognized among the most promising and effective passive sensing technologies.[1,2] PBR based localization is accomplished using bistatic target range measurements obtained collecting the target echoes resulting from the signals transmitted by one/multiple illuminators of opportunity.[2,3] To this end, PBR receivers are equipped with two receiving channels per transmitter of oppurtunity: one is used to acquire the direct path signal from the selected emitter, the other to gather the induced echoes. Each pair of measurements allows to localize the target over an ellipse and the intersection of multiple ellipsoids due to different illuminator-passive receiver pairs paves the way for target positioning.[2,3]

This paper proposes an innovative approach for elliptic location with reference to a 2D PBR receiver exploiting multiple transmitters of opportunity. At the design stage, it is assumed that the receiving antenna has a 2D directional beampattern and specific constraints to account for its main-beam size are enforced to the target localization process. Therefore, positioning is formulated as a constrained Least Squares (LS) estimation whose optimal solution provides the Cartesian coordinates of the target. The resulting non-convex optimization problem can be efficiently solved resorting to the theory of Generalized Trust Region Subproblems (GTRSs),[4] already exploited for the case of hyperbolic location.[5] At the analysis stage, the performance of the proposed algorithm is assessed in comparison with the unconstrained LS positioning solution, the Two-Step Estimation (TSE) procedure[6] which estimates the target position via an iterative algorithm, and the CRLB on the target Cartesian coordinates that provides a benchmark on the achievable estimation accuracy.

1.1 Notation

We adopt the notation of using boldface for vectors $\boldsymbol{a}$ (lower case), and matrices $\boldsymbol{A}$ (upper case). The n-th element of $\boldsymbol{a}$ and the (m, l)-th entry of $\boldsymbol{A}$ are respectively denoted by a_n and $\boldsymbol{A}_{m,n}$. The symbols $(\cdot)^T$ indicates the transpose operator, while $\text{tr}(\cdot)$ is the trace of its matrix argument. $\boldsymbol{A}^\dagger$ represents the Moore-Penrose inverse of the matrix $\boldsymbol{A}$. $\boldsymbol{I}$ and $\boldsymbol{0}$ denote respectively the identity matrix and the matrix with zero entries (their size is determined from the context). $\mathbb{R}^N$, $\mathbb{R}^{N,M}$, and $\mathbb{S}^N$ are respectively the sets of N-dimensional vectors of real

---

Further author information: (Send correspondence to A. De Maio)
A. De Maio: E-mail: ademaio@unina.it, Telephone: +39 081 76 83101

numbers, of $N \times M$ real matrices, and of $N \times N$ symmetric matrices. $\mathbf{diag}\,(\boldsymbol{a})$ indicates the diagonal matrix whose i-th diagonal element is the i-th entry of $\boldsymbol{a}$. The curled inequality symbol $\succeq$ (and its strict form $\succ$) is used to indicate generalized matrix inequality: for any $\boldsymbol{A} \in \mathbb{S}^N$, $\boldsymbol{A} \succ \boldsymbol{0}$ means that $\boldsymbol{A}$ is a positive definite matrix. $\lambda_1(\boldsymbol{X}), \lambda_2(\boldsymbol{X}), \ldots, \lambda_N(\boldsymbol{X})$, with $\lambda_1(\boldsymbol{X}) \geq \lambda_2(\boldsymbol{X}) \geq \ldots \geq \lambda_N(\boldsymbol{X})$, denote the eigenvalues of $\boldsymbol{X} \in \mathbb{S}^N$, arranged in decreasing order. Furthermore, given $\boldsymbol{B} \succ \boldsymbol{0}$ and $\boldsymbol{A} \in \mathbb{S}^N$, the generalized eigenvalues of the matrix pair $(\boldsymbol{A}, \boldsymbol{B})$ are given by $\lambda_i(\boldsymbol{A}, \boldsymbol{B}) = \lambda_i(\boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2})$, $i = 1, \ldots, N$. The Euclidean norm of the vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|$. The letter i often serves as index. For any optimization Problem $\mathcal{P}$, $v(\mathcal{P})$ represents its optimal value.

2. SYSTEM MODEL

Consider a 2D passive bistatic radar exploiting multiple transmitters of opportunity. Then, denote by:

- $(x_p, y_p) \in \mathbb{R}^2$ the target position;

- $(x_0, y_0) \in \mathbb{R}^2$ the receiver position (without loss of generality, it is assumed to coincide with the reference system origin, i.e., $(x_0, y_0) = (0,0)$);

- $(x_{t_i}, y_{t_i}) \in \mathbb{R}^2$ the position of the i-th transmitter of opportunity, $i = 1, \ldots, N$;

- $L_i = \sqrt{(x_{t_i} - x_0)^2 + (y_{t_i} - y_0)^2}$ the distance between the i-th transmitter and the receiver.

- $\boldsymbol{p} = [x_p, y_p]^T$ the target position vector.

At the receiver side, after the classic PBR cross-correlation based processing, the following N delay/range measurements are available

$$\tau_i = \tilde{\tau}_i + n_i, \ i = 1, \ldots, N, \tag{1}$$

where, for $i = 1, \ldots, N$,

$$\tilde{\tau}_i = \frac{1}{c}\left(\|\boldsymbol{p}\| + \sqrt{(x_p - x_{t_i})^2 + (y_p - y_{t_i})^2} - L_i\right), \tag{2}$$

with c the speed of light and $n_1, \ldots, n_N$ statistically independent zero-mean (usually assumed Gaussian distributed, even if this assumption is not mandatory for the present paper) random variables with variance $\sigma_1^2, \ldots, \sigma_N^2$. In particular,

$$\sigma_i = \frac{\sqrt{2}}{B_i\sqrt{\mathrm{SNR}_i}}, \ i = 1, \ldots, N, \tag{3}$$

where B_i is the frequency bandwidth of the i-th signal of opportunity and SNR_i is the Signal to Noise Ratio (SNR) of the i-th bistatic pair (i.e., receiver/i-th transmitter of opportunity) evaluated according to the bistatic radar range equation.[7,8]

Now, elaborating on (2), it is possible to get an equivalent form which is fundamental for the proposed estimation algorithm (referred to in the following as Angular Constrained Least Square (ACLS)). To this end, denoting by

$$b_i = \tilde{\tau}_i c + L_i \quad i = 1, \ldots, N, \tag{4}$$

equation (2) can be recast as

$$b_i - \sqrt{x_p^2 + y_p^2} = \sqrt{(x_p - x_{t_i})^2 + (y_p - y_{t_i})^2}, i = 1, \ldots, N, \tag{5}$$

or equivalently as

$$\begin{cases} b_i^2 + r^2 - 2b_i r = r^2 + r_i^2 - 2x_{t_i}x_p - 2y_{t_i}y_p, \\ r \leq b_i \qquad\qquad\qquad\qquad i = 1, \ldots, N, \\ r = \sqrt{x_p^2 + y_p^2} \end{cases} \tag{6}$$

with $r_i = \sqrt{x_{t_i}^2 + y_{t_i}^2}$, $i = 1, \ldots, N$. All the relationships described in (6) can be also framed in a more compact matrix form

$$\begin{cases} A\bar{p} - g = 0 \\ \bar{p}^T B\bar{p} = 0 \\ \bar{p}_3 \leq b_i, \ i = 1, \ldots, N \end{cases} \tag{7}$$

where

- $\bar{p} = [p^T, r]^T \in \mathbb{R}^3$;

- $A^T = [a_1, a_2, \ldots, a_N] \in \mathbb{R}^{3,N}$, with $a_i = [-2x_{t_i}, -2y_{t_i}, 2b_i]^T \in \mathbb{R}^3$, $i = 1, \ldots, N$

- $g = [b_1^2 - r_1^2, b_2^2 - r_2^2, \ldots, b_N^2 - r_N^2]^T \in \mathbb{R}^N$;

- $B = \mathbf{diag}\{[1, 1, -1]\} \in \mathbb{R}^{3,3}$.

2.1 Receiver Beampattern Constraint

To perform the search process, conventional 2D PBR systems employ a scanning antenna (in the azimuth domain) characterized by a specific beampattern and in particular a given mainlobe width. In this subsection, some constraints able to capitalize such a-priori information are formalized with the goal of improving target localization reliability. To this end, let us denote by:

- $\bar{\theta}$ the receiving (half-side) antenna beamwidth (with respect to the boresight), with $0 \leq \bar{\theta} < \pi/2$;

- $\theta \in]-\pi, \pi[$ the squint angle of the antenna boresight with respect to the x-axis;

- For $(x, y) \neq (0, 0)$, $\theta_p = \text{atan2}(y, x)$ the target angle of arrival, where $\text{atan2}(\cdot)$ is the four-quadrants inverse tangent.

Suppose, now, that the target is within the receiving antenna main-beam, i.e., that

$$\theta - \bar{\theta} \leq \theta_p \leq \theta + \bar{\theta} \tag{8}$$

or equivalently

$$-\bar{\theta} \leq \theta_p - \theta \leq \bar{\theta}. \tag{9}$$

Otherwise stated, it is assumed that the detector performs correctly its task, i.e., the target triggering the detection resides in the antenna main-beam.

Equation (9) coupled with the assumption $0 \leq \bar{\theta} < \frac{\pi}{2}$ can be equivalently rewritten as

$$-\tan\bar{\theta} \leq \tan(\theta_p - \theta) \leq \tan\bar{\theta} \tag{10}$$

and

$$|\theta_p - \theta| < \frac{\pi}{2}. \tag{11}$$

Now, the following relationships on the tangent function hold

$$\tan(\theta_p - \theta) = \frac{\sin(\theta_p - \theta)}{\cos(\theta_p - \theta)} = \frac{\sin(\theta_p)\cos(\theta) - \cos(\theta_p)\sin(\theta)}{\cos(\theta_p)\cos(\theta) + \sin(\theta_p)\sin(\theta)} = \frac{y_p\cos\theta - x_p\sin\theta}{x_p\cos\theta + y_p\sin\theta}, \tag{12}$$

where the last equality stems from $\sin\theta_p = y_p/\sqrt{x_p^2 + y_p^2}$ and $\cos\theta_p = x_p/\sqrt{x_p^2 + y_p^2}$.

Exploiting (12), (10) can be written in a more useful form i.e. as

$$-\tan\bar{\theta} \leq \frac{y_p \cos\theta - x_p \sin\theta}{x_p \cos\theta + y_p \sin\theta} \leq \tan\bar{\theta}. \tag{13}$$

Let us now manipulate the previous inequalities introducing a new reference system, say (x_1, y_1), obtained rotating the actual one (i.e., the (x, y)-coordinates system) such that the x_1-axis is aligned with the receiving antenna boresight. Precisely, denoting by $\bar{R}(\theta)$ the rotation matrix of an angle θ clockwise, i.e.,

$$\bar{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{14}$$

the new (x_1, y_1)-coordinates system is related to the previous (x, y)-coordinates system by means of the following transformation

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \bar{R}(\theta) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{bmatrix}. \tag{15}$$

Consequently, (13) becomes

$$-\tan\bar{\theta} \leq \frac{y_{1p}}{x_{1p}} \leq \tan\bar{\theta}. \tag{16}$$

From the relationship

$$|\text{atan2}(y_{1p}, x_{1p})| = |\text{atan2}(y, x) - \theta| < \frac{\pi}{2}, \tag{17}$$

it directly follows that $x_{1p} > 0$. Hence, the additional constraints to consider at the estimator design stage as induced by the antenna beamwidth are given by

$$\begin{cases} -x_{1p}\tan\bar{\theta} \leq y_{1p} \leq x_{1p}\tan\bar{\theta} \\ x_{1p} > 0 \end{cases}. \tag{18}$$

Finally, to take in account also the point with coordinates $(x_{1p}, y_{1p}) = (0, 0)$ and letting $\gamma = \tan\bar{\theta}$, the constraints in (18) can be expressed as

$$\begin{cases} -x_{1p}\gamma \leq y_{1p} \leq x_{1p}\gamma \\ x_{1p} \geq 0 \\ \begin{bmatrix} x_{1p} \\ y_{1p} \end{bmatrix} = \bar{R}(\theta) \begin{bmatrix} x_p \\ y_p \end{bmatrix} \end{cases}. \tag{19}$$

3. PROBLEM FORMULATION

Starting from (7), it can be observed that the first equation holds only approximately in practical situations due to measurement errors (1). A viable mean to overcome this shortcoming is to resort to the constrained LS framework and formalize the estimation problem as

$$\begin{cases} \min_{\bar{p}} & \|A\bar{p} - g\|^2 \\ \text{s.t.} & \bar{p}^T B \bar{p} = 0 \\ & 0 \leq \bar{p}_3 \leq b \end{cases} \tag{20}$$

with $b = \min_{i=1,\ldots,N} b_i$. Notice that, with a slight abuse of notation, the model parameters in (20), i.e., A, g, and b, are computed exploiting the actual measurements τ_i, $i = 1, \ldots, N$, in place of $\tilde{\tau}_i$, $i = 1, \ldots, N$.

In the current form the considered estimation problem (20) does not account for the receive antenna main-beam constraints introduced in 2.1. Thus, to proceed further, let us introduce the target position vector in the rotated (x_1, y_1)-coordinates reference system, indicated in the following as $\tilde{\boldsymbol{p}}$. Precisely, the position vector in the rotated coordinate system is given by $\tilde{\boldsymbol{p}} = \boldsymbol{U}^T \bar{\boldsymbol{p}}$, where

$$\boldsymbol{U} = \begin{bmatrix} \bar{\boldsymbol{R}}(\theta)^T & 0 \\ 0 & 1 \end{bmatrix} \tag{21}$$

with $\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{I}$. As a consequence, elliptic localization problem with the addition of the beampattern constraints (19), is formulated as

$$\mathcal{P} \begin{cases} \min_{\tilde{\boldsymbol{p}}} \left\| \tilde{\boldsymbol{A}} \tilde{\boldsymbol{p}} - \boldsymbol{g} \right\|^2 & \\ \text{s.t. } \tilde{\boldsymbol{p}}^T \boldsymbol{B} \tilde{\boldsymbol{p}} = 0 & \text{(22a)} \\ 0 \leq \tilde{p}_3 \leq b & \text{(22b)} \\ -\tilde{p}_1 \gamma \leq \tilde{p}_2 \leq \tilde{p}_1 \gamma & \text{(22c)} \\ \tilde{p}_1 \geq 0 & \text{(22d)} \end{cases}$$

where $\tilde{\boldsymbol{A}} = \boldsymbol{A} \boldsymbol{U}$ and $\boldsymbol{U}^T \boldsymbol{B} \boldsymbol{U} = \boldsymbol{B}$.

Problem $\mathcal{P}$ is a non-convex optimization problem apparently difficult to solve. However, leveraging the special structure of the objective function/constraints and resorting to the theory of GTRSs,[4] a closed-form optimal solution to $\mathcal{P}$ is now derived. A first step toward this goal is provided by the following lemma

LEMMA 3.1. *Any feasible point $\boldsymbol{x} \neq (0, 0, 0)$ is regular for the optimization Problem $\mathcal{P}$.*

The following proposition establishes a procedure leading to a closed-form global optimal solution of the constrained estimation problem $\mathcal{P}$. This represents the main technical contribution of the present paper from an optimization theory point of view.

PROPOSITION 3.2. *An optimal solution to $\mathcal{P}$ belongs to the following finite set of feasible points (whose cardinality is at most thirteen):*

1. $\bar{\boldsymbol{x}}_0^\star = \boldsymbol{0}$.

2. $\bar{\boldsymbol{x}}^\star(\eta_h) = \left(\tilde{\boldsymbol{A}}^T \tilde{\boldsymbol{A}} + \eta_h \boldsymbol{B} \right)^{-1} \tilde{\boldsymbol{A}}^T \boldsymbol{g}$, *where η_h, $h = 1, \ldots, N_r$ $(N_r \leq 4)$, are the roots of the fourth order equation*

$$\bar{\boldsymbol{x}}^\star(\eta)^T \boldsymbol{B} \bar{\boldsymbol{x}}^\star(\eta) = 0 \tag{23}$$

with

$$\eta \in \left(-\frac{1}{\lambda_2 \left(\boldsymbol{B}, \tilde{\boldsymbol{A}}^T \tilde{\boldsymbol{A}} \right)}, +\infty \right) - \left\{ -\frac{1}{\lambda_1 \left(\boldsymbol{B}, \tilde{\boldsymbol{A}}^T \tilde{\boldsymbol{A}} \right)}, -\frac{1}{\lambda_3 \left(\boldsymbol{B}, \tilde{\boldsymbol{A}}^T \tilde{\boldsymbol{A}} \right)} \right\} \tag{24}$$

such that

$$\begin{cases} 0 < \bar{x}_3^\star(\eta_h) < b \\ -\gamma \bar{x}_1^\star(\eta_h) \leq \bar{x}_2^\star(\eta_h) \leq \gamma \bar{x}_1^\star(\eta_h) \\ \bar{x}_1^\star(\eta_h) > 0 \end{cases} . \tag{25}$$

3. $\bar{\boldsymbol{x}}^\star(\beta_h) = [\tilde{\boldsymbol{q}}^\star(\beta_h)^T, b]^T$ *with*

$$\tilde{\boldsymbol{q}}^\star(\beta_h) = \left(\tilde{\boldsymbol{A}}_1^T \tilde{\boldsymbol{A}}_1 + \beta_h \boldsymbol{I} \right)^{-1} \tilde{\boldsymbol{A}}_1^T \left(\boldsymbol{g} - \tilde{\boldsymbol{a}}_3 b \right),$$

where $\tilde{\boldsymbol{A}} = \left[\tilde{\boldsymbol{A}}_1, \tilde{\boldsymbol{a}}_3 \right]$ and β_h, $h = 1, \ldots, N_{r_1}$ $(N_{r_1} \leq 4)$ are the roots of the fourth order equation

$$\tilde{\boldsymbol{q}}^\star(\beta)^T \tilde{\boldsymbol{q}}^\star(\beta) = b \tag{26}$$

such that

$$\begin{cases} \beta_h \geq -\lambda_{max}\left(\tilde{A}_1^T \tilde{A}_1\right) - \left\{-\lambda_{min}\left(\tilde{A}_1^T \tilde{A}_1\right)\right\} \\ -\gamma \tilde{q}_2^\star(\beta_h) < \tilde{q}_1^\star(\beta_h) < \gamma \tilde{q}_2^\star(\beta_h) \\ \tilde{q}_1^\star(\beta_h) > 0 \end{cases} . \tag{27}$$

4. $\bar{x}_{4_i}^\star = \left[b/\sqrt{1+\gamma^2}, (-1)^{i+1}\gamma b/\sqrt{1+\gamma^2}, b\right]^T$, *for* $i = 1, 2$.

5. $\bar{x}_{5_i}^\star = \alpha_\pm^\star \left[1, (-1)^{i+1}\gamma, \sqrt{1+\gamma^2}\right]^T$, $i = 1, 2$, *with*

$$\alpha_\pm^\star = \max\left(0, \frac{v_{\pm,1}g_1 + v_{\pm,2}g_2 + v_{\pm,3}g_3}{v_{\pm,1}^2 + v_{\pm,2}^2 + v_{\pm,3}^2}\right).$$

and $v_\pm = \tilde{A}\left[1, \pm\gamma, \sqrt{1+\gamma^2}\right]^T$.

4. PERFORMANCE ANALYSIS

This section is devoted to the performance assessment of the proposed target localization algorithm for a PBR exploiting multiple transmitters of opportunity. As case study, a localization scenario comprising $N = 3$ omnidirectional broadcast transmitters of opportunity is analyzed. As to the specific geometric configuration, the transmitters are located at the vertices of an equilateral triangle whose barycenter is the position of the receiver, that coincides with the origin of the reference system. Precisely, the distance of each transmitter of opportunity from the receiver is $L_i = 10$ km, $i = 1, 2, 3$. The considered setup is graphically illustrated in Figure 1.
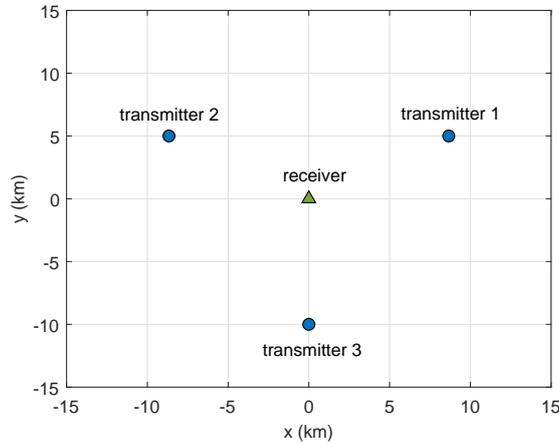
Figure 1. Geometric configuration composed of a receiver at the reference system origin and $N = 3$ transmitters of opportunity located at the vertices of an equilateral triangle ($L_i = 10$ km, $i = 1, 2, 3$).

Modelling the measurement errors according to (3), the SNR of the $N = 3$ bistatic pairs (receiver-transmitter of opportunity) is set as[7]

$$\text{SNR}_i = \text{SNR}_0 \frac{\|q_0\|^2}{\|p\|^2} \frac{\|q_0 - t_1\|^2}{\|p - t_i\|^2}, \quad i = 1, 2, 3, \tag{28}$$

where $t_i = [x_{t_i}, y_{t_i}]^T$, $i = 1, 2, 3$, are the locations of the transmitters of opportunity, p is the true target position, and SNR_0 is a reference SNR value computed via the bistatic radar range equation[7,8] related to a reference point $q_0 = [\bar{x}, \bar{y}]^T$ and considering the transmitter located at t_1 as radiation source.

The performance of the devised localization algorithm is assessed considering as figure of merit the Root Mean Square Error (RMSE) of the target position estimate. To this end, due to the lack of a closed-form expression for the RMSE, Monte Carlo simulation method is employed, performing 1000 independent runs. The analyses are conducted also in comparison with the LS and TSE[6] algorithms, and with the CRLB benchmark, i.e. $\sqrt{\text{tr}\left(\text{FIM}^{-1}\right)}$ with FIM the Fisher Information Matrix.

The following study considers a target located at $(x_p, y_p) = (r\cos\theta_p, r\sin\theta_p)$ with $r = 40$ km and different values of θ_p, $\theta_p = 0, 7, 9$ deg. Moreover, the receiving antenna is assumed steered at $\theta = 0$ deg with a main-beam width of $\bar{\theta} = 10$ deg.

The obtained results are reported in Figure 2, where the RMSE is plotted versus SNR_0 for $B_i = 150$ kHz (representative of FM radio stations), $i = 1, 2, 3$. Therein, SNR_0 coincides with the actual SNR, regardless of the target position. Indeed, it is assumed that the reference point $\boldsymbol{q}_0$ involved in the SNR computation coincides with the actual target position $\boldsymbol{p}$, therefore $\boldsymbol{q}_0$ changes among the subplots.

(a) $\theta_p = 9°$.

(b) $\theta_p = 7°$.

(c) $\theta_p = 0°$.
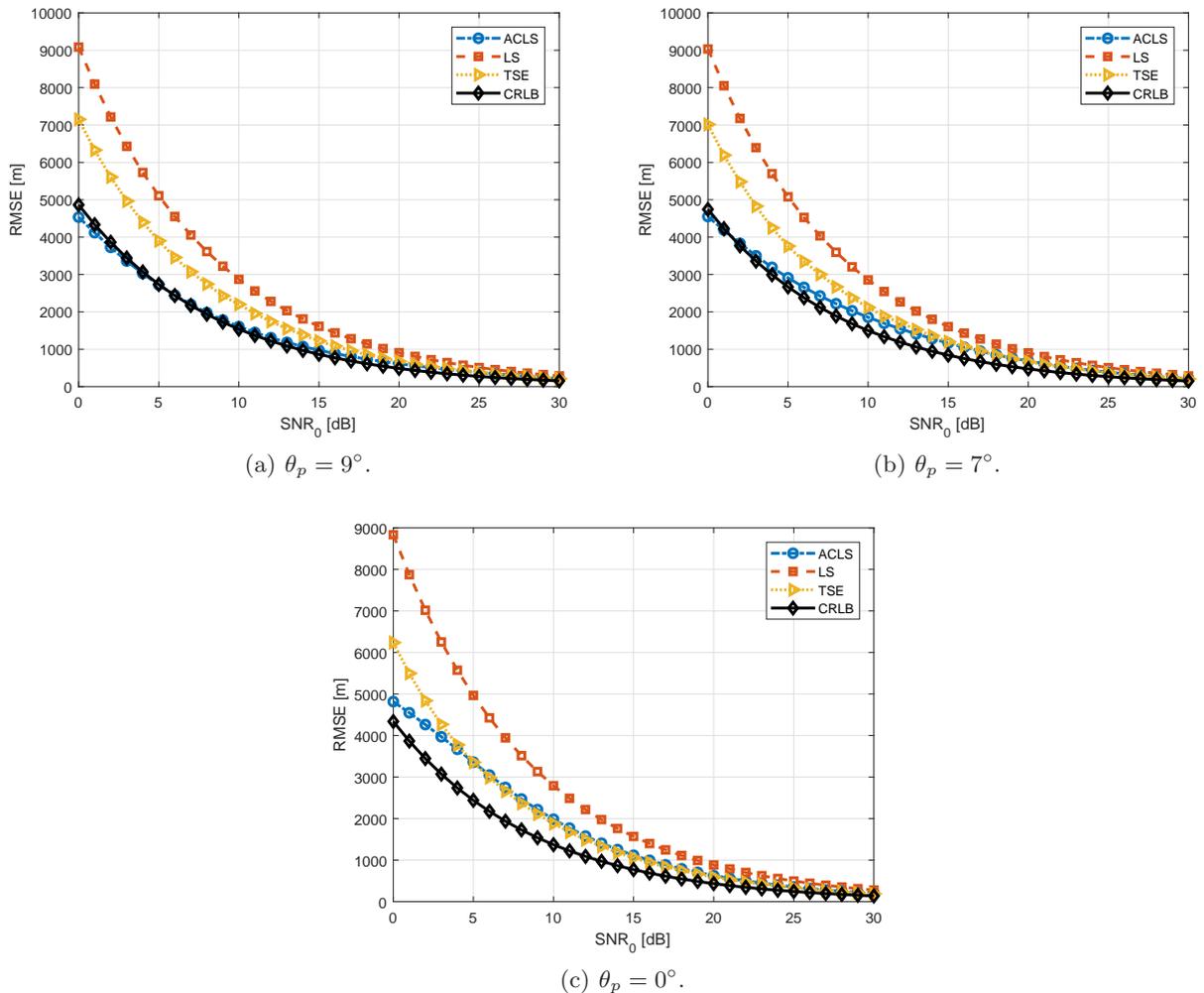
Figure 2. RMSE versus SNR_0 assuming $\boldsymbol{p} = \boldsymbol{q}_0$ considering $B_i = 150$ kHz, $i = 1, 2, 3$.

The curves show that for low SNR values, the ACLS provides better performance than the unconstrained LS and TSE counterparts, clearly revealing the effectiveness of the new procedure to exploit a-priori information on the beampattern main-beam size. Specifically, the RMSE values achieved by ACLS are close to the CRLB. Interestingly, as the SNR increases, all the considered positioning algorithms, i.e., ACLS, unconstrained LS, and

TSE, improve their location estimation performance providing RMSE values converging to the CRLB benchmark. Finally, it is worth observing that the closer the target position to the beampattern boundary the better the performance of the ACLS method as compared with the counterparts. This is not surprising since the a-priori information on the beampattern extent becomes more and more valuable when the target lies in proximity of the boundary region.

5. CONCLUSIONS

In this paper an innovative approach for the elliptic location has been proposed with reference to a PBR receiver that exploits multiple transmitters of opportunity. To this end, at the design stage, some constraints have been enforced to the target localization process in order to exploit a-priori information on the receiving antenna mainbeam size. Therefore, the problem has been formulated as a constrained LS estimation whose optimal solution provides the Cartesian coordinates of the target. To handle the resulting non-convex optimization problem, an efficient solution technique able to provide a global optimal point has been provided.

The conducted analyses have demonstrated the effectiveness of the proposed algorithm also in comparison with other counterparts available in the open literature especially for low SNRs. Precisely, for the considered study cases, the curves have highlighted that the RMSE values attained by the novel technique are usually lower than those of the competitors, therefore validating the benefits provided by the additional constraints on the receiving beampattern. As a matter of fact, the obtained results have shown values very close to those of theoretical CRLB.

REFERENCES

[1] Griffiths, H. D. and Baker, C. J., [*An Introduction to Passive Radar*], Artech House (2017).

[2] Klemm, R., Nickel, U., Gierull, C., Lombardo, P., Griffiths, H., and Koch, W., eds., [*Novel Radar Techniques and Applications*], vol. 1, Scitech publishing (2017).

[3] Melvin, W. L. and Scheer, J. A., [*Principles of Modern Radar: Radar Applications*], Schitech Publishing (2014).

[4] More', J. J., "Generalizations of the Trust Region Subproblem," *Optim. Methods Softw.* **2**, 189–209 (August 1993).

[5] Beck, A., Stoica, P., and Li, J., "Exact and Approximate Solutions of Source Localization Problems," *IEEE Transactions on Signal Processing* **56**, 1770–1778 (May 2008).

[6] Shen, J., Molisch, A., and Salmi, J., "Accurate Passive Location Estimation Using TOA Measurements," *IEEE Transactions on Wireless Communications* **11**, 2182–2192 (June 2012).

[7] Richards, M. A., Scheer, J. A., and A., H. W., [*Principles of Modern Radar: Basic Principles*], Schitech Publishing (2010).

[8] Anastasio, V., Farina, A., Colone, F., and Lombardo, P., "Cramer - Rao Lower Bound with Pd < 1 for Target Localisation Accuracy in Multistatic Passive Radar," *IET Radar, Sonar and Navigation* **8**, 767–775 (August 2014).

Photonics applied to Coherent Radar Networks for Border Security

S. Maresca*[a], G. Serafino[b], F. Scotti[b], L. Lembo[ac], P. Ghelfi[b] and A. Bogoni[ab]

[a]TeCIP Institute, Scuola Superiore Sant'Anna, Pisa, Italy;
[b]PNTLab, Consorzio Nazionale Interuniversitario per le Telecomunicazioni (CNIT), Pisa, Italy;
[c]Vallauri Institute, Centro di Supporto e Sperimentazione Navale (CSSN), Livorno, Italy.

ABSTRACT

This paper investigates the target detection and localization capabilities of a centralized radar network with widely separated antennas in the framework of maritime border security. The proposed system takes advantage from photonic technology for signal generation/detection and signal distribution between the central and the remote radar nodes. Such a solution allows to achieve the necessary level of time and phase synchronization to benefit from the coherent multiple input multiple output (MIMO) paradigm for very high cross-range resolution.

The proposed photonics-based radar network is developed within the NATO-funded "Multistatic and multi-band coherent radar fleet for border security (SOLE)" project. Currently, the radar network architecture consists of two transmit/receive radar front-ends, for a total of four coherent virtual channels, employing linear frequency modulated pulses in time diversity at 9.7 GHz with 100 MHz bandwidth. The system has been operated in a preliminary down-scaled outdoor scenario for detecting two collaborative closely-spaced moving targets. With respect to the nominal aperture of the antenna, the cross-range resolution has been improved by a factor of five. The preliminary results demonstrate the impact of photonics applied to coherent centralized MIMO radars and suggest the possibility to soon endow the radar network with high-performance radar imaging capabilities.

Keywords: Maritime border security, centralized radar network, MIMO radar, coherence, widely-distributed antennas, photonics-based radar, radio-over-fiber system.

1. INTRODUCTION

Motivated by the ever-evolving operative conditions and number/variety of possible threats, radar-based border surveillance applications, especially in the maritime domain, see a predominance of solutions based on sensor fusion. Standalone monostatic/bistatic radars cannot meet all the required mission tasks. They are subject to strong target fluctuations, i.e. the spatial/temporal variations of the measured target radar cross section (RCS). For complex or stealth targets, small variations in the target range and orientation with respect to the sensor could result in large RCS variations (e.g. 20 dB), thus inevitably leading to significant degradations of the signal-to-noise ratio (SNR).

Multistatic radars, which employ multiple closely and/or widely separated transmitting and receiving radar nodes to monitor a common surveillance area [1], can capture a larger percentage of the power backscattered from the target, by virtue of different transmitter-target-receiver geometries (i.e. aspect diversity). However, due to realization costs and strong sensor deployment constraints, the radar nodes seldom cooperate with each other. In fact, they typically provide the central node, i.e. the one responsible for information fusion, with their pre-elaborated pictures of the surveyed area (i.e. decentralized processing). This local pre-processing limits the synchronization procedure only to time and reduces the amount of data to be shared between the remote nodes and the fusion center. However, the information extraction process is not truly maximized.

The multiple input multiple output (MIMO) approach for networked radars grants better coordination and information extraction than the multistatic radar approach [2]. The main difference between the two radar architectures is in the capability of MIMO radars to jointly process the data streams coming from all the remote

---

For further author information, send correspondence to Salvatore Maresca:
E-mail: salvatore.maresca@santannapisa.it, Telephone: +39050882233.
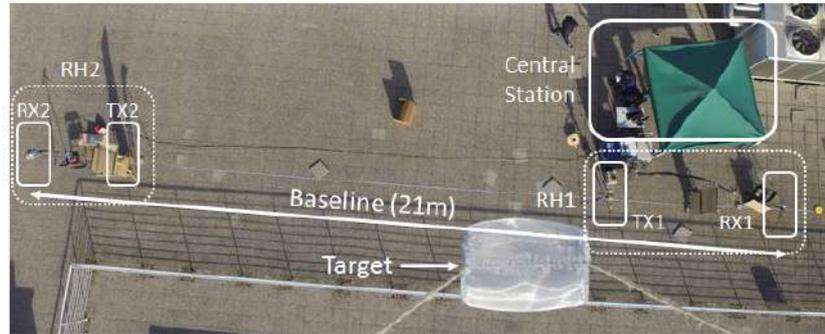
Figure 1. Aerial picture of the test field, taken by one of the drones carrying the targets. The target is the metal net cylinder hanging from the drone on a twine.

nodes, which directly transfer the raw received signals to the central node for data fusion (i.e. centralized processing). This solution, which requires high-bandwidth communication links, ideally maximizes the informative content that can be extracted from the raw data. The structural differences between multistatic and MIMO radar systems together with a comprehensive performance comparison can be found in [3].

Unfortunately, the pratical implementation of MIMO radars with widely separated antennas is hindered by two main problems. The first one is the necessity of precise time/phase synchronization among the carrier signals at the radar nodes. Since networked radars are typically operated with physically different local oscillators, each of them suffering from statistically independent phase offsets, proper solutions must be necessarily found [4]. The second problem is how to guarantee a reliable bi-directional large-bandwidth long-range signal distribution among the remote nodes and the fusion centre. Up to now, these two issues represent the main limitation factor to the development of such systems in real operative scenarios [5].

Recently, photonic technology has been investigated to implement the main functionalities of microwave systems. The goal is to overcome the actual limitations of electronic devices and subsystems and, possibly, to replace them. In this sense, photonics has demonstrated several properties that are fundamental for developing future radar systems [5–7]: low phase noise due to the excellent spectral purity of lasers, generation of ultra-wide bandwidth signals (i.e. up to tens of GHz with carriers in the mm-waves), easy tunability of lasers and optical devices (e.g. filters) enabling waveform diversity, low loss/distortion signal propagation via optical fiber links, immunity to electromagnetic interference and eavesdropping. The possibility to guarantee long-time phase stability and frequency/phase coherence between the transmitted and received signals [8], as well as the possibility to distribute the signals by means of optical fiber links, assures high-quality/coherence and the needed long-range large-bandwidth connectivity between the radar remote nodes [9]. These features make photonics very appealing for the development of future coherent radar networks with widely separated antennas [5].

In this work, the proposed radar network is based on photonics for the local oscillator generation and for the peripherals remoting, exploiting radio-over-fiber (RoF) techniques. The first coherent radar system based on photonics [10] and the results presented in [11] have fostered the development of the first photonics-based radar network demonstrators in outdoor environments [12, 13]. These experiments have been conducted within the SOLE project, with the final aim of designing and implementing the demonstrator of a fully-operational multiband centralized radar network based on photonic technology.

This paper focuses on presenting the main achievements from the radar network side and the proposed coherent data fusion strategy applied to real experimental data acquired in a controlled outdoor scenario on the rooftop of our lab, see Fig. 1. Finally, starting from the promising results presented in [12, 13], a modified version of the cell-averaging constant false alarm rate (CA-CFAR) detection algorithm is applied in which the decision threshold is evaluated from the bistatic range cells around the cell-under-test (CUT) [14]. Far from being exhaustive, this paper aims at considering and addressing some of the main architectural and signal processing issues that will rise from a real operative scenario, in which a multiband coherent MIMO radar based on photonic technology will be operated in a real maritime surveillance scenario within SOLE.
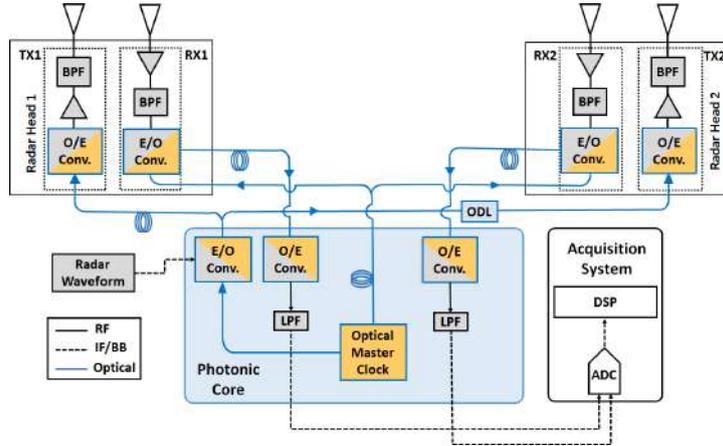
Figure 2. Architecture of the proposed Photonic Radar Network. DSP: Digital Signal Processing; ADC: Analog-to-Digital Converter; RF: Radio Frequency; LPF: Low-Pass Filter; E/O: Electro-Optical; ODL: Optical Delay Line; O/E: Opto-Electrical; IF/BB: Intermediate Frequency/Base Band; BPF: Band-Pass Filter.

2. PHOTONIC RADAR NETWORK ARCHITECTURE

For an initial validation of the proposed radar network architecture, a preliminary demonstrator has been developed. The demonstrator, depicted in Fig. 2, is based on a photonics-based radar transceiver (i.e. the photonic core) for signal generation/detection, a co-located acquisition system and two radar heads (RHs) with one transmitter (TX) and one receiver (RX) each. The RHs are remoted by means of radio-over-fiber (RoF) links.

Photonics-Based Radar Transceiver

The main building element of the photonic core is the optical master clock [10–13], which consists in a solid-state mode-locked laser (MLL), i.e. a pulsed laser whose optical spectrum is a combination of several phase-locked lines. These lines are spaced by the MLL pulse repetition frequency $f_{MLL} = 400$ MHz. The MLL allows to generate multiple radio-frequency (RF) carriers with very low phase noise, upon the opto-to-electronic (O/E) conversion of the MLL laser in a photodiode (PD) [8]. The same MLL is used for both signal up- and down-conversion, ensuring a perfect synchronization of the RHs and preserving the coherence of the signals.

In up-conversion, the radar signal, digitally generated at intermediate frequency (IF), is modulated on the MLL spectrum by an electro-to-optic (E/O) modulator. Then, it is photo-detected, generating replicas at $kf_{MLL} \pm f_{IF}$, with k positive integer and with $f_{IF} = 100$ MHz [9,10]. The desired RF output carrier frequency f_{RF} can be selected by means of a band-pass microwave filter (BPF), centered at f_{RF} and with bandwidth larger than $B = 100$ MHz. In down-conversion, the RF signal is similarly modulated on the MLL, then the O/E conversion at the PD generates again replicas of the RF signal down-converted at $kf_{MLL} \pm f_{IF}$, including $k = 0$ [9]. The large MLL optical spectrum ensures high efficiency, while the wide E/O and O/E bandwidth of the available modulators and PDs allows for the flexible management of RF signals up to several tens of GHz [8].

In [15], a coherent dual-band system operating at 2.5 GHz in (S-band) and 10 GHz (X-band), exploiting the same f_{MLL} and with 20 MHz signal bandwidth, was described and its performance evaluated. The transmitter generated signals with a time jitter of 15 fs (integrated from 100 Hz), exhibiting a spurious free dynamic range (SFDR) > 70 dB over 20 MHz, and a SNR of 73 dB/MHz. The receiver subsystem had an overall timing jitter of 10 fs in the same integration interval, enabling a SFDR of 50 dB and a noise figure (NF) of 8 dB. The minimum detectable signal was -122 and -124 dBm, in the S- and X-band respectively. In [10,11], the proposed photonic radar demonstrator was compared with an X-band commercial radar system, achieving comparable performance, with the additional feature of frequency flexibility to cope with sprectum crowding.

Radio-over-Fiber Signal Distribution

Since the RF signals are loaded on an optical carrier, the TXs/RXs at the RHs are connected with the photonic core through spans of single-mode fiber (SMF), depicted by the blue coils in Fig. 2. Thus, a RoF system is implemented [16]. Optical fibers grant small propagation losses (i.e. in the order of 0.2 dB/km), absence of electromagnetic interference, and preservation of signal coherence [9]. The optical delay line (ODL), implemented using a 1 km SMF spool inserted before the TX of RH2, generates the time diversity between RH1 and RH2.

Commercial RoF systems ensure an analog bandwidth up to 20 GHz with a gain of −25 dB (external modulation and no external RF amplification), SNR up to 146 dB/Hz, a SFDR up to 109 dBHz, and a NF of 38 dB. Otherwise, if external RF amplification is employed, a gain of 18 dB and a NF of 19 dB can be obtained, at the expenses of SNR and SFDR that slightly decrease to 140 dB/Hz and 101 dBHz, respectively. The linearity of the link can be improved by employing linearized or dual-output E/O modulation, and/or by keeping the power at suitable levels to avoid non-linear phenomena in the optical fiber. However, in RoF applications for radar systems, the intrinsic chromatic dispersion (CD) of optical fibers can represent an additional issue. This can be easily cancelled, e.g. by inserting in the link a span of CD-compensating fiber, or by employing single side-band modulation. A comprehensive analysis of the aforementioned issues and solutions can be found in [7] and references therein.

Radar Heads

Once the E/O converted radar waveform is delivered to a TX, this operates the O/E conversion, by which implies the band-pass filtering the signal at the desired RF carrier. The employed BPFs are multi-cavity filters are centered at f_{RF} = 9.7 GHz. Finally, after the amplification stage, the radar signal is transmitted by the antenna. The employed antennas at the TXs are ultra-wideband Vivaldi-shaped horn antennas with about 50° aperture and 12 dBi maximum gain. The RXs are equipped with very similar antennas. The detected radar echoes received by the RHs antennas are amplified, pass-bandband-pass filtered and E/O converted. Then, the received signals are transmitted back to the photonic core and O/E converted. After this operation, the signals from each RX are low-pass filtered and fed into a two-channel acquisition system, where they are digitized by an analog-to-digital converter (ADC) at 400 MS/s per channel.

3. SIGNAL PROCESSING SCHEME

The proposed radar network architecture allows to achieve excellent phase coherence, thus making the proposed radar network architecture ideal for coherent MIMO processing. First, the same optical clock (i.e. the MLL) is used in transmission for the up-conversion and in reception for the down-conversion of all the signals. Second, the employed optical oscillator is a kind of laser that is usually characterized by very low phase noise [8]. If the angular jitter of the overall architecture is lower than 10^{-1} rad, the negative effects on MIMO coherent detection due to sidelobe increase can be neglected [17]. Otherwise, phase jitter could negatively affect not only target detection performance, but also increase the sidelobe level, thus introducing additional false alarms. Finally, the temporal jitter of the considered system architecture, which corresponds to the integration of the oscillator phase noise for the offset frequencies in the interval [20 Hz, 200 MHz], is in the order of 10^{-12} s, while the angular jitter is in the order of 10^{-2} rad, lower than the limit reported in [17].

3.1 Radar Signal Model and MIMO Processing

In the general case of a radar network composed by M TXs and N RXs, the photonic core shall have N ADCs which digitize the received signals. These signals, down-converted at f_{IF}, are the linear combination of the M echo signals from each TXs. The subscript k, l indicates the generic virtual channel, determined by the kth TX and l^{th} RX, for $k = 1, \ldots, M$ and $l = 1, \ldots, N$. Each of the N digitized signals is split into the M individual bistatic channels for data processing. To maximize the SNR, pulse compression is performed on the complex received base-band (BB) signals, as detailed in [11–13]. Thus, the signal $r_{k,l}(t)$ can be written as:

$$r_{k,l}(t) = a_{k,l}(x,y) \cdot s_k\left(t - \tau_{k,l}(x,y)\right) e^{j[\theta(t - \tau_{k,l}(x,y)) - \theta(t)]} + n_{k,l}(t), \tag{1}$$

where $s_k(t)$ is the signal transmitted by the kth TX, $a_{k,l}(x,y)$ is an amplitude factor, while $\tau_{k,l}(x,y)$ is the time delay proportional to the bistatic distance. Both are functions of the target location (x,y) and the positions of

on the kth TX and l^{th} RX positions in the Cartesian plane. For simplicity, the term $n_{k,l}(t)$ is modelled as an additive white Gaussian noise (AWGN) stochastic process, while the term $\theta(t)$ takes into account the phase shift caused by the oscillator instability. Finally, the following log-likelihood function represents the "fused" coherent MIMO output:

$$ln\,[f\,(r\,(t)\,|\,(x,y))] = c' \cdot \left| \sum_{k=1}^{M} \sum_{l=1}^{N} e^{-j2\pi f_{IF}\tau_{k,l}(x,y)} \cdot \int r_{k,l}^{BB*}(t) \cdot s_k^{BB}\,(t - \tau_{k,l}\,(x,y))\,dt \right| + c''. \qquad (2)$$

According to eq. (2), for each possible target location with coordinates (x,y), the decision statistic is computed determining, for all the M TXs and N RXs, the correlation between the received and transmitted baseband equivalent signals, denoted by $r_{k,l}^{BB}(t)$ and $s_k^{BB}(t)$ respectively, while $*$ is the complex conjugation operator. In addition to this, the exponential terms $e^{-j2\pi f_{IF}\tau_{k,l}(x,y)}$ re-align the phases of the signals after they have travelled different TX-target-RX paths. After this re-phasing, all the $M \times N$ correlation contributes are summed together coherently. For details about c' and c'', and a complete mathematical description, please refer to [17].

3.2 CFAR Detection applied to MIMO Radar

Target detection based on the CFAR paradigm applied to MIMO radars has been to subject of many works, some of which are reported in [18]. In [19], the optimal detector in the Neyman-Pearson sense is derived in white Gaussian noise with known variance, together with the uniformly most powerful invariant (UMPI) detector having the CFAR property in the case of unknown noise variance. The problem of detection in non-Gaussian clutter has been considered in [20–22], while multi-pulse detection schemes were presented in [23]. However, to satisfy the CFAR property, adaptive threshold techniques are necessary. Among these, the most common one is the CA-CFAR, which estimates the mean noise power from a set of reference cells. The CA-CFAR strategy is the optimal detector only when the input noise to the square law detector is Gaussian and homogeneous, while it suffers performance degradation in non-homogeneous environments [24].

Here, following the methodology described in [18], we consider a 2D detection problem in which the statistic Y is the MIMO cross-ambiguity function described by eq. (2) in the Cartesian space. The detection threshold is usually proportional to the estimate of total noise power and it is obtained by processing the contents of K reference cells surrounding the CUT, except for N_G guard cells, which are not considered in the calculation. In this way, the statistic Y is compared with a threshold such that, if $Y \geq T \cdot Z$, we decide for the "target present" hypothesis (i.e. H_1); we decide for the "target absent" (i.e. H_0) otherwise.

The detection threshold is the product of two factors: T is the scale factor that depends on the desired constant false alarm probability, (i.e. P_{FA}), while Z is the noise power estimate. This procedure is repeated for all the (x,y) coordinates into which the Cartesian plane is discretized. As mentioned, instead of the classic rectangular window centered around the CUT, we consider the union of the isorange ellipsoids of each TX-RX radar node. This choice is motivated by the possible presence in the final MIMO cross-ambiguity function of sidelobes distributed along the isorange ellipsoids. Especially when M and N are small, these ellipsoids may lead to very high sidelobes and, thus, cause a significant increase of false alarms [14].

4. EXPERIMENTAL RESULTS

In our experiment, we further investigate the network presented in [14] and we test it in a real outdoor scenario, on the rooftop of our lab, with multiple closely-spaced targets, see Fig. 1. The network is composed of two TXs and two RXs, connected with the photonic core through optical fiber links. At the actual status of development, the photonic core generates a 100 MHz-bandwidth linear frequency modulated upsweep chirp signal with 100 ns duration, pulse repetition interval (PRI) of 50 μs and $B = 100$ MHz centered at $f_{RF} = 9.7$ GHz, corresponding to a range resolution of 1.5 m. The two transmitted waveforms are separated in the time domain with an ODL on the path between the photonic core and RH2. The network is deployed, according to the scheme in Fig. 3, with the four antennas aligned over a 21 m baseline. These are oriented upwards, to mitigate clutter and multipath returns due to the surrounding structures, buildings and vegetation, and to ensure the simultaneous illumination of both targets. The output power from the TX antennas is ≈ 100 mW.
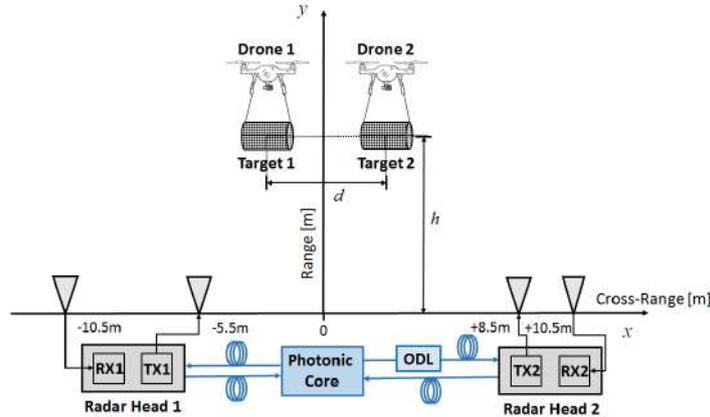
Figure 3. Geometry of the coherent photonics-based MIMO radar network and in-field experimental setup.

The targets consist of two cylinders, with 17 cm radius and 50 cm height, made of a tight-mesh metal net and hanging from two commercial mini-drones (Fig. 1). The height h of the two targets was controlled and kept always between 15 and 20 m, while the two drones were always kept hovering at the same height with respect to each other and separated by a cross-range distance $d = 3$ m to avoid possible collisions. Unfortunately, it was not possible to univocally determine the exact position of the drones.

Let us consider a 2D detection problem, in which the output of eq. (2) undergoes CA-CFAR detection in the search space. The reference cell parameters are $K = 20$ and $N_G = 20$, for both the rectangular window and the proposed methods, respectively in Fig. 4(a) and Fig. 4(b). As it will be discussed in the following, by using the proposed range-gating method, the effect of the constant bistatic-range ellipsoids intersecting at target pixels results strongly mitigated. Thus, we expect a reduction of the number of false alarms originated by the target.
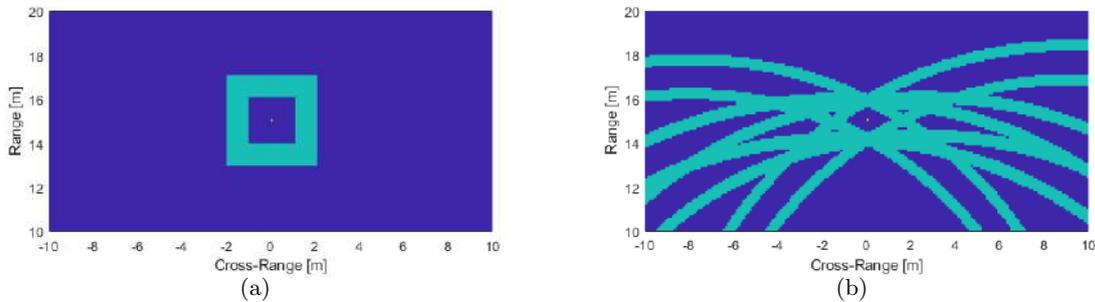
Figure 4. Reference cells (green) for the CUT (yellow) at $(0, 15)$ m in the Cartesian plane: (a) the classic rectangular window, (b) the proposed method based on the combination of bistatic ellipsoids. Detector parameters: $K = 20$, $N_G = 20$.

A 25 ms integration time (i.e. 500 PRIs) has been considered for the coherent MIMO processing output, which is evaluated in a search space of 151×301 samples in the interval $[5, 20]$ m and $[-15, +15]$ m (i.e. 10 cm spacing), along range and cross-range respectively. Results of the coherent MIMO processing and the CA-CFAR detection strategy are depicted in Fig. 5(a) and 5(b), respectively. As we can observe in Fig. 5(a), the two targets can be separated in cross-range only by processing coherently the data. The presence of high sidelobes in the monitored area is due not only to the presence of clutter and multipath, but also to the reduced MIMO configuration (i.e. few TXs and RXs). In this sense, a larger number of TXs and RXs would significantly reduce the intensity of the peaks and grant a finer cross-range resolution. The presence of a long spool of fiber (i.e. the ODL) simulates

a distance of about 1 km between the two RHs. However, the results demonstrate the low attenuation and negligible phase distortion introduced by the fiber. The employed network is characterized by widely distributed antennas, since the four channels are spatially decorrelated, according to the definition proposed in [2]. Thus, as we can observe, aspect diversity is a key parameter to cope with target RCS fluctuations; it represents the main difference between MIMO radars and phased arrays having the same number of elements [2].

Fig. 5(b) depicts the output of the modified CA-CFAR detection algorithm. With respect to the nominal aperture of the antenna (e.g. given the distance from the baseline of 18 m and the 50° antenna aperture, the expected monostatic cross-range resolution should be about 15.7 m), the cross-range resolution has been improved by a factor of about 5. Unfortunately, due to the low emitted power and the small drone RCS, it is almost impossible to detect the two drones without applying also Doppler processing. In conclusion, given the reduced radar array configuration and the non-idealities introduced by the environment and target RCS fluctuations, the results achieved in terms of cross-range resolution can be truly considered a remarkable achievement, aimed to design and implement future coherent MIMO radars with widely separated antennas based on photonic technology.
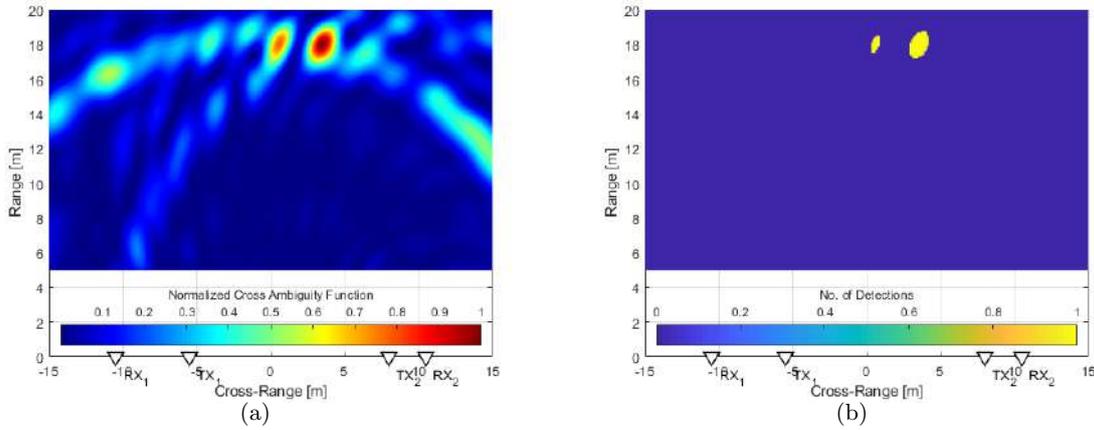
Figure 5. (a) Coherent MIMO output and (b) CA-CFAR detection output, based on the combination of bistatic ellipsoids. Detector parameters: $K = 20$, $N_G = 20$, $P_{FA} = 10^{-4}$.

5. CONCLUSIONS

In this paper, a coherent MIMO radar network demonstrator based on photonic technology has been presented and tested in a real outdoor environment within the NATO-funded SOLE Project. Two cooperative closely-spaced targets have been considered during the trials to prove the validity of the proposed architecture. They have been successfully resolved in cross-range by applying coherent MIMO processing and correctly detected by means of a modified version of the CA-CFAR detection strategy. With respect to the nominal aperture of the antenna, the cross-range resolution has been improved by a factor of five.

The contribution of photonics in preserving signal coherence among the TX and RX elements and in granting large-bandwidth long-range undistorted signal distribution over fiber is apparent. The robust features in terms of time/phase coherency allow to fully advantage from coherent MIMO processing in combination with the system geometry. These confirm that photonics is an enabling technology for coherent MIMO radars with widely distributed antennas. However, the results presented in this paper must be considered just as a first milestone, aimed to guide us through the development of a more complex radar network, with a significantly larger number of TX and RX antennas, and specifically tailored to real-time surveillance tracking and imaging applications.

ACKNOWLEDGMENTS

REFERENCES

[1] Chernyak, V. S., [*Fundamentals of Multisite Radar System*], Gordon & Breach Science Publ., London (1998).

[2] Haimovich, A. M., Blum, R. S., and Cimini, L. J., "MIMO radar with widely separated antennas," *IEEE Signal Processing Magazine* **25**(1), 116–129 (2008).

[3] Gorji, A. A., *Localization, tracking, and antenna allocation in multiple-input multiple-output (MIMO) radars*, PhD thesis, McMaster University, Hamilton, Ontario (2012).

[4] Yang, Y. and Blum, R. S., "Phase synchronization for coherent MIMO radar: Algorithms and their analysis," *IEEE Transactions on Signal Processing* **59**, 5538–5557 (Nov. 2011).

[5] Bogoni, A., Ghelfi, P., and Laghezza, F., [*Photonics for Radar Networks and Electronic Warfare Systems*], IET SciTech Publishing, London (2019).

[6] Lembo, L., Ghelfi, P., and Bogoni, A., "Analysis of a coherent distributed MIMO photonics-based radar network," in [*European Radar Conference*], 170–173 (Sep. 2018).

[7] Serafino, G., Scotti, F., Lembo, L., Hussain, B., Porzi, C., Malacarne, A., Maresca, S., Onori, D., Ghelfi, P., and Bogoni, A., "Toward a new generation of radar systems based on microwave photonic technologies," *Journal of Lightwave Technology* **37**, 643–650 (Jan 2019).

[8] Serafino, G., Ghelfi, P., Perez-Millan, P., Villanueva, G. E., Palaci, J., Cruz, J. L., and Bogoni, A., "Phase and amplitude stability of EHF-band radar carriers generated from an active mode-locked laser," *Journal of Lightwave Technology* **29**, 3551–3559 (Dec. 2011).

[9] Williams, P. A., Swann, W. C., and Newbury, N. R., "High-stability transfer of an optical frequency over long fiber-optic links," *J. Opt. Soc. Am. B* **25**, 1284–1293 (Aug. 2008).

[10] Ghelfi, P., Laghezza, F., Scotti, F., Serafino, G., Capria, A., Pinna, S., Onori, D., Porzi, C., Scaffardi, M., Vercesi, V., Lazzeri, E., Berizzi, F., and Bogoni, A., "A fully photonics-based coherent radar system," *Nature* **507**, 341–345 (Mar. 2014).

[11] Laghezza, F., Scotti, F., Serafino, G., Banchi, L., Malaspina, V., Ghelfi, P., and Bogoni, A., "Field evaluation of a photonics-based radar system in a maritime environment compared to a reference commercial sensor," *IET Radar, Sonar Navigation* **9**(8), 1040–1046 (2015).

[12] Lembo, L., Maresca, S., Serafino, G., Scotti, F., Amato, F., Ghelfi, P., and Bogoni, A., "In-field demonstration of a photonic coherent MIMO distributed radar network," in [*IEEE Radar Conf.*], (Apr. 2019).

[13] Maresca, S., Serafino, G., Scotti, F., Amato, F., Lembo, L., Bogoni, A., and Ghelfi, P., "Photonics for coherent MIMO radar: an experimental multi-target surveillance scenario," in [*Int. Radar Symp.*], (2019).

[14] Maresca, S., Bogoni, A., and Ghelfi, P., "CFAR detection applied to MIMO radar in a simulated maritime surveillance scenario," in [*European Radar Conference*], (Oct. 2019).

[15] Ghelfi, P., Laghezza, F., Scotti, F., Onori, D., and Bogoni, A., "Photonics for radars operating on multiple coherent bands," *Journal of Lightwave Technology* **34**, 500–507 (Jan 2016).

[16] Beas, J., Castanon, G., Aldaya, I., Aragon-Zavala, A., and Campuzano, G., "Millimeter-wave frequency radio over fiber systems: A survey," *IEEE Communications Surveys Tutorials* **15**(4), 1593–1619 (2013).

[17] Lehmann, N. H., Haimovich, A. M., Blum, R. S., and Cimini, L., "High resolution capabilities of MIMO radar," in [*2006 Fortieth Asilomar Conference on Signals, Systems and Computers*], 25–30 (Oct. 2006).

[18] Janatian, N., Modarres-Hashemi, M., and Sheikhi, A., "CFAR detectors for MIMO radars," *Circuits, Systems, and Signal Processing* **32**, 1389–1418 (Jun. 2013).

[19] Fishler, E., Haimovich, A., Blum, R. S., Cimini, L. J., Chizhik, D., and Valenzuela, R. A., "Spatial diversity in radars—models and detection performance," *IEEE Trans. on Sig. Proc.* **54**, 823–838 (March 2006).

[20] Cui, G., Kong, L., Yang, X., and Yang, J., "The rao and wald tests designed for distributed targets with polarization mimo radar in compound-gaussian clutter," *Circ., Sys., and Sig. Proc.* **31**, 237–254 (Feb 2012).

[21] Chong, C. Y., Pascal, F., Ovarlez, J., and Lesturgie, M., "Mimo radar detection in non-gaussian and heterogeneous clutter," *IEEE Journal of Selected Topics in Signal Processing* **4**, 115–126 (Feb 2010).

[22] Xun Chen and Blum, R., "Non-coherent mimo radar in a non-gaussian noise-plus-clutter environment," in [*2010 44th Annual Conference on Information Sciences and Systems (CISS)*], 1–6 (March 2010).

[23] Sheikhi, A. and Zamani, A., "Temporal coherent adaptive target detection for multi-input multi-output radars in clutter," *IET Radar, Sonar Navigation* **2**, 86–96 (April 2008).

[24] Gandhi, P. P. and Kassam, S. A., "Analysis of CFAR processors in nonhomogeneous background," *IEEE Transactions on Aerospace and Electronic Systems* **24**, 427–445 (Jul. 1988).

A Marine Radar Dataset for Multiple Extended Target Tracking

Jaya Shradha Fowdur[a,b], Marcus Baum[b], Frank Heymann[a]

[a]German Aerospace Centre (DLR), Neustrelitz, Germany;
[b]University of Göttingen, Göttingen, Germany

ABSTRACT

The marine radar remains one of the most extensively used sensor for maritime surveillance. Owing to improved technologies, it can nowadays be exploited to gain information about the extents of targets, since multiple measurements can be obtained from a single target. This paper introduces an open multitarget marine radar-based dataset subjected to a linear-time joint probabilistic data association (JPDA) filter for tracking extended targets using ellipsoidal approximations. A tailored version of the Multiplicative Error Model-Extended Kalman Filter* (MEM-EKF*) algorithm is used for estimating the orientations and kinematic properties of multiple targets recursively. Using the automatic identification system (AIS) information as ground truth, the positional errors are evaluated using the optimal sub pattern assignment (OSPA) metric and the performance of the algorithm for orientation estimation is rationally discussed with respect to the ground truth and dataset scenario. The dataset proposed is intended for comparing different algorithms for the purpose of multiple extended target tracking (METT).

Keywords: Multiple extended target tracking, ellipsoidal shapes, marine radar dataset, linear-time JPDA

1. INTRODUCTION

The past few decades have been witness to the establishment and steady growth of oceanic trade connections worldwide. To ensure safety of the goods and crew onboard, a reliable situational awareness is necessary to, for instance, provide an early detection of and avert potential collisions, detect abnormal activities such as smuggling and to maintain a constant monitoring of specific routes. The marine radar is still predominantly utilised and with improvement in sensor technologies and image processing techniques, it provides several noisy spatially distributed measurements that arise from the target(s) of interest at each observation step. The target tracking problem is, in this context, to estimate the extent (geometric shape) properties in addition to their kinematic ones simultaneously given the measurements' distribution.

In scenarios of high measurement noises, approaches of extended target tracking (ETT) have been developed where the target extent is estimated using basic geometric shapes for example, ellipses or rectangles. Most of the radar measurements acquired from generic real-world maritime environments are highly noisy, rendering it difficult to accurately model the true underlying physical extent and contour of vessels. Going by this perspective, an ellipsoidal approximation of targets aids in getting suitable insight and awareness on the traffic situation under an observation region. The random matrix approach models an extended target by its kinematic state represented by a Gaussian distribution and an extent matrix represented by a symmetric and positive definite matrix which provides an ellipsoidal approximation of the target's extent.[1,2] The approach has for instance been employed for ETT using converted X-band radar measurements that analysed extent estimation with respect to different measurement models.[3] The multiplicative error model (MEM)[4] approach models the spatial distribution of measurements from the target's surface with multiplicative noise. This is accounted for in the measurement equation by the use of a scaling factor. Targets can thus be tracked using ellipsoidal approximations,[5,6] where an ellipse is parametrised with the semi-axis lengths and orientation. In contrast to the random matrix method,

---

Jaya Shradha Fowdur and Frank Heymann are with the Institute of Communications and Navigation at the German Aerospace Centre (DLR). E-mail: {jaya.fowdur, frank.heymann}@dlr.de. Marcus Baum is with the Institute of Computer Science at the University of Göttingen. E-mail: marcus.baum@cs.uni-goettingen.de.

this parameterisation permits a custom modelling of the uncertainty of each individual parameter where the state of a target of known dimension can be tracked.[6]

METT is to simultaneously estimate the number of targets, kinematic state and extent properties of each target respectively. To solve the same, multiple target tracking (MTT) algorithms have been incorporated within the random matrix framework such as the probabilistic multi-hypothesis tracking filter[7] and variants of the probabilistic data association filter.[8,9] Vivone and Braca presented an ellipsoidal METT method based on the JPDA filter, with an explicit ellipse parameters estimation[10] where the measurement vector comprises features extracted from a specific clustering approach. In addition to the data association and multi-hypothesis approaches, the random finite set (RFS) one has been applied to solve METT. Some of its recent methods include the (generalised) labelled multi-Bernoulli (LMB)[11] and the Poisson multi-Bernoulli mixture (PMBM) filters,[12] along with their respective gamma Gaussian inverse Wishart (GGIW) variants.[13,14]

In this work we introduce the MANV dataset, which is the second marine radar dataset from the German Aerospace Centre's (DLR) benchmark scenarios to be made open to the tracking and signal processing community. Furthermore, we present initial tracking results using the linear-time extended target JPDA filter[9,15] together with the tailored MEM-EKF*.[6] AIS information has been used to initialise both the kinematic and extent related information. Based on the assumptions made with regard to the radar measurements and dataset scenario, the results are presented and discussed.

The rest of the paper is organised as follows. Section 2 gives an outline of the dataset with appropriate plots. In Section 3, the target state parametrisation, the measurement model, and a summary of the approach adopted for METT are covered. Section 4 carries the tracking and evaluation results which is then followed by a conclusion.

2. MANV DATASET

The MANV (**man**oeuvres) dataset, depicted in Figure 1 was procured from a measurement campaign held in the Baltic sea. The dataset is one of the proposed radar-based real-world benchmark scenarios[16] that are going to be made available to the tracking community to be used for METT, but not limited to it, leaving options open for ETT and other related applications*. MANV has been regarded in particular as it contains dynamic manoeuvres that pose a challenge to tracking algorithms.

For the observation period considered, that is, from step $k = 1$ up to step $k = 1000$, the scenario consists of six known targets in total, four of which are anchored – Target 1 is the own vessel/observer – while Targets 4, 5, and 6 were buoys fitted with radar reflectors during the campaign and the other two (Targets 2 and 3 of dimension 29.04m×6.7m and 23m×6m respectively) are engaged in manoeuvres. Target 2 had a moderate speed over ground (SOG) on average, starting at about 4kn, going up to 6kn before decelerating after $k = 800$. Target 3 started at about 10kn, fluctuating down to 2kn throughout most of the period till $k = 800$, then went around 6kn towards the end.

In the radar recording of Target 1 shown in Figure 3, an unknown random vessel underway has been observed which could have been either a leisure craft or fishing vessel that was not equipped with any AIS transponder. The radar measurements arising from each target in the dataset are in fact quite dense and noisy at almost all steps. During the image processing step adopted,[6] a good proportion of clutter had already been filtered out and as resolution, a pixel unit represented 6m. The original radar measurements were extracted from radar images, individually expressed in range r^j and bearing ψ^j with respect to the own vessel whose position was retrieved from the onboard AIS. The extracted measurements occur either in clusters or may occur individually, their origins being unknown. They have been then converted to the Cartesian East North Up (ENU) coordinates (up is assumed negligible) prior to being fed to the filter in Section 3 following the standard conversion scheme.[17] A measurement vector $\mathbf{z}^j$ hence follows as

$$\left[\begin{array}{c} z_{\tilde{e}}^j \\ z_{\tilde{n}}^j \end{array} \right] = \left[\begin{array}{c} r^j \sin(\psi^j) \\ r^j \cos(\psi^j) \end{array} \right]. \tag{1}$$

The error between an extracted polar measurement and its true one is assumed to follow a zero-mean independent Gaussian measurement noise having standard deviations σ_r and σ_ψ are 40m of 4°.

---

*The dataset can be requested for by contacting the authors.
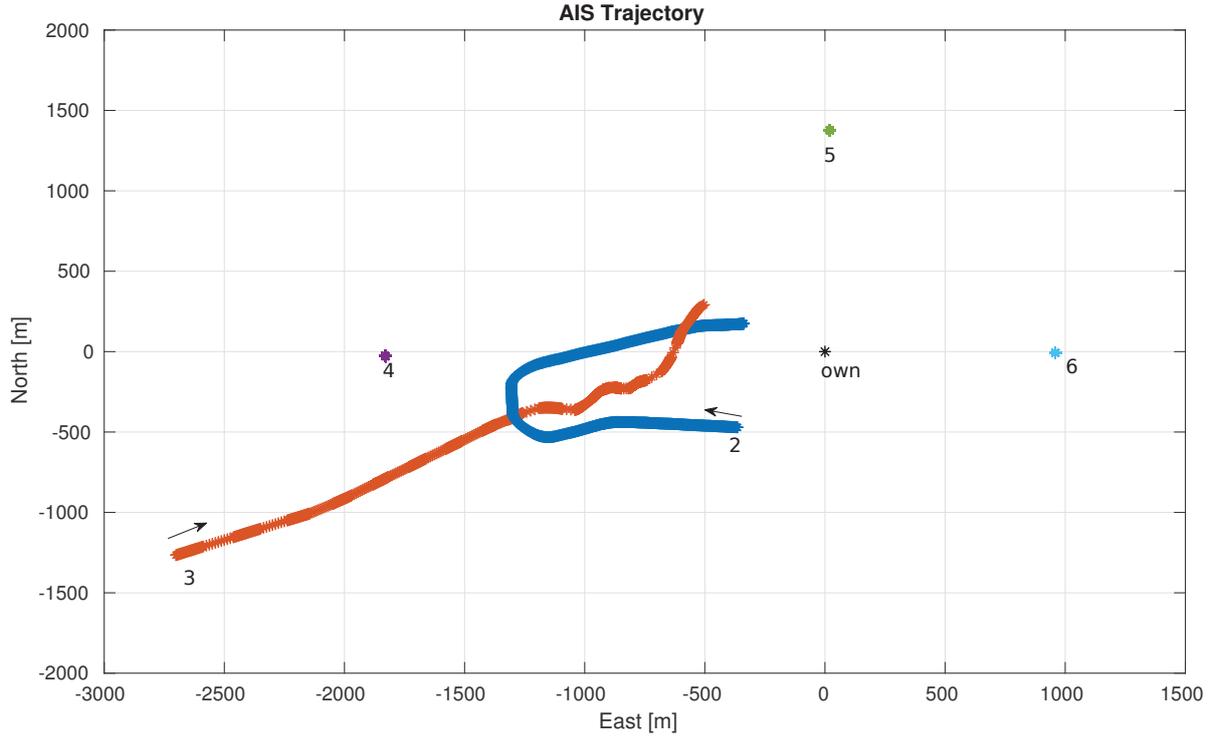
Figure 1. The plot corresponds to the AIS measurements obtained for the targets in the dataset. The targets are labelled, including the own vessel and the arrows indicate their respective courses at the beginning of their trajectories.

3. MULTIPLE EXTENDED TARGET TRACKING

At each observation step k, a set of measurements $\mathbf{Z}_k = \{\mathbf{z}_k^j\}_{j=1}^{M_k}$ becomes available. There is also a set of target states, defined $X_k = \{\mathbf{x}_k^i\}_{i=1}^{N_k}$. Formally, the METT problem is to jointly estimate the states – each individual target state comprises kinematic properties such as the position, velocity, or acceleration and the extent parameters – from the measurements.

3.1 State Model

The tailored MEM-EKF* algorithm[6] approximates the extent of a target of interest i as an ellipse with a multiplicative noise to link a measurement to its state, $\mathbf{x}_k^i$, such that $\mathbf{x}_k^i = \left[\mathbf{r}_k^i{}^T, \alpha_k^i\right]^T$, which consists of both the kinematic state vector $\mathbf{r}_k^i$ and the orientation parameter α_k^i. $[.]^T$ denotes the transpose operator. α_k^i is the property of interest from the shape parameter vector $\mathbf{p}_k^i$, visualised in Figure 2. For our adaptation, the kinematic state vector

$$\mathbf{r}_k^i = [t_{k,\tilde{e}}, t_{k,\tilde{n}}, \psi_k, \upsilon_k] \tag{2}$$

comprises the position $(t_{k,\tilde{e}}, t_{k,\tilde{n}})$ in the local Cartesian ENU coordinates, the course over ground (COG) ψ_k and the SOG υ_k. The shape parameter vector is given by

$$\mathbf{p}_k^i = [\alpha_k, l_1, l_2], \tag{3}$$

where α_k is the orientation of the ellipse taken along the ellipse's major axis, and l_1 and l_2 represent the lengths of the semi-axes of the ellipse respectively.
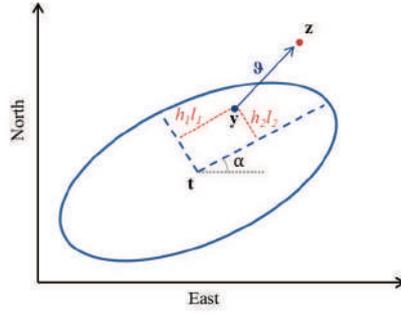
Figure 2. The adapted measurement model[5,6] where the extent is parameterised as $\mathbf{p} = [\alpha, l_1, l_2]$, showing the association of measurement $\mathbf{z}$ to source $\mathbf{y}$ with noise $\boldsymbol{\vartheta}$. $\mathbf{z}$ is related to $\mathbf{p}$ with multiplicative noise $\mathbf{h} = [h_1, h_2]^T$. The centre $\mathbf{t} = [t_{\tilde{e}}, t_{\tilde{n}}]^T$ denotes the target's position in the ENU coordinates.

3.2 Measurement Model

Each measurement $\mathbf{z}_k^j = [z_{\tilde{e},k}, z_{\tilde{n},k}]^T$ originates from a source $\mathbf{y}_k^j$ on a target's extent corrupted by an additive noise $\boldsymbol{\vartheta}_k$. The measurement model is illustrated in Figure 2. The measurement equation is given by[5]

$$\mathbf{z}_k^j = \mathbf{t}_k + \boldsymbol{\mathcal{R}}(\alpha_k) \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} h_{k,1}^j \\ h_{k,2}^j \end{bmatrix} + \boldsymbol{\vartheta}_k \tag{4}$$

where the measurement noise is $\boldsymbol{\vartheta}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k^j)$, the rotation matrix is $\boldsymbol{\mathcal{R}}(\boldsymbol{\alpha_k})$, and $\mathbf{h}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{C^h})$ is a multiplicative error term, where $\mathbf{C^h} = \frac{1}{4}\mathbf{I}_2$ that corresponds to the covariance of the uniform distribution an ellipse.

3.3 Dynamic Motion Model

The temporal evolution of the kinematic state of each target is based on a motion model of the following form

$$\begin{aligned} \mathbf{r}_k &= \boldsymbol{a}_k(\mathbf{r}_{k-1}, \boldsymbol{\omega}_k^{\mathbf{r}}) \\ &= \begin{bmatrix} t_{k-1,\tilde{e}} + \sin(\psi_k) \cdot \Delta T \cdot \upsilon_k \\ t_{k-1,\tilde{n}} + \cos(\psi_k) \cdot \Delta T \cdot \upsilon_k \\ \psi_k \\ \upsilon_k \end{bmatrix} + \boldsymbol{\omega}_k^{\mathbf{r}} \end{aligned} \tag{5}$$

where $\boldsymbol{\omega}_k^{\mathbf{r}} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{\mathbf{r}}^{\boldsymbol{\omega}})$ and $\Delta T = T_k - T_{k-1}$ is the time interval, assigned 1s for the dataset. (The target index's superscript i is left out.) The evolution of the orientation parameter follows a linear model represented by

$$\alpha_k = \alpha_{k-1} + \omega_k^\alpha \tag{6}$$

with process noise $\omega_k^\alpha \sim \mathcal{N}(0, Q_\alpha^\omega)$ driving the orientation.

3.4 Data Association and State Estimation

The main objective of data association is to deal with the unknown correspondence of a measurement to its respective origin, which can be either a target or clutter. Gilholm and Salmond have introduced the concept of extended targets having the possibility to be each represented by a spatial probability distribution modelled as a Poisson process.[18,19] With a Poisson model representing the underlying measurements, multiple measurements can result from a single target.[15] The same concept has been applied in this work to track the orientations and the kinematic properties of multiple targets of known ellipsoidal extents present within an observation region. The multitarget tracker employed is a linear-time JPDA filter[9] based on the aforementioned measurement model and the Poisson model. For the state estimation, the individual target states are updated following a sequential probabilistic data association (PDA) scheme described in Section III of Yang et al.'s work[9] by taking into account the marginal association probabilities.
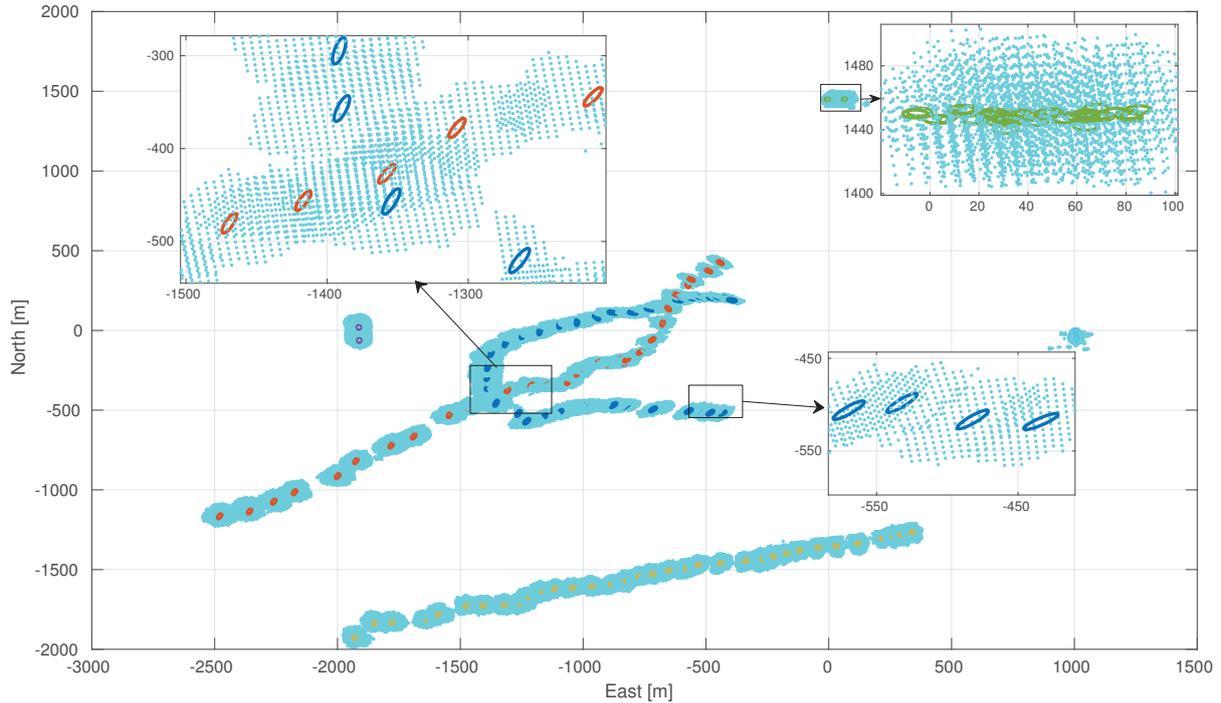
Figure 3. The radar measurements (dots) and the extent estimates (ellipses) of the targets. The estimates have been plotted at intervals of every 25 observations steps. Notice the unknown target starting from ENU coordinates $[-2000, -2000]^T$.

Table 1. Filter Parameter Settings

Parameter	Target 2	Target 3	Unknown Target	Targets 4-6
Initial orientation parameter, α_0	277°	62°	45°	50°
Initial orientation process noise variance Q_α^ω	9	25	25	16
Measurement Rate, λ	232	270	277	152, 84, 26

4. RESULTS

The linear-time JPDAF for extended targets described earlier was applied to the dataset, and the results are as shown in Figure 3. The number of targets in the scenario is assumed to be constant and known apriori for the scenario considered. We point out that this work was realised by virtue of two main assumptions. Firstly, a target of interest is not necessarily aligned along its course, which can be explained by external factors influencing vessel motion such as tide and weather conditions. Secondly, due to the limited perspective of the targets' bodies visible in the own radar's field of view, it is assumed that the targets' extents are symmetric and can be approximated using ellipses. The dynamic model adopted for every target is the one defined in (5) for simplicity.

While the track of each target was initialised based on AIS information where available – more specifically the position, COG, SOG, and dimensions, the initialisation of the unknown was based on guesses. Further target-specific parameter settings as applied are summarised in Table 1 with the predetermined rate parameter for each target given their individual measurements.

The clutter density ρ over the observation region was calculated as 8×10^{-9} and the parametric model has been used to predetermine the probability mass function of the clutter measurements. The clutter rate was taken as 5. The process noise standard deviations for the kinematic states, $\mathbf{Q}_\mathbf{r}^\omega$ of all targets was $\mathrm{diag}(30, 30, 0.02, 5)$. For the validation gate, the threshold yielding to a gating probability of 0.65 was used to validate the measurements.
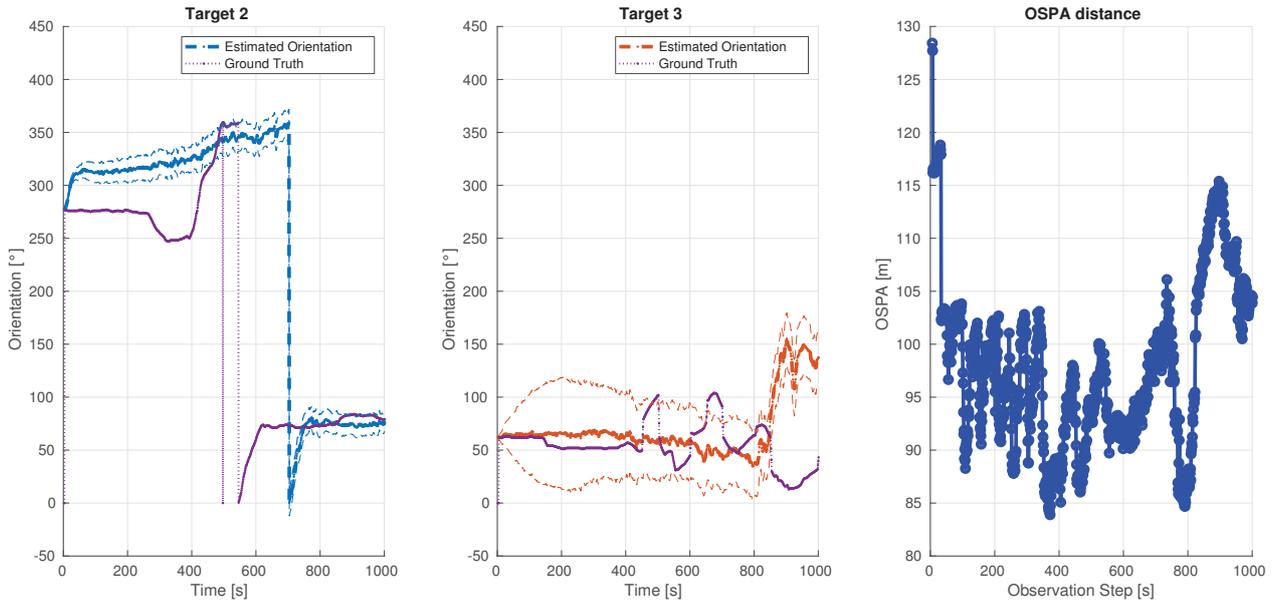
Figure 4. The estimated orientations with their corresponding standard deviations, against the AIS true heading measurements for Target 2 (left) and Target 3 (centre). (Right) The OSPA results from the kinematic state estimates evaluated against the AIS positions.

Figure 3 illustrates the extent estimates of the targets at every 25 steps along with the radar measurements displaying the rather high spreading of the measurements for farther distances with respect to the own. The results have been evaluated using the OSPA metric with AIS as ground truth as shown in Figure 4 with order setting of $p = 2$ and cut-off of $c = 150$m. The estimated orientations of the dynamic targets are also presented together with their ground truths (unavailable for the unknown one).

In Figure 4, the estimated orientations for Target 2 appear to be sensitive to the measurement spread emanating from the reflective side of the vessel due to the own's perspective. Those at the manoeuvring turns were not always well estimated, although the results do improve as the vessel gets closer to the own. Target 3's orientation estimates were rather smooth from up to $k = 450$, as observed. Later, it appears to have missed the sharper turns due to the own's perspective between $k = 450$ to $k = 750$. After $k = 800$, the target has been subjected to partial obstruction due to Target 2 which explains the estimates being off. Based on the aforementioned assumptions, a point to bring out regarding the own sensor's perspective aspect is that in highly dynamic and complex multitarget scenarios, using a single sensor will be insufficient to yield precise estimations.

5. CONCLUSION

In this paper, a JPDA filter based on the Poisson model has been employed to track the orientations of multiple targets based on a special version of the MEM-EKF* algorithm on a marine radar-based dataset. The OSPA metric was used to calculate the positional error, the vessels' estimated orientations were visualised together with the AIS-based true heading measurements of the vessels. The performance of the algorithm was fettered by the perspective of the own vessel, as expected, and seemed to be somewhat sensitive to the direction of the measurement spreads being tracked. The results would most likely be significantly improved should measurements from multiple sensors be used. In the future, we shall investigate further multiple extended target tracking algorithms using the dataset for a standard comparison.

REFERENCES

[1] Feldmann, M., Fränken, D., and Koch, W., "Tracking of Extended Objects and Group Targets Using Random Matrices," *IEEE Transactions on Signal Processing* **59**, 1409–1420 (April 2011).

[2] Lan, J. and Li, X. R., "Tracking of Extended Object or Target Group Using Random Matrix: New Model and Approach," *IEEE Transactions on Aerospace and Electrical Systems* **52**, 2973–2988 (December 2016).

[3] Vivone, G., Braca, P., Granström, K., Natale, A., and Chanussot, J., "Converted Measurements Random Matrix Approach to Extended Target Tracking using X-band Marine Radar Data," in [*2015 18th International Conference on Information Fusion (Fusion)*], 976–983 (July 2015).

[4] Baum, M., Faion, F., and Hanebeck, U. D., "Modeling the Target Extent with Multiplicative Noise," in [*2012 15th International Conference on Information Fusion*], 2406–2412 (July 2012).

[5] Yang, S. and Baum, M., "Tracking the Orientation and Axes Lengths of an Elliptical Extended Object," *IEEE Transactions on Signal Processing* **67**, 4720–4729 (September 2019).

[6] Fowdur, J. S., Baum, M., and Heymann, F., "Tracking Targets with Known Spatial Extent Using Experimental Marine Radar Data," in [*22nd International Conference on Information Fusion (FUSION 2019)*], (July 2019).

[7] Monika Wieneke, W. K., "Probabilistic Tracking of Multiple Extended Targets using Random Matrices," *Proceedings of SPIE - The International Society for Optical Engineering* **7698** (April 2010).

[8] Schuster, M., Reuter, J., and Wanielik, G., "Probabilistic Data Association for Tracking Extended Targets under Clutter using Random Matrices," in [*2015 18th International Conference on Information Fusion (Fusion)*], 961–968 (July 2015).

[9] Yang, S., Thormann, K., and Baum, M., "Linear-Time Joint Probabilistic Data Association for Multiple Extended Object Tracking," in [*2018 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2018)*], (July 2018).

[10] Vivone, G. and Braca, P., "Joint Probabilistic Data Association Tracker for Extended Target Tracking Applied to X-Band Marine Radar Data," *IEEE Journal of Oceanic Engineering* **41**, 1007–1019 (October 2016).

[11] Vo, B., Vo, B., and Phung, D., "Labeled Random Finite Sets and the Bayes Multi-Target Tracking Filter," *IEEE Transactions on Signal Processing* **62**, 6554–6567 (December 2014).

[12] Granström, K., Fatemi, M., and Svensson, L., "Poisson Multi-Bernoulli Conjugate Prior for Multiple Extended Object Estimation," *ArXiv* **abs/1605.06311** (2016).

[13] Beard, M., Reuter, S., Granström, K., Vo, B., Vo, B., and Scheel, A., "Multiple Extended Target Tracking With Labeled Random Finite Sets," *IEEE Transactions on Signal Processing* **64**, 1638–1653 (April 2016).

[14] Granström, K., Fatemi, M., and Svensson, L., "Gamma Gaussian Inverse-Wishart Poisson Multi-Bernoulli Filter for Extended Target Tracking," in [*2016 19th International Conference on Information Fusion (FUSION)*], 893–900 (July 2016).

[15] Baum, M., "Linear-time JPDAF based on Many-2-Many Approximation of Marginal Association Probabilities," *Electronics Letters* **51**, 1526–1528(2) (September 2015).

[16] Fowdur, J. S., Baum, M., and Heymann, F., "Benchmark for Radar-based Multiple Target Tracking in Maritime Environments." In Prep.

[17] Lerro, D. and Bar-Shalom, Y., "Tracking with Debiased Consistent Converted Measurements versus EKF," *IEEE Transactions on Aerospace and Electronic Systems* **29**, 1015–1022 (Julys 1993).

[18] Gilholm, K., Godsill, S., Maskell, S., and Salmond, D., "Poisson Models for Extended Target and Group Tracking," *Proceedings of SPIE - The International Society for Optical Engineering* **5913** (August 2005).

[19] Gilholm, K. and Salmond, D., "Spatial Distribution Model for Tracking Extended Objects," *IEEE Proceedings - Radar, Sonar and Navigation* **152**, 364–371 (October 2005).

Ship length estimation using common radar field entries

Eftychios Protopapadakis*[a], Matthaios Bimpas[a], Dimitrios Dres[a]

[a] *Telesto Technologies, Cholargos 15561, Greece*

ABSTRACT

In this paper, we focus on identifying the length of a ship given the track history. The problem entails to a traditional classification problem, on which multiple Bayesian optimized classifiers are tested. The input values, i.e. speed, course, and position, are provided by a data fusion algorithm from multiple sources, including over-the-horizon radar and AIS. Results indicate that using track history, spanning a five-minute period or less, can provide adequate information to estimate correctly the length of a ship.

Keywords: Classification, ship length, hyperparameter optimization

1. INTRODUCTION

Maritime Situational Awareness (MSA) is a challenging research field. Marine routes represent a huge portion of commercial and human trades; therefore, surveillance, security and environmental protection themes are gaining increasing importance. There are various criteria that can provide significant information regarding the behavior of a ship: these criteria can be related to position, speed, course, existence of AIS or other type of attributes.

Existing literature employs various computer vision algorithms for the ship detection. These approaches are RGB based [1], multispectral based [2], or SAR based [3]. Many of these approaches utilize more than one sources of data as inputs. There are, also, radar based approaches: Over-the-horizon high frequency doppler radars [4], compact high frequency radar [5] or other types, e.g. MIMO radars [6]. The former case, i.e. vison based, allows for more details regarding the ship's characteristics. The latter case, i.e. radar based, provides less detailed information on a larger scale.

This case study investigates the refinement possibilities of common radar field entries, for the extraction of specific ship's traits. Thus, in this paper we emphasize on ship's length estimation. Such information can be beneficial in many ways, and can be easily incorporated to advanced systems for ship behavior analysis [7].

2. PROPOSED METHODOLOGY

In this paper, we present a simple and efficient method to estimate the category of a ship's size, given the track history and common fields provided by any type of radar. The proposed methodology exploits fussed track data; i.e. data entries originating from multiple sensors [8]. These data entries are stored locally. Once a specific number of track entries is gathered, a low-level feature extraction occurs and then the feature values are fed to a classifier.

Let as denote as $x_i^{(t)} = \{ln_i, lt_i, c_i, s_i\}^{(t)}$, the common filed values, i.e. longitude, latitude, course and speed, for ship i at a time t. Then, the input to the classifiers has the form of $x_i = [x_i^{(t)}, x_i^{(t-1)}, ..., x_i^{(t-l)}]$, $l = 1, ..., \ell$, where ℓ denotes the number of past moments to be considered. All values are normalized to [0,1]. Let us also denote as $y_i \in \mathbb{Z}^{k \times 1}$, a vector indicating the corresponding class C_k, i.e. the range category for the length. We would like to learn a function $f: \mathcal{X} \to \mathcal{Y}$ so that f is expected to be a good predictor on future data.

2.1 Employed classifiers

We have scrutinized the effectiveness of a series of well-known classifiers in ship's length recognition, using track history data entries. The techniques considered for constructing linear and non-linear models were:

Naïve Bayes classifier, a family of simple probabilistic classifiers based on applying Bayes' theorem with strong independence assumptions between the features [9]. Abstractly, naive Bayes is a conditional probability model: given a

problem instance to be classified, represented i.e. vector x representing some $\ell \times 4$ features (independent variables), it assigns to this instance probabilities $p(C_k|x_1, ..., x_n)$ for each of K possible outcomes or classes C_k.

Linear discriminant analysis, a statistical analysis method useful in determining whether a set of variables is effective in predicting category membership. Discriminant analysis (Discr) classifiers assume that different classes generate data based on different Gaussian distributions so that $p(x|y = C_k) \sim N(\mu_k, S), k = 1, ..., K$. In order to train such a classifier, we need to estimate the parameters of a Gaussian distribution for each class. Then, to predict the classes of new data, the trained classifier finds the class with the smallest misclassification cost: $\hat{y} = arg\max_k \left\{ \left(x - \frac{\mu_k}{2} \right)^T \beta_k + \log \pi_k \right\}$, where $\beta_k = S^{-1}\mu_k$.

k-nearest neighbors, a non-parametric method used for classification [10]. A majority vote of its neighbors classifies an object, with the object being assigned to the class most common among its k nearest neighbors; it is, therefore, a type of instance-based learning, where the function is only approximated locally, and all computation is deferred until classification.

Classification trees use a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. In classification tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels [11]. Each internal (non-leaf) node is labeled with an input feature. The arcs coming from a node labeled with a feature are labeled with each of the possible values of the feature. Each leaf of the tree is labeled with a class or a probability distribution over the classes.

Support vector machines, a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap (margin) that is as wide as possible [12]. New examples are then mapped into that same space and predicted to belong to a category based on which side of the margin they fall on. The mappings used by SVM schemes are defined through a kernel function $K(x, y)$ selected to suit the problem.

The above methods all lead to standalone models. In addition, ensembles are also used, which are based on combinations of multiple models resulting from a single learning algorithm. For the purposes of this analysis, we consider three ensemble algorithms, namely boosting and random forests using classification trees and discriminant analysis as the base learning algorithms.

2.2 Parameterization of learning algorithms

Typically, machine learning models have a set of parameters that need to be defined before the learning process (i.e. training using available data) begins. All these parameters are known as hyperparameters. Some indicative examples are the reqularization parameter c in SVMs, the number of k nearest neighbors and the number of children (max depth) in decision trees. In this work, the hyperparameter optimization process, for each of the classifiers, was based on grid-search optimization algorithm, which was limited to three hours [13]. Grid search approach is an exhaustive search through a specified subset of hyperparameters. Given the relatively small amount of training data (see sec. 3.1) and the complexity of the models, grid search was a viable solution for parameters fine tuning.

3. EXPERIMENTAL EVALUATION

The length of a ship is estimated using five categories: i) <10 meters, ii) 10-15 meters, iii) 15-20 meters, iv) 20-25 meters and v) >25 meters. The number of feature values, used as inputs to the classifiers, varied, depending on how many past track records we use. The following scenarios applied in our case:

1) 2 past moments: 1.2 minutes duration, 8 values per ship

2) 4 past moments: 2.6 minutes duration, 16 values per ship

3) 6 past moments: 4 minutes duration, 24 values per ship

4) 8 past moments: 5.3 minutes duration, 32 values per ship

The data transmission times had an approximately 40 seconds interval rate.

3.1 Dataset description

The utilized dataset pertains to approximately 7 hours of data captured from the Mediterranean coast of southern France. Only ships that have AIS were considered, since we need the actual length to use it as ground truth for performance evaluation. Despite the large number of ships in the vicinity, only a small subset provided the actual length values. Also, the number of ships with length less that 20 meters was limited, resulting in unbalanced dataset. The final training sets (after the data balancing approach) are shown in Table 1.

Table 1. Train set length distribution

Class / Past Moments	2	4	6	8
<10 meters	14	12	12	7
10-15 meters	42	25	19	14
15-20 meters	43	40	33	23
20-25 meters	49	41	35	22
>25	45	40	40	25

3.2 Performance scores

The results are analyzed through standard measures of predictive performance for binary classification tasks. First, the confusion matrices (i.e. 5×5 matrices) are formed for each of the suggested combinations of the classifier algorithm with the past moments to use. If we apply the one against all paradigm, we can divide the original 5×5 to five 2×2 matrices. The elements of the confusion (2×2) matrices in that case involve: (1) the number of ships, with length different that the investigated one, correctly classified as ships with different length (true negatives, TN), (2) the number of ships in the investigated length category misclassified as different length category (false negatives, FN), (3) the number of ships outside investigated length ranges, misclassified as within investigated range (false positives, FP), and (4) the number of ships within the investigated range correctly classified (true positives, TP). Using these elements, four well-known classification performance indices are calculated:

1. Accuracy (ACC) represents the percentage of correct classification for all classes. It is defined as:

$$ACC = (TP + TN) / (P + N)$$

2. Recall (Re) indicates the fraction of the specific length cases identified by a model. It is defined as:

$$Re = TP / P$$

3. Precision (Pr) indicates the correct ship's length predictions. It is defined as:

$$Pr = TP / (TP + FP),.$$

4. F1-score (F1) is the harmonic mean of precision and recall. It is defined as:

$$F_1 = 2 \cdot \frac{Pr \cdot Re}{Pr + Re}$$

Figure 1 illustrates the confusion matrices generated by different type of classifiers. An ideal scenario would be a diagonal matrix. In that case, all ships would be classified correctly. Depicted cases achieve overall accuracy score of, approximately, 90%. Yet, a closer look indicates that the capability of identifying small length ships is less than 50%.
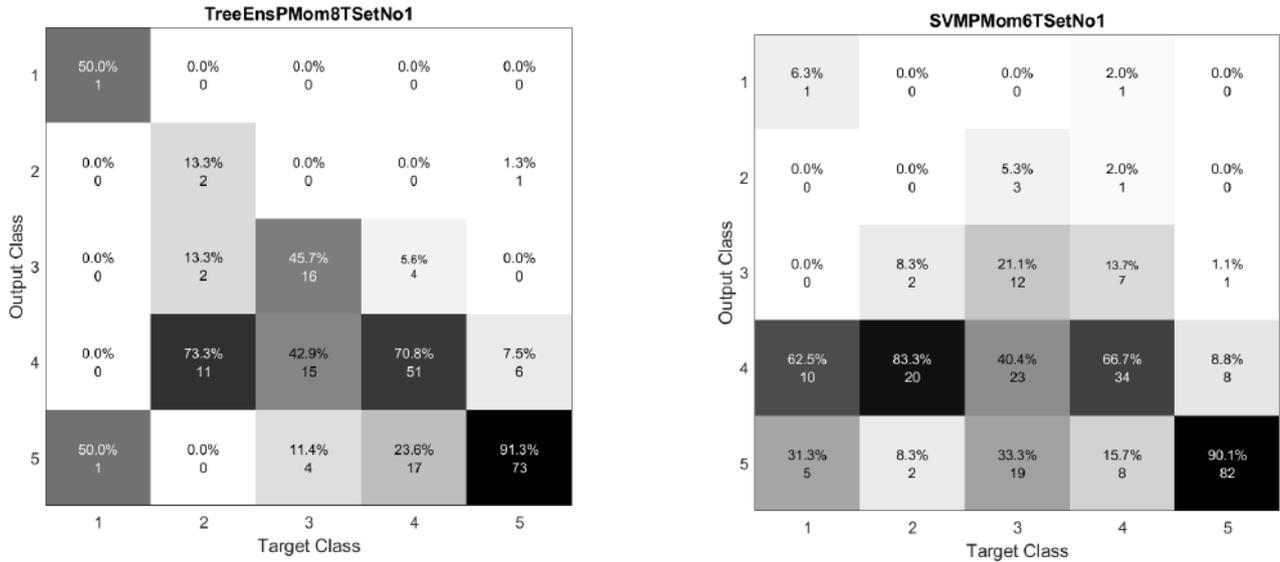
Figure 1: (Left) Tree ensemble classifier confusion matrix, over test set number 2, using as inputs 8 past moments. (Right) SVM classifier confusion matrix, over test set number 1, using as inputs 6 past moments.
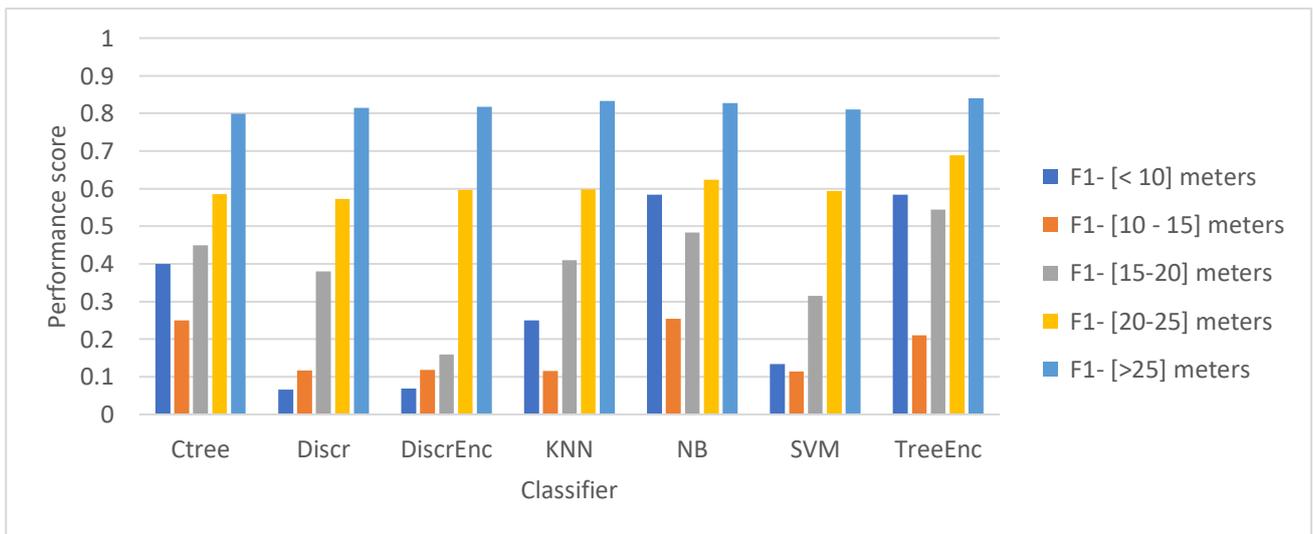
Figure 2. Average F1 scores, for each of the investigated categories.

Figure 2 provides details on the classifier's suitability for the length estimation. All classifiers achieve a good performance for the over 25 meters class. As the length of the ship declines, so does the performance score. Less than 10-meter ships are a peculiar case, for which the ensemble of classification trees appear to be the best solution.
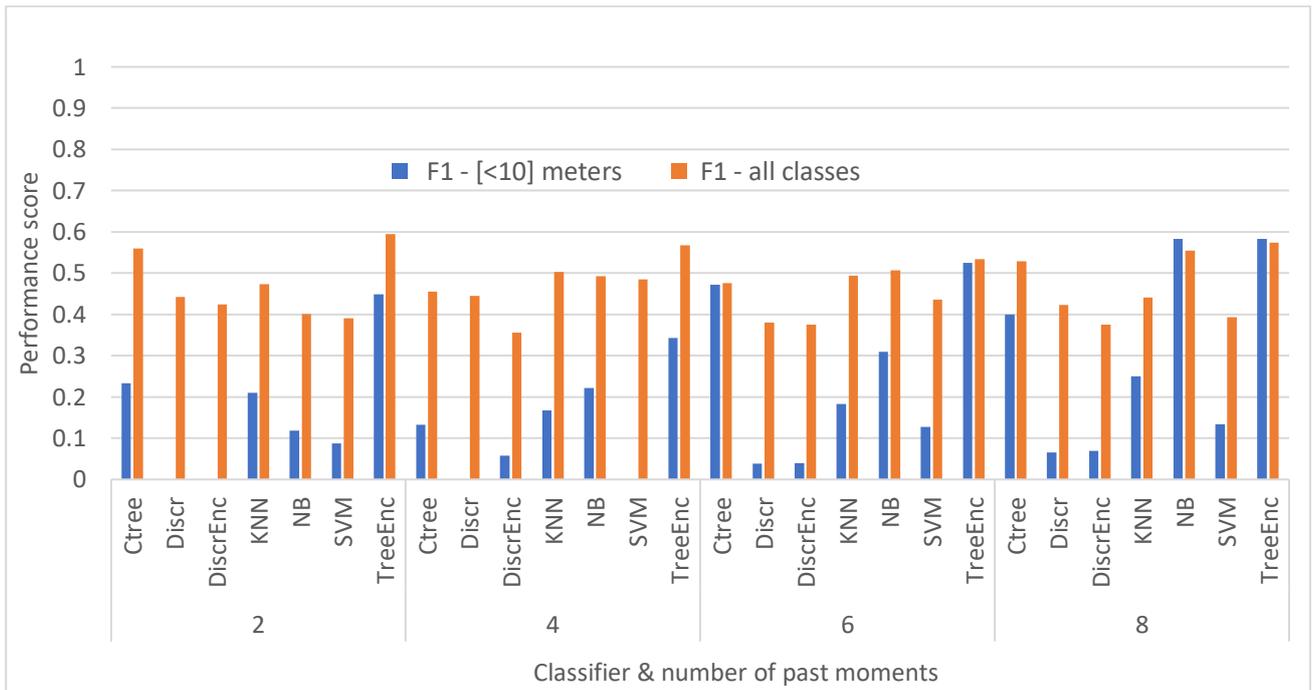
Figure 3. Analytical assessment for the detection capabilities when we have less than 10 meters size of ships.

Figure 3 provides a further insight on the less than 10 meters case. There are two points worth mentioning. At first, increasing the number of past moments (i.e. more input values) can lead to an increase in detection performance. Secondly, there is a trade-off between the detection in less than 10 meter class and the rest. An overall suggestion would be the utilization of a tree ensemble classifier for the length estimation problem.

4. CONCLUSIONS

The applicability of traditional classification schemes for the identification of a ship's length has been investigated. Results indicate that the track history of a vehicle, in terms of location, speed and course can provide adequate information for the estimation of its length. Despite the relatively small length sizes, i.e. four out of five categories were below 25 meters, an ensemble of tree classifiers using five-minute track history, provide an average of 60% in F1 score. The impact of additional values and how the time intervals among radar transmission affects the performance, will be further investigated.

ACKNOWLEDGEMENTS

REFERENCES

[1] K. Makantasis, E. Protopapadakis, A. Doulamis, and N. Matsatsinis, "Semi-supervised vision-based maritime surveillance system using fused visual attention maps," *Multimed. Tools Appl.*, vol. 75, no. 22, pp. 15051–15078, 2016.

[2] M. Gianinetto *et al.*, "OBIA ship detection with multispectral and SAR images: A simulation for Copernicus security applications," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1229–1232.

[3] F. Mazzarella, M. Vespe, and C. Santamaria, "SAR Ship Detection and Self-Reporting Data Fusion Based on Traffic Knowledge," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 8, pp. 1685–1689, Aug. 2015.

[4] F. R. G. Cruz *et al.*, "An Interactive and Plotted Data Visualization of High Frequency Doppler Radar Ship Detection and Tracking Software for Maritime Surveillance," in *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM)*, 2018, pp. 1–6.

[5] S. Park, C. J. Cho, B. Ku, S. Lee, and H. Ko, "Compact HF Surface Wave Radar Data Generating Simulator for Ship Detection and Tracking," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 6, pp. 969–973, Jun. 2017.

[6] J. N. Bathurst, M. Hefnawi, J. R. Bray, and Y. M. M. Antar, "Design considerations for a shipboard MIMO radar for surface target detection," in *2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0316–0321.

[7] E. Protopapadakis, A. Voulodimos, A. Doulamis, N. Doulamis, D. Dres, and M. Bimpas, "Stacked autoencoders for outlier detection in over-the-horizon radar signals," *Comput. Intell. Neurosci.*, vol. 2017, 2017.

[8] G. Soldi *et al.*, "Heterogeneous Information Fusion for Multitarget Tracking Using the Sum-product Algorithm," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 5471–5475.

[9] A. C. Antonakis and M. E. Sfakianakis, "Assessing naïve Bayes as a method for screening credit applicants," *J. Appl. Stat.*, vol. 36, no. 5, pp. 537–545, May 2009.

[10] N. Bhatia and Vandana, "Survey of Nearest Neighbor Techniques," *ArXiv10070085 Cs*, Jul. 2010.

[11] W. Stuetzle, "Estimating the Cluster Tree of a Density by Analyzing the Minimal Spanning Tree of a Sample," *J. Classif.*, vol. 20, no. 1, pp. 025–047, May 2003.

[12] S. Abe, *Support Vector Machines for Pattern Classification*. Springer, 2010.

[13] E. Protopapadakis, D. Niklis, M. Doumpos, A. Doulamis, and C. Zopounidis, "Sample selection algorithms for credit risk modelling through data mining techniques," *Int J Data Min. Model. Manag.*, vol. 11, no. 2, pp. 103–128, 2019.

5

Anomaly detection and behavioural analysis

Contents

Traffic density measures for mapping maritime patterns of life

Knut Landmark[a] and Morten Aronsen[a]

[a]Norwegian Defence Research Establishment, PO box 25, 2027 Kjeller, Norway

ABSTRACT

We describe a systematic approach to operational ship traffic density mapping. A ship track is represented as the set of space-time cells it intersects and the visit time for each cell. A density measure is interpreted as the expected number of vessels to be observed in an area at a given instant of time. The cells form an equal-area index mesh of the Earth's ellipsoid. Tracking data have variable spatio-temporal resolution and frequent errors; we therefore implement methods for removing wrong positions and predicting paths between reported positions. Data compression and aggregation are key steps for obtaining good performance in online processing services.

Keywords: data visualization, maritime situational awareness, patterns of life, tracking systems, traffic density

1. INTRODUCTION

Cooperative vessel tracking and identification systems, such as the Automatic Identification System (AIS), are key technologies for maritime surveillance, safety, security, and situational awareness. With the advent of satellite-borne AIS receivers, coverage is now global and no longer limited by the short range of ground receiver stations (~ 50 nm). As military analysts and maritime authorities realize, vessel tracking data accumulated over time are instrumental to understanding maritime patterns of life, i.e., the regular traffic patterns and normal activities that occur in a given geographical area, and how they vary over time.[1] This awareness in turn makes it simpler to detect anomalous events and conditions, be it illicit activities, heightened risk of accident, or hostile electronic spoofing and manipulation.

AIS employs an open-standard, unencrypted, VHF radio-based messaging system that in principle enables anyone to harvest or manipulate the data. Processed data, in particular traffic density maps, are available from several commercial service providers today. However, military users may have to avoid such services due to the risk of disclosing sensitive information about their own modus operandi and areas of interest. There are also concerns about the integrity, veracity, and accuracy of raw and processed AIS data. For example, filtering and error detection, data reduction, interpolation and aggregation influence the products, but the effects are not necessarily transparent. Finally, there is the question of how, precisely, to interpret the visual and numerical information conveyed in ship traffic density maps, and how to apply this information in analysis and decision-making accordingly. In our experience and interactions with operative communities, this has been a pertinent issue.

In this paper we outline the data processing underpinning operational density mapping services developed for the Norwegian Joint Headquarters, the Norwegian Coast Guard, and others. The method has also been re-implemented in the BarentsWatch[2] maritime information and monitoring system (restricted, governmental part). In Sec. 2 we consider a fundamental measure of traffic density, discuss its interpretation and how to derive related measures. Lossy data compression, determined by the basic spatio-temporal resolution, is implemented to speed up the generation of maps in operational services. A global spatial subdivision scheme is introduced for indexing and aggregating data and estimating and displaying density measures. We also propose a solution for handling missing data and variable temporal resolution in raw data, in particular a method for predicting the path of a vessel between two consecutive reported positions. Example maps generated with our method are provided in Sec. 3.

Further author information:
M.A.: E-mail: morten.aronsen@ffi.no, Telephone: +47 63 80 76 69
K.L: E-mail: knut.landmark@ffi.no, Telephone: +47 63 80 76 83

2. METHODOLOGY

2.1 Density measure

While taking Earth's ellipsoidal geometry into account in Sec. 2.2–2.3, we assume for the moment that a surface patch Ω can be subdivided into rectangular cells with physical size $\Delta X \times \Delta Y$. This defines the spatial resolution, in rectangular coordinates (x, y), where x is easting/longitude and y is northing/latitude. Clock time, t, is also discretized with a fixed resolution ΔT. A vessel position (x, y, t) is associated with a unique space-time cell with volume $\Delta X \Delta Y \Delta T$ and integer index $(x, y, t) \mapsto \mathcal{I}(x, y, t) \in \mathbb{Z}$. Conversely, cell index $c \in \mathbb{Z}$ corresponds to a volume $\mathcal{I}^{-1}(c) \subset \Omega \times [T_1, T_2]$, where $[T_1, T_2]$ is the time span of the raw data. When a ship track intersects a space-time cell, the ship identification number (MMSI), the estimated time τ spent in the cell, and certain other attributes like maximum speed over ground (SOG) are recorded in a database table.

The traffic density is computed from the time fraction $\tau / \Delta t$. Suppose $N(c)$ ship tracks intersect cell $c \in \mathbb{Z}$ with visiting times $\tau_1, \tau_2, \ldots, \tau_{N(c)}$, respectively. The space and time-dependent density is defined as

$$\rho(x, y, t) = \frac{1}{\Delta X \Delta Y \Delta T} \sum_{k=1}^{N(c)} \tau_k, \text{ for } (x, y, t) \in \mathcal{I}^{-1}(c). \tag{1}$$

Without resorting to probability theory here, the aggregated time fraction

$$\rho_T(x, y, t) = \frac{1}{\Delta T} \sum_{k=1}^{N(c)} \tau_k = \rho(x, y, t) \Delta X \Delta Y \tag{2}$$

can be interpreted as follows: If someone took a photograph of the area of cell c at time t, then the expected number of ships in the image is $\rho_T(x, y, t)$. The inverse value $1/\rho_T$ is the number of photographs that must be taken around time t to observe on average a single ship (any activity at all) in the cell.

Various density maps can be obtained by averaging ρ over space and time. For arbitrarily fine resolution, averaging is accomplished by integrating ρ with respect to space and/or time. Due to the finite discretization, the integrals are replaced by sums. The density at position (x, y), averaged over the interval $[T_1, T_2]$, is

$$\rho(x, y; T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \rho(x, y, t) \, \mathrm{d}t. \tag{3}$$

The interpretation of $\rho(x, y; T_1, T_2) \Delta X \Delta Y$ is the same as for Eqn. (2) above, and in fact equivalent if we choose $\Delta T = T_2 - T_1$. Similarly, the density averaged over area patch Ω at time t is

$$\rho_\Omega(t) = \frac{1}{\text{area}(\Omega)} \int_\Omega \rho(x, y, t) \, \mathrm{d}A. \tag{4}$$

If Ω corresponds to K basic cells $\mathcal{I}(x_1, y_1, t), \ldots, \mathcal{I}(x_K, y_K, t)$, then $\rho_\Omega(t) = 1/\text{area}(\Omega) \sum_{k=1}^{K} \rho_T(x_k, y_k, t)$. Hence $\rho_\Omega(t)\text{area}(\Omega) = \int_\Omega \rho(x, y, t) \, \mathrm{d}A$ is the expected number of ships seen if we observe Ω at instant of time t. Averaging over both space and time yields $\rho_\Omega(T_1, T_2)$, defined such that $\rho_\Omega(T_1, T_2)\text{area}(\Omega)$ is the expected number of ships observed in a single snapshot of Ω during the interval $T_2 - T_1$.

2.2 Indexing, compression, and storage

For our operational services, we have chosen a basic spatio-temporal resolution of $\Delta X = \Delta Y = 0.1 \, \text{nm} \approx 185 \, \text{m}$ and $\Delta T = 1 \, \text{h}$. Hence the Earth's reference ellipsoid is subdivided into rectangular cells of dimension $0.1 \, \text{nm} \times 0.1 \, \text{nm}$; tracks with low resolution are interpolated and up-sampled to the basic resolution (Sec. 2.3). This yields 108.000 cells in the latitudinal direction while the number of cells in the longitudinal direction is maximum at the equator (216.000) and decreases towards the poles. Cells can be indexed using row and column indices in the range $I_{\text{row}} \in [-108.000, 108.000]$ and $I_{\text{col}} \in [-54.000, 54.000]$, as shown in Fig. 1.

Raw and interpolated data are parsed in one-hour (ΔT) batches and a database record is written for each unique space-time cell and MMSI combination. The following items are recorded: cell index, MMSI, date/time,
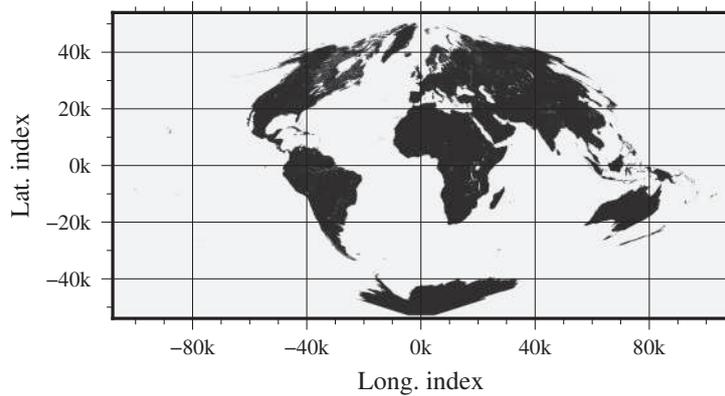
Figure 1: Global index grid with spatial resolution $0.1\,\mathrm{nm} \times 0.1\,\mathrm{nm}$. By coloring every cell that intersects dry nodes in the global land mask (Sec. 2.3) we obtain an everywhere-equal-area world map.

visit time [sec], navigation status, sensor type(s), SOG, observation type (raw/estimated). This table is a (lossy) compressed representation of the vessel data. The grid is sparse, i.e., only a small number of space-time cells are visited by ships, so only non-zero (visited) cells are stored in the database.

By using a one-hour temporal resolution, time-of-day type products can be made, such as day/night plots of general ship traffic and time-of-day differences in, e.g., ferry traffic. In addition, by using the registered maximum speed, a time-of-day difference in speed of the general ship traffic is a possible product. Combining ship density map values ordered by nation and applying vessels of interest lists, such as the Paris MoU black list, areas more frequently visited by nations with a poor performance score are easily visualized. These types of density maps are being used by the Norwegian Joint Headquarters.

2.3 Filtering, prediction, and interpolation

The rate at which messages are received is situation dependent and system dependent. The AIS transmission rate is variable and speed dependent; the Vessel Monitoring System (VMS) stipulated message rate is minimum once per hour, and the Long Range Identification and Tracking (LRIT) rate is between once per 15 minutes and 6 hours. Interruptions may occur for a number of reasons: a transceiver may pass through a terrain shadow; satellite-borne receivers may cover any particular surface patch only about 10–15 minutes per orbit. Satellite receivers in particular also suffer from channel capacity limitations since the time-domain multiple access (TDMA) scheme of AIS allows at most 2250 messages per minute per frequency. In addition, reported positions may be wrong due to, e.g., erratic navigation equipment or the (wrongful) use of non-unique MMSI numbers, causing apparent position jumps. To avoid false paths between ships reporting identical MMSI numbers, the estimated and reported speeds between two positions must be consistent, and care is taken to associate each received position with the nearest track in time and space.

To compute the traffic density with the desired spatio-temporal resolution, we need to estimate ship tracks between reported positions and to filter out messages with evidently wrong positions. At the outset, trajectories are estimated using simple linear interpolation (Fig. 2), but this approach fails if a land area obstructs the direct path, as shown in Fig. 3a. Therefore, a mask is used to reject any reported or interpolated position that occur on dry land. We generated a global land mask using data primarily from the 30 m Global Land Cover dataset from the US Geological Survey (USGS).[3] As shown in Fig. 1, this land mask may be stored as a binary index grid matrix for quick look-ups during data parsing. The grid has 216000×108000 pixels and is stored as a 30 MB GeoTiff file using internal compression and one bit per pixel.

It turns out that about 15 % of all data points are rejected using the land mask filter. We are therefore working on a more accurate alternative to linear interpolation, which is to estimate the actual travel distance and path between two observations. We let $\boldsymbol{r} \in \mathbb{R}^2$ denote a spatial point in some coordinate system of a surface patch Ω. For a fixed source point, $\boldsymbol{r}_0$, the travel time map is a function $\tau : \Omega \rightarrow [0, \infty)$ such that $\tau(\boldsymbol{r})$ is the time
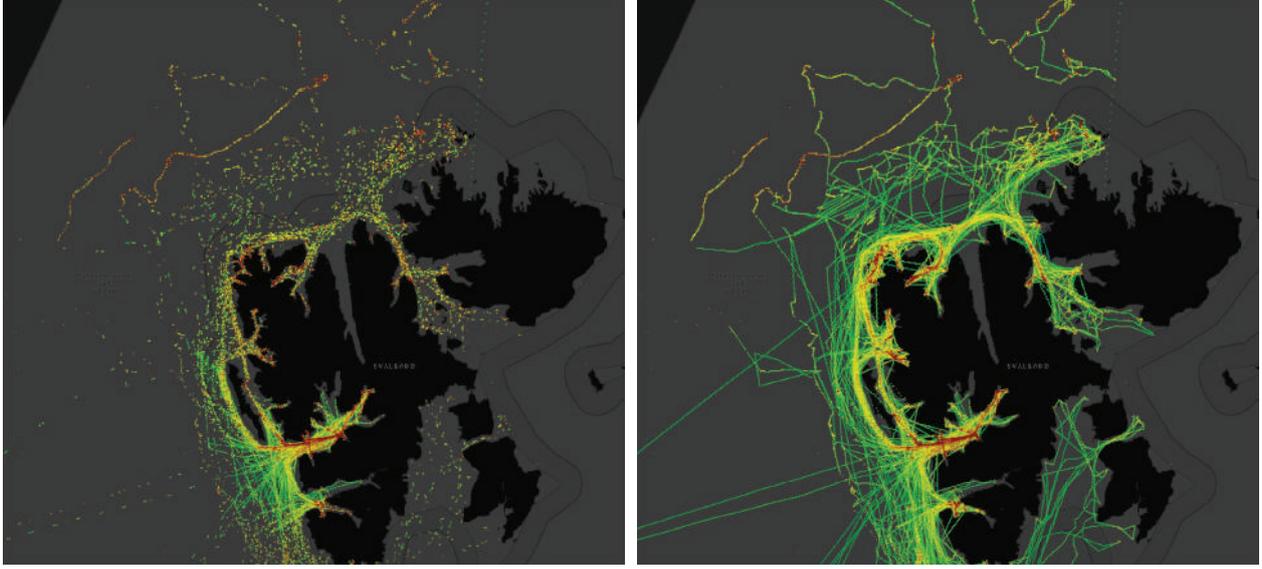
(a) Original data   (b) Data with naive interpolation

Figure 2: Ship traffic density in vicinity of the Svalbard archipelago, based on data from June 2015. The map (a) is based exclusively on reported positions, whereas (b) includes predicted positions using linear interpolation. Upsampling the data yields a better understanding of the actual ship movements and transit patterns. However, linear interpolation also introduces artifacts and false positions that can be amended using a combination of error detection, a land mask, and path prediction.

to reach $\boldsymbol{r}$ along a minimum-time, continuous trajectory from $\boldsymbol{r}_0$. The travel time map satisfies the non-linear differential equation

$$|\nabla \tau(\boldsymbol{r})|^2 = \frac{1}{|\nabla \tau(\boldsymbol{r})|^2}(\nabla \tau)^T G(\boldsymbol{r}) \nabla \tau, \tag{5}$$

where G is the metric tensor for travel time displacements. Assuming isotropic speed, $c(\boldsymbol{r})$, and using geodetic coordinates, $x = \varphi$ (latitude) and $y = \lambda$ (longitude), the metric tensor is

$$G(\boldsymbol{r}) = \frac{1}{c^2(\boldsymbol{r})} \begin{bmatrix} a^2 \cos^2 \varphi/d & 0 \\ 0 & a^2(1-e^2)^2/d^3 \end{bmatrix}, \tag{6}$$

where a is the ellipsoid's semi-major axis, e is the eccentricity, and $d = 1 - e^2 \sin^2 \varphi$. If we use a flat-earth approximation and rectangular coordinates, G is the identity matrix and Eqn. (5) reduces to $|\nabla \tau(\boldsymbol{r})|^2 = 1/c^2(\boldsymbol{r})$, which is the eikonal equation for a linear wave in an isotropic medium.

Equation (5) is a form of Hamilton-Jacobi equation. Using a fast marching (FM) method[4, 5] which admits anisotropic metrics, a finite difference solution can be obtained in running time $\mathcal{O}(N \log N)$, where N is the number of grid points in the (local) spatial domain. Equation (5) is not solved for dry nodes, so the running time depends on the relative extent of land and sea areas. Setting $c(\boldsymbol{r}) \equiv 1$ yields the physical, minimum distance from $\boldsymbol{r}_0$ provided routes are restricted to sea areas. This is shown with the color plot in Fig. 3a. The solution of (5) also yields the normal vector field defined by the expanding contours ("wavefronts") of constant travel time from the source point. By tracing this vector field, the most likely routes can be obtained and sampled.

If the set of coastline pixels is used as source points we obtain a function $D : \Omega \to [0, \infty)$ such that $D(\boldsymbol{r})$ is the distance to the nearest land area (obstacle) with respect to $\boldsymbol{r}$ (Fig. 3b). Using the function $D(\boldsymbol{r})$, possibly combined with other relevant static information such as bathymetry and, crucially, accumulated raw position data, a modified place-dependent speed function $c(\boldsymbol{r})$ can be constructed (Fig. 3c). This approach obtains more realistic paths by giving preference for sea lanes, maintaining safe distance to obstacles, etc. Finally, by tracing the spatially inverted vector field from the target point we obtain an estimate of the vessel path (Fig. 3d).

(a) Wavefront propagation

(b) Distance transform

(c) Relative speed

(d) Predicted path/travel time
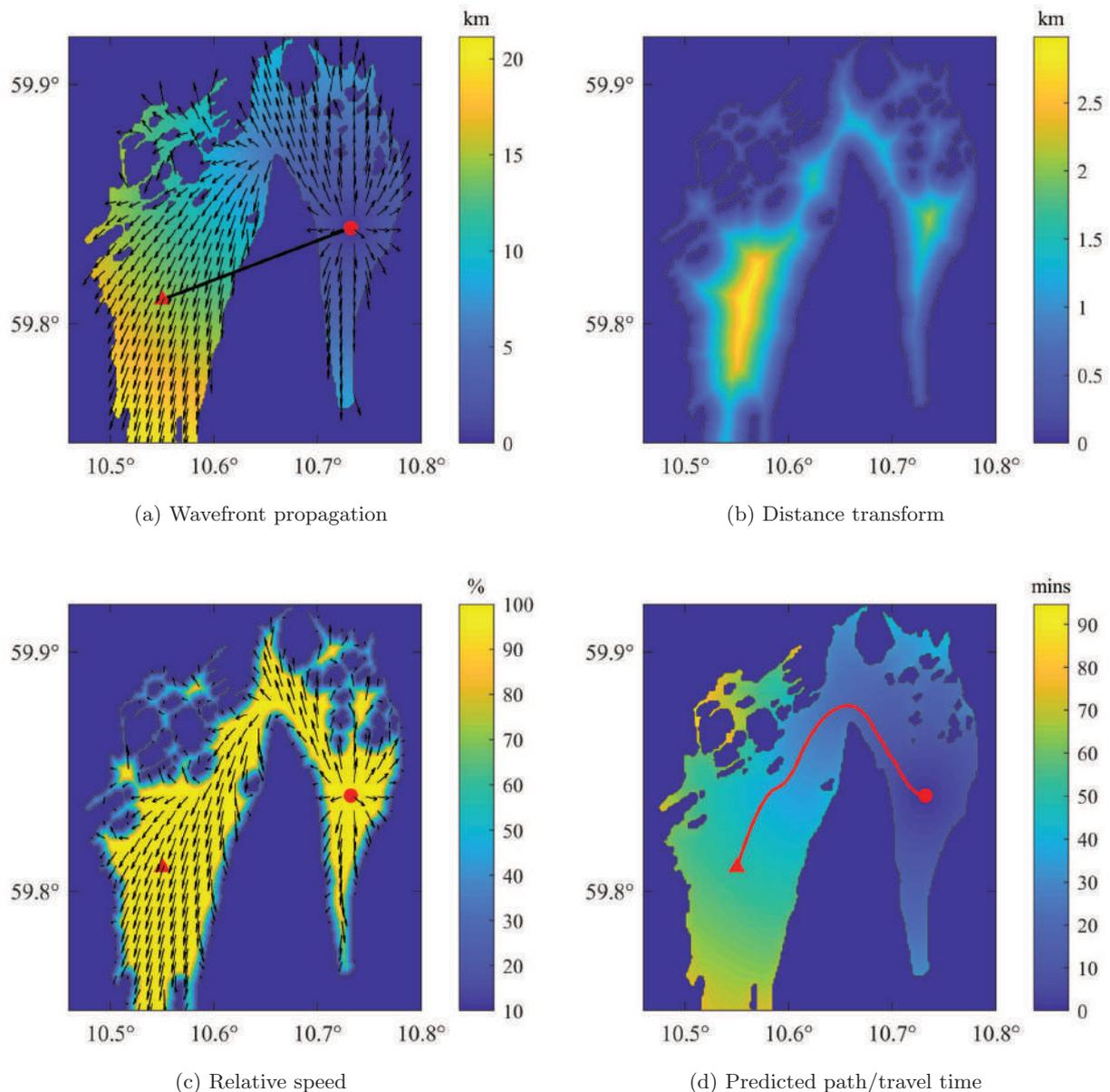
Figure 3: Path prediction, Inner Oslofjord: (a) Simple linear interpolation (black line) between reported positions (red markers) fails wherever the path is obstructed by a land area (dark blue). A more accurate trajectory can be obtained by solving (5) with source point r_0 (red circle) to find the actual travel distance and wavefront normal vector field (black arrows). (b) Eqn. 5 may also be used to pre-compute the distance-to-obstacles function D, which may be used (in combination with other data) to construct metrics for more realistic path predictions (c). Tracing the inverted vector field from the target point yields the path estimate as well as a travel time estimate for a given nominal speed (here: 10 kts.) (d).

3. PRODUCTS AND EXAMPLES

The data sources used in this study are: Norway's land-based AIS network, the satellites AISSat-1 and AISSat-2, the International Space Station AIS receiver, VMS information from the Norwegian Directorate of Fisheries, and
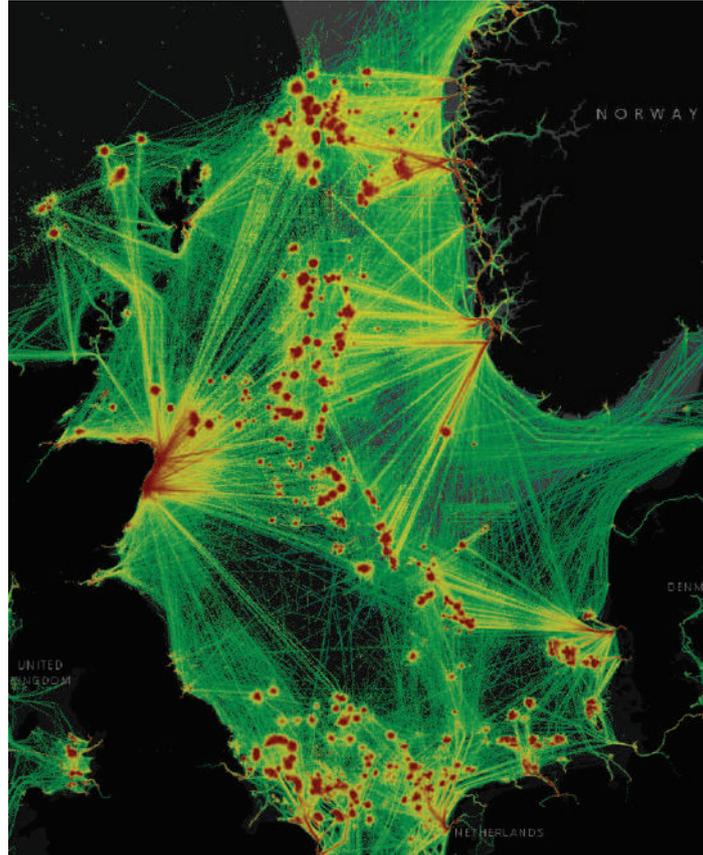
Figure 4: Density map of offshore activity, North Sea, based on 12 months of data (2015). Scale factor: $S = 5$; Statcode 5 type: B2/Offshore. Note: Jenks optimization is not applied here.

LRIT data from the Norwegian Coastal Administration. Ships are classified according to the industry-standard IHS Fairplay Statcode 5 system.

Depending on the extent of the area of interest, basic cells are aggregated to reduce the size of the resulting dataset. With respect to Eqn. (4) we average over patches, Ω, that consist of $S \times S$ basic cells, where S is referred to as the scale factor and typically varies from $S = 1$ (local to reginal maps) to $S = 5$ (large regions, like Europe) to $S = 10$ (world maps). We may also choose to aggregate only cells that are non-empty; this implies less smoothing and generates a different graphics primitive (different polygon). Both methods have advantages and disadvantages. More examples and details can be found in a technical report.[6]

Density maps are generated in vector format (using points and polygons) rather than raster format. This is feasible due to the sparsity of the data. Vector data are scalable so that a single product can be displayed at arbitrary zoom levels, and it is straightforward to attach metadata to each map object. A great variety of maps can be generated by selecting different attributes and ship/activity types and spatio-temporal scales; one example is shown in Fig. 4. To accentuate patterns in the data we form a partition of the density value range using Jenks natural breaks optimization,[7] the objective of which is to minimize the variance within each interval while maximizing the variance between intervals. Colors from the selected color map can be assigned according to this classification. The effect of Jenks optimization can be seen in the example in Fig. 5.
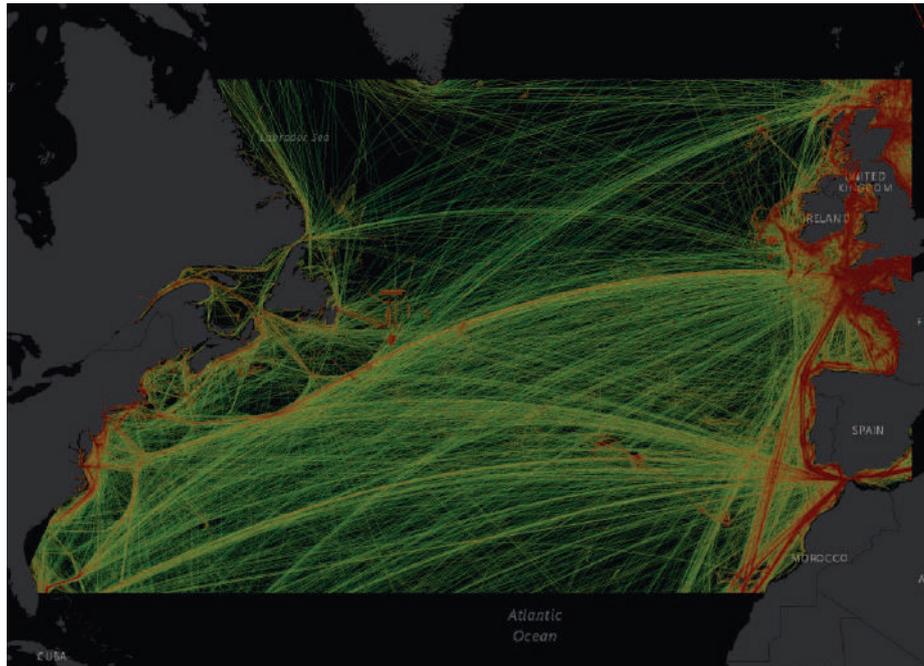
(a) Equi-distant levels

(b) Jenks-optimized levels

Figure 5: Density map of North-Atlantic crossings, based on data from August 2017. The color scale in (a) is divided into fixed-length traffic density intervals, whereas in (b) Jenks optimization has been used to partition the range of density values.

4. DISCUSSION AND CONCLUSION

Density maps are not mere visual patterns. The density defined by Eqns. (1)–(4) has a straightforward numerical interpretation that may inform various analyses and decisions. The density $\rho_\Omega(t)$ defined by (1) and (4) is such that $\rho_\Omega(t)\mathrm{area}(\Omega)$ is the expected number of ships observed in area Ω at instant of time t. A series of $N_{\mathrm{obs}} = N_{\mathrm{ship}}/\rho_\Omega(t)\mathrm{area}(\Omega)$ observations ("photographs") is needed on average for N_{ship} detections in Ω; these detections are not necessarily of distinct ships. The time-averaged density $\rho_\Omega(T_1, T_2)$ has the same interpretation as $\rho_\Omega(t)$ for observations in the interval $T_1 \le t \le T_2$.

The mapping of maritime traffic patterns and activities is useful in many settings, e.g., when planning military exercises, for estimating underwater ambient noise levels, for safe navigation purposes, or for detecting anomalous events. In our approach, a vessel track is represented as the set of space-time cells it intersects and the amount of time it visits each cell. Cells can be aggregated in large-scale maps to increase performance and reduce file size. These data reduction steps are key to obtaining good performance in operational services. With data pre-processing as proposed in this paper, it is also straightforward to map related quantities, such as the maximum observed speed, and more measures will be investigated.

The proposed solution for non-linear path prediction and up-sampling (Sec. 2.3) is computationally more expensive than linear interpolation. However, it need only be executed in the special circumstances where a vessel must navigate around an obstacle in the time between two reported positions; otherwise simpler methods are adequate. A pre-computed distance-to-obstacles map, as in Fig. 3b, may be used to decide when this condition occurs, effectively a simple look-up operation. Moreover, path prediction is required in related work, including online anomaly detection, risk assessment and safe navigation,[8] and optimum path planning. Therefore, the data ingestion step in density mapping may share computing resources and results with other online/real-time systems.

REFERENCES

[1] Pallotta, G., Vespe, M., and Bryan, K., "Vessel pattern knowldege discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy* **15**, 2218–2245 (June 2013).

[2] "BarentsWatch." `https://www.barentswatch.no/en/` (2019). Last accessed: 2019-05-23.

[3] Hansen, M. et al., "High-resolution global maps of 21st-century forest cover change," *Science* **342**(6160), 850–853 (2013).

[4] Sethian, J. A. and Vladimirsky, A., "Ordered upwind methods for static Hamilton-Jacobi equations," *PNAS* **98**(20), 11069–11074 (2001).

[5] Tsitsiklis, J. N., "Efficient algorithms for globally optimal trajectories," *IEEE Trans. Autom. Control* **40**(9), 1528–1538 (1995).

[6] Aronsen, M. and Landmark, K., "Density mapping of ship traffic," FFI-rapport 16/02061, Forsvarets forskningsinstitutt, Kjeller (Nov. 2016).

[7] Jenks, G. F., "The data model concept in statistical mapping," in [*International Yearbook of Cartography*], **7**, 186–190 (1967).

[8] Messel, E. and Bjørndal, M. G., "Behavioral analysis for maritime safety," in [*Proceedings of the 1st Maritime Situational Awareness Workshop, CMRE, La Spezia, 8–10 October 2019*], Braca, P. and Jousselme, A.-L., eds. (2019). (This volume).

A data driven approach to maritime anomaly detection

Zissis D.[a,b], Chatzikokolakis K. [b], Vodas M. [b] , Spiliopoulos G. [b] and Bereta K. [b],
[a] University of the Aegean, Syros, Greece
[b] MarineTraffic, Athens, Greece;

ABSTRACT

The operational community has long identified anomaly detection systems as vital for increasing the effectiveness of maritime surveillance systems, since the huge quantities of data produced today quickly reduce their effectiveness in the field. The limited range of currently available maritime anomaly detection systems rely heavily on expert knowledge and hardcoded rules; thus, limiting their scope and effectiveness only to known situations and patterns of behaviour. By contrast, data driven systems learn from the data itself and can thus generalise well to new tasks and previously unseen situations. In this paper, we present an overview of the anomaly detection system developed in the context of the European Commission H2020 funded project BigDataOcean, and specifically the "Maritime Security and Anomaly Detection" pilot. Through a combination of unsupervised machine learning methods and behavioural analytics, the developed system is capable of i) automatically modelling shipping routes at a global scale; ii) constructing vessel specific profiles and class baselines; and iii) detecting deviations from patterns of normalcy in real time.

Keywords: Anomaly Detection, Machine Learning, Big Data, Maritime Situational Awareness

1. INTRODUCTION

Anomaly detection in the maritime domain has been identified by the operational community as an important aspect requiring intensified research efforts and development [1][2]. Commonly, surveillance operators have to search and predict emerging conflict situations, for example, potential collisions, dangerous vessels, or suspicious activities from a large number of vessels within geographical regions. Today there are more than 23 mandatory ship reporting systems (e.g. the Automatic Identification System) and numerous surveillance systems (e.g. coastal radars), which produce constant streams of high-speed and high-volume data while tracking vessels at sea. Exploring and monitoring the data manually is a demanding task, not only due to the complexity and heterogeneous nature of the data itself but also due to other factors like uncertainty, fatigue, cognitive overload, or other time constraints. Early detection of dangerous situations provides critical time to take appropriate action, possibly before potential incidents evolve [2]. Increasing automation through advanced machine learning and artificial intelligence methods enables the system and the operator to spot complex situations by correlating various events from all surveillance sensors and classify them into important incidents.

As such, Maritime Situational Awareness (MSA) involves developing the ability to identify patterns emerging within huge amounts of data, fused from various uncertain sources and generated from monitoring thousands of vessels a day, so as to act proactively and minimise the impact of possible threats. At the core of this process is data mining, an essential step in the process, consisting of applying data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns over the data. Vessels conducting illegal behaviour often try to hide their intentions and follow a set of patterns depending on the activity they perform: deviation from standard routes, unexpected AIS activity, unexpected port arrival, close approach, and zone entry [3]. Within this context anomaly detection can be understood as a method that supports situational assessment by indicating objects and situations that, in some sense, deviate from the expected, known or "normal" behaviour.

Essential for effective anomaly detection is building an accurate model of normalcy. The understanding of the complex maritime environment and a vessel behaviour though, cannot be limited to simply connecting vessel positions as they travel across the seas. As such "Patterns of Life" are understood as observable human activities that can be described as patterns related to a specific action (e.g. fishing) taking place at a specified time and place. Essentially, vessel-based maritime activity can be described in space and time, while classified to a number of known activities at sea (e.g., fishing, dredging, etc.). The spatial element describes recognised areas where maritime activity takes place; thus, including ports, fishing grounds, offshore energy infrastructure, dredging areas and others. The transit paths to and from

these areas also describe the spatial element, e.g. commercial shipping and ferry routes etc., while the temporal element often holds additional information for categorising these activities (e.g. fishing period, time of year).

This paper summarises the work completed in the context of the BigDataOcean H2020 project, which developed a fully-automatic method for first defining and then classifying vessel activities to "Patterns Of Life" in a data driven way, with no reliance on external information or need for a priori knowledge. The proposed framework is capable of dealing with big data volumes such as those produced in the maritime domain and specifically by the Automatic identification System at a global scale. In sum, the steps followed include first detecting port and activity areas, then trade routes and traffic patterns between these, before detecting deviations in true real time. The system has been developed in a step-wise approach and is currently in its final development phase. It relies on a number of state of the art and beyond state of the art frameworks for big data processing such as Spark[1] for batch and Akka[2] for stream processing, which are combined into a modified purpose built Lambda architecture [18]. The following sections present a short literature overview before describing the overall architecture and each element of the framework. In the final section a short discussion of results regarding real work incident detection is provided.

2.   RELATED WORK

In recent years there has been a growing output of scientific papers and systems dealing with anomaly detection in the maritime domain and attempting to apply automated techniques to anomaly discovery. Throughout the literature anomalies are studied at different levels; both at a vessel information level (such as that provided in relevant registries and databases) and at a mobility (positional/trajectory) level. Often these two approaches are complementary.

In the first case, any available information about a vessel that can be found is used to build a behavioural profile for the ship, chartering company, etc. Such information is often static, unvarying or slowly changing over time. Two classes of solutions are dominant in this perspective; the ones relying on probabilistic risk assessment and the ones using fuzzy logic as a relaxation approach to the definite boundaries of probabilistic approaches. Probabilistic risk assessment has been introduced as a solution for the assessment of risk in the maritime domain in [6]. In [7] the authors applied a Bayesian simulation for the occurrence of situations with accident potential and a Bayesian multivariate regression analysis of the relationship between factors describing these situations and expert judgments of accident risk, to perform a full-scale assessment of risk and uncertainty. A fuzzy approach that evaluates the maritime risk assessment when applied to safety at sea and more particularly, the pollution prevention on the open sea is introduced in [8].

In the second case the focus is on mobility analytics. Some of the typical mining tasks in the spatiotemporal context include, frequent pattern discovery, trajectory pattern clustering, trajectory classification, forecasting, and outlier detection. Specifically, trajectory classification, includes constructing a model capable of predicting the class labels of moving objects based on their trajectories and other features [4]. The majority of works here focus on a data driven definition of normalcy from AIS data which then is used as baseline information to detect critical deviations. A number of publications rely on statistics to generate simple analytics of ship traffic and frequencies [5] [6]. More complex approaches can be categorised into (i) grid-based methods and (ii) methods of using vectorial representations of traffic. In grid-based approaches, the area of coverage is split into cells characterised by the motion properties of the crossing vessels to create a spatial grid (e.g., [16-21]). For example in [7], a data-driven methodology is proposed to estimate the vessel times of arrival in port areas. Grid-based anomaly detection algorithms include Fuzzy ARTMAP [8] , Holst Model [9], [10], Support Vector Machine and others. [11]. In the second category, vessel trajectories are modelled as a set of connected waypoints. Thus, vessel motions in large areas (e.g., at a global scale) can be managed thanks to the high compactness of the waypoint representation [12][13]. For example, the authors in [14] apply a Bayesian vessel prediction algorithm based on a Particle Filter (PF) on AIS data. The authors of [17] present an approach for detecting route deviations based on the Ornstein-Uhlenbeck (OU) mean-reverting stochastic process.

3.   OVERALL APPROACH AND ARCHITECTURE

An efficient anomaly detection tool needs to support practitioners and operators along the complete lifecycle of monitoring situations at sea, from the observation of elements in the environment up to detection of anomalies and aids to planning. Following a number of dedicated workshops with representatives from the operational community a number of system requirements were identified, including but not limited to: (i) Providing a robust understanding of what to

---

[1] https://spark.apache.org/
[2] https://akka.io/docs/

expect in an area. (e.g. Patterns of Life), (ii) the system needs to prioritise information to the operator, filtering down data to a manageable cohort that may be considered for further investigation, using intelligent algorithms, (iii) real-time vessel risk profiling and activity classification should be provided, and (iv) anomaly detection should be performed in real-time, providing operators with justifiable alerts of critical unfolding incident and activities.

To fulfil this array of requirements we developed a number of components and novel algorithms, coupled together in a modified Lambda architecture. The architecture provides both support for batch processing of large volumes of spatiotemporal data and real time streaming capabilities. Each component uses the output of the previous step in an incremental fashion to produce its results. An overview of the workflow is given in the figure below
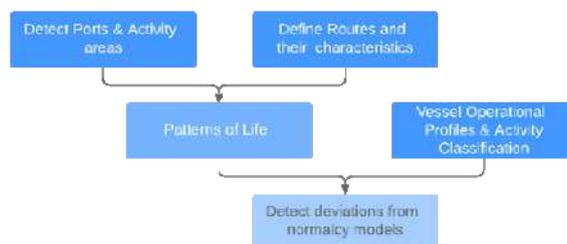

Fig. 1 BigDataOcean framework for detecting deviations from normalcy

In the following sections we discuss the related parts of the framework which build up the holistic systems behaviour.

3.1 Port Area and Activity Detection

In our previous work, in collaboration with CMRE [15], [16], we developed a practical data driven implementation of a distributed method for calculating port operational areas from large amounts of spatiotemporal data. It is often an overlooked fact that seaport areas do not remain static over time; such spatial units often evolve according to environmental and seasonal patterns in size, connectivity, and operational capacity. Thus, accurately defining a seaport's exact location, operational boundaries, capacity, connectivity indicators, environmental impact and overall throughput is a challenging task. An adaptation of the well-known KDE algorithm to the map-reduce paradigm was developed and implemented on the distributed Spark framework so that it could cope with massive amounts of spatiotemporal data [15], [16]. In the context of the BigDataOcean project, the proposed approach was extended to other types of stationary areas besides ports: such as off-shore platforms, anchorage areas, and fishing grounds.

3.2 Route Definition

Sea going vessels follow specific paths when travelling across the vast blue ocean; these "roads" connecting major ports are some of the busiest places on earth, often only a few kilometres wide, scattered with many physical constraints (e.g., reefs), where enormous vessels perform risky manoeuvres under constantly changing environmental conditions (e.g., wind, sea currents). These waterways form a global maritime exchange network. More than often these connections are not direct lines (e.g., the shortest distance from the point of departure to destination), but "climatological routes" along which higher speeds can be achieved due to the existence of currents or the prevalence of wind, sea or swell. Sea roads though are not paved in concrete, as the location of the connector, its width and its content, can vary significantly over space and time, under the influence of various trade and carrier patterns, but also due to large infrastructure investments (e.g., canal expansions), climatic changes (e.g., global warming), new traffic restrictions (e.g., Emission Control Areas), political events and other international incidents (e.g., increase of piracy in specific regions).

In this context, a maritime big data modelling approach was developed, capable of accurately identifying the spatiotemporal dynamics of ship routes and most crucially their characteristics (such as route variable width, types of vessels, direction of travel etc.), thus deriving the maritime "patterns of life" at a global scale, without the reliance on any additional information sources or a priori knowledge. The proposed approach is based on the MapReduce programming model. The algorithm first filters and sorts all the positional data by assigning each position to a trip before "clustering" (reducing) each trip's data. In this sense the algorithm operates initially at a micro level, assigning data to trips, before developing a global network. The entire network can be produced on request in a few hours by using big data technologies (Spark and MapReduce) on a cluster of distributed computing nodes so as to depict any changes in the network due to external factors. In terms of performance and accuracy, experimental results on real world noisy datasets

confirm an achieved overall accuracy of approximately 80%, while the entire processing duration is less than 3 hours for a terabyte of data.

3.3 Vessel operational profiles and activity classification

Every vessel has a unique Operational Profile. Each profile is a continuous learning model that employs machine learning to interpret behaviour in real-time. To construct the vessel profiles, we based our analysis on a number of features i.e., static vessel characteristics that, when combined, can give a representative view of a behaviour of a vessel. We observed that when the values of these features deviate considerably from the average values of the total vessel population, this can be an indication that the vessel could have an increased risk.

We collected data about vessels that have been involved in illegal activities, such as smuggling, illegal fishing or have been detained or even banned and we used them as ground truth. We observed that for a number of features, they deviate considerably from the respective average values for the vessel population of their category. Data driven methods are used to discover anomalies by learning statistical properties of the data and finding which data points do not fit.

Some of the features that we used are the following: Number of flag changes and name changes, number of transhipments, number of night port calls, number of port visits in countries with no anti-terrorism measures, the age of the vessel, number of days of coverage for a year, number of different countries visited, etc. In some cases, the importance of a feature makes sense when applied to the global fleet (e.g. number of flag changes or name changes) while in other the market segment plays significant role. For instance, number of transhipments may differ significantly for various vessel types (e.g., cargo or tankers) or completely irrelevant for other (e.g., passenger vessels or tugs). Thus, a subset of the features used have been calculated for each vessel type separately taking into account the different operations of shipping segments. Then, we employed a Random Forest Regressor to estimate the importance of each feature (shown in the figure below) and calculate the vessel's risk indicator.
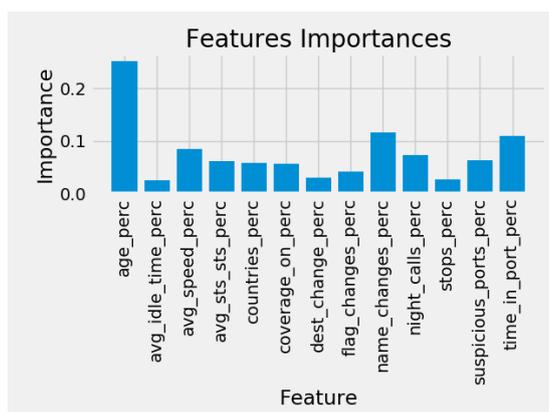
Fig. 2 Random Forest Regressor features' importance

3.4 Detecting deviations from normalcy models in real time

Time critical computations need to be completed in real time. Such systems have been described as systems which "control an environment by receiving data, processing it and returning the results sufficiently quickly to affect the environment at that time". In traditional system architectures the increasing volume of data ingested obviously negatively affects computation time. In our approach we introduce various architectural optimisations in order to achieve enhanced performance and low latency of computation. We have deployed a modified Lambda architecture, as this scheme allows the decoupling of batch processing (usually performed upon historical data) and real-time analysis, which typically exploits the knowledge extracted from the batch processing. The batch layer performs the analysis of historical positional data of vessels and extracts the required Patterns of Life. This is a long-running process which takes several hours to complete. Each Pattern of Life is a set of polygons that connect a departure port with a destination port for a specific vessel type. Taking into account the type of a vessel when producing the Patterns of Life is crucial, considering that vessels of different type may follow different routes for various reasons (e.g., due to vessel size limitations, sea depth, etc.). Once completed, the PoLs, together with the vessel operational profiles and activity classifications, are fed into the real-time layer in order to accommodate detection of vessel anomalies in real-time. Upon detection those incidents are

displayed to the end user through the service layer. Furthermore, historical data are sent from this layer back to the batch layer at specific time intervals defined from the seasonality of the data, thus replacing previous processed patterns with new ones.

In the real-time layer, queries are performed on streaming and previously unseen data, thus enabling detection of security incidents in real time. More specifically, streaming vessel's positions and voyage-related AIS data are combined with static datasets and data mining models in real-time to detect anomalous events.

Route deviations: A route deviation is triggered when a ship is found travelling outside the "normal route" it is expected to follow in a given area and time, out of normal limits (e.g. speed patterns) or there is a mismatch in its type of activity reported for the given area or time period (e.g. fishing in a forbidden area). The first case (i.e., vessel travelling outside the "normal route") is triggered when a spatial intersection query of the vessel's position and the geometry of the PoL that the vessel should follow (based on its departure and destination ports) is empty. The second case (i.e., vessel travelling out of normal limits) is triggered when the first case is not triggered and in addition the vessel does not follow the PoL's statistical information in terms of SoG, CoG, Heading or Travelling Time. Finally, the third case is triggered when the vessel's identified activity (e.g. fishing) spatially intersects with an area that is semantically annotated as forbidden for that type of activity.

Proximity events: A proximity event is triggered when at least two vessels are detected sailing within close proximity to each other. This could be interpreted as an imminent collision, Ship to Ship Transfer or other. For each new message received at time *t* the real-time layer filters out former messages that are neither close spatially (i.e., messages with position more than 200 m away), nor temporally (i.e., messages with timestamp more than 10 minute before *t*), and projects the rest at time t based on their reported SoG and CoG. Then, for each of the projected positions we assume an area of probable movement (based on its SoG and CoG). If any of those areas intersects spatially with the area of probable movement of the vessel, then a proximity event is triggered.

Sailing in shallow/dangerous waters. Sailing in shallow waters are events that occur when a vessel is travelling in areas where the sea depth is less than the ship's draught. This information is correlated with bathymetry data and vessel's draught reported through AIS to increase the accuracy of detected groundings. Additionally, these events are fired when a ship is travelling in previously reported dangerous areas (such as high piracy areas etc.). In some cases, sailing in shallow waters or in dangerous areas would also mean that the vessel is sailing out of its Pattern of Life, and one or multiple route deviation events could also be triggered. Furthermore, when investigating potential groundings, relying only on bathymetry data may lead to poor accuracy, as most of the available dataset offer a coarse grain analysis of the sea depth providing sea depth values for large areas (i.e., several km^2) and do not take into account tides or seasonality that affect the sea depth. Thus, in order to increase the confidence level of our algorithm the real-time layer takes into account also the navigational status that is reported through AIS and vast decrease in vessel's SoG and CoG which may indicate that the vessel has ran aground.

Frequent or vessel specific AIS field Changes. Alerts are generated for vessels that highly exceed the number of frequent changes in the static parameters of AIS or for user defined vessels. The vessel may report through the navigation status field of AIS messages that is not be under command, or its ability to manoeuvre is restricted, or even explicitly report it is aground; thus an alert is generated.

Loss of AIS signal. A loss of AIS signal happens when the service stops receiving data from a ship whilst the ship is within the network coverage of AIS receivers, suggesting a possible "dark target". The reporting rate of an AIS transponder varies from 2 seconds to 3 minutes depending on the class of the transponder, the vessel's speed and changes of speed and/or course. Thus, using a time threshold the real-time layer can roughly estimate when to expect a message from each vessel and based on its last known position, SoG and CoG it can also roughly estimate where to expect that message. Thus, if the estimated location is within network coverage (i.e., we have received messages from other vessels in the vicinity) at the estimated timestamp and no message from the vessel has been received, the real-time layer yields a Loss of AIS Signal event.

Vessels of Interest (high risk ships). Ships that have been previously identified as High Risk or Vessels of Interest, trigger a number of alerts as they visit ports, user defined areas or conduct activities at sea. These are displayed on a map as a layer providing insights to end-users for potential illegal activities.

4. CONCLUSION

Situational awareness and anomaly detection tools are of outmost importance for the operational community in the maritime domain. In the context of the BigDataOcean project we developed a framework and system prototype capable of detecting maritime anomalies at a global scale in a data driven way, thus without the requirement of user defined rules or a priori knowledge. This paper provides an overview of the work conducted in the context of this project, highlighting important findings and breakthroughs.

ACKNOWLEDGMENTS

REFERENCES

[1]   E. Martineau and J. Roy, "Maritime Anomaly Detection: Domain Introduction and Review of Selected Literature," Defence Research and Development Canada, 2011.

[2]   M. Riveiro, G. Pallotta, and M. Vespe, "Maritime anomaly detection: A review," Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 8, no. 5, p. e1266, 2018.

[3]   "Maritime Anomaly Detection Based on Mean-Reverting Stochastic Processes Applied to a Real-World Scenario - Semantic Scholar." . Available: https://www.semanticscholar.org/paper/Maritime-Anomaly-Detection-Based-on-Mean-Reverting-d'Afflisio-Braca/9606a9958d70ce5686e3c285dd07b3681fd604b0. [Accessed: 16-Mar-2019].

[4]   J.-G. Lee, J. Han, X. Li, and H. Gonzalez, "TraClass: Trajectory Classification Using Hierarchical Region-based and Trajectory-based Clustering," Proc VLDB Endow, vol. 1, no. 1, pp. 1081–1094, Aug. 2008.

[5]   L. Zhang, Q. Meng, and T. Fang Fwa, "Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters," Transp. Res. Part E Logist. Transp. Rev., Aug. 2017.

[6]   Q. Meng, J. Weng, and S. Li, "Analysis with Automatic Identification System Data of Vessel Traffic Characteristics in the Singapore Strait," Transp. Res. Rec. J. Transp. Res. Board, vol. 2426, pp. 33–43, Sep. 2014.

[7]   A. Alessandrini, F. Mazzarella, and M. Vespe, "Estimated Time of Arrival using Historical Vessel Tracking Data," IEEE Trans. Intell. Transp. Syst., pp. 1–9, 2018.

[8]   N. A. Bomberger, B. J. Rhodes, M. Seibert, and A. M. Waxman, "Associative Learning of Vessel Motion Patterns for Maritime Situation Awareness," in 2006 9th International Conference on Information Fusion, 2006, pp. 1–8.

[9]   R. Laxhammar, "Anomaly detection for sea surveillance," in 2008 11th International Conference on Information Fusion, 2008, pp. 1–8.

[10]   B. Ristic, B. L. Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in AIS Data: Anomaly detection and motion prediction," in 2008 11th International Conference on Information Fusion, 2008, pp. 1–7.

[11]   L. Wu, Y. Xu, Q. Wang, F. Wang, and Z. Xu, "Mapping Global Shipping Density from AIS Data," J. Navig., vol. 70, no. 01, pp. 67–81, Jan. 2017.

[12]   Y. Li, R. W. Liu, J. Liu, Y. Huang, B. Hu, and K. Wang, "Trajectory compression-guided visualization of spatio-temporal AIS vessel density," in 2016 8th International Conference on Wireless Communications & Signal Processing (WCSP), 2016, pp. 1–5.

[13]   M. Fiorini, A. Capata, and D. D. Bloisi, "AIS Data Visualization for Maritime Spatial Planning (MSP)," Int. J. E-Navig. Marit. Econ., vol. 5, pp. 45–60, 2016.

[14]   F. Mazzarella, V. F. Arguedas, and M. Vespe, "Knowledge-based vessel position prediction using historical AIS data," in 2015 Sensor Data Fusion: Trends, Solutions, Applications (SDF), 2015, pp. 1–6.

[15]   L. M. Millefiori, D. Zissis, L. Cazzanti, and G. Arcieri, "A distributed approach to estimating sea port operational regions from lots of AIS data," in IEEE Int. Conference on Big Data (Big Data), 2016, pp. 1627–1632.

[16]   L. M. Millefiori, D. Zissis, L. Cazzanti, and G. Arcieri, "Scalable and Distributed Sea Port Operational Areas Estimation from AIS Data," in 16th International Conference on Data Mining Workshops (ICDMW), 2016, pp. 374–381.

[17]   E., P. Braca, L. M. Millefiori, P. Willett: Detecting Anomalous Deviations From Standard Maritime Routes Using the Ornstein-Uhlenbeck Process. IEEE Trans. Signal Processing 66(24): 6474-6487 (2018)

[18]   Amazon Web Services, Lambda Architecture for Batch and Stream Processing on AWS, May 2015

Verification of Sensor Data in a Maritime Multi-Sensor Network

Martina Broetje, Giulia Battistello, Martin Ulmke

Dept. Sensor Data and Information Fusion
Fraunhofer FKIE
D-53343 Wachtberg, Germany

ABSTRACT

The reliability of sensor data is a prerequisite for sensor data fusion. In practice, sensor data can be faulty or even manipulated. Thus, automatic verification and if applicable automatic calibration of sensor data are important steps to guarantee robust performance of a surveillance system exploiting fusion of multiple heterogeneous sensors. In this paper we discuss a concept for step-by-step verification and fusion of sensor data within two applications. The first one addresses a harbor protection scenario. The goal is to detect suspicious or potentially hazardous behavior of targets and warn the user over a human machine interface. The second application deals with an acute threat case in a maritime environment. The goal is to clear up this special situation by adding additional sensor data to an existing sensor system. In both applications the additional information extracted from sensor data needs to be weighted by the reliability of sensors in the network.

Keywords: Sensor data, verification, fusion, tracking

1. INTRODUCTION

The combination of sensor data from multiple heterogeneous sources and from different locations can be necessary to clear up a critical situation. Sensor data may be available either at measurement (i.e. detections/plots) or track level. Due to the abundance and complexity of information in such a situation a human operator needs assistance by automatic data evaluation processes to make a decision.

To get the best out of the fusion from multiple sources, the sensors in a surveillance network usually need to be trained, calibrated and synchronized, see e.g.[1] However, if we add a faulty (uncalibrated or even jammed) sensor to an existing network, this can even degrade the performance of the whole system. Faced to a critical situation, we want to combine information from all available sensors but maintain reliability of the combined surveillance system.

In this paper we discuss a concept for automatic verification of sensor data within a sensor network. If a faulty sensor is detected, it is either removed from the fusion process or, if possible, corrected.

The detection and estimation of error sources in sensor data is discussed in literature in detail. From this we distinguish between different error sources. Generally established in tracking literature is the model of additive Gaussian measurement noise, which summarizes a large number of random errors due to the signal, hardware and signal processing characteristics, see e.g.[2] Such *random errors* can be caused, e.g., by imprecision in the position, velocity and heading of the receiving antenna. For robust target tracking it is necessary that the imprecision in these parameters are correctly reflected in the tracking sensor model.

Deterministic errors arise due to simplified modeling assumptions, e.g. the measurement or the target prediction equation. By modification of the measurement equation in tracking (improving the model of the relationship between target state and measurement) this bias can be compensated, see[3] and.[4] Deterministic errors typically depend on the unknown target state and are handled in target tracking directly.

Systematic errors do not depend on the target state but are characteristic for the considered sensors. Typical systematic error sources in target tracking are summarized in.[5] It is differentiated between *calibration*

Send correspondence to martina.broetje@fkie.fraunhofer.de

error/measurement bias, *location bias*, *attitude/orientation bias* and *temporal bias*. For multi-sensor multi-target tracking operations, the impact of systematic errors is discussed in.[6] There, the focus lies on derivation of a bias estimation algorithm at track-to-track fusion level. The coupling of bias estimation and track-to-track association is handled in.[7] Systematic errors can be reduced by (i) exploiting the measurements of reference targets,[8] (ii) using tracks of fixed targets[9] or (iii) aligning tracks from multiple sensors (e.g.[10]). The first two methods can even be applied if only measurements of a single sensor are available and make sense if the prior knowledge about the targets exists. This can e.g. be information about the existence of a static target (like a ground-based radar with rotating antenna[9]). In[11] the prior knowledge in form of air lanes is exploited for bias estimation.

In order to test sensor data, we develop a multi-stage verification scheme, which resorts on the following steps:

1. Continuity check: this verification step is applied on single sensor measurements. The test is fulfilled if the extracted sensor features (tracks or static features) are continuous in time;

2. Plausibility check: it is also performed at single sensor level. It compares tracks or features from the previous verification stage with prior knowledge (e.g. map data);

3. Consistency check: it is executed within the multi-sensor fusion. Tracks from the first verification stage are compared with other sensor data.

2. MODELLING

The output of a data fusion system is typically in Cartesian coordinates, thus we define the target state by its position and velocity, $\mathbf{x} = (x, y, z, \dot{x}, \dot{y}, \dot{z})$. The sensor data can be either tracks in Cartesian coordinates or plots in measurement coordinates. We further assume that for each sensor i there exists a measurement function $\mathbf{h}_i$, which describes the relationship between the ideal target state $\mathbf{x}$ and the measurement coordinates $\mathbf{z}$, i.e.: $\mathbf{z} = \mathbf{h}_i(\mathbf{x})$. This assumption is necessary to process heterogeneous sensor data in a general framework. The measurements can thus be available e.g. as range and azimuth or Cartesian positions. The target state is propagated over time by exploiting the target propagation equation, i.e. $\mathbf{x}(t_k) = f(\mathbf{x}(t_{k-1}, t_k - t_{k-1}))$, where $\mathbf{x}(t_\ell)$ describes the target state at time t_ℓ.

As typical in target tracking we model random measurement errors by Gaussian noise, i.e.

$$\hat{\mathbf{z}} \sim \mathcal{N}(\mathbf{h}_i(\mathbf{x}), \mathbf{R}), \tag{1}$$

where $\mathbf{R}$ represents the covariance matrix of the measurement error. For simplicity we neglect the dependency on time, when measurement and target state are considered at the same time instance.

Also the target propagation model is described by

$$\mathbf{x}(t_k) \sim \mathcal{N}(f(\mathbf{x}(t_{k-1}), t_k - t_{k-1}), \mathbf{Q}(t_k - t_{k-1})), \tag{2}$$

with process noise matrix $\mathbf{Q}$. Additional errors can be incorporated by expanding the measurement model, as shown in the following examples related to heading, time and registration errors:

- Heading error: A heading error influences the azimuth measurement component φ and can be modeled by an additive constant bias Δ_φ, i.e.

$$\hat{\varphi} \sim \mathcal{N}(\mathbf{h}(\mathbf{x}) + \Delta_\varphi, \sigma_\varphi^2) \tag{3}$$

with $\mathbf{h}(\mathbf{x}) = \mathrm{atan}(x - r_x, y - r_y)$ and where $\mathbf{r} = (r_x, r_y)$ describes the position of the receiving antenna.

- Time error: The time error is modeled as a constant bias Δ_t. As consequence the target state and the measurement are evaluated at different times. To correct this, a time shift of the target state is modeled by using the target motion equation $\mathbf{f}(\mathbf{x}, \Delta_t)$, thus we have:

$$\hat{\mathbf{z}} \sim \mathcal{N}(\mathbf{h}(\mathbf{f}(\mathbf{x}, \Delta_t)), \mathbf{R}). \tag{4}$$

Please note that by utilizing the target motion model the uncertainty in the motion model (described by $\mathbf{x}$) is considered as an additional random error, which is relevant for large Δ_t.

- Sensor registration error: The measurement equation is not only dependent on the target, but also on the sensor state (e.g. position $\mathbf{r}$). To model this error, the sensor state has to be added as a parameter of the measurement equation, i.e.

$$\mathbf{z} = \mathbf{h}_i(\mathbf{x}, \mathbf{r}^{(i)}). \tag{5}$$

The model is expanded by a systematic error (Δ_r) and/or a random error (modeled by covariance $\mathbf{S}$). Thus, the relation between the true sensor state $\mathbf{r}$ and estimated sensor state $\hat{\mathbf{r}}$ is modeled by

$$\mathbf{r} \sim \mathcal{N}(\hat{\mathbf{r}} + \Delta_r, \mathbf{S}). \tag{6}$$

From this we obtain:

$$\hat{\mathbf{z}} \sim \mathcal{N}(\mathbf{h}(\mathbf{x}, \hat{\mathbf{r}} + \Delta_r), \mathbf{R} + \hat{\mathbf{S}}_h), \tag{7}$$

where $\hat{\mathbf{S}}_h$ models the additional random error, which is introduced by $\mathbf{S}$, see e.g.[3] Sec. 4 for techniques to approximate this term.

In practice, there are multiple bias components influencing the estimation procedure. Thus, the estimation of different components is coupled. Robust estimation of different bias terms is an important practical task, which is discussed e.g. in[12], where the importance of the choice of an appropriate coordinate system for bias compensation is stressed.

3. CONTINUITY CHECK

Let $\mathcal{Z}^{1:k} = \{\mathcal{Z}_1, \dots, \mathcal{Z}_k\}$ be the collection of k measurements from the sensor under test. By application of standard tracking and data association techniques (see e.g.[13]), the measurements are associated with tracks. Let $\mathcal{F}$ be the collection of N tracks from the sensor under test, with $\mathcal{F} = \{T_1, \dots, T_N\}$ and the track $T_i = \{\mathbf{z}_1^{(i)}, \dots, \mathbf{z}_{n_i}^{(i)}\}$ be characterized by the n_i measurements associated to it. Since, per assumption, a measurement is associated to at most one track, the collection of measurements can be divided into disjoint sets, i.e.: $\mathcal{Z}^{1:k} = \{T_1, \dots, T_N, \bar{T}\}$, where $\bar{T}$ is the set of measurements that are not associated with any track.

To score the sensor under test we calculate the likelihood ratio (LR) $p(\mathcal{F}|h_1)/p(\mathcal{F}|h_0)$, where h_1 is the hypothesis that there are measurements that can be associated to targets and h_0 that there are none.

$$p(\mathcal{F}|h_1) = \int_X p(\mathcal{F}|X, h_1)p(X|h_1)dX \tag{8}$$

$$= \sum_{n=1}^{N} \int \cdots \int p(\mathcal{F}|(n, (\mathbf{x}_1, \dots, \mathbf{x}_n)))p((n, (\mathbf{x}_1, \dots, \mathbf{x}_n)))d\mathbf{x}_1 \dots d\mathbf{x}_n, \tag{9}$$

where $X = (n, (\mathbf{x}_1, \dots, \mathbf{x}_n))$ describes the number and states of n hypothetical targets. Note that, for h_1, we need not consider the case with zero targets ($n = 0$) nor undetected targets ($n > N$).

We define the association vector α_n with dimension n by $\alpha_n(k) = \ell$ if target at $\mathbf{x}_k$ is associated to T_ℓ. Thus,

$$p(\mathcal{F}|X) = \sum_{\alpha_n} p(\mathcal{F}, \alpha_n|X) = \sum_{\alpha_n} p(\alpha_n) \prod_{\ell=1}^{n} p(T_{\alpha_n(\ell)}|\mathbf{x}_\ell)f_C(\mathcal{F} \setminus \{T_{\alpha_n(\ell)}\}_{\ell=1}^{n}), \tag{10}$$

where $f_C(\mathcal{G}) = p(\mathcal{G}|h_0)$ denotes the probability that all measurements in $\mathcal{G}$ are false alarms. Assuming that f_C is Poisson distributed (parametrized by the false alarm intensity ρ_F) and each association α_n to be a-priori equally likely we have[14]:

$$\frac{p(\mathcal{F}, \alpha_n|X)}{f_C(\mathcal{F})} = \prod_{\ell=1}^{n} LR(T_{\alpha_n(\ell)}|\mathbf{x}_\ell), \tag{11}$$

where $LR(T_i|\mathbf{x}_k)$ is the likelihood ratio with respect to the measurements contained in track T_i. Following,[15] it is sequentially calculated by multiplication of the factor

$$\begin{cases} \frac{P_D(\mathbf{x})}{\rho_F(\mathbf{z})}p(\mathbf{z}|\mathbf{x}) & \text{in case of a detection of target } \mathbf{x} \text{ with measurement } \mathbf{z} \\ 1 - P_D(\mathbf{x}) & \text{in case of a miss-detection.} \end{cases} \tag{12}$$

where

- $P_D(\mathbf{x})$ is the probability of detection of the sensor for a target with state $\mathbf{x}$;

- $\rho_F(\mathbf{z})$ is the false alarm intensity value for a measurement $\mathbf{z}$ and

- $p(\mathbf{z}|\mathbf{x})$ is the single-detection likelihood function, which is here described by $\mathcal{N}(\mathbf{z}, \mathbf{h}(\mathbf{x}), \mathbf{R})$, see Sec. 2.

We further assume that X is the realization of a poisson point process, see e.g.,[16] i.e.

$$p(n, (\mathbf{x}_1, \ldots, \mathbf{x}_n)) = \frac{1}{n!}\exp(-\mu)\prod_{\ell=1}^{n}\lambda(\mathbf{x}_\ell), \tag{13}$$

with intensity λ and $\mu = \int \lambda(s)ds$, thus

$$\frac{p(\mathcal{F}|h_1)}{f_C(\mathcal{F})} = \exp(-\mu)\sum_{n=1}^{N}\frac{1}{n!}\sum_{\alpha_n}\prod_{\ell=1}^{n}\int LR(T_{\alpha_n(\ell)}|\mathbf{x}_\ell)\lambda(\mathbf{x}_\ell)d\mathbf{x}_\ell \tag{14}$$

We define $LR(T_s) = \int LR(T_s|\mathbf{x}_\ell)\lambda(\mathbf{x}_\ell)d\mathbf{x}_\ell$, which is the LR calculated sequentially over time by exploiting additionally an appropriate target motion model.[15] Since, due to the integration, any permutation of α_n delivers the same scoring, we define the association vector β_n that ignores associations which are equal except permutation. Thus,

$$\frac{p(\mathcal{F}|h_1)}{f_C(\mathcal{F})} \propto \sum_{n=1}^{N}\sum_{\beta_n}\prod_{\ell=1}^{n}LR(T_{\beta_n(\ell)}) \tag{15}$$

The continuity is checked by comparing the above, sequentially updated, likelihood ratio with a threshold. The continuity check aims to the extraction of features of a single sensor. This can be tracks of moving targets (classical tracking approach), but also stationary features (e.g. clutter) can be interesting for further evaluation. Both features are called tracks in this paper.

4. PLAUSIBILITY CHECK

As shown in the previous section, to calculate the LR the sensor model has to be defined by specifying $P_D(\mathbf{x})$, $\rho_F(\mathbf{z})$ and $p(\mathbf{z}|\mathbf{x})$.

Thus, for scoring the sensor we already use the prior knowledge about the sensor itself. This means that the test includes a check of the available prior information. Since in the first stage we only want to verify that the sensor contains any useful information, it is recommended to choose only general values describing the expected performance of a sensor of this type, i.e. using fixed P_D and ρ_F. A fixed P_D of 0.5 is e.g. in general adequate to extract tracks of a single sensor in those regions that are illuminated by the sensor. In the plausibility stage we can specify this prior knowledge either due to specifications of the sensor manufacturer or by external context information e.g. from maps. A geographic map is shown in Fig. 1 as example. The depth of the water is used to model the residence probability of ships, which in turn has influence on the model of P_D for maritime targets. Using the new model for calculating the LR will produce high values for tracks at sea and low values for tracks at land. Thereby a unreliable sensor can be detected.
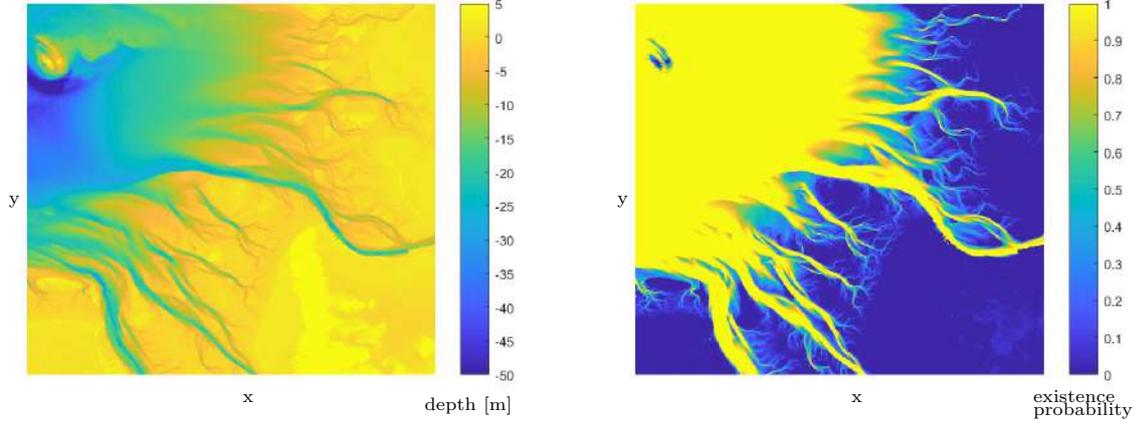
Figure 1: Extraction of external knowledge from map information. The map on the left hand side is a composition of landmass and bathymetry information (extracted from data provided by the Service Centre of the Federal Government for Geo-Information and Geodesy `http://www.geodatenzentrum.de` [2018] and GeoSeaPortal of the "Bundesamt fuer Seeschifffahrt und Hydrographie" (BSH) `http://www.geoseaportal.de`[2018]). The map on the right hand side (extracted from the previous one) describes the area where observation of ship traffic is possible (positive values).

5. CONSISTENCY CHECK

Let $\mathcal{G}$ be the collection of tracks from an additional sensor (or even multi-sensor tracks), with $\mathcal{G} = \{S_1, \ldots, S_L\}$.

5.1 Goodness of fit

In analogy to Sec. 3 we can calculate:

$$p(\mathcal{F}|\mathcal{G}) = \int_X p(\mathcal{F}, X|\mathcal{G})dX = \int_X p(\mathcal{F}|X)p(X|\mathcal{G})dX \tag{16}$$

The posterior density function $p(X|\mathcal{G})$ is thereby defined by the tracks state and covariance. Thus, for scoring the sensor output we calculate the likelihood of the sensor observations given the track input from additional sensors. However, this is only applicable if the tracks of the additional sensors adequately describe the full scenario.

5.2 Correlation coefficient

As an alternative measure we can calculate the correlation coefficient according to:

$$\frac{p(\mathcal{F}, \mathcal{G}|h_1)}{p(\mathcal{F}|h_1)p(\mathcal{G}|h_1)} = \frac{\int_X p(\mathcal{F}, \mathcal{G}|X, h_1)p(X|h_1)dX}{p(\mathcal{F}|h_1)p(\mathcal{G}|h_1)}. \tag{17}$$

For $X = (n, (\mathbf{x}_1, \ldots, \mathbf{x}_n))$ we define the association variable β_n by a $2 \times n$ matrix, where $\beta(: k) = (\ell, m)$ means for $\ell > 0$ and/or $m > 0$ target k to be associated with track T_ℓ from sensor A and/or track S_m from sensor B. $\ell = 0$ and/or $m = 0$ represent a miss-detection from the respective sensor. Analogous to Sec. 3 we have:

$$p(\mathcal{F}, \mathcal{G}|X) = \sum_{\beta_n} p(\mathcal{F}, \mathcal{G}, \beta_n|X). \tag{18}$$

Since we assume false tracks from different sensors to be uncorrelated, i.e. $p(\mathcal{F}, \mathcal{G}|h_0) = p(\mathcal{F}|h_0)p(\mathcal{G}|h_0)$, we can expand the fraction in (17) by $p(\mathcal{F}|h_0)p(\mathcal{G}|h_0)$ and use for calculation the LR according to:

$$\frac{p(\mathcal{F}, \mathcal{G}|h_1)}{p(\mathcal{F}, \mathcal{G}|h_0)} = \sum_{n=1}^{N+L} p(n) \sum_{\beta_n} \prod_{k=1}^{n} LR(T_{\beta_n(1,k)}, S_{\beta_n(2,k)}). \tag{19}$$

6. VERIFICATION SCHEME

The described verification scheme is embedded in the sensor data fusion processing, as depicted in Fig. 2. This relies on a distributed tracking strategy. The continuity and plausibility checks are applied to single sensor tracks, whilst the consistency check is embedded in the correlation stage of tracks from different sensors.
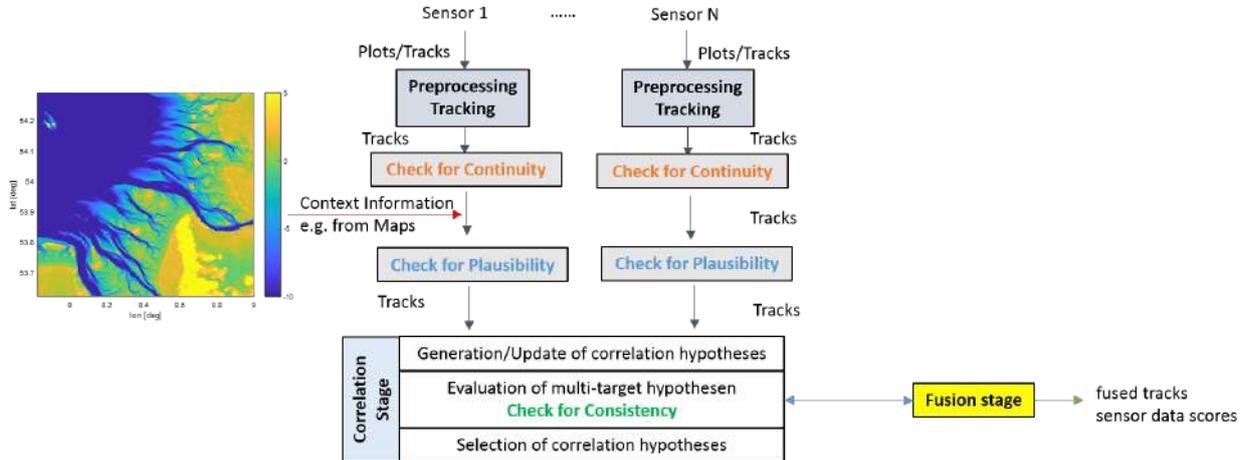

Figure 2: Sensor verification scheme on different levels of the data fusion architecture

7. COMPENSATION OF SYSTEMATIC SENSOR ERRORS

The verification scheme checks if a sensor delivers (as expected) a contribution to a surveillance picture. If a deviation is detected the sensor should be either excluded from the fusion process or if possible corrected. As motivated in the introduction, the compensation of systematic sensor errors is discussed in detail in literature. Important issues are estimation of errors from different sensors (see e.g.[6]) and addressing the coupling of different errors see.[12] In this paper for illustration estimation of a systematic heading offset is addressed, see Sec. 2. The offset is calculated first individually for each possible correlation hypothesis (see Fig. 2) and from this the global sensor offset is calculated during the multi-target optimization step.

8. SIMULATION RESULTS

Fig. 3 shows a simulation scenario with three maritime targets (ships) and an additional slow moving small target (e.g. drifting container). AIS, radars and cameras sensors are considered. Specifically, the AIS provides positional information about the ship, but no information about the container. The two radar sensors are characterized by a high probability of detection $P_D = 0.8$ for ships and a lower $P_D = 0.4$ for the container. The two cameras operates with a $P_D = 0.5$ for ships and a lower $P_D = 0.3$ for the container. For one camera sensor a heading offset of 10° was simulated.

Radar sensors generate single sensor tracks in Cartesian coordinates while cameras provide tracks in polar coordinates. All sensors pass the continuity and plausibility checks. In Fig. 4(a) tracking results of a standard centralized multi hypothesis tracker (MHT) are shown for the container target. In this case, tracks from one radar and from both cameras represent the inputs for the the multi sensor fusion stage. Fig. 5 shows the sensor consistency score (red line) of the camera under test. The score has been calculated through the correlation coefficient by exploiting AIS and camera sensor data. The blue line shows as reference the sensor score for a correct heading. In Fig. 4(b) the tracking results of the centralized MHT are shown after compensation of the heading offset. Localization is obviously improved. The tracking results are shown here for central MHT, while the scoring of the sensor uses a distributed strategy as displayed in Fig 2. For 10 Monte Carlo runs, the probability for track detection $P_{D,t}$ has been calculated for both approaches. For the centralized strategy, we obtain $P_{D,t} = 0.6$ while for the distributed strategy $P_{D,t} = 0.3$. This demonstrates the importance of both
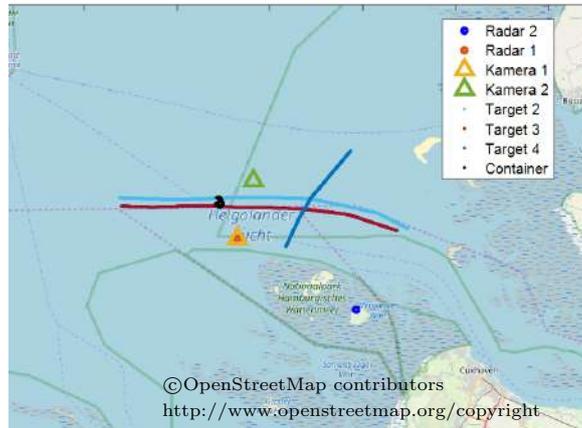
Figure 3: Simulation scenario with three ships and a floating container
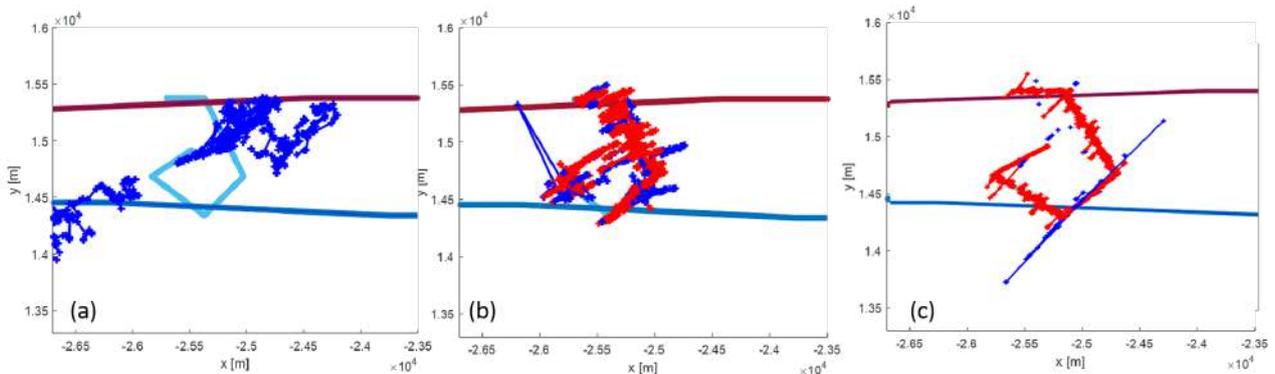
Figure 4: Simulation results: Tracking results are blue and red '+'. The color indicates if the track has been associated (red) with the container in a subsequent evaluation step. (a) no Bias Compensation, only one radar sensor used, (b) Bias Compensation, only one radar sensor used, (c) no Bias Compensation, both radar sensors used.

strategies: the centralized tracking in order to achieve good performance for corrected data and a distributed tracking strategy to verify and correct sensor data. In Fig. 4(c) results of the centralized MHT are shown, when using both radar sensors, where $P_{D,t}$ increases to 0.8.

9. CONCLUSION

The paper discusses concepts for the verification of sensor data on different levels of a multi-sensor data fusion architecture. The method is based on sequential likelihood ratio calculations by which different measures of data integrity can be checked during tracking and data fusion. First evaluations with simulated data are presented for the detection and tracking of a weak target, including misalignment detection and correction. Future work will focus on demonstrating the approach in real maritime surveillance scenarios.

ACKNOWLEDGMENTS
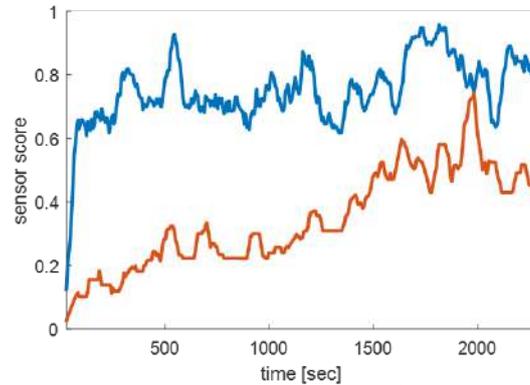
Figure 5: Sensor score with (blue) and without (red) misalignment correction

REFERENCES

[1] Dana, M. P., "Registration: A prerequisite for multisensor tracking," in [*Multitarget-Multisensor Tracking: Advanced Applications*], Bar-Shalom, Y., ed., ch. 5, Norwood, MA: Artech House (1990).

[2] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., [*Estimation with Application to Tracking and Navigation*], Wiley-Interscience (2001).

[3] Daun, M. and Ehlers, F., "Tracking algorithms for multistatic sonar systems," *EURASIP Journal on Advances in Signal Processing* **2010** (2010).

[4] Song, X., Willett, P., and Zhou, S., "Range bias modeling for hyperbolic frequency modulated waveforms in target tracking," in [*Sensor Array and Multichannel Signal Processing Workshop (SAM), 2012 IEEE 7th*], 249–252 (2012).

[5] Crouse, D. F., Bar-Shalom, Y., and Willett, P., "Sensor bias estimation in the presence of data association uncertainty," in [*Proceedings of the SPIE*], **7445** (2009).

[6] Taghavi, E., Tharmarasa, R., Kirubarajan, T., Bar-Shalom, Y., and Mcdonald, M., "A practical bias estimation algorithm for multisensor-multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems* **52**, 2–19 (February 2016).

[7] Papageorgiou, D. J. and Holender, M., "Track-to-track association and ambiguity management in the presence of sensor bias," in [*2009 12th International Conference on Information Fusion*], 2012–2019 (July 2009).

[8] Jonsdottir, I. and Hauksdottir, A., "Integrity monitoring and estimation of systematic errors in radar data systems," in [*Radar Conference, 1995., Record of the IEEE 1995 International*], 310–316 (1995).

[9] Bar-Shalom, Y., "Airborne GMTI radar position bias estimation using static-rotator targets of opportunity," *Aerospace and Electronic Systems, IEEE Transactions on* **37**(2), 695–699 (2001).

[10] Blackman, S. S. and Bah, N. D., "Track association using correction for bias and missing data," in [*Proceedings of the SPIE*], **529** (1994).

[11] Ong, H.-T., Ristic, B., and Oxenham, M., "Sensor registration using airlanes," in [*Information Fusion, 2002. Proceedings of the Fifth International Conference on*], **1**, 354–360 vol.1 (2002).

[12] Zhu, H. and Han, S., "Track-to-track association based on structural similarity in the presence of sensor biases," *J. Applied Mathematics* **2014**, 294657:1–294657:8 (2014).

[13] Blackman, S., [*Multiple-Target Tracking with Radar Applications*], Artech House, Inc. (1986).

[14] Reid, D., "An algorithm for tracking multiple targets," *IEEE Journal AC* **24**, 843–854 (Dec 1979).

[15] Van Keuk, G., "Sequential track extraction," *IEEE Journal AES* **34**, 1135–1148 (Oct. 1998).

[16] Streit, R. L., [*Poisson Point Processes: Imaging, Tracking, and Sensing*], Artech House (2010).

Aggregated risk assessment from multi-source data fusion

Filippo Daffinà[a], Torbjorn Stahl[a], Dino Quattrociocchi[a], Massimo Zavagli[a], Simon Chesworth[b],
Roberta Migliorini[c], Guy Sear[c]

[a]e-Geos S.p.a., via Tiburtina 965, Roma, Italy [b]exactEarth Europe Ltd., Harwell, Didcot,
Oxfordshire, OX11 0QR, UK [c]IHS Markit Maritime & Trade, 25 Ropemaker Street, London EC2Y
9LY, UK

ABSTRACT

The increasing maritime trade and the considerable availability of huge amounts of maritime data, gathered by multiple sensors and sources, provides new approaches to allow efficient monitoring of the maritime traffic for different application domains such as security and intelligence, environmental protection, marine resource management, market analysis and emergency response. Gathering and analyzing these large volumes of maritime data is challenging because of frequent changes in regulations, security conditions, ship features, crew, operators and registration. In this paper, a new approach is proposed for the early identification of maritime threats and assessing the risk of each vessel, by analyzing and fusing dynamic data (e.g. anomalous behavior, visited ports and geographical areas), collected from multi-sensor satellite images and AIS messages, with static information (e.g. the vessel type, the presence of the vessel's IMO/Company Name/Manager Name in black- or sanction lists). The process is assigning customizable weights on each identified risk type in order to adapt the results to the application in hand. The evaluation of the aggregated risk assessment allows for an optimization of allocated resources by allowing attention to be focused on the vessels reporting a high risk profile.

Keywords: Data fusion, risk assessment, remote sensing, AIS, vessel detection, anomaly detection

1. INTRODUCTION

According to the latest UNCTAD and Eurostat reports, the global seaborne trade is expanding quickly and progressively, currently confirming the fastest growth in the last five years and the total volumes reaching 10.7 billion tons (Figure 1). The vast majority of global trade travels through the open ocean, which is a relatively lawless zone with very limited possibility (in terms of costs and efficiency) to be systematically monitored with traditional assets. Over 184,000 ships exceeding 100 gross tons navigate the seas between more than 10,000 ports worldwide, carrying more than 1 million seafarers and 20 million shipping containers - not counting vessels for fishing, research, recreation, military and other purposes.
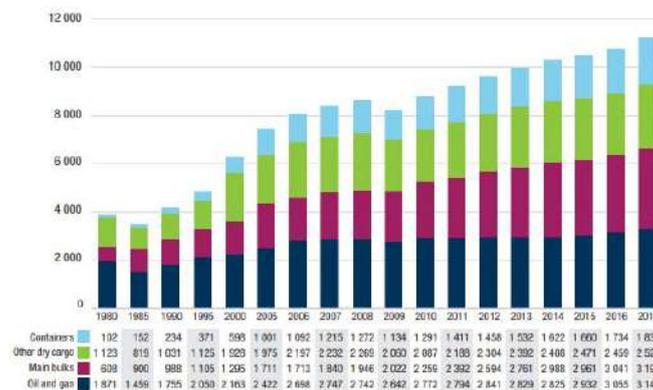

Figure 1. International seaborne trade volumes per cargo type, reported in millions of tons, over the last years.

At the same time, new sensors and information sources have been introduced for improving the Maritime Picture and the Maritime Situational Awareness (MSA) activities. In the space domain, the introduction of new constellations of satellites, with different size and performances, allows systematic monitoring of any area of interest around the world, thus allowing for the detection and monitoring of vessels and their behavior, in all sea and weather conditions, during the day and night:

- There are new constellations of satellites that are able to gather AIS messages and quickly forward the information to the ground stations using satellite-to-satellite communication, thereby reducing the latency in the messages delivery to less than 1 minute.

- In the context of Earth Observation (EO), new or extended satellite constellations provide both free and open data (e.g. the Sentinel satellites managed and operated in the context of the European space program Copernicus) and commercial on-demand data (e.g. COSMO-SkyMed and COSMO-SkyMed Second Generation, SAOCOM, ICEYE, BlackSky, Planet). These new space assets, equipped with SAR and/or optical sensors, together with the new processing technologies, are ready to: (i) be federated for granting systematic monitoring of global areas, (ii) better contextualize the detected anomalous events (e.g. the triggering of fast tasking of VHR optical acquisitions as a result of the anomaly detection process from systematic SAR monitoring), (iii) reduce the reaction time and analysts'/operator's allocated effort.

The vessel identification information coming from "cooperative systems" (such as AIS, VMS or LRIT) is also used for the identification and tracking of vessels detected in the satellite images and to highlight the presence of potentially dark vessels.

The MSA activities can be aimed at many different applications such as security and intelligence, environmental protection, marine resource management, market analysis and emergency response. Government agencies track vessels, their movements and related entities, searching for anomalous behavior, whether for strictly trade-related reasons or for law and regulatory enforcement, anti-smuggling or defense initiatives.

One approach for identifying which vessels to track, is to make a risk assessment and assign a risk value to each vessel around the world. In the maritime environment there are a number of different risk factors that can be considered and used to compute an aggregated risk value. It is possible to customize the aggregation of the different factors depending on the aim of the analysis (e.g. giving more relevance to the security or to the environmental aspects). If the aggregated risk value supersedes a certain threshold, the vessel is of interest for further ad-hoc analysis and/or monitoring.

The risk factors can be divided into two groups: **dynamic** and **static** risk factors. Possible dynamic risk factors to be assessed and evaluated with different weights can be related to:
- anomalous behavior, e.g. engagement in meeting at sea, switching off the AIS transponder or loitering,
- current and historical routes,
- list of visited ports,
- crossed national waters, Exclusive Economic Zones (EEZ) or critical geographical areas.

Using the vessels' IMO code, there are various static risk factors that can be calculated, based on for example:
- vessel and/or owner is present on a black list (e.g. Paris MoU, IUU vessel list or UN Security Council Blacklists),
- vessel has been involved in previous detentions.

2. RISK ASSESSMENT APPROACH

The increased availability of data from heterogeneous sources (satellite and in situ) and the related big data technologies are creating information superiority for various maritime applications. In particular the availability of new satellite constellations enables new approaches for enriching the maritime picture; the efficient federation of the different space assets permits to extend the continuous monitoring above all over remote areas of interest where it is difficult (or too expensive) to maintain a persistent monitoring by traditional assets only.

In the presented work, the European space program Copernicus is used to enable sustainable continuous monitoring of global sea thanks to the free and open satellite images acquired by multiple sensors (*Sentinel-1* and *Sentinel-2*), while exactEarth's new Real-Time Satellite AIS Constellation, *exactView™ RT*, is used to acquire and access the global AIS messages (Figure 2). In case there is a need for ad-hoc and on-demand acquisitions over the monitored areas (e.g. in case of detected anomalies), the commercial EO satellites (e.g. COSMO-SkyMed) can be tasked for providing better knowledge of the detected event, with the possibility to acquire VHR images with pixel resolution < 1m.
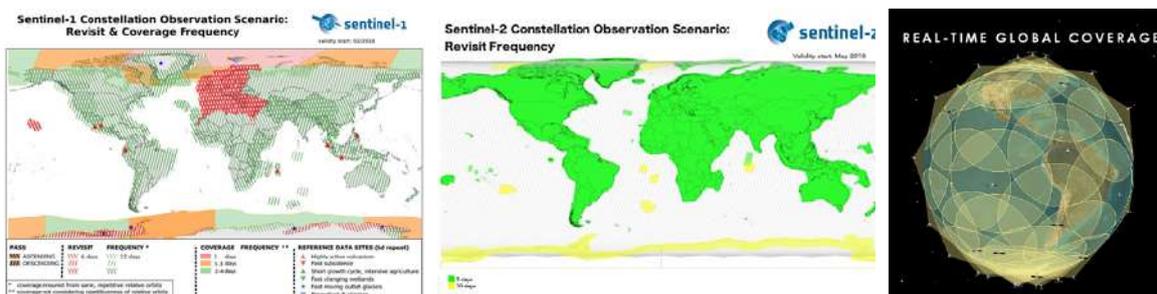
Figure 2. Examples of current coverage of Earth Observation satellites (Sentinel-1 on the left, Sentinel-2 in the middle - source ESA) and Satellite AIS (on the right, source exactEarth)

The continuous monitoring of the seas by collecting images, acquired by space-borne SAR and optical sensors, and by real-time global AIS, is one of the most complete approaches for the systematic monitoring of the sea allowing the automatic detection of the vessels and the main maritime features (e.g. vessels, offshore platforms, wind farms) and the extraction of maritime patterns of life (see Figure *3. Example of systematic monitoring over the Mediterranean sea by fusing Sentinel-1 and AIS data for detecting vessels and extracting maritime patterns of life.*Figure 3).
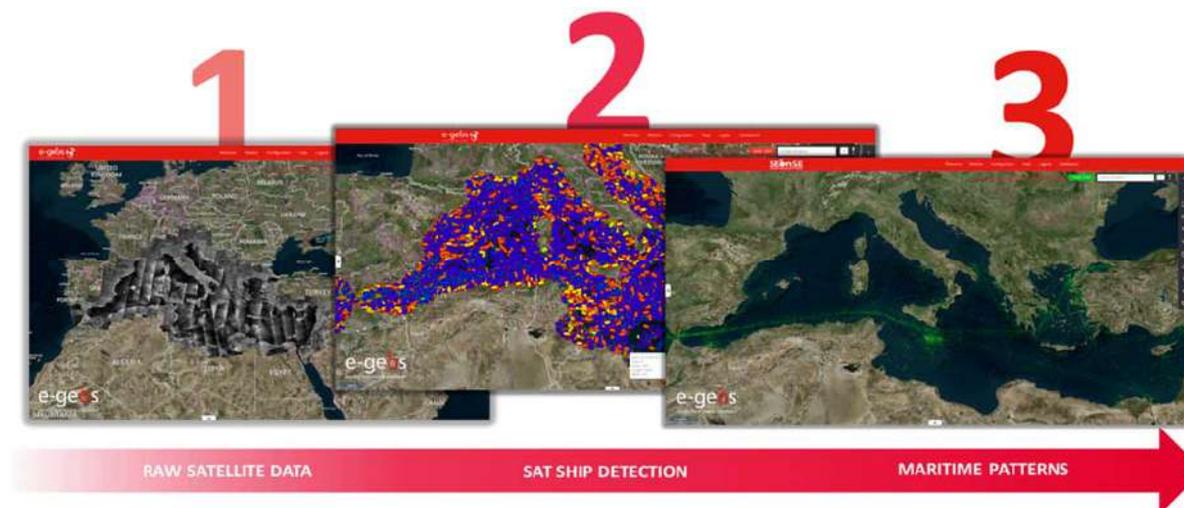
Figure 3. Example of systematic monitoring over the Mediterranean sea by fusing Sentinel-1 and AIS data for detecting vessels and extracting maritime patterns of life.

The collected datasets are subject to an anomaly detection process (see Figure 4) and, in case of positive matching of the defined anomalous behaviour rules and thresholds, the automatic workflow supports the decision making process by:

- automatically generating a satellite acquisition plan over the areas of the detected anomaly taking into account the closest (in time and space) satellite passage, required image resolution (e.g. considering the dimensions of involved vessels) and weather conditions (for optical satellites);
- forwarding an automatic notification to the team of experts/operators for focusing their attention on the detected event;
- triggering a mission planning activity for requiring the tracking of the detected event by the Authorities' assets (e.g. boats, airplanes, UAV).
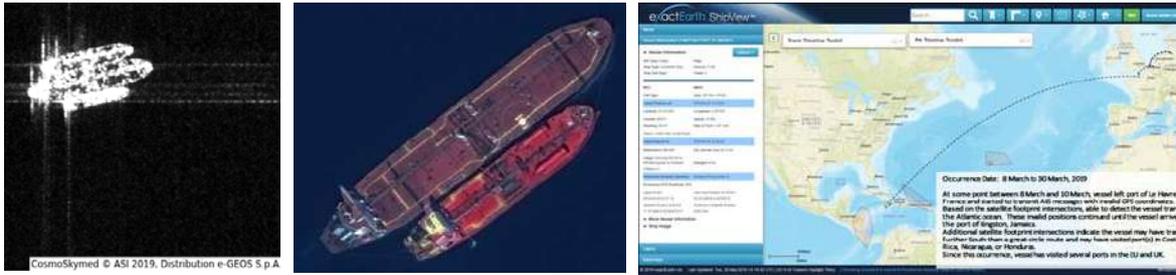
Figure 4. Examples of anomaly detection: meetings at sea as detected by SAR satellite (on the left) and optical satellite (in the middle), AIS position spoofing as detected by satellites (on the right).

The data fusion between satellite detected vessels and AIS messages allows to classify all the vessels of interest as **cooperative** (detected by satellite imagery and identified by the correlation with AIS, offering the possibility to analyse also historical tracks and gaps of AIS transmissions) or **dark** (detected only by satellite imagery).

The correlation between vessels detected by EO data analysis and AIS data is based on a global optimal algorithm that minimizes the error in the linking between AIS signals and detected vessels, taking into account all the AIS data and detected vessels in the satellite scene and in the relevant temporal range (across the acquisition time).

The data fusion algorithm consists of the following steps:

1. Retrieve AIS ship positions and velocities in a specified area and at times of the satellite image acquisition.

2. Transform AIS ship positions and velocities in image coordinates in order to compare them with positions and velocities derived from satellite images.

3. Interpolate AIS positions and velocities at the time of the satellite image acquisition by using cubic spline interpolation.

4. Global matching of positions, velocities and dimensions between AIS and satellite-based vessel detections. The matching is based on the Maximum Matching Bipartite Algorithm (based on network flow optimization).

5. Assign an assessment on the matching error e_h to each match h ($h=1, ..., N$) with the formula:

$$\epsilon_h = \frac{w_{\delta p}\delta p + w_{\delta v}\delta v + w_{\delta s}\delta s}{w_{\delta p} + w_{\delta v} + w_{\delta s}}$$

where δp, δv and δs are the percentage differences (percentage relative errors; e.g. $\delta p = \frac{|\bar{p}_{AIS} - \bar{p}_{EO}|}{|\bar{p}_{EO}|}$) respectively for the positions, velocities and sizes between AIS and EO derived vessel parameters; $w_{\delta p}$, $w_{\delta v}$ and $w_{\delta s}$ are weights used respectively for δp, δv and δs.

6. Assign an empirical reliability index r_h to each match h (h=1, …, N) ranging from 0 (low confidence) to 1 (high confidence):

$$r_h = e^{-k\epsilon}$$

where k is opportunely calibrated in order to have a reasonable dynamic of the confidence behavior.

The correlation algorithm has been extensively tested on different cases and its behaviour is completely in agreement with the aim of the objective function of the global optimal algorithm.

The **dynamic** component of the risk is based on an assessment of the provided information in terms of:
- analysis of suspicious behaviour of each vessel, based on the set of defined rules, such as rendezvous at sea (Figure 4), changing of identity during the voyage (e.g. name, IMO) or gaps of AIS transmission,
- comparing the provided information in the AIS message (e.g. position, dimension) w.r.t parameters estimated from satellite images,
- identifying position spoofing, comparing the transmitted position w.r.t. the receiving satellite's position, Figure 4,
- analysis of the historical route of each vessel, e.g. presence in ports or national waters of countries with a high risk factor (using e.g. JWC Oceanic Zones, JWC Territorial Zones and Sanctioned Countries as sources).

Threats at sea are pervasive, from geopolitical crises and global money laundering to cargo theft, piracy and crew kidnappings. These risks pose a huge challenge to marine insurers, finance institutions, banks, governments and any business with global supply chains. A complete profiling of the ship security risks can be obtained only extending the dynamic risk assessment with the integration of maritime and risk event databases to provide a single maritime risk mitigation solution. Typical information sources for assessing this **static** risk component are:

(i) past casualties, piracy and pollution events,
(ii) vessels/companies in blacklists (e.g. Office of Foreign Assets Control - OFAC, sanctioned countries layer, Paris MoU),
(iii) number (and frequency) of changes of the ship name, flag or owner. In the proposed work, the global vessel database from the IHS Markit *Maritime Intelligence Risk Suite* is used for assessing the static component of the risk, Figure 5 gives some examples of what information can be extracted and fused in the final estimation of the risk.
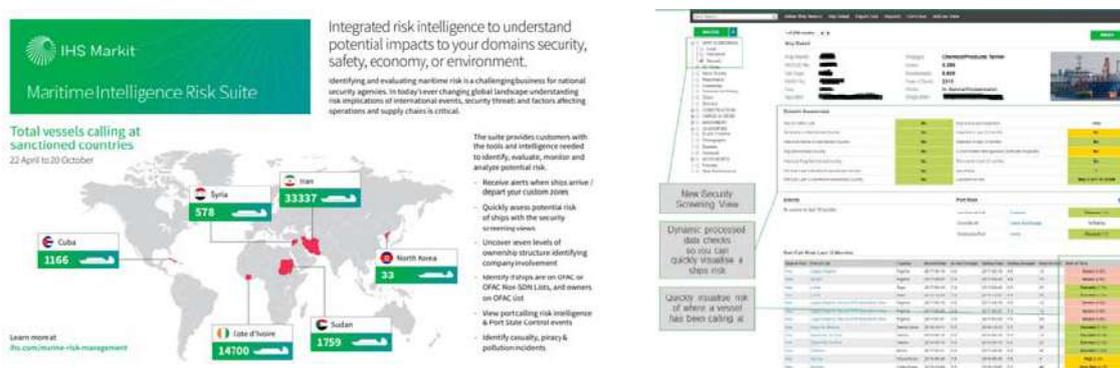
Figure 5. Information for assessing this static risk component

The dynamic and static risk factors are then combined, using a specialized weight assigned to each risk component, customized according to the application at hand:

$$r = \frac{\sum_{i=1}^{n} w_i r_i}{\sum_{i=1}^{n} w_i}$$

where r is the combined risk factor for a vessel, w_i is the weight assigned to a certain risk i and r_i is the binary value if that risk is present or not for the vessel at hand.

The fusion of the dynamic and static risk components, managing the weight of each single parameter for specializing the analysis to a specific domain (e.g. security and intelligence, marine resource management, environmental protection), makes it possible to quickly adapt the algorithm to the different users' interest and use cases (e.g. oil smuggling, illegal fishing or pollution). The methodology of combining the dynamic and static risk factors is promising and the results are under validation process. Figure 6 contains an example of the calculation of each vessel's risk and illustrates how the analysis of maritime traffic and the identification of potential maritime threats is improved by reducing the number of vessels to track and focusing the attention on those reporting a medium and high risk profile.
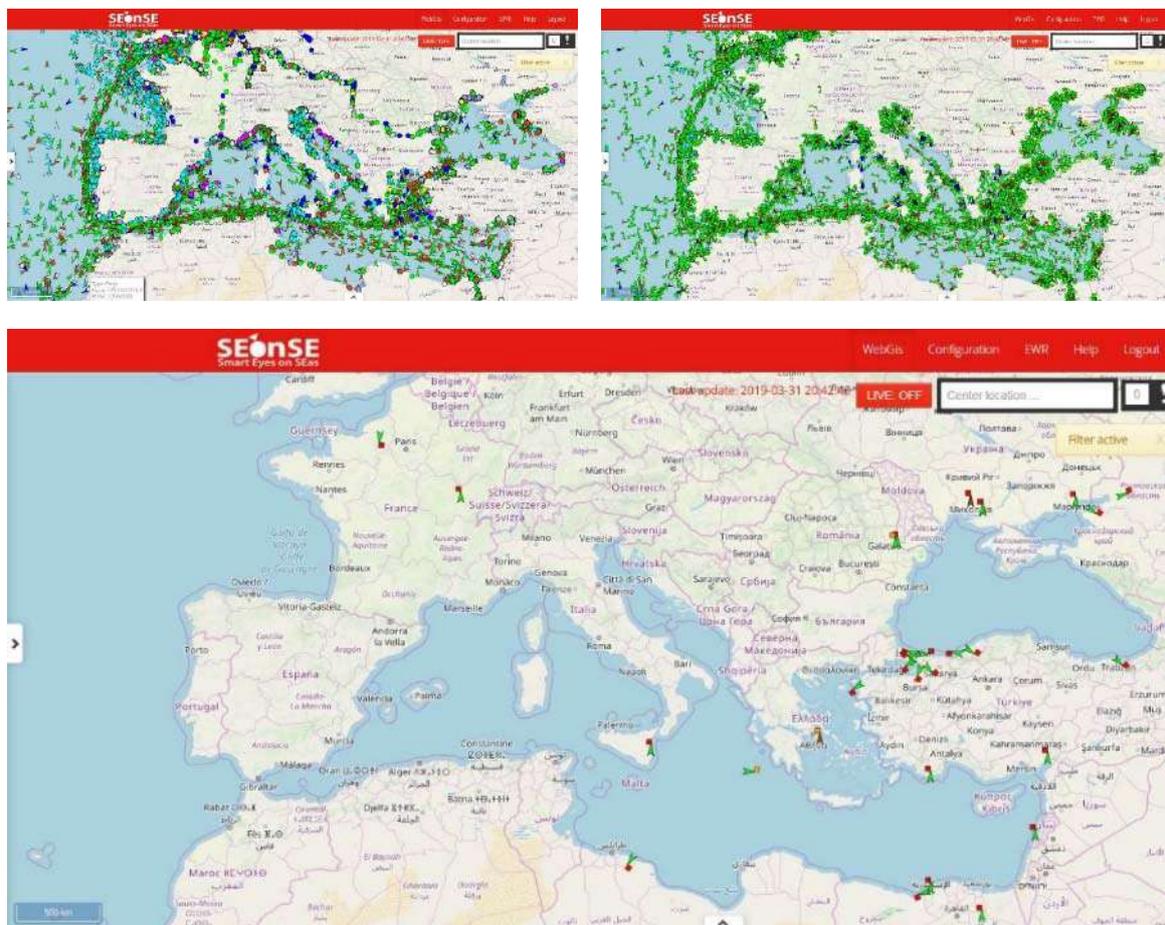
Figure 6. Example of maritime traffic monitoring by analyzing the latest position by AIS in one day (on the top-left), the corresponding risk assessment (on the top-right) and focus only on vessels reporting medium-high risk profile (bottom).

3. CONCLUSIONS

The proposed risk assessment methodology is based on the fusion of information coming from multiple and heterogeneous sources (including multi-sensor space assets), and consider the characterization of both static and dynamic factors of the risk, allowing the definition of various risk profiles by the customization and application of different weights for each risk factor according to the user needs. The evaluation of the aggregated risk assessment allows for an optimization of resources allocated on MSA activities by reducing the number of vessels of interest, focusing the attention of the analysts and

operators no longer on the entire maritime traffic but only on the vessels reporting a high risk value. The described approach has been finalized and integrated in SEonSE [9] to be validated in the context of OCEAN2020 [18].

REFERENCES

[1] United Nations Conference on Trade and Development (UNCTAD), "Review of Maritime Transport 2018", <https://unctad.org/en/PublicationsLibrary/rmt2018_en.pdf> (2018).

[2] Eurostat DS-022469, "International trade in goods by mode of transport", <https://ec.europa.eu/eurostat/statistics-explained/index.php/International_trade_in_goods_by_mode_of_transport> (2017).

[3] Moyseenko, S.S., Meyler, L.E. and Bondarev, A., "Risk Assessment for Fishing Vessels at Fishing Grounds", TransNav International Journal on Marine Navigation and Safety of Sea Transportation vol. 9 nr. 3, 351-355 (2015).

[4] Kristiansen, S. [Maritime Transportation: Safety Management and Risk Analysis], Elsevier (2010).

[5] Lane, R. O., Nevell, D. A., Hayward, S. D. and Beaney, T. W., "Maritime anomaly detection and threat assessment," 13th International Conference on Information Fusion, Edinburgh, 2010, 1-8. (2010).

[6] Agenzia Spaziale Italiana, "COSMO-SkyMed mission and products description", <https://archives.asi.it/sites/default/files/attach/bandi/cosmo-skymed_mission_and_products_description_update_2_1.pdf> (2016).

[7] ESA, "Sentinel-1 Observation Scenario", <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario> (2018).

[8] ESA, "Sentinel-2 Revisit and Coverage", <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/revisit-coverage> (2018)

[9] e-GEOS, "SEonSE" <https://www.e-geos.it/seonse>.

[10] exactEarth, "Optimised Real-Time Satellite AIS Constellation Whitepaper", <http://main.exactearth.com/exactview-rt-whitepaper>.

[11] IHS Markit, "Maritime Intelligence Risk Suite (MIRS)", <https://ihsmarkit.com/products/Maritime-intelligence-risk-suite.html>.

[12] Galil, Z., "Efficient algorithms for finding maximum matching in graphs". ACM Computing Surveys (CSUR), 18(1), 23-38. (1986).

[13] Alessandrini, A., Alvarez, M., Greidanus, H., Gammieri, V., Fernandez Arguedas, V.; Mazzarella, F., Santamaria, C., Stasolla, M., Tarchi, D. and Vespe, M., "Mining Vessel Tracking Data for Maritime Domain Applications", IEEE 16th International Conference on Data Mining Workshops (ICDMW). (2016).

[14] Alvarez, M., Fernandez Arguedas, V., Gammieri, V., Mazzarella, F., Vespe, M., Aulicino, G. and Vollero, A., "AIS event-based knowledge discovery for Maritime Situational Awareness", 19th International Conference on Information Fusion (FUSION) (2016).

[15] Avolio, C., Costantini, M., Di Martino, G., Iodice, A., Macina, F., Ruello, G., Riccio, D. and Zavagli, M., "A method for the reduction of ship-detection false alarms due to SAR azimuth ambiguity", IEEE Geoscience and Remote Sensing Symposium (2014).

[16] Riveiro, M., Pallotta, G. and Vespe, M., "Maritime anomaly detection: A review", Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8(4):e1266, May 2018 (2018).

[17] Xiao, Z., Fu, X., Zhang, L. and Goh, R. S. M., "Traffic Pattern Mining and Forecasting Technologies in Maritime Traffic Service Networks: A Comprehensive Survey", in IEEE Transactions on Intelligent Transportation Systems (2019).

[18] OCEAN2020 project, funded from the European Union's Preparatory Action on Defence Research under grant agreement No 801697 <https://ocean2020.eu/>

Behavioral analysis for maritime safety

Espen Messel and Martin G. Bjørndal

Norwegian Defence Research Establishment, Instituttveien 20, Kjeller, Norway;

ABSTRACT

This paper contains theory, ideas and examples for anomaly detection in the coastal zone. Increased levels of ship traffic calls for automatic surveillance methods in an effort to ease the workload of a human operator overseeing movement and communication in the maritime domain. The anomaly detector compares live ship tracks to local historical data using both statistical and geographical methods. The goal of the anomaly detector is to detect anomalous behaviour related to unwanted situations such as groundings and collisions.

Keywords: Anomaly detection, behavioural analysis, maritime safety, grounding

1. INTRODUCTION

As vessel traffic in Norwegian waters is expected to increase in the near future, there is a growing need for services to help prevent unwanted situations such as collisions and groundings. This paper provides a theoretical basis for an ensemble of methods that may be included in such a system.

The maritime traffic situation is monitored by the Vessel Traffic Service (VTS) centres at all times. Increased traffic means a higher workload on the VTS operators. In order to aid the operators, we propose software to detect anomalous behaviour. The goal is to allow the operator to focus mainly on vessels that behave in an unexpected way, or otherwise different from the historical vessel data at, or close to, its location. Such vessels can be reported by the system as anomalous, and thereby further investigated by the human operator. Factors that contribute to anomaly triggering includes the following:

- Sailing in uncharted waters (especially close to land). This is assumed to be a grounding risk.

- Sailing in an unusual direction at a given location. This is assumed to be both a collision and grounding risk.

- Extreme speed values. This is assumed to be abnormal behaviour.

A combination of the above is then used for anomaly detection, along with other criterions. Anomaly detection is based on revealing behaviour that deviates from the historic ship traffic. This traffic consists of messages from Automatic Identification System (AIS) transmitted by the ships and stored in our database.

2. DATA PREPARATION

2.1 Sensor Data

Automatic Identification System (AIS) is an anti-collision system primarily intended to let vessels see, and be seen by, other marine traffic in the area. The International Maritime Organization (IMO) requires that all maritime vessels exceeding 300 gross tonnes on international travels, all vessels exceeding 500 gross tonnes and all passenger ships are to carry an AIS transceiver. AIS signals are standardized and structured, and divided into message types depending on type of information. As identification for each transceiver AIS use the Maritime Mobile Service Identity (MMSI) number. Only the position report message types are treated in this work. We may thus extract data from many years of history, providing data in the range of billions of AIS-messages. For this paper we have collected data for the last two years for an area of interest.

Further author information: (Send correspondence to E.M.)
E.M.: E-mail: espen.messel@ffi.no, Telephone: +47 63 80 76 07
M.G.B.: E-mail: martin-gjesdal.bjorndal@ffi.no, Telephone: +47 63 80 71 33
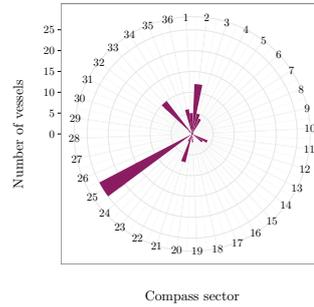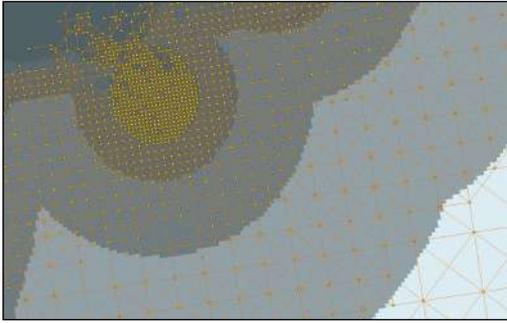
(a) The grid graph. Nodes drawn as yellow dots, edges as orange lines. The shaded areas are the buffer zones described in Section 2.3

(b) The compass rose for one vessel type in a node.

Figure 1: These images show the area and the different resolution zones for the graph.

2.2 Grid Graph

The data described in the previous section is used to build a graph. A graph G is a pair consisting of a set of nodes V and a set of edges E, so formally $G = (V, E)$. G will serve as the underlying data structure on which all functionality is based.

The nodes are created based on the positions in the AIS messages. All positions are rounded off, so that observations close together may end up with an identical position. A node u is created at every position where at least one rounded position exists. This way, observations are clustered together at their closest node. Hence, we can calculate node statistics that are relevant to the area, since all data originates from the same general location. This area is the grid cell where the node u is at the center.

A node set is created for each ship type to permit the anomaly detector to easily compare a new observation with historical data from similar ships. The ship type categories used in this preliminary analysis are "passenger ferry", "car ferry", "cargo", "tanker", "tug", "small ship" and "other". In the following, they will be referred to as $V_{pf}, V_{cf}, V_c, V_{ta}, V_{tu}, V_s$ and V_o, respectively. As more data are added we may be able to split the ship categories into more precise bins. At this point, ferries are treated rather specially. We use more detailed type information on ferries for two main reasons. The first is that ferries operate pretty much constantly, so the data output of a single ship is huge. The other is that such vessels move along roughly the same track with similar characteristics day after day. Some passenger crafts need to be separated from roll on-roll off crafts mainly due to the extreme difference in speeds. High speed passenger vessels often achieve over 20 knots, while car ferries make about 10 knots. The vertex set in G is the union of the type-specific node sets, i.e.

$$V = \bigcup_{i=1}^{7} V_i,$$

for the above described node sets numbered from 1 to 7.

An edge e_{ij} is drawn from node u_i to u_j if a ship sails directly from the area covered by u_i to the area covered by u_j. An area with nodes and edges plotted on a map is showed in Figure 1a.

As with the nodes, we create an edge set corresponding to each ship type, because some ship types travel in another fashion and along different edges than others. Correspondingly, the union of all edge sets may be expressed as

$$E = \bigcup_{i=1}^{7} E_i.$$

We have now defined the node and edge sets of $G = (V, E)$.
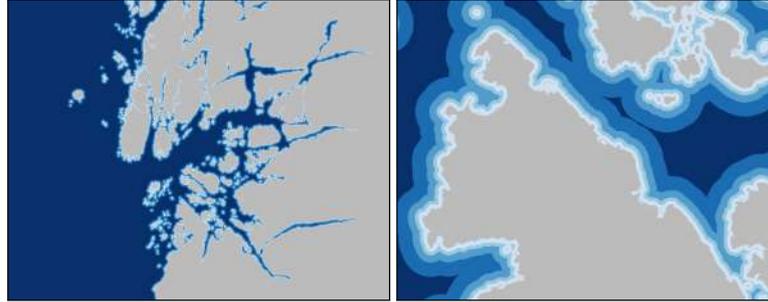
(a) The area of interest.   (b) Close up on the Stavanger area.

Figure 2: These images show the area and the different resolution zones for the graph.

2.3 Cell size

The cell size for the graph is based on Equation (1). We will use five different cell sizes for our graph, with size depending on distance from land.

$$f_g\left(n\right) = 10 \cdot 2^n|_{n=0,1,2,3,4} \tag{1}$$

We will create the graph by defining buffer areas with different distances from the contour of 10 meters depth. This is done to allow more accurate computations close to shore, where the risk of grounding is higher. The distance to land will follow the same pattern as for the Equation (1) but the distance will be multiplied by 10. We will create four buffer zones and let the rest of the area have the largest cell size given by the formula above. All distances are in meters.

Figure 2 shows what the buffer zones look like from two different map scales.

3. ANOMALY DETECTION

One of the cornerstones of this system is to a produce point-by-point anomaly score for each vessel observation. The intention of the pointwise anomaly score is to serve as a basis for more sophisticated services.

3.1 Pointwise Scoring

Recall from Section 2 that historical data in the test area from two years is collected and grouped into bins corresponding to nodes in G (see Section 2.2 for details on the construction of G). Each node $u \in V$ has a known geographical position. For each vessel observation, we may find the closest node u'.

At this point we can produce some indicators of normal behaviour for a node relative to historic ship traffic. Consider a vessel observation u at a position (x_u, y_u) with speed v_u and course θ_u. The most local approach is to compare u directly to u'. Alternatively, we can draw a circle with radius r depending on the grid size drawn around u. The radius is defined as $r = 10 \cdot f_g(n)$, where f_g is the function from (1). Let B_u be the set of all points inside the circle. Now we can calculate some statistics regarding the set of all vessel observations p with position inside B_u, namely the set $P_u = \{p_1, p_2, \ldots, p_k\}$, where $p_i = (x_{p_i}, y_{p_i}) \in B_u$. The geometric center of P_u, c_P is calculated. Further, let μ_v and σ_v be the mean and standard deviation of speed, respectively, in P_u. Let $N = |P_u|$ be the number of vessel observations in B_u, and n_i is the number of observations at node number i with position $(x_i, y_i) \in B$. We denote by μ_n and σ_n the mean number and standard deviation of the number of observations at each node. Course over ground are separated into bins to form a compass rose (see Figure 1b for an example compass rose).

All of the statistics named above provide information on the (historical) normal situation in B_u. Deviations from this normal situation may be considered anomalies. A key concept is to produce an anomaly score based on deviations from the normal situations. This score is described in the following.

3.2 Scoring Function

An anomaly score is calculated for a single observation by comparing it to the normal situation established in the preceding. For preliminary analysis, we produce four separate anomaly scores, based on speed, position and bearing. All four scores are numbers between 0 and 1. Like before, let u be an observation with u' as its closest node in the grid and let c_P be the geometric center of P_u.

1. The anomaly score for speed A_1 is based on the speed measurements standardized in each node to have zero mean and unit variance. This means that each speed measurement is transformed to

$$v = \frac{v_u - \hat{\mu}_v}{\hat{\sigma}_v}, \tag{2}$$

where $\hat{\mu}_v$ and $\hat{\sigma}_v$ are the mean and estimated standard deviation in speed at node u'. In other words, the score represents the number of estimated standard deviations by which an observation deviates from the node mean. In order to compare all subscores, v is transformed to a fixed scale. This is done by min-max transforming v to $[0, 1]$ as follows:

$$A_1 = \frac{v - v_{\min}}{v_{\max} - v_{\min}}, \tag{3}$$

where $v_{\min}$ and $v_{\max}$ are the min and max of the standardized speeds recorded at node u', respectively.

2. Recall that c_P is the centroid of the set P_u introduced above. The spatial anomaly score A_2 is a function of the (Euclidean) distance from u to c_P and the buffer zone containing u:

$$A_2 = \frac{d(u, c_P)}{10 \cdot f_g(n)}. \tag{4}$$

The radius of the circle is $10 \cdot f_g(n)$ meters, so $A_2 \in [0, 1]$ when we divide it by this number.

3. The directional anomaly score A_3 is based on the compass rose explained above. The bearing of u is categorized to its sector k. The score is the sum of the number of observations in this sector and the adjacent sector on each side, divided by the total, i.e

$$A_3 = \frac{n_{k-1} + n_k + n_{k+1}}{N}, \tag{5}$$

where n_k is the number of vessels headed in a direction contained in sector k, and $k = \{1, 2, \ldots, 36\}$.

4. The angle-distance anomaly score A_4 is an attempt at exploiting the speed vector in combination with the map. We define a 15-degrees wide sector symmetrically around the course over ground. This will form a triangle R with one corner being the position of the vessel. Define the height of the triangle to be $f_g(n)$. Let d_R be the distance from u to closest node u^* inside R. If R is empty, $d_R = f_g(n)$. Let n_{u^*} be the number of observations in u^*. The angle-distance anomaly score is defined as

$$A_4 = \frac{d_R \cdot l}{\log(n_{u^*})}, \tag{6}$$

where l is the ship length. The vessel's course sector and all the nodes are shown in Figure 3.

The score function s is a combination of the four aforementioned components. As a first try, it is defined as

$$s = \max\left\{\frac{A_1 + A_2 + A_3}{3}, A_4\right\}. \tag{7}$$

The intention of (7) is to allow different forms of hazardous behaviour trigger the alarm. A_1, A_2 and A_3 may be considered "stationary" anomaly indicators in the sense that they evaluate the behaviour at exactly the time and place where the vessel is located. In contrast, A_4 relies on the vessel course to be near constant for some distance.
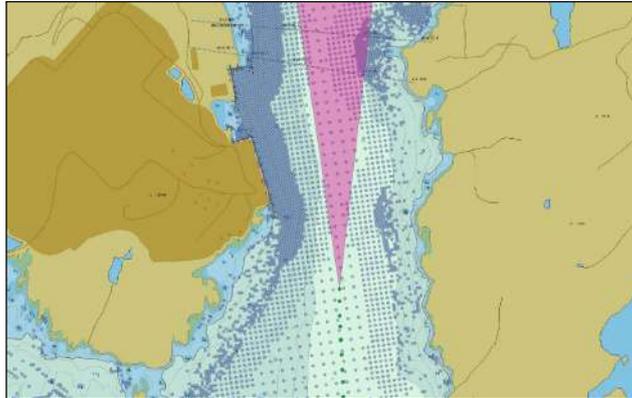
Figure 3: The 15-degree sector around the vessel's course is shown in pink color in this figure. This example shows a common course but when a vessel is in an area without any historic positions it will not have many points within the pink sector.

3.3 False Alarm Rate

The main goal is to provide reliable alarms when there is danger, but this is of no use to the operator if the total number of alarms is too high: then it would be impossible to attend carefully to every single one. As the score is distributed across an interval, the alarm sensitivity may be predefined and tuned by the user. In any case, it is important to minimize the false alarm rate. A few special situations are treated to keep the number of false alarms low:

1. Slow speeds: A ship being on a collision course is not necessarily an accident waiting to happen. For example, a ferry is regularly on a collision course with the landing area. A feature common to most of these situations is that speed over ground is low. For this reason, both the directional anomaly score A_3 and the angle-distance anomaly score A_4, which is meant to alarm bad navigation, is turned off if the speed is less than 1 knot.

2. Sparse data: Far from land, there is less risk of both grounding and colliding, The angle-distance anomaly score A_4 is simply turned off if the distance from land and shallows is large. A course vector is plotted 2.5 kilometres ahead of the vessel, regardless of its speed and direction. If the course vector intersects the innermost buffer zone, i.e. within 100 meters from all known 10 meter shallows, then A_1 is turned on again.

4. EXAMPLES

4.1 Sjernarøy

Consider the track from Figure 4a. This shows the passenger ferry Sjernarøy on October 12th, 2014, when it grounded off the small island Nord-Hidle in Boknafjorden. The coloring is made in accordance with the anomaly score shown in Figure 5a. High score is indicated with red color, low with green. Clearly, the colouring turns abruptly from green through yellow to orange after the ship sails past the landing area and proceeds towards shore. This situation can be investigated further by inspecting the anomaly scores in Figure 5a. Note that the score increase abruptly at about 8.15 PM, roughly when the ship grounded. The vessel score service does not react in due time for this case to be averted. The reason for this can be found in the data. As this is a ferry route, there is plenty of ship traffic at the spot, especially close to the landing area. Direction of travel is also widely distributed across the compass. This explains the slow reaction: close to the landing area, this ship does not exhibit any anomalous behavior. It is unusual first after passing the landing area and steering for the shore. Indeed, this is when the anomaly score spikes.
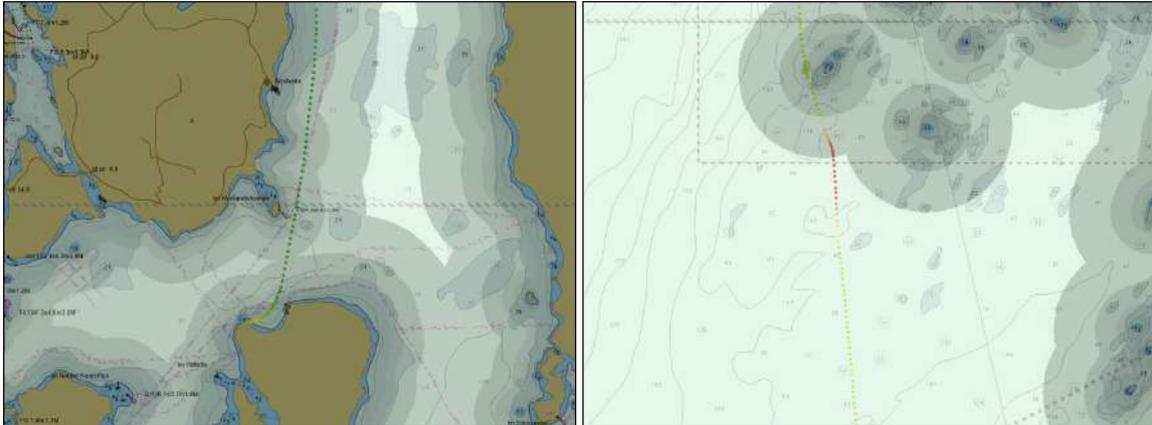
(a) The track of the passenger ferry Sjernarøy from October 12th, 2014.

(b) The track of the fishing vessel Christina E from October 22th, 2018.

Figure 4: The shaded areas drawn from the raster described in subsection 2.3.
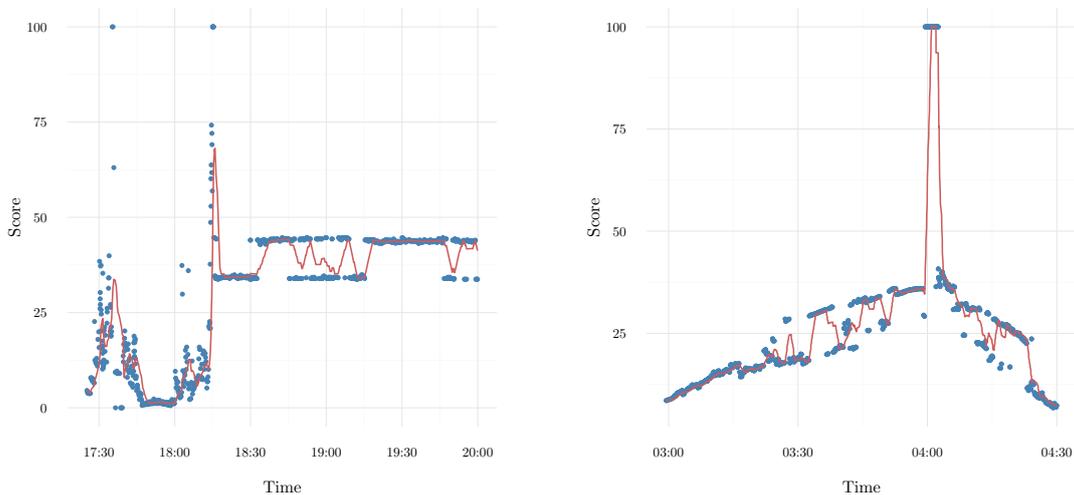
(a) Vessel: Sjernarøy.

(b) Vessel: Christina E.

Figure 5: Pointwise Anomaly score (blue) with moving average (red).

4.2 Christina E

Consider the track from Figure 4b. This depicts the fishing vessel Christina E heading for a shallow of 7 meters, while the vessel itself has a registered depth of 7.8 meters. Clearly, this is not a desirable situation. A vigilant VTS operator spotted the ship and instructed it to turn port to escape the shallow. The track coloring in Figure 4b and the anomaly scores in Figure 5b tells us that the vessel score service reacts well before the ship turns. The alarm situation is obvious if we inspect Figure 5b. At approximately 4 AM, there is a large spike in the score. This is due to false-alarm-controlling mechanism described above: the angle-distance score is large before this time, but due to the 2.5 kilometre course vector not intersecting the 10 meter zone, the angle-distance score is turned off. When the vector intersects the area, the score is turned on, causing a large total anomaly score.

5. CONCLUDING REMARKS

This work is an attempt at exploiting the vast amounts of historical data on ship traffic for anomaly detection. The anomaly detector is generally capable of detecting ships headed for shallows or shore. In addition, it is

capable of producing soft alarms for some anomalous behaviour, such as ships sailing in disagreement with the leads.

Most calculations are based on the grid graph described in Section 2.2. This data structure is well equipped for the task due to its ability to give preprocessed local context in a simple way, e.g. by investigating the statistics at neighbouring nodes. It also enables us to easily exploit the varying levels of node resolution discussed in Section 2.3.

5.1 Future Work

The simplest idea for improvement is to add more data. The current graphs contain data from the year 2017 and 2018, so there are plenty more to take from. A drawback is the computation time for the preprocessing steps. With the amount of data used now, the time spent for preprocessing and constructing the graphs is in the order of days on a rather regular desktop computer. In addition to increased accuracy, more data may facilitate a finer partitioning of ship types and properties. This may in turn increase the accuracy further.

Other concepts to be further developed include:

- More sophisticated weighting of subscores in (7).

- Better analysis of close nodes in the angle-distance score.

- The use of geographical features in the anomaly score, such as score discounts at known landing areas and score penalties in non-meeting zones.

Context-enhanced maritime surveillance optimization

Steven Horn

Defence Research and Development Canada (DRDC)

ABSTRACT

A method for the optimization of wide area surveillance given contextual information in the form of intelligence about existing or potential maritime threats is presented. Given the complexity of the surveillance parameter space, exhaustive search is typically not possible. A Bayesian search using posterior sampling is proposed whereby a near optimal surveillance plan can be generated with reduced computational requirements. Given any new intelligence over time, the surveillance plan can then be quickly updated to integrate the new information. The method also allows the incorporation of past surveillance activities including nil observations.

Keywords: Surveillance, Pattern of Life, Modelling and Simulation, Bayesian Inference

1. INTRODUCTION

The problem of surveillance resource allocation and optimization given constraints and objectives is a widely studied area of research.[1,2] The goal of surveillance resource optimization is to maximize the available information and to minimize the operational risk. Typically, past research has focused on the management of a single sensor,[3] or a set of sensors[4] for performance optimization of the sensor or sensors. Additionally, attention has been given to the net objective of surveillance in relation to the reduction of operational risk.[5] Furthermore, the optimization process can be dynamic with repeated optimizations providing adjustments on short *tactical* minute-to-minute time-scales such as when a sensor is employed *in-situ*, or the optimization can be on a longer *operational* time-scale where the adjustments are days in advance.

The task of optimization of a wide-area surveillance plan which aims to conduct *operational* time-scale planning is considered in this paper. In contrast to some existing work, the optimization discussed here is focused on the singular detection (*i.e.* detected at least once) of one or more potential maritime threats and not necessarily the optimization of the tracking of a threat. Figure 1 presents an illustration of the operational wide area surveillance planning problem.

This paper is organized as follows: Section 2 presents the mathematical framework for describing the problem, Section 3 presents the proposed optimization technique, and Section 4 presents some simulated results for a notional scenario.

2. MODELLING OF SURVEILLANCE, THREATS, AND CONTEXT

The operational wide area surveillance problem is described mathematically as follows. A surveillance planner has a set of n surveillance resources $\mathcal{R} = \{R_i\}$, $i = 1 \ldots n$, which are each subject to a set of k constraints $\mathcal{C}_i = \{C_j\}$, $j = 1 \ldots k$. A surveillance plan is generated for the employment of a subset of $\mathcal{R}$, and each configured with parameters $\mathcal{P}_i$ which include the locations and times of surveillance subject to $\mathcal{C}_i$. The choice of which subset of $\mathcal{R}$ is employed in the surveillance plan is such that the total cost of surveillance should be minimized while maximizing the effectiveness of the plan for detecting threats.

To model the multiple types of concurrent threats for which to conduct surveillance against, we define a set of threats $\mathcal{T} = \{T_x\}$ each with a set of characteristics with uncertain values. The following characteristics are considered here for each threat: the origin (T_o), destination (T_d), speed (T_s), and size (T_z) of a threat such that the characteristics of T_x are $\{T_{x,o}, T_{x,d}, T_{x,s}, T_{x,z}\}$.

Each surveillance resource R_i is described by a sensor detection model d_i which is a function of the ability to detect a threat T_x with given characteristics such that $d_i = f(R_i, C_i, T_{x,z})$.[6,7] The simplest approach is if we
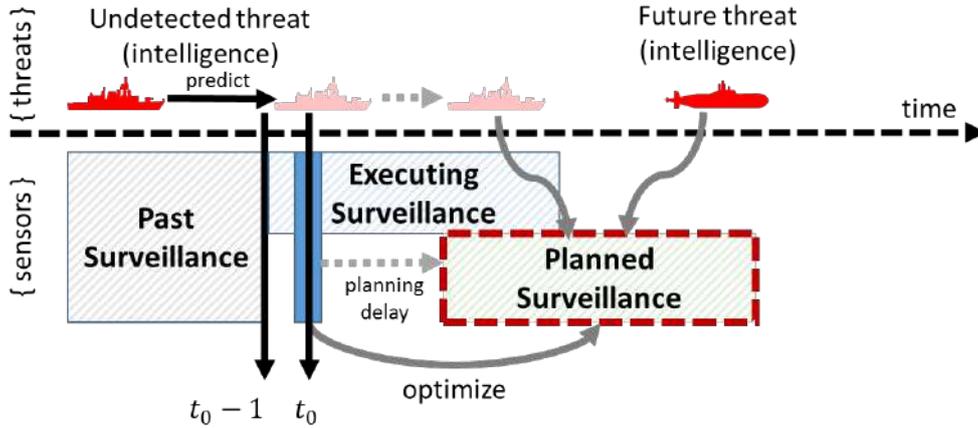

Figure 1. Illustration of the operational wide area surveillance planning problem. Surveillance optimization leverages past surveillance information, current information, and intelligence to plan future sensor allocations.

assume that each sensor is independent of one another. However, this assumption is not always valid in reality: for example, if one sensor is cross-cuing another sensor. Even so, it is not a bad assumption for many surveillance scenarios and sensors.[8] Consideration for sensor dependence can be included in a more advanced model but will not be discussed in this paper.

Elaborating on surveillance shown in Figure 1, there are three aspects to the surveillance model: 1) information about surveillance which has already occurred; 2) surveillance which is committed and will execute; and, 3) surveillance which is planned but not yet committed. The past surveillance is defined as the set of surveillance resources which were executed and the resulting combined information generated from that surveillance including nil detections. This is the *a priori* state of detection of threats. Executing surveillance is defined as the set of surveillance resources which are providing information at present. For example, this could consist of real-time radar information, satellite collections, or patrols (air and surface).

The route of the threat transiting the surveillance region can be contextually modelled using intelligence derived from the pattern of life[9] given T_o and T_d. Through simulation of the surveillance plan, the posterior probability distribution is estimated for the detection of threats ($f_{d|R,T}$) after the threats transit through the surveillance zone given the resources $\mathcal{R}$ and threat characteristics.

Figure 2 illustrates how the use of multiple independent sensors improves the effect of surveillance. For illustrative purposes, $\mathcal{R}$ consists of three fake sensors with parameters shown in Table 1. These were simulated against a threat transiting from $x_0 = 0 \pm 2$ to $x = 1000$ nautical miles with speed $v = U(10, 15)$ knots, starting at $t_0 = U(0, 100)$ hours, where U is the uniform distribution. The distribution is derived from the simulation using kernel density estimation with a Gaussian kernel with a bandwidth derived from Scott's rule.[10] The score for surveillance using one type of sensor independently shown in Figure 2 (a-c) versus when all of the sensors are used (d) illustrates the coordination effect of surveillance. Equation 1 shows how this score is calculated for threat x. In general, for all threats, the overall surveillance score is $s = \sum_x s_x$. Depending on when and where each sensor is in relation to the threats, the overall surveillance score will vary as the surveillance parameters are varied. This score function is a form of linearly weighted sum of the distribution which increases for improved expected detection probability.

$$s_x = \int_0^1 p \cdot f_{d(p)|R,T_x} \, dp \, . \tag{1}$$

For each sensor parameter free from constraints, the objective is to then find a surveillance plan which results in a maximum value of s given the surveillance parameters. The next section presents a method to search the parameter space to optimize the surveillance plan.
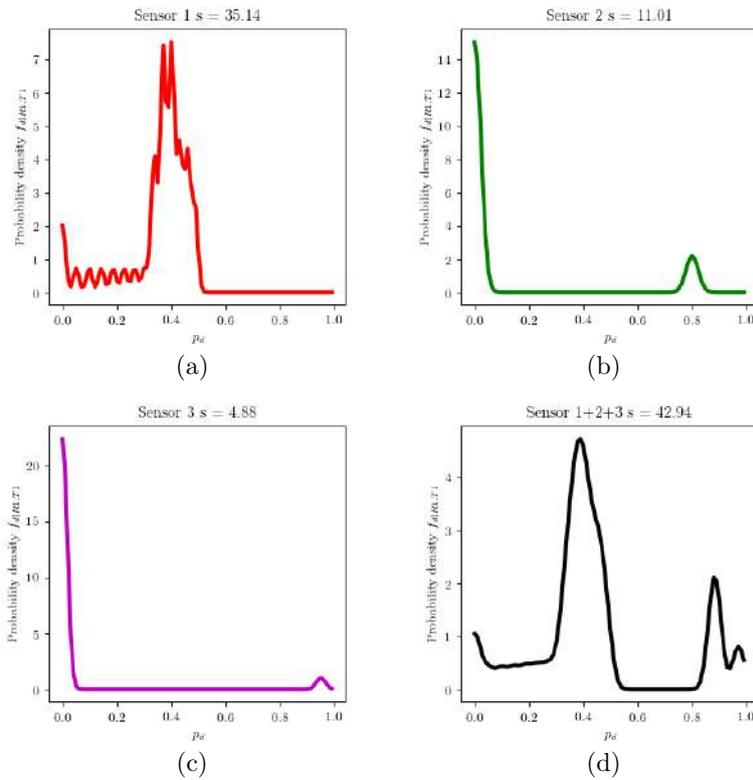
Figure 2. For a given threat T_1, the combined detection performance for three sensors each individually (a-c), and all sensors combined (d). The surveillance score is indicated by s.

Table 1. Surveillance parameters for three simulated sensors.

Sensor	Description	Parameters
1	Surveys a large area relatively often with low probability of detection. This is representative of a constellation of satellites.	$P_d = 0.04$, $dt = 2$ h, $x = [50, 300]$
2	Surveys a small area (50 nm wide) every 24 hours. This is representative of a long range patrol aircraft.	$P_d = 0.8$, $dt = 2$ h, $dx = 50$
3	Surveys a specific box every 20 hours. This is representative of a surface patrol.	$P_d = 0.95$, $dt = 2$ h, $x = [900, 1000]$

3. OPTIMIZATION THROUGH POSTERIOR SAMPLING

The objective is to select and configure the resources $\mathcal{R}$ by varying the surveillance parameters P_n subject to constraints $\mathcal{C}_R$ to maximize the detection of the set of threats, as represented by the surveillance score s. Given a surveillance configuration with a set of parameters, one can calculate through simulation the probability distribution for the detection of a threat $f_{d|R1,T1}$, as shown in Figure 2. The brute force approach would be to simulate and compute the posterior distribution of all possible surveillance options. However, the search space for possible surveillance configurations, even under constraints, is extremely large. Moreover, changes to the configuration are highly non-linear and finding an optimal (or better) solution is not straightforward. Depending on the complexity of the surveillance optimization, there can be numerous parameters (hundreds), and evaluating the combinations of each parameter through simulation is not tractable.

Proposed here is a Bayesian approach using posterior sampling[11] with Gibbs method[12] to *learn* and predict with uncertainty the effect of surveillance in order to explore more informative simulations. For each surveillance option parameters which satisfy the constraints, a distribution of the surveillance score s is defined conditioned on those parameters. A scaled score over the search space can be represented by a multivariate distribution of the surveillance parameters under the test that the score is maximized given the parameters

$$f(s = s_{max}|P_1, \ldots, P_n), \tag{2}$$

which is initialized as per standard Bayesian methods. Here, one can choose an uninformed (uniform) prior for each parameter P_n to be optimized. The objective of the optimization is to find the tuple of P which maximizes s. This optimization is equivalent to identifying the maximum after inference of the distribution in Equation 2

Step 0: Initialization

The distribution in Eq. 2 is initialized with an uninformative prior such that

$$\int_{P_1} \cdots \int_{P_n} f(s_{max}|P_1, \ldots, P_n) \, dP = 1. \tag{3}$$

Step 1: Sampling

The selection of the tuple of surveillance parameters P to simulate is made using collapsed Gibbs sampling.[13] The objective in making this choice is to evaluate the simulation with parameters which is expected to provide useful information for the inference of $f(s_{max})$, while still exploring the parameter space which has yet to be evaluated. To select a value for a parameter P_i, the search space is collapsed through integration respecting the constraints C.

$$f(s_{max}|P_i) = \int_{P|C} f(s_{max}|P_{1\ldots n}) \, dp_{-i} \tag{4}$$

Where $f(s_{max}|P_i)$ is the collapsed marginal distribution, and the differential dp_{-i} represents integration over all parameters except i. The calculation of Eq. 4 is done numerically. Since the distribution of $f(s_{max})$ is normalized as in Eq. 3, and the values are all positive, then the maximum of the cumulative distribution is 1. The choice of value for P_i is then derived from a random number $0 < r < 1$, and the parameter is derived from the cumulative distribution:

$$r = \int_{min(P_i)}^{P_i} f(s_{max}|P_i = x) \, dx. \tag{5}$$

The calculations in Eqs. 4-5 are repeated for each parameter dimension to choose the tuple of surveillance parameters R_p.

Step 2: Simulation

The surveillance score is calculated as per Eq. 1 for the sampled R_p, resulting in a calculated score $\hat{s}$. Recall that this surveillance score now considers the uncertainties in the threat, and the plan with R_p is simulated against the threat intelligence (context).

<center>(a)</center>

<center>(b)</center>
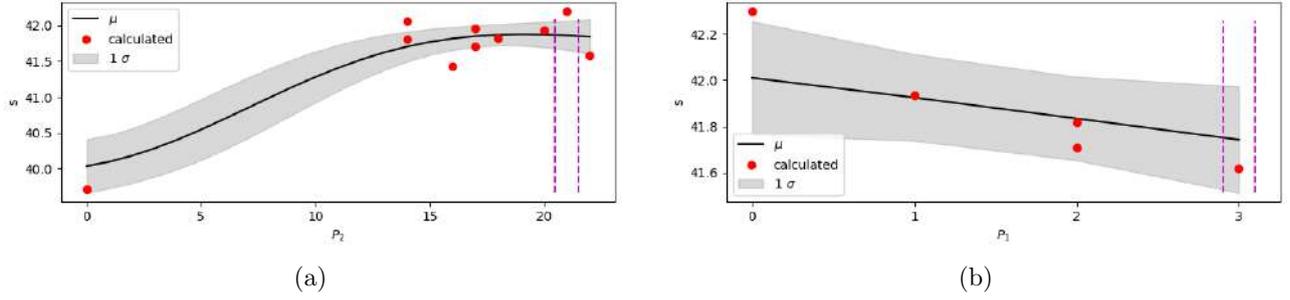
Figure 3. Visualization of the local Gaussian Process Regression for the estimation of the likelihoods for two parameters. The latest iteration is shown between the vertical dashed lines.

Step 3: Calculate posterior

The joint posterior distribution of the form in Eq. 2 is estimated using Bayes method.

$$f(\hat{s} = s_{max}|P) = \frac{\prod_{i \in R_p} p(\hat{s}|P_i, R_{p-i})}{p(\boldsymbol{s})} \cdot f(s_{max}|P) \tag{6}$$

The fraction in Eq. 6 is the Bayes factor where the top term is the joint likelihood, and the bottom term is the marginal likelihood. It is difficult to calculate the exact value of these terms, and so here it will be approximately calculated using numerical non-parametric methods. For the parameters which were simulated in the tuple R_p, the likelihood of the simulation result $\hat{s}$, given a parameter $P_i = q$, is $p(\hat{s}|P_i, R_{p-i})$. This likelihood is estimated using local regression, which has been shown to be effective[14] when sampling from a computationally expansive process such as the simulations in the problem here.

The local regression method used here is a Gaussian Process Regression on each parameter. The benefit of this particular regression is that it itself is a Bayesian estimator, and the posterior of it's estimate provides a distribution with a variance, which when applied here models the level of ignorance from lacking evidence from yet to be computed simulations. Given the regression on parameter P_i, the likelihood is:

$$p(\hat{s}|P_i, R_{p-i}) = 1 - \int_0^{\hat{s}} N(s|\mu_i(q), \sigma_i(q)^2) \, ds, \tag{7}$$

where $\mu_i(q)$ and $\sigma_i(q)$ are the regression outputs for the Gaussian as a function of the parameter value. Figure 3 shows an example of a local regression for the optimization problem which will be presented in the next section. A radial-basis function kernel was used with a length parameter of 2 for the regression.

Given the set of simulation results $\{s\}$ for the simulations, the marginal likelihood $p(\boldsymbol{s})$ in Eq. 6 is estimated using kernel density estimation (KDE).[15] A standard Gaussian KDE was used with a bandwidth derived using Scott's method.[10] This integration is illustrated graphically in Figure 4. Finally, since the estimation methods used for the likelihoods are not-exact, the posterior is re-normalized after applying the Bayesian update.

Step 4: Iterate

Steps 1-3 are repeated until a stopping condition is reached. This stopping condition can be a function of the update such as when the information gain from each simulation is no longer significant. Once the information gained from additional simulations for the update of the multi-dimensional surface $f(s|P)$ is low (*e.g.* below a threshold), sampling can be stopped, and the best discovered parameters R_p are taken as the best available surveillance plan.

Figure 5 shows the Kullback-Leibler (KL) divergence mean and variance over multiple iterations for 10 runs of the same optimization problem. The quickest optimizations in this plot were completed after just 8 iterations. One can note that as the algorithm converges on the optimal parameter set, the KL divergence tends to zero.
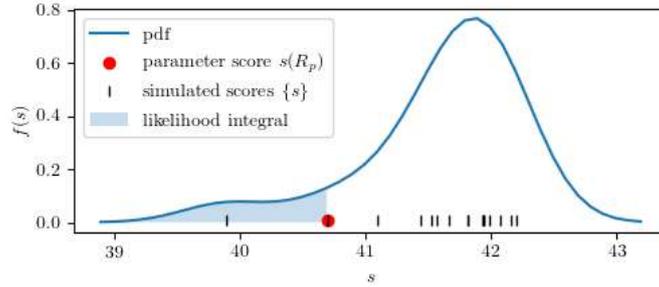
Figure 4. Example of the estimation of the likelihood that a given score for parameters is an optimum given a set of previous simulations.
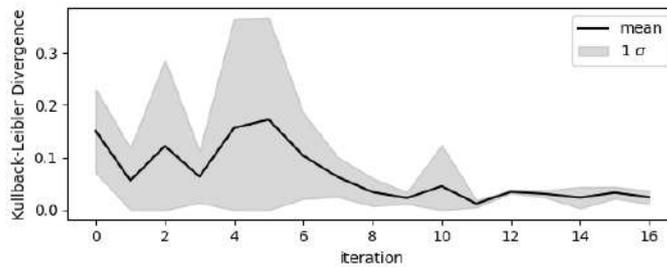
Figure 5. The Kullback-Leibler divergence of the posterior to the prior on updates from multiple simulation steps.

The steps with the higher divergence indicate locations where the information from new simulation runs was higher.

It must be noted that this algorithm does not guarantee a globally optimum result, but since the computational cost of evaluating each surveillance option is high, a good solution arrived by exploring as few as possible parameter variations.

4. RESULTS

To demonstrate the posterior sampling approach, the scenario shown in Section 2 was used while the parameters for sensor 2 were varied under arbitrary constraints for demonstration. These are not necessarily realistic constraints, but they serve to create a parameter space to demonstrate searching. The constraints used are:

1. $dx = 50$ (the surveillance box is 50 nm wide)

2. $x_0 = [0, 50, 100, 150]$ (the surveillance is constrained to one of these 4 boxes)

In Figure 6a, the posterior distribution is visualized for progressive solution search steps. One can note that the peak in the likelihood $f(s_{m}ax|P)$ moved and narrowed during the search procedure, as it arrived at a solution. The best solution found was with $P_1 = 1$ and $P_2 = 23$. Figure 6b shows the distribution function of this distribution for threat detection before and after optimization of sensor 2.

5. CONCLUSIONS

A method to quickly search a large parameter space for surveillance planning is proposed and was demonstrated on a simulated simple scenario. These types of searches are computationally costly, but the method here may make the optimization tractable. Future work includes exploring different Markov Chain Monte Carlo sampling techniques for sampling the prior, and applying the method to realistic high fidelity models of sensors and platforms.
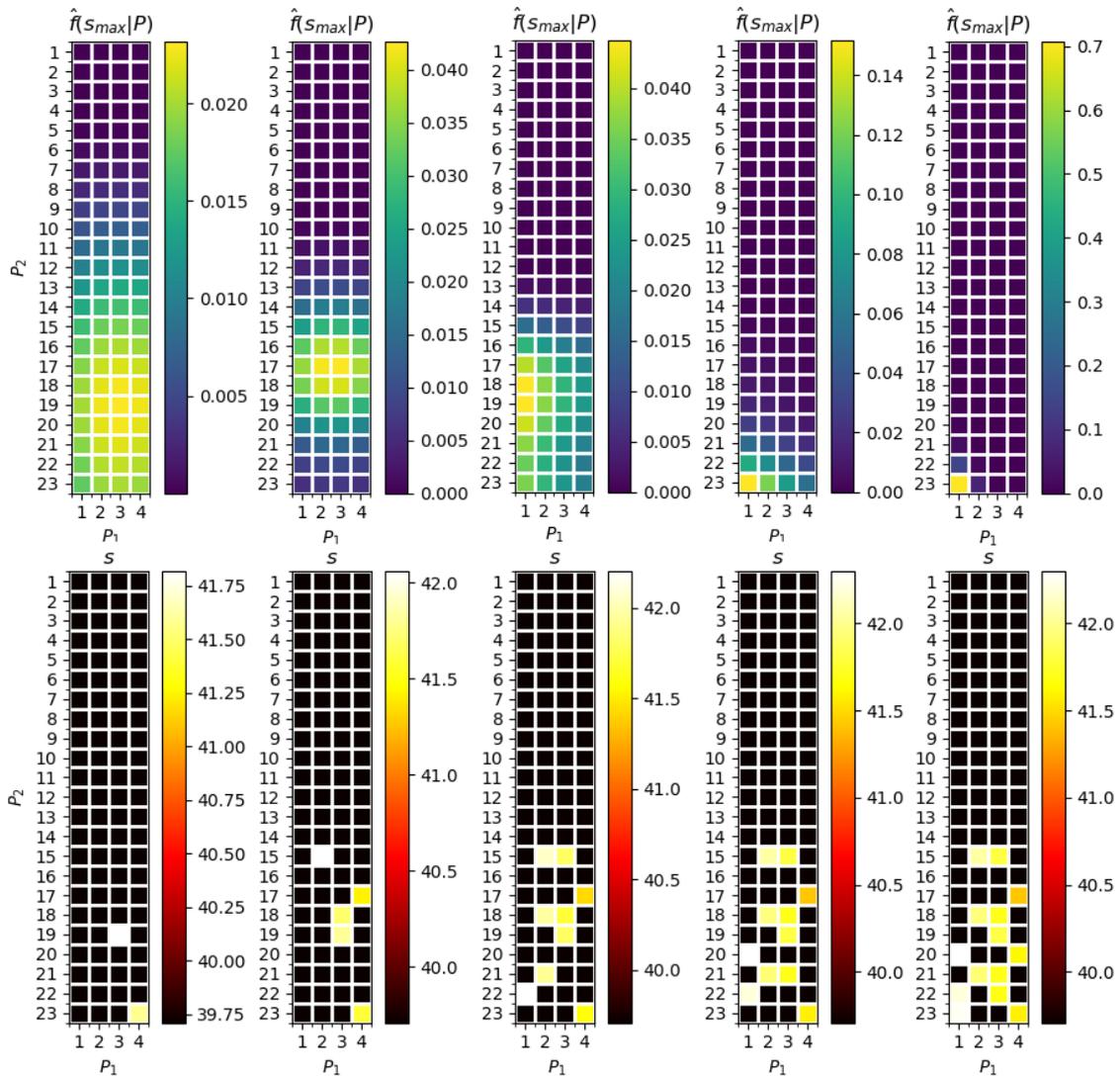
(a)
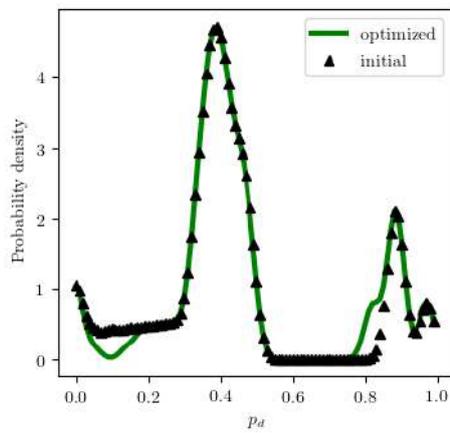
(b)

Figure 6. Progression of the posterior distribution during search of surveillance parameters (a), and distribution of detection probabilities for the optimized surveillance plan (b).

REFERENCES

[1] Dridi, O., Krichen, S., and Guitouni, A., "A multi-objective optimization approach for resource assignment and task scheduling problem: Application to maritime domain awareness," in [*2012 IEEE Congress on Evolutionary Computation*], 1–8 (June 2012).

[2] Berry, P., Pontecorvo, C., and Fogg, D., "Optimal employment of space surveillance resources for maritime target tracking and re-acquisition," in [*International Conference on Information Fusion*], 719–725 (2003).

[3] Hero, A. O. and Cochran, D., "Sensor management: Past, present, and future," *IEEE Sensors Journal* **11**, 3064–3075 (Dec 2011).

[4] Ponsford, A., "Multi-sensor resource management system for navy applications," (2019). DRDC-RDDC-2018-C240.

[5] Papageorgiou, D. and Raykin, M., "A risk-based approach to sensor resource management," in [*Advances in Cooperative Control and Optimization*], 129–144, Springer (2007).

[6] Horn, S., "Near real time estimation of surveillance gaps," in [*Proceedings of the 16th International Conference on Information Fusion*], 1871–1877 (July 2013).

[7] Horn, S. A., Mason, D., Zegers, A., and Edmond, E. J., "A bayesian approach to estimating detection performance in a multi-sensor environment," Tech. Rep. DRDC-RDDC-2014-R56, Defence Research and Development Canada CORA (2014).

[8] Papa, G., Horn, S., Braca, P., Bryan, K., and Romano, G., "Estimating sensor performance and target population size with multiple sensors," in [*2012 15th International Conference on Information Fusion*], 2102–2109 (July 2012).

[9] Millefiori, L. M., Braca, P., Bryan, K., and Willett, P., "Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction," *IEEE Transactions on Aerospace and Electronic Systems* **52**, 2313–2330 (October 2016).

[10] Scott, D. W., [*Multivariate density estimation: theory, practice, and visualization*], John Wiley & Sons (2015).

[11] Russo, D. J., Van Roy, B., Kazerouni, A., Osband, I., Wen, Z., et al., "A tutorial on thompson sampling," *Foundations and Trends® in Machine Learning* **11**(1), 1–96 (2018).

[12] Casella, G. and George, E. I., "Explaining the gibbs sampler," *The American Statistician* **46**(3), 167–174 (1992).

[13] Griffiths, T. L. and Steyvers, M., "Finding scientific topics," *Proceedings of the National academy of Sciences* **101**(suppl 1), 5228–5235 (2004).

[14] Conrad, P. R., Marzouk, Y. M., Pillai, N. S., and Smith, A., "Accelerating asymptotically exact mcmc for computationally intensive models via local approximations," *Journal of the American Statistical Association* **111**(516), 1591–1607 (2016).

[15] Thijssen, B. and Wessels, L. F., "Approximating multivariate posterior distribution functions from monte carlo samples for sequential bayesian inference," *arXiv preprint arXiv:1712.04200* (2017).

Context information analysis from IMM filtered data classification

David Sánchez Pedroche, Daniel Amigo Herrero, Jesús García Herrero, José Manuel Molina López

Group GIAA, University Carlos III of Madrid, Spain

Email: { davsanch, damigo, jgherrer } @inf.uc3m.es, molina@ia.uc3m.es

ABSTRACT

On this paper the authors study a set of maritime trajectory classifiers that classify the type of the ship or the kind of movement that is doing. All with the objective of deciding the viability of using the variables set introduced to the algorithm and the classification result as context information useful for a maritime vigilance system that could estimate the ship trajectories with the help of those classification results. To address the classification problem a decision tree and a support vector machine algorithms have been tested and compared. The set of variables to evaluate is obtained from the movement estimation of an Interacting Multiple Model (IMM) filter and its mode probabilities changes. Through the experimentation is evaluated the viability of the problem with the proposed variables, concluding those that could be useful for the proposed problem.

Keywords: AIS, anomaly detection, classification, data mining, IMM filter, maritime surveillance, trajectory reconstruction

1. INTRODUCTION

Currently there is a large amount of technologies that allow the localization and tracking of moving targets with the objective of using the resulting values in higher level systems like the surveillance systems.

For its better operation the tracking algorithms need a parameter adjustment in which the knowledge about the operational context could involve a great improvement. Context can be defined as any information surrounding a situation that is helpful to understand it, predict its evolution but it is not part of the situation [1]. For instance, in this domain the ship category, intended type of motion, or operational rules applied in a certain area is contextual information which can be helpful to estimate and predict the ship trajectories.

This paper proposes a context information analysis in the maritime traffic field with the objective of evaluate the capability of classify the movement or the type of a ship obtaining context information about the targets that could improve their tracking.

Between all the areas in which a context analysis for tracking improvement could be used this study is going to focus into the maritime surveillance area, this area is vital to ensure the safety and security of the maritime traffic and to prevent illegal activities [2].

To this study, due to the ubiquity of AIS-equipped ships worldwide, a set of AIS data is filtered to obtain a smoothed trajectory estimation that could be used by a context learning process, that process aims to provide useful knowledge to describe the behavior of ships and estimate more precisely their positions and cinematic states.

That behavior could be useful as context information for specific estimation systems that use the context to obtain better trajectory information to the maritime surveillance systems.

The first approach of the study [3] was the utilization of a binary tree classification algorithm to classify two possible multivalue classes, the ship type and the ship maneuver movement, all with the objective of create a useful context information source to the problem by the classification results.

With this approach also is possible to study the different filter configurations and the usefulness of the different used variables.

However, the conclusions also shown the necessity of improve the classification success rates to achieve a usable context information source that will provide class information useful for a theoretical position estimation system that works specifically adapted to the classes.

To achieve that objective, in this paper a binary class (an instance is classified as the class or not as the class) analysis of the predominant classes of the previous study is made to obtain new information about the classes and the variables of the used environment.

This paper is organized as follows: In section 2 the state-of-art methods in classification of maritime vehicles tracks are analyzed. In section 3 is explained the source of the data used for the investigation and in section 4 is related the implemented system and in 5 results of the work are shown. Finally, the conclusions and perspectives for future works are presented in section 6.

2.   RELATED WORK

There are numerous studies on AIS trajectory data. Specifically, in the area of maritime surveillance there are previous works addressing the search of traffic patterns to enhance Situational Awareness in maritime domain, especially to organize (cluster), reconstruct and classify trajectories, including prediction of activities and anomaly detection.

The work [4] processes AIS messages with deep learning framework (recurrent neural networks with latent variables) to address real aspects such as noisy data and irregular time- sampling for tasks of trajectory reconstruction, anomaly detection and vessel type identification.

The work in [5] has a proposal for a representation of routes as spatial grids built with AIS data to model the navigational patterns. It is extended in [6] to perform trajectory classification and anomaly detection in a system named as Traffic Route Extraction and Anomaly Detection (TREAD) based on extraction of frequent routs to classify real-time trajectories and trigger anomaly detection.

A survey of techniques proposed for mining trajectory data in multiple domains is provided in [7], focusing on data preparation, preprocessing, management and mining tasks (pattern mining, outlier detection, and trajectory classification), while a specific survey of maritime anomaly detection is provided in [8], distinguishing available data, methods, systems and user aspects. In [9], an analysis of AIS trajectory clustering is presented, with appropriate distance measures and dimensionality reduction. Aspects related with efficiency and scalability are dealt in [10], with data organization based on quad trees and modelling with Gaussian Mixtures.

Other approach is proposed in [11] and [12], where the computed dynamic parameters of a IMM filter are processed to perform segment classify missiles based on their dynamics and to divide Air Traffic Control data in homogenous segments and reconstruct the trajectories for evaluation purposes, respectively.

In [3] and in this paper, the objective is to analyze the ability to separate multiple categories of ships based on a IMM filter kinematic output and his filter information.

3.   DATA SOURCE

This study needs a reliable dataset from which obtain enough information to create classifiers that could learn correctly. For this task, a fundamental element would be obtaining pre-labeled data to use supervised classification algorithms, providing to the classifier a class to learn how to classify it. Also, is required to have a measurement of the position of the vessel in each instant, allowing the IMM filter to estimate and generate new reliable information for the classifier.

The selected AIS data is a repository provided by Danish Maritime Authority [13], in which millions of raw AIS contacts detected on the coast of Denmark and surrounding areas are available every day from 2006 to the present. This source provides a practically unlimited amount of information, enough to generate a useful dataset for the objective of the study. Also, in the raw AIS plots, two possible classes for the classification process are detected: a "ship type" category that could be useful to obtain information about the ship based on data about its movement, and the "navigational status" category which is useful to classify the type of maneuver that a ship is doing in a specific moment.

4.   DEVELOPED SYSTEM

The approach of this study aims to use the information of a state estimator to obtain trajectories from the sensor measurements and use them in a context analysis.

To achieve that the first step implies a strong preprocessing of the raw data, to clean it and extract only the useful information for the analysis and also preparing it as a viable input for the state estimator.

This approach uses an Interacting Multiple Model (IMM) filter as the state estimator for its well proven performance in problems of trajectory processing and smoothing by the reduction of the influence of the atypical sensor measurements.

The output of the estimation is prepared by a segmentation of trajectories to allow a comparable classification process (some trajectories have thousands of plots while others only have around one hundred) that provides new information useful for the context analysis.

4.1 Data preprocessing

The process starts with a file that contents all the plots of one day, around ten million of AIS plots. The first step consists in dividing the file in multiple files, each of one containing the information of one specific ship using the MMSI identifier.

Later, the MMSI contacts that doesn't provide useful information for these concreted objectives are discarded. These ones are contacts sent by a base station, and contacts that does not contain any of the classes for all the plots.

4.2 Track division

Once the tracks of a ship are obtained, the following process is to provide a correct input for the IMM filter. As the dataset used is filled by real information, some parts present some inconsistencies like big time gaps between two plots of the same vessel. These inconsistencies would make the filter return bad results that would compromise the classification outcomes.

Therefore, the solution for this task is to divide all the plots for every vessel into different tracks, each of one containing a minimum number of sequential plots with a maximum time gap between them. Also, this minimum whole track must be labeled with the same maneuver type to simplify the classification problem. This solution ensures that the IMM filter is capable of estimate correctly all the tracks and later be a correct and useful inputs to the classifier.

With the trajectories obtained as a result of these algorithm is possible to proceed to the next stage, where the IMM filter is going to process these trajectories to obtain some new information and to smooth the possible outliers in the position measurement.

4.3 Interacting Multiple Model (IMM)

The IMM filter is a tracking technique that allows adaptation to the target pattern of movement by the combination of the state estimation of multiple filter models into one common estimation.

On this study, the IMM filter use one mode for the representation of linear movement and another for the representation of the target maneuvers. Both implemented with an Extended Kalman filter (EKF) with the only difference of their Q matrix.

4.4 Trajectories segmentation

To allow the classification needed to the context analysis is necessary to transform the filtered trajectories into a useful input for the classification algorithm. To achieve this the first step is the segmentation of trajectories in equal size blocks to make them comparables into the classification.

An analysis of the classes values shows that there are instances with non-instanced values as "Reserved for future amendment [HSC]", "-" or "Other" that are cleaned with the minority classes with less than 0.0005% instances.

As a final step, is necessary to extract the information of each segment into data usable by the classification algorithm, using inputs that represent all the segment information and not only specific plots.

To achieve that the inputs are statistical values like average, mode, standard deviation, maximum, minimum and the 3 quartiles. All of them applied over the movement information of the segment:

- Speed of the target.
- Speed variation within the segment.
- The length of the movement.
- The heading variation.
- The time duration of the movement.

The attributes of the classifier instances would be like "the average speed of the target along the segment"

Also, as the IMM filter allows the location of maneuver movements with the analysis of mode probabilities (μ_j) a group of descriptors are defined to categorize the changes in the type of movement:

- Descriptor 1: Linear movement probability over 0.9 (the other one less than 0.1).
- Descriptor 2: Linear movement probability between 0.9 and 0.6 (the other mode between 0.1 and 0.4)
- Descriptor 3: Both probabilities between 0.6 and 0.4.
- Descriptor 4: Maneuver movement probability between 0.9 and 0.6 (the other one between 0.1 and 0.4).
- Descriptor 5: Maneuver mode probability over 0.9 (the other one less than 0.1).

With the information of the descriptor, is possible to explore the changes between type of movement in the segment by counting them, providing information about the maneuver movement model activation.

Figure 1 show the changes between descriptor with the variation of the linear movement mode probability. The information of the descriptors would be the number of measures in each descriptor (in the figure there are only two measures for the descriptor 3) and changes between descriptors (in the figure there are two changes from descriptor 2 to descriptor 1).
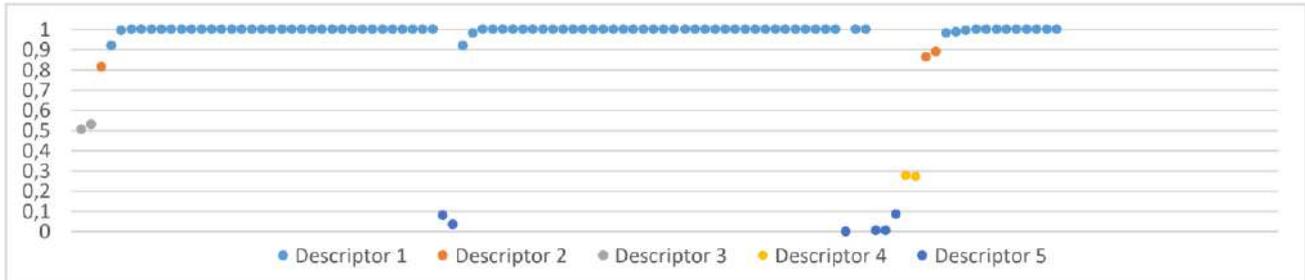

Figure 1. Descriptors over the linear movement probability

5. EXPERIMENTS AND RESULTS

In this paper four predominant classes from the ship type and the maneuver (four each one) are selected to made binary classifiers (classified as the class or not classified as the class) with a decision tree algorithm. The success rate comes from each segment (the classification results) or from the trajectory (the mode class of all the segments).

Table 1 shows the success rate of those classifiers to the ship type and maneuver classes.

Table 1. Classifiers success rates

SHIP TYPE CLASSIFIERS			MANEUVER CLASSIFIERS		
Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE	Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE
isCargo	73.6183 %	77.7693 %	isEngagedInFishing	75.9482 %	77.0614 %
isFishing	73.3423 %	73.4934 %	isRestrictedMovement	96.6902 %	96.6807 %
isPassenger	81.1167 %	85.1827 %	isSailing	94.8493 %	94.0119 %
isTanker	84.6864 %	86.0532 %	isUsingEngine	71.3486 %	68.5958 %

That results seems quite satisfactory in some of the classes but with the analysis of the confusion matrix is possible to observe an unbalancing problem, as the classifiers prefer to classify most of the variables as no members of each minoritarian class, obtaining false negatives instead of true positives.

As an example, is possible to compare the "*isRestricted*" class with the "*isUsingEngine*" class, the first one obtains the best results of all the classifiers whereas the second one obtains the worst results. In Table 2, the confusion matrices of the restricted movement and the is using engine classes are shown.

Table 2. Confusion matrices for example classifiers

Real Class / Assigned Class	ISNOT	IS
isNotRestrictedMovement	113819	2424
isRestrictedMovement	1474	54
isNotUsingEngine	10082	13217
isUsingEngine	20526	73946

The results clearly show that for the restricted movement classifier is preferable to classify most of the variables as *NotRestrictedMovement*, not finding the *restrictedMovement* class (which is the objective of the classifier).

In addition, the classifier "Using Engine" (which is the one with most of the instances) find is class but with poorer results.

To balance the classifiers is possible to reduce the "ISNOT" instances to the same number of the "IS" instances, taking as an example the *isRestrictedMovement* class, the balanced classifier will work only with around 1528 instances for other classes.

Also, is necessary to ensure the presence of all the classes, avoiding taking all of the predominant class. To achieve this is possible to obtain the percentage of presence of each class and calculate their instances.

In Table 3, there are the results for the balanced classification.

Table 3. Success rates of the balanced classifiers

SHIP TYPE CLASSIFIERS			MANEUVER CLASSIFIERS		
Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE	Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE
isCargo	71.0087 %	86.2541 %	isEngagedInFishing	63.4969 %	88.5977 %
isFishing	83.1454 %	93.1127 %	isRestrictedMovement	60.5168 %	98.5556 %
isPassenger	62.5543 %	89.8125 %	isSailing	61.6474 %	97.6468 %
isTanker	66.7262 %	94.7293 %	isUsingEngine	71.5179 %	72.8047 %

Is important to remark that the trajectory success rate with the balanced classification have a lot of improvement because the balanced data have an important reduction in the number of trajectories and as a result obtain better results which are not the objective of this part of the study.

The new results have less success rate in comparison with the not balanced classifier, but as is shown in Table 4, the confusion matrices are better at the class prediction which is the objective of this classification.

In other words, with the new confusion matrix our classifier reduces the false negatives (improving the true positives that are the objective of those classifiers) but in exchange there is an increase of false positives obtaining lowest success rates.

Table 4. Balanced classifier for maneuver class

Real Class \ Assigned Class	ISNOT	IS
isNotRestrictedMovement	869	660
isRestrictedMovement	547	981

The results of these classifiers show the same conclusion of the previous multivalue classification, being a difficult task for the classification to divide the instances into different classes.

To analyze that is possible to view the attributes changes through all the classes, by the creation of some characteristic vectors for each class (calculating the average of each attribute between all the instances classified as a specific class).

With the results of the classifiers is possible to aggregate all the instances classified as a class making a representative vector which dimensions are the average of the attributes of the instances set.

Also, to ensure that all the dimensions of the vectors are equivalent, the classification is made with normalized values, making all the attributes comparable.

In Figure 2 could be seen that the vectors of the four ship type classes have the same variations with the balanced classifier, whereas in Figure 3 the results of the maneuver classifier are shown.
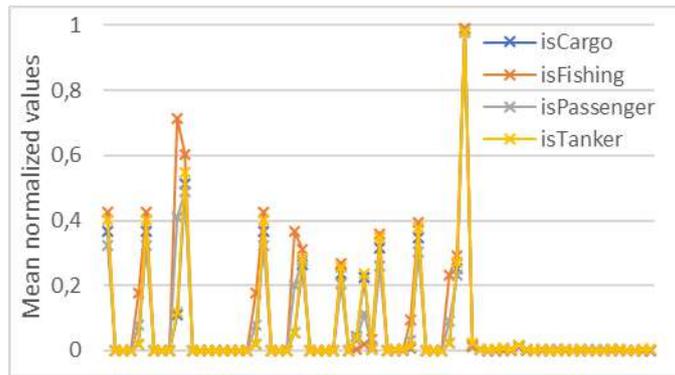
Figure 2. Ship type classes representative vectors

As it can be seen, most attributes have close to zero value and all the vectors shown really little variations between classes stating that the attributes hardly change between classes.
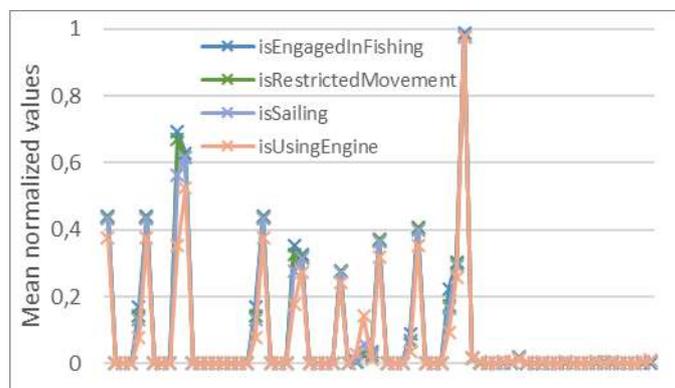
Figure 3. Maneuver classes representative vectors

An observation of the attributes shows that all of those with more variation comes from the movement variables of the direction variation or the time.

The figures from Figure 4 to Figure 7 show an example of those variables among the instances classified as the different classifiers, being possible to observe the real variation of those variables in the different classifiers.
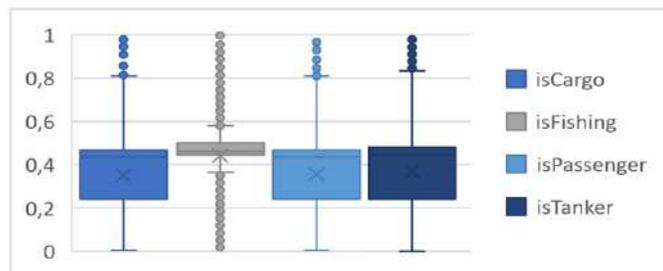
Figure 4. Ship type classes time duration
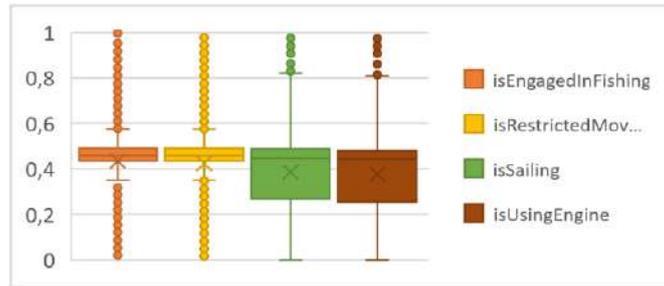
Figure 5. Maneuver classes time duration
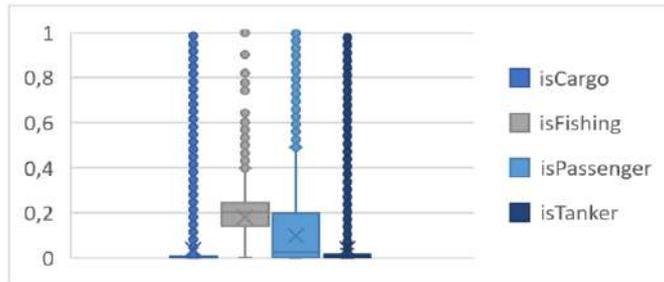
Figure 6. Ship type classes direction variation
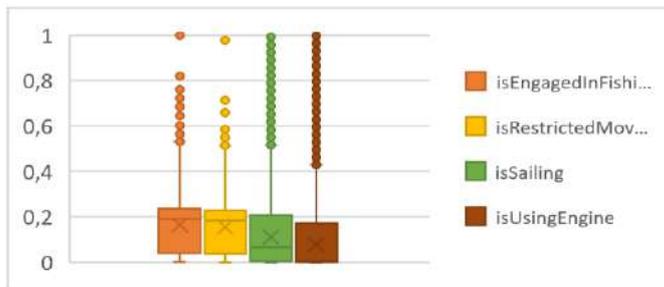
Figure 7. Maneuver classes direction variation

These results show that between these attributes (those with variation between classes the difference is minimal), the ship type classes barely differ between the *isTanker* and the *isCargo* and only the *isFishing* show enough differences.

Between the maneuver classes only the time create a difference and only with two groups whose members are also difficult to differentiate.

To a mayor improvement of the classification a new algorithm of support vector machines has been tested also as binary classifiers (classified as the class or not classified as the class), showing the results some improvement in the success rates as it can be seen in Table 5 and in Table 6.

Table 5. Support vector machines classifiers success rates

SHIP TYPE CLASSIFIERS			MANEUVER CLASSIFIERS		
Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE	Classifier	TEST SEGMENT SUCCESS RATE	TEST TRAJECTORY SUCCESS RATE
isCargo	79.1626%	88.2629%	isEngagedInFishing	75.1039%	89.8508%
isFishing	83.6765%	92.6727%	isRestrictedMovement	58.0635%	98.4695%
isPassenger	62.6371%	89.1429%	isSailing	68.3383%	97.9434%
isTanker	63.7878%	91.6109%	isUsingEngine	80.3062%	81.8251%

Table 6. SVM classifier confusion matrix

Real Class	Assigned Class	ISNOT	IS
isNotRestrictedMovement		688	841
isRestrictedMovement		441	1087

6. CONCLUSIONS AND PERSPECTIVES

As a conclusion of the obtained results, most of the introduced attributes not give enough differentiation to achieve an optimal classification. The experiments show that only the time duration of each segment, and the direction variation seems to have enough variability between the different classes to consider it useful to the proposed classification problem.

With that fact in mind is obvious that is necessary to find in future works other variables that have the required differentiation between classes to achieve an optimal classification that could be useful for the maritime surveillance systems as information of the context of operation.

Also, the SVM algorithm has clearly superior results than the decision trees algorithm so in future works is possible to study different classification algorithms to a major improvement of the success rate obtained from the algorithms shown on this paper.

7. ACKNOWLEDGEMENT

8. REFERENCES

[1]   J. Llinas, L. Snidaro, J. García, and E. Blasch, "Context-Enhanced Information Fusion," in *Context Enhanced Information Fusion*, Springer, 2015, pp. 3–23.

[2]   M. A. McNicholas, B. Wilson, and S. D. Genovese, *Maritime Security: An Introduction*. 2016.

[3]   D. Amigo, D. Sánchez Pedroche, J. García, and J. M. Molina, "AIS trajectory classification based on IMM data," *Int. Conf. Inf. Fusion*, 2019.

[4]   D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, "A Multi-task Deep Learning Architecture for Maritime Surveillance using AIS Data Streams," 2018.

[5]   V. F. Arguedas, G. Pallotta, and M. Vespe, "Maritime Traffic Networks : From Historical Positioning Data to Unsupervised Maritime Traffic Monitoring," vol. 19, no. 3, pp. 722–732, 2018.

[6]   G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.

[7]   M. Riveiro, G. Pallotta, and M. Vespe, "Maritime anomaly detection: A review," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 5, 2018.

[8]   Y. U. Zheng, "Trajectory data mining: An overview," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 3, p. 29, 2015.

[9]   H. Li, J. Liu, R. W. Liu, N. Xiong, K. Wu, and T. H. Kim, "A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis," *Sensors (Switzerland)*, vol. 17, no. 8, 2017.

[10]   A. Graser and P. Widhalm, "Modelling massive AIS streams with quad trees and Gaussian Mixtures," *21st Int. Conf. Geogr. Inf. Sci. (AGILE 2018)*, pp. 1–5, 2018.

[11]   K. Yoo, J. Chun, and J. Shin, "Target Tracking and Classification for Missile Using Interacting Multiple Model (IMM)," *2018 Int. Conf. Radar, RADAR 2018*, no. Imm, pp. 1–6, 2018.

[12]   J. García, J. A. Besada, J. M. Molina, and G. De Miguel, "Model-based trajectory reconstruction with IMM smoothing and segmentation," *Inf. Fusion*, vol. 22, pp. 127–140, 2015.

[13]   Danish Maritime Authority, "AIS data." [Online]. Available: https://www.dma.dk/SikkerhedTilSoes/ Sejladsinformation/AIS/Sider/default.aspx#.

6

Deep learning and classification

Contents

Vessel identification using Convolutional Neural Networks

Alex Reche Martinez[a] and Alberto Baldacci[a]

[a]MARSS, 9 Avenue d'Ostende, Monaco;

ABSTRACT

Recent advances in the field of object classification by He et al.,[1] Xie et al.[2] and Hu et al.,[3] have allowed us to implement a vessel recognition system that is able to correctly identify a vessel from a database of 2300 vessels. In this paper we present a convolutional neural network that, when presented with a picture of a vessel, computes the MMSI code of the vessel in the picture. We also introduce a potential improvement to generalize the algorithm by being able to identify any number of vessels by directly comparing two images of them.

Keywords: Vessel Identification, Artificial Intelligence, Machine Learning, Convolutional Neural Networks

1. INTRODUCTION

In this paper we present a method for vessel identification, that can be used to verify the MMSI code of a vessel when presented with only a picture of it. Our method consists of a convolutional neural network that, when given an image of a vessel returns a 2300 dimensions vector with the probabilities per class, where each class corresponds to one MMSI number. In addition, we present our approach to generalize the network by computing a signature for each vessel, as a 64 dimensions vector in an Euclidean space. This allows us to compare the signatures of two different vessels as a simple Euclidean distance, and use a threshold to decide if the images are from the same vessel or not.

To our knowledge, no other method exists to compare and recognize vessels from pictures alone, in a way that can be scale to an unknown number of vessels. Today's techniques for vessel identification are mostly based on the AIS automatic tracking system. However, this system is limited to a certain size of vessel, depends on the actual radio emitter being switched on and is open to hackers modifying the device to emit the wrong data.

There are already a multitude of approaches for image classification that learn to identify a certain number of classes from images [He et al.,[1] Xie et al.[2] and Hu et al.[3]]. Our network uses an image classification network with an output of 64 dimensions and a fully connected layer of 2300 dimensions to identify each vessel. The 64 dimensions vector, also called bottleneck, is intended to force the network to generalize by carefully selecting the 64 most important features to be observed.

2. PREVIOUS WORK

Our approach is a purely data driven method which learns to describe a vessel directly from the pixels of an image. Rather than trying to engineer the important features of a vessel in an image, we used a relatively large and privately owned dataset of images (40000 images for 2300 vessels), a classification convolutional neural network and multiple techniques of data augmentation. This allowed us to correctly classify the vessels and attain a signature that is invariant to illumination, pose, optics, weather conditions, and any other conditions.

We selected the classification network Squeeze and Excitation[3] because of its performance compared to other classification networks, despite a small computational cost. Squeeze and Excitation networks extend previous convolutional neural networks like Inception,[4] ResNet,[1] ResNext,[2] by exploiting channel dependencies.

To our knowledge no previous work has been done in the field of vessel identification, however our work has been highly inspired by works on the field of face recognition, where a vast corpus of research exists for face verification and recognition.

Further author information: (Send correspondence to A.R.M.)
A.R.M.: E-mail: alex.reche@marss.com, Telephone: +33.(0)6.21.23.64.93
A.B.: E-mail: alberto.baldacci@marss.com, Telephone: +39.328.844.5068
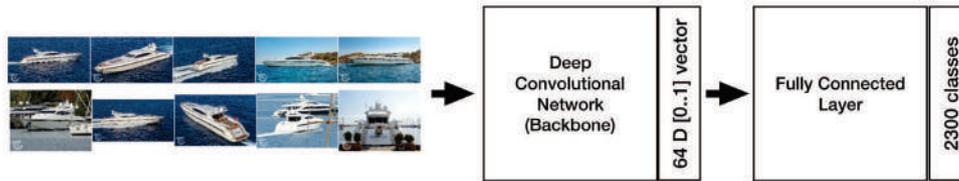
Figure 1. Architecture used to classify 2300 vessels while at the same time computing the 64 dimensions signature.

The works of Taigman et al.[5] and Sun et al.[6] employ a complex system of multiple stages, that combines the output of a convolutional neural network with principal component analysis for dimensionality reduction and an support vector machine for classification.

Taigman et al.[5] propose a multi-stage approach that aligns faces to a general 3D shape model. A multi-class network is trained to perform the face recognition task on over four thousand identities. The authors also experimented with a so called Siamese network where they directly optimize the L1-distance between two face features. Their best performance on the Labeled Faces in the Wild dataset (97.35%) stems from an ensemble of three networks using different alignments and color channels. The predicted distances (non-linear SVM predictions based on the X 2 kernel) of those networks are combined using a non-linear support vector machine.

Sun et al.[6] introduce the concept of FaceID, a bottleneck layer introduced in a classification network with the purpose to computing a generalized description of a face.

Schroff et al.[7] augmented further the performance by introducing a triplet embedding training using three images (a pivot, a positive and a negative match) and augmenting and reducing the distance, in an Euclidean space, between the computed face embedding of each image.

3. METHOD

Our system uses a convolutional neural network. We use as a backbone an SENet,[3] which is intended for image classification. We configured the output of the SENet to 64 classes using a sigmoid activation function and then added a new fully connected layer in order to expand these 64 classes to the 2300 corresponding to the vessels in our dataset. This network compression (or bottleneck) is intended to force the network to better select what features to observe in an image by essentially creating a signature of a vessel as a 64 dimensional vector. This will allow us in the future to generalize the algorithm.

Fig. 1 shows the architecture used to classify the vessels in our dataset, while at the same time computing the 64 dimensions signature.

3.1 Dataset and data augmentation

To train our model we used a privately owned dataset consisting of more than 60000 images of 6000 vessels. Unfortunately the dataset was very unbalanced with most of the vessels having very few images and some vessels having lots of images. In addition, only a portion of the images showed the outside of the vessel. We finally selected around 40000 images of 2300 vessels. Despite this careful selection of images and vessels, the final dataset was still very unbalanced (See Fig. 3). Consequently, the convolutional neural network ignored vessels with a small number of images, concentrating more on the ones that had many images to improve its accuracy.

To avoid this we had to artificially balance our dataset with repetition, however, a simple repetition of the same image resulted in the over-fitting of the network for the vessels whose images were repeated too often. We choose to heavily augment images by simulating conditions that the system will encounter in real life deployment like random rotations, different optics, weather conditions, blurred images due to being out of focus or motion, signal noise, etc. Fig. 2 shows some of the augmented images used for training.

The resulting artificially balanced dataset had close to 500000 images (Fig. 3).

Figure 2. Two series of data augmented images from our validation dataset. On each series the first image is the original image from the validation dataset. In the left series our final network was able to recognize all of the images, whereas in the right series only 7 out of 10 were correctly identified. The ones not correctly identified are shown in black boxes.
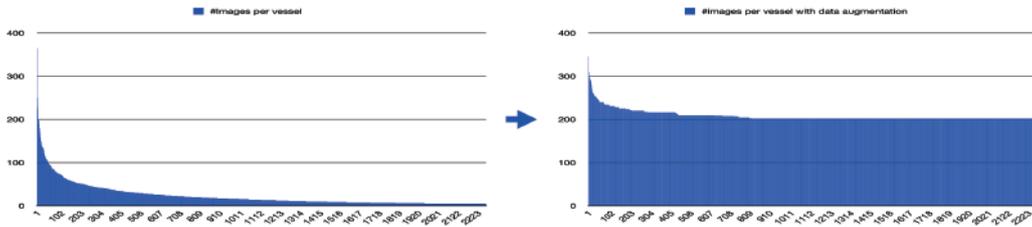
Figure 3. Number of images per class on the 2300 selected classes before and after data augmentation.

We divided the dataset into two sets, one for training and one for validation, 75% and 25% of the images of each vessel respectively. Both sets have samples of the 2300 vessels but data repetition and augmentation was applied only during training.

Using this setup we obtained the results presented in Table 1. We do not reach the Top-1 and Top-5 error rates obtained with the state of the art algorithms, but we suspect the large amount of classes in our dataset as being the reason for this.

Table 1. Results comparison between the state of the art SE-ResNext-101 and our method.

Network	Classes	Top-1 Error	Top-5 Error
SE-ResNeXt-101	1000	21.2	5.6
Ours	2300	26.4	16.2

3.2 Using the signature and Future works

The 64 dimensions bottleneck introduced in the network forces it to very carefully select the features observed in the images, enabling it to assign the correct class in the last layer.

While we do not know exactly what each value of the 64 dimensions vector means, in Fig 4 we can clearly see that the network configures itself to compute the same vectors for the same vessel independently of view point.

As of today, we have not used the signature calculated by the neural network to generalize the algorithm, however, we can see that the signatures of multiple images of the same unknown vessel are all similar (Fig 5). We are planing to use a technique similar to FaceNet[7] to increase the difference between signatures of different vessels and make the signatures of the same vessel even closer. For that, we will use the current network to create three signatures (a pivot, a positive match and a negative match). The loss function used during the optimization will compute the Euclidean distance between the pivot and the positive match, trying to get it as close to zero as possible, and the Euclidean distance between the pivot and the negative match, trying to make it as big as possible.
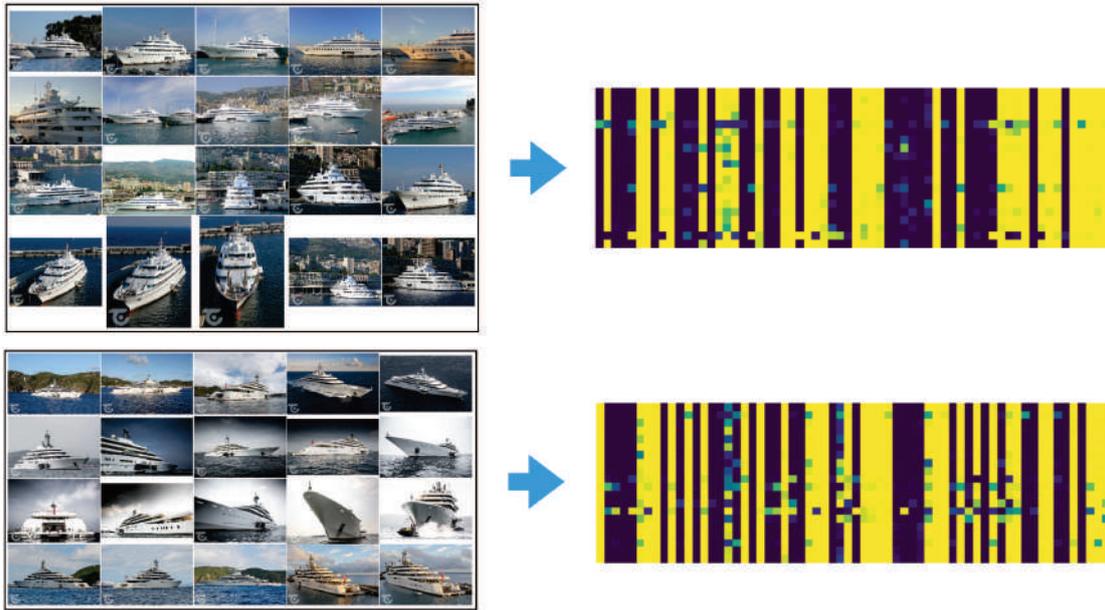
Figure 4. Computed signatures for a set of images of two known vessels from the validation set. The images on the right represent the signatures computed from the images on the left, each row representing a different signature.
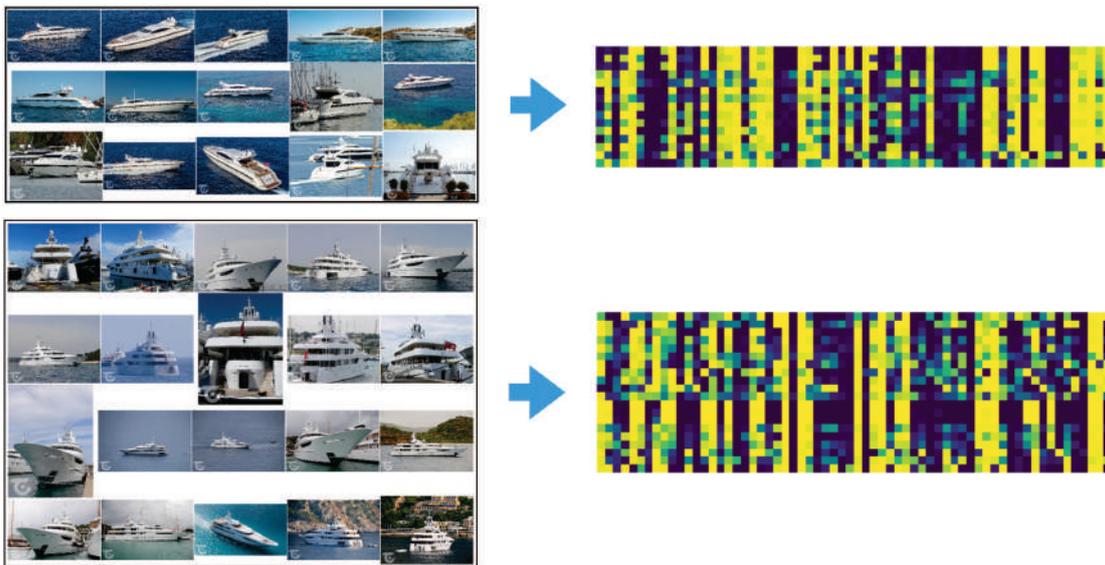
Figure 5. Computed signatures for a set of images of two vessels unknown during training.

ACKNOWLEDGMENTS

REFERENCES

[1] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , 770–778 (2016).

[2] Xie, S., Girshick, R. B., Dollár, P., Tu, Z., and He, K., "Aggregated residual transformations for deep neural networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , 5987–5995 (2017).

[3] Hu, J., Shen, L., and Sun, G., "Squeeze-and-excitation networks," in [*IEEE Conference on Computer Vision and Pattern Recognition*], (2018).

[4] Szegedy, C., Ioffe, S., and Vanhoucke, V., "Inception-v4, inception-resnet and the impact of residual connections on learning," in [*AAAI*], (2016).

[5] Taigman, Y., Yang, M. W., Ranzato, M., and Wolf, L., "Deepface: Closing the gap to human-level performance in face verification," *2014 IEEE Conference on Computer Vision and Pattern Recognition* , 1701–1708 (2014).

[6] Sun, Y., Wang, X., and Tang, X., "Deep learning face representation from predicting 10,000 classes," *2014 IEEE Conference on Computer Vision and Pattern Recognition* , 1891–1898 (2014).

[7] Schroff, F., Kalenichenko, D., and Philbin, J., "Facenet: A unified embedding for face recognition and clustering," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , 815–823 (2015).

Operationalizing Ship Detection Using Deep Learning

Torgeir Brenn[a, b], Lars-Petter Gjøvik[a], Gunnar Rasmussen[a], Tony Bauna[a], Michael Kampffmeyer[b], Robert Jenssen[b], and Stian Anfinsen[b]

[a]Kongsberg Satellite Services, Prestvannveien 38, Tromsø, Norway;
[b]UiT The Arctic University of Norway, Hansine Hansens veg 18, Tromsø, Norway

ABSTRACT

In this work, we showcase use of an extensive radar earth observation image dataset based on a long history of vessel detection services, with the goal of operationalizing a next generation service using deep learning. The extent of the data represents an opportunity for unprecedented performance, but the data was not gathered with deep learning in mind, and challenges of weak labels and extremely unbalanced classes must be overcome. We achieve this through a mix of techniques, including a novel approach within the framework of weakly supervised learning. Additionally, we recognize that the image augmentations which have become ubiquitous in computer vision is often not physically valid in the radar domain. We ask whether such augmentations can still be *pragmatically* valid, i.e. if they can be justified by increased model performance.

Keywords: Deep learning, vessel detection, weakly supervised learning, extremely unbalanced dataset, image augmentation, earth observation, synthetic aperture radar

1. BACKGROUND AND MOTIVATION

Kongsberg Satellite Services (KSAT) has an extensive history of delivering near real-time (NRT) vessel detection services based on synthetic aperture radar (SAR) earth observation (EO) images.* This 24/7 service is centered on human verification of automatically detected vessel candidates. Due to the SmallSat revolution, we anticipate an exponential increase in available data in the near future. It is not feasible to increase staffing proportionally, i.e. the automated part of the processing chain must be improved to detect a bigger proportion of the vessels while raising fewer false alarms.

2. DATASET CHALLENGES

These historical vessel detections are recorded as point labels (single-pixel), and they are not necessarily perfectly centered at the ships center of mass. A typical SAR image can easily contain $> 10^9$ pixels, and in our example dataset there are on average ~ 10 vessels per image. With single pixel labels, the class imbalance on the pixel level becomes extremely high, in the order of 10^8. We view this as two separate challenges: How to automatically account for this weak labeling, and how to adapt to such extreme class imbalance.

Further author information: (Send correspondence to T.Br.)
T.Br.: E-mail: torgeir.brenn@ksat.no
L.P.G.: E-mail: larspg@ksat.no
G.R.: E-mail: gunnar.rasmussen@ksat.no
T.Ba.: E-mail: tonyb@ksat.no
M.K.: E-mail: michael.c.kampffmeyer@uit.no
R.J.: E-mail: robert.jenssen@uit.no
S.A.: E-mail: stian.normann.anfinsen@uit.no
*The advantages of SAR for surveillance and situational awareness purposes include invariance to clouds and sunlight, giving a reliable all-weather all-night capability.
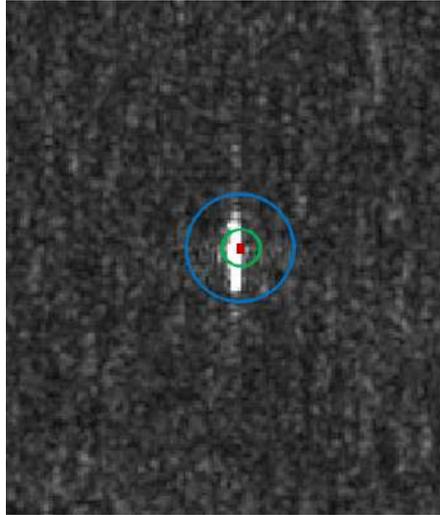
Figure 1. Our labeling scheme: The red dot is a typical label we would encountered: single-pixel and slightly off-center. That label is grown to a small circle (green). During training, all loss from area which is in the band between the small and large circles is ignored, to account for variability in ship sizes.

2.1 Dealing with Weak Labels

Weakly supervised learning (WSL) is a sub-field of deep learning which is gaining momentum and receiving increased attention.[1–4] In general, WSL is used to train models on data where the labeling is of a lower level that what would be ideal - for instance an image-level notation ("there is at least one car in this image") while asking for bounding boxes of all cars in that image. The promise of WSL is to reduce the manual effort required compared to the size of the dataset. KSAT is in the quite typical situation that such techniques are crucial in order to leverage our historical dataset for deep learning.

Figure 1 demonstrates our novel WSL strategy. Originally, the dataset contains only single-pixel labels (red), and often there's an offset from the center of mass due to time pressure in the service delivery. From an operational standpoint, this is a perfectly fine delivery; it is usually not interesting to get a segmentation mask, when it is the vessels position which is of interest.[†] These labels are grown into small circular regions which are intended to match small vessels, and as such depends on the satellite and image type used. To account for larger ships, we also use a bigger region in training, where all non-vessel pixels are ignored by zeroing their loss. In other words, we require that all pixels within the small circle be classified as vessel pixels, and all pixels outside the big circles be classified as background, but the network is free to designate the pixels between the circles as either vessel or background, since we admit to not knowing the vessel size due to our weak labels.

2.2 Counteracting the Extreme Data Imbalance

Recall that the original dataset has around 10^8 times more background pixels than vessel pixels. To overcome this, we employ several tools:

- Label growing as described above. This reduces the background-to-vessel ratio by a few orders of magnitude.

- The weighted focal loss.[5]

- The discarding of a proportion of tiles containing no background.

- Augmenting only those tiles where a vessel is present.

[†]KSAT also estimates vessel size and heading, but for relatively low resolution SAR this is often difficult and inaccurate. Estimating these quantities and also vessel type is future work w.r.t. deep learning at KSAT.
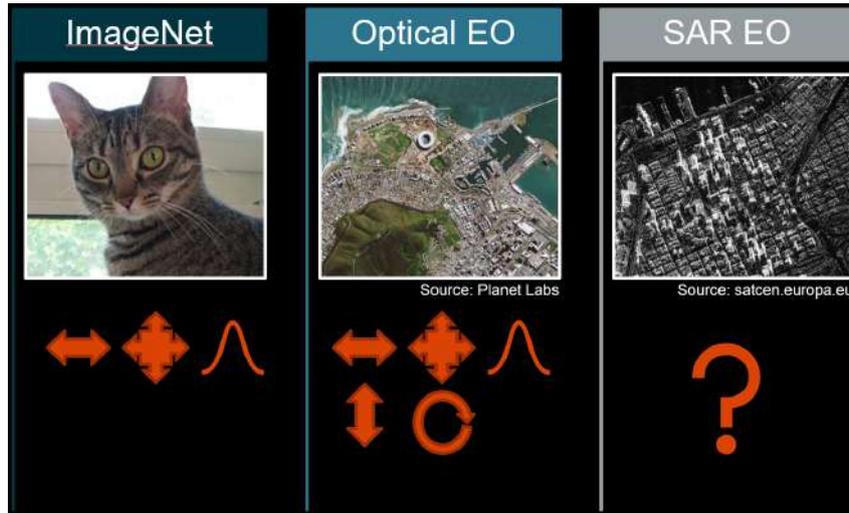
Figure 2. The physics of image augmentation: Human perspective images can be zoomed/scaled, flipped horizontally and augmented with additive Gaussian noise, while still being physically valid. Optical EO images can also be rotated and flipped vertically, but what about SAR images?

E.g., a model trained trained for 20 epochs will have seen augmentations of each detection 20 times, but each background tile perhaps only once on average. Additionally, the weighted focal loss focuses the training on the hard examples (typically vessels and vessel look-alikes), also giving the loss much more weight when the true class is "vessel". This combination of approaches allows us to train the model outlined in section 4.

3. SAR IMAGE AUGMENTATION

Image augmentation is a popular way to inflate a dataset by transforming the training images in different ways each time they are presented to the neural network.[6] A question which, to our knowledge, has yet to be properly addressed is the physics of SAR image augmentation, and how this relates to model performance. As figure 2 demonstrates, it does not make sense to vertically flip an image taken horizontally (unless of course we are expecting upside-down cats in our test data). For optical EO images the flexibility is increased, since accurate corrections for incidence angle means that any flip or rotate operation corresponds to the same operation on either the sensor or the scene, both of which are physically valid.

Radar on the other hand, works in a fundamentally different way, as it is an active rather than passive sensor. What this means is that a rotation does not correspond to *either* a rotation of the sensor or the scene, as the axes have distinct physical interpretations. Flipping the image in the cross-range direction[‡] represents opposite satellite flight direction and as such is physically valid, while the other flip is dubious at best due to the complexities of SAR range corrections. Even zooming is problematic, as the active nature of the SAR sensor yields a "noisy texture" called speckle, meaning that zooming is not equivalent to changing the distance between sensor and scene. This same property also excludes Gaussian additive noise, although other noise schemes might be valid.

Nonetheless, these physics observations do not automatically disqualify SAR image augmentations for training operational deep learning models, we must be pragmatic enough to do whatever gives the best possible model. This pragmatic approach might put physics in the back seat, especially for smaller datasets.

4. SATELLITE MODES, MODEL AND RESULTS

We designed a fully convolutional segmentation model with learned upsampling and skip connections, see figure 3. By now this is a quite standard approach, popularized by e.g. U-net.[7]

[‡]A SAR image has range and cross-range axes, with the former denoting distance from the sensor.
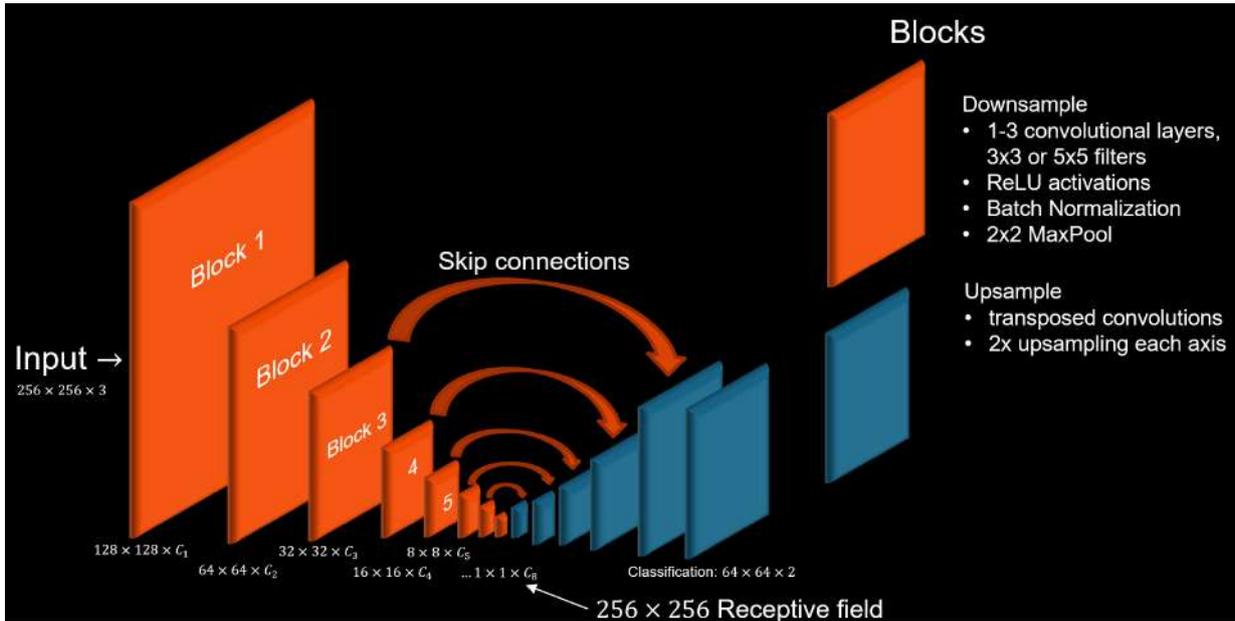
Figure 3. Our model: This a work in progress so none of the hyperparameters are final. Due to the fully convolutional nature of the network, the input can be any multiple of 256×256. Note that the receptive field of the most downsampled layer is an important design choice as it dictates the size of the region which can contribute to the detection decision for each pixel. The final two upsampling blocks required to recover the input resolution have been omitted, since our labels are weak we see no operational downside to a $4\times$ downsampled classification output.

4.1 Preliminary Results

Based on an assessment of the available data and operational needs, we chose the Copernicus Sentinel 1 mission, satellites A and B, IW GRDH mode, as our pilot. That dataset was divided into 1358 training images and 129 validation images, containing respectively 19316 and 927 manually verified vessel detections. Each image is several gigabytes, meaning that training is highly time-consuming.

Table 1. A comparison between KSATs in-house classical constant false alarm rate (CFAR) ship detector and a preliminary version of our convolutional neural network (CNN) ship detector.

	CFAR	CNN
Detections	879	803
Misses	48	124
False Alarms	1021	18

Table 1 compares the deep learning approach with the currently used classical ship detector. Since these algorithms both are input to a human-in-the-loop which manually assesses each vessel candidate, removing a false alarm represents significantly less work than finding a missed vessel. However, for potential fully automatic services these results are particularly encouraging, especially considering this is only a pilot study on a part of our dataset, with limited training time, no model optimization etc.

4.2 Future Work

To put KSAT in a leading position in terms of developing maritime situational awareness solutions, we are currently building an infrastructure of software and hardware capable of training and deploying models based on neural networks. Concurrently, we must develop novel methods to use our vast but weakly labelled dataset,

and conduct research into the validity and effectiveness of SAR image augmentation and other fundamental questions.

REFERENCES

[1] Han, J., Zhang, D., Cheng, G., Guo, L., and Ren, J., "Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning," *IEEE Transactions on Geoscience and Remote Sensing* **53**(6), 3325–3337 (2015).

[2] Bilen, H. and Vedaldi, A., "Weakly supervised deep detection networks," in [*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*], 2846–2854 (2016).

[3] Zhang, D., Han, J., Cheng, G., Liu, Z., Bu, S., and Guo, L., "Weakly supervised learning for target detection in remote sensing images," *IEEE Geoscience and Remote Sensing Letters* **12**(4), 701–705 (2015).

[4] Papadopoulos, D. P., Uijlings, J. R., Keller, F., and Ferrari, V., "Training object class detectors with click supervision," *arXiv preprint arXiv:1704.06189* (2017).

[5] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P., "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018).

[6] Simard, P. Y., Steinkraus, D., Platt, J. C., et al., "Best practices for convolutional neural networks applied to visual document analysis.," in [*Proceedings of International Conference on Document Analysis and Recognition*], **3** (2003).

[7] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," in [*International Conference on Medical image computing and computer-assisted intervention*], 234–241, Springer (2015).

Monitoring Marine Protected Areas using Data Fusion and AI Techniques

Konstantina Bereta[a], Aristides Millios[b], Konstantinos Chatzikokolakis[a], Dimitris Zissis[c]

[a]MarineTraffic, Athens, Greece;
[b]Dalhousie University, Halifax, Canada
[c]University of the Aegean, Ermoupolis, Syros, Greece

ABSTRACT

In this paper we describe a workflow for monitoring maritime activity in marine protected areas which makes use of open non-collaborative surveillance data. We developed a workflow that automatically searches for satellite images that cover marine NATURA 2000 areas which are then classified using a Convolutional Neural Network that is able to detect vessels in the areas of interest. Our neural network is trained using AIS data as ground truth.

Keywords: Maritime awareness, AIS, Convolutional Neural Networks, Classification

1. INTRODUCTION

Recent initiatives by the European Commission study the effect of maritime activities in protected areas, such as NATURA 2000 sites.[1–3] In this context, EU working groups, such as the European Economic Interest Group, have published studies aiming at raising awareness regarding the fact that certain maritime activities, such as fishing vessels that carry specific gear can sometimes have a negative impact on the conservation of biodiversity in protected areas (e.g., trawlers can cause damage to corals). Since the effects of vessel activities differ from area to area and depend on the respective species under protection, each EU member state is responsible for establishing its own regulations. In most cases, for each member state, the regulations come at the aftermath of severe loss of biodiversity due to certain maritime activities.

Maritime Situational Awareness (MSA) is defined as the effective monitoring of all activities in the maritime domain that could affect maritime safety and navigation. Sensor based data represent the main pillar of MSA and can be divided into two broad categories: cooperative and non-cooperative systems. Cooperative systems rely on the vessels crews collaboration to identify and report the vessels information, while non-cooperative systems are designed to detect and track vessels that do not provide such information voluntarily.

The most popular Cooperative system, is the Automatic Identification System (AIS) that allows for exchanging messages between ships equipped with AIS transponders, coastal stations and other ships, equipped with AIS receivers. Using AIS, both static and dynamic information about the navigation of vessels can be exchanged. Static messages contain information about the vessel, such as its MMSI, name, type, dimensions, draught, destination, estimated time of arrival, etc. Dynamic messages provide navigational information about the vessel, such as the current position speed, heading, course, and its navigational status. AIS was first introduced in order to enable vessels exchange navigational information with nearby vessels in order to avoid collisions. Since 2004, AIS is considered mandatory for all commercial vessels over 299 Gross Tonnage (GT) and for all passenger vessels regardless of their GT that travel internationally to carry a Class A AIS transponder (which transmits and receives AIS data) aboard. Smaller vessels can also be equipped with a Class B AIS transponder. This resulted in a large amount of information transmitted from vessels all around the world about their navigational status and therefore AIS is considered nowadays as one of the most powerful tools employed to increase the

maritime domain awareness in global scope. Despite the fact that AIS is a valuable source of information, it is often insufficient, as (i) there are areas with limited AIS coverage, (ii) most vessels involved in illegal activities turn off their transponder, and (iii) AIS is mandatory for only certain types of vessels, so there is part of the global fleet that cannot be monitored through AIS.

Non-cooperative systems which are are designed to detect and track vessels that do not voluntarily provide information, include coastal and high-frequency (HF) radar, active and passive sonar, ground- or vessel-based cameras (e.g. thermal), and satellite and airborne Earth Observation (EO) systems. EO systems can be divided into optical (generally visual and near-infrared) and Synthetic Aperture Radar (SAR) systems.

Copernicus* is a programme funded by the European Space Agency (ESA) with the aim of sending constellations of new generation satellites, named Sentinels, into orbit in order to monitor the Earth. The Sentinels, in collaboration with global network of thousands of land-, air- and marine-based sensors collect approximately 12 TB of data daily. In this way, Copernicus can be considered as the largest space data provider in the world. Copernicus images are publicly and freely accessible from the Sentinel Data Hub† and mirror sites. Sentinel satellites are equipped with different kinds of sensors. Sentinel 1 satellites are a pair of satellites equipped with SAR-C radar, providing high resolution radar images such that they are not affected by weather conditions.

Combining multi-origin information to determine relationships among the data is critical for MSA, as it improves the understanding of a current complex environment. The fusion of data produced by cooperative and non-cooperative systems contributes towards better coverage and robustness to failure, thus improving the reliability and quality of the situational picture.

Our approach is based on the fusion of different data sources, mostly AIS and satellite data, and on the employment of AI techniques such as neural networks. In the context of the work described in this paper, we will integrate AIS information combined with Sentinel 1 images. The motivation of the work described in this paper is the facilitation and automation of data analysis tasks for monitoring marine traffic in protected areas.

2. AUTOMATIC MONITORING OF PROTECTED AREAS USING AIS AND SATELLITE DATA

In this section we describe our prototype implementation that is able to automatically detect vessels in protected areas. In the scope of this paper we only consider NATURA 2000 sites as protected areas but other datasets could be supported as well. In a nutshell, we present a fully automatic approach that is able to monitor protected areas by searching, downloading and classifying images with respect to whether they contain vessels in protected areas. For the classification task, we employed a Convolutional Neural Network (CNN). For the training task, we implemented a fully automatic workflow which is based on the automatic annotation of satellite images using AIS data.

In this section, we first describe the data sources that are used in our approach, and then we describe in more detail the system architecture and workflows.

2.1 Data sources

A description of the data sources used in our workflow is provided below.

- Europe Coastline. A dataset provided by the European Environment Agency (EEA) that contains detailed geometries of the european coastline in shapefile format‡.

- NATURA 2000 areas. We used a dataset that contains information and geometry boundaries of European NATURA 2000 sites. The dataset is available in Shapefile format by EEA§. We refined this dataset using

---

*https://www.copernicus.eu/en

†https://scihub.copernicus.eu/

‡https://www.eea.europa.eu/data-and-maps/data/eea-coastline-for-analysis-1/gis-data/europe-coastline-shapefile

§https://www.eea.europa.eu/data-and-maps/data/natura-10/natura-2000-spatial-data/natura-2000-shapefile-1

the European boundaries, so we created a dataset that contains only Marine NATURA 2000 sites, i.e., sites that are located in or intersect with the sea.

- AIS data. We used AIS data provided by MarineTraffic[¶]. We used a subset of the attributes provided through AIS from each vessel: the vessel location at the time the satellite image was acquired, the speed, the heading, the course and the vessel type (e.g., cargo, tanker, etc.).

- Satellite images. We used Sentinel 1 images that are available online through the Sentinel Hub[‖]. Sentinel 1 images are equipped with a C-SAR sensor, so it provides radar images in high resolution regardless of the weather conditions.
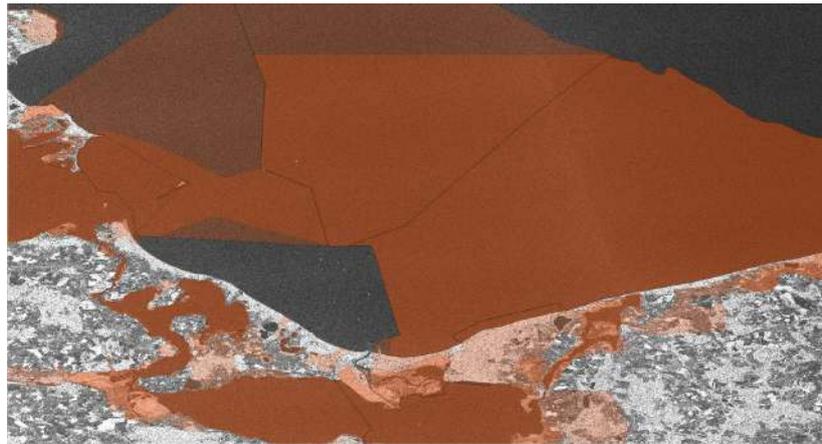
Figure 1. NATURA 2000 site in the Baltic Sea overlayed by a Sentinel 1 image, depicting vessels, some of which are located inside the protected area visualised in red colour

Figure 1 shows a Sentinel 1 image that covers an area in the Baltic sea. The NATURA 2000 sites located in this area are highlighted in red. Vessels can be seen in the satellite image as white spots. It can be observed that some of them are located into the protected area. AIS can provide auxiliary information about these vessels (e.g., type of vessel, speed, how long it stayed in the area, etc.)

2.2 System Architecture and workflow
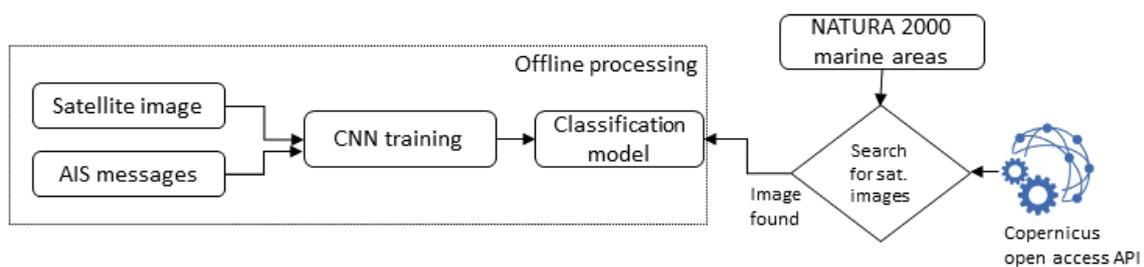
Figure 2. High level overview of the system workflow

We have created a system to monitor protected areas using Satellite images and AIS data. The workflow of the system is depicted graphically in Figure 2 and it consists of two main tasks: (a) the offline processsing task and (b) the online search and classification task. The workflow of the offline processing task consists of the following steps:

[¶]http://www.marinetraffic.com/

[‖]https://www.sentinel-hub.com/

1. First, Sentinel 1 imagery is downloaded from Sentinel portals, such as the Copernicus Data Hub and it is divided into tiles.

2. Next, a land masking step is performed. This means that all pixels depicting land get removed from the image, so that the reflection of the vessels will not get confused with the reflection coming from the land.

3. Then, the Sentinel image is correlated with AIS data. More specifically, we consider vessel positions that were located in the area depicted in the Sentinel image at approximately the same time that the image was acquired (We use a temporal window of five seconds for that).

4. The output of the correlation step is the automatic annotation of Sentinel image tiles: each image tile in the training set gets labelled with whether it contains a vessel or not.

5. Given the training set produced automatically from the previous step, a Convolutional Neural Network is trained using the Keras framework[**]. We experimented with a variety of different CNN architectures, and resulted to a CNN that consists of 3 convolutional layers and 2 fully-connected layers, each one followed by a ReLU activation layer. We also used Nadam optimiser, which is a variation of Adam optimiser enhanced with Nesterov momentum. The fully-connected layers are followed by drop-out layers in order to avoid model over-fitting.

The online classification task consists of the following steps:

1. First, the system automatically searches for Satellite images covering areas that overlap with the geographical boundaries marine NATURA 2000 sites. To implement this we use the Sentsat API[††]. Then, we automatically download the image and the metadata and divide the image into tiles.

2. Next, we proceed with a refinement step in which we filter out the image tiles that are not located in protected areas.

3. Last, each image tile is classified using the trained model produced by the offline processing task.

3. DISCUSSION

One of the main components of our system architecture is the convolutional neural network. Convolutional neural networks are neural networks that are known for their efficiency in classification tasks using images as input, that stems from the fact that they are able to capture relevant features from an image on different levels, in a way that is similar to the human brain. For this reason, there are a lot of efforts in literature that employ CNNs to classify satellite images.[4–16] The approach described in this paper is built upon these methods.

One of the limitations of the approach described in this paper is the fact the training set is automatically generated by correlating AIS and satellite data. However, vessels with no AIS signal transmitted might appear in the satellite image, creating errors in the automatically created training set, which if not corrected, might create more false negatives in the prediction step. In the context of these work, we used satellite images in areas with good AIS coverage to decrease the number of undetected vessels in the training set by eliminating cases when the vessels are out of coverage (and have not necessarily switched off their transponder on purpose). To eliminate the un-detected vessels in the training set even more, one can use the approach in a semi-automatic way and correct the respective cases manually by examining the tiles that appear to have no vessel.

In the context of this work we used NATURA 2000 dataset and Sentinel 1 images. However, other data sources (e.g., protected areas datasets published by governmental portals) can be used interchangeably. Sentinel 2 images can also be used instead of Sentinel 1 images. Since Sentinel 2 images are optical, instead of radar, they (i) provide more visual information about a vessel (i.e., colour), and (ii) they are able to depict moving objects more accurately, so the processing workflow does not need to take the azimuth shift into account. On the other hand, Sentinel 2 images are not weather-independent, in contrast to the Sentinel 1 images. The efficiency of the proposed workflow can also be improved by using Satellite images from commercial satellites (e.g, micro-satellites) as input.

[**]`https://keras.io/`
[††]`https://github.com/sentinelsat/sentinelsat`

4. CONCLUSIONS AND FUTURE WORK

In this paper we described a fully automatic approach for monitoring protected areas using AIS and satellite imagery. Our approach is based on data fusion from multiple sources (i.e., Sentinel 1 images, AIS, NATURA 2000 sites, European coastline) and also on the implementation of a Convolutional Neural Network that is used to classify the images that cover protected areas.

In future work, we plan to improve our data analysis workflow in order to be able to provide an estimation about the risk associated with the activity of certain vessels in the respective protected areas, taking into account the kinds of species that are protected in each NATURA 2000 site as well as the local regulations.

ACKNOWLEDGMENTS

REFERENCES

[1] Group, E. E. I., "Overview of the potential interactions and impacts of commercial fishing methods on marine habitats and species protected under the eu habitats directive." `http://ec.europa.eu/environment/nature/natura2000/marine/docs/Fisheries\%20interactions.pdf` (2013). Online; accessed 11 June 2019.

[2] Group, E. E. I., "Review of fisheries management measures in natura 2000 sites." `http://ec.europa.eu/environment/nature/natura2000/marine/docs/Review\%20of\%20fisheries\%20management\%20measures\%20in\%20Natura\%202000\%20sites.pdf` (2018). Online; accessed 11 June 2019.

[3] European Commission, "Guidelines for the establishment of the natura 2000 network in the marine environment. application of the habitats and birds directives." `http://ec.europa.eu/environment/nature/natura2000/marine/docs/marine_guidelines.pdf` (2007). Online; accessed 11 June 2019.

[4] Vieira, F. M., Vincent, F., Tourneret, J., Bonacci, D., Spigai, M., Ansart, M., and Richard, J., "Ship detection using SAR and AIS raw data for maritime surveillance," in [*24th European Signal Processing Conference, EUSIPCO 2016, Budapest, Hungary, August 29 - September 2, 2016*], 2081–2085 (2016).

[5] Mazzarella, F., Alessandrini, A., Greidanus, H., Alvarez, M., Argentieri, P., Nappo, D., and Ziemba, L., "Data Fusion for Wide-Area Maritime Surveillance," (06 2013).

[6] Guerriero, M., Willett, P., Coraluppi, S., and Carthel, C., "Radar/AIS data fusion and SAR tasking for Maritime Surveillance," in [*2008 11th International Conference on Information Fusion*], 1–5 (June 2008).

[7] Huo, W., Huang, Y., Pei, J., Zhang, Q., Gu, Q., and Yang, J., "Ship detection from ocean sar image based on local contrast variance weighted information entropy," *Sensors* **18**, 1196 (04 2018).

[8] Paes, R., Lorenzzetti, J., and Gherardi, D., "Ship detection using terra-sar-x images in the campos basin (brazil)," *Geoscience and Remote Sensing Letters, IEEE* **7**, 545–548 (07 2010).

[9] Grover, A., Kumar, S., and Kumar, A., "Ship detection using sentinel-1 sar data," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **IV-5**, 317–324 (11 2018).

[10] Mazzarella, F., Vespe, M., and Santamaria, C., "Sar ship detection and self-reporting data fusion based on traffic knowledge," *IEEE Geoscience and Remote Sensing Letters* **12**, 1685–1689 (Aug 2015).

[11] Iervolino, P., Guida, R., Lumsdon, P., Janoth, J., Clift, M., Minchella, A., and Bianco, P., "Ship detection in SAR imagery: A comparison study," (07 2017).

[12] Achiri, L., Guida, R., and Iervolino, P., "Sar and ais fusion for maritime surveillance," (09 2018).

[13] Carthel, C., Coraluppi, S., and Grignan, P., "Multisensor tracking and fusion for maritime surveillance," in [*2007 10th International Conference on Information Fusion*], 1–6 (July 2007).

[14] Lang, H., Wu, S., and Xu, Y., "Ship classification in sar images improved by ais knowledge transfer," *IEEE Geoscience and Remote Sensing Letters* **PP** (01 2018).

[15] Stasolla, M. and Greidanus, H., "The exploitation of sentinel-1 images for vessel size estimation," *Remote Sensing Letters* **7**(12), 1219–1228 (2016).

[16] Wang, Y., Wang, C., and Zhang, H., "Ship classification in high-resolution sar images using deep learning of small datasets," *Sensors* **18**, 2929 (09 2018).

Custom state-of-the-art CNN algorithm for ship detection and segmentation

Gianfausto Bottini[a], Marco Corsi[a], Filippo C. Daffinà[a], Simone Tilia[a], Torbjorn Stahl[a],
Dino Quattrociocchi[a]

[a]e-Geos S.p.a., via Tiburtina 965, Roma, Italy

ABSTRACT

The use of Deep Neural Networks (DNN) for the detection and segmentation of specific objects from satellite images is a topic of great interest today. This paper presents the results of the application of a CNN (Convolutional Neural Network) state-of-the-art algorithm, Mask R-CNN [4], for the detection and segmentation of ships from Very High Resolution (VHR) and High Resolution (HR) optical images starting from open-source and proprietary annotated VHR data-sets. After having tried several CNN models as Unet and SegNet, the result is that the Mask-R-CNN shows better results both for the detection of ships in the open sea and for detecting ships near the coasts or in case of difficult conditions such as presence of clouds and waves. We are improving the Mask-R-CNN method's detection capability by extending the training dataset using Data Augmentation rotating, flipping, cropping and zooming out images to enlarge the training dataset both on VHR and on HR.

Keywords: Deep Learning, Mask R-CNN, SegNet, Unet, Data Augmentation, Vessel Segmentation, Maritime Domain Awareness

1. INTRODUCTION

The increasing availability of satellite imagery by a heterogeneous set of sensors, results in the strong need to automatize the processing of the huge amount of input raw data, selecting the fastest and most accurate approach. In this context it is very difficult to find a unique solution able to maintain the same level of performance and accuracy over VHR images as well as over HR images. The adoption of already existing Deep Neural Networks, opportunely trained with specific training datasets, is showing the best results over the variety of data provided as input. The aim of the presented activity is to be able to automatically segment every type of vessel from optical images from different satellites using both VHR (Very High Resolution) and HR (High Resolution) sensors.

The original reference dataset on which the deep learning algorithms were driven consists of about 200,000 VHR 768x768 pixels images, 1.5 m resolution, containing carefully labeled ships of every type, and with a test set of about 15,000 specimens to which add more samples from other datasets. This dataset is one of the few freely available regarding ship segmentation. Having assessed the execution of three different existing CNN (SegNet [2], U-Net [3] and [4] Mask R-CNN) at the end Mask R-CNN was selected due to its accuracy boost in case of cloud, waves and coasts in ship detection and segmentation.

The dataset has been expanded with custom samples of HR sentinel2 images carefully labeled by our mapping team and ad hoc data augmentation by means of zoom out technique and the model target is recognizing the shape, the color and the background of the vessels despite the image sensor, HR or VHR.

The final workflow of the software consists of downloading an image, which can be from Sentinel2, SPOT or BlackSky, making predictions on it, detecting and segmenting vessels and producing a digital report containing several information for each of them like:

- length and width of the vessel and of the bounding box around the vessel
- Orientation

- Vessel area in square meters
- coordinates in longitude, latitude
- Classification of the ship in four categories A-B-C-D, depending on its size
- The date of acquisition

The results of the algorithm are stored in a KML file along with the final GeoTiff clip containing the predictions.

2. METODOLOGY AND RESULTS

As already mentioned in the previous section, 3 CNN models have been considered and implemented. SegNet [2] and Unet [3] are considered as the standard architectures of Convolutional Neural Networks with regard to the classification of images, i.e. class recognition plus object segmentation. Both SegNet and Unet are fully convolutional neural networks and provide pixel wise classification that for our target, from a visual analysis, are reaching satisfactory results in favorable conditions.

Mask R-CNN is a regional Convolutional Neural Network at the state-of-the-art and reaches very high accuracy both on training and validation sets and, from a visual analysis on the test set, in comparison with the previous two methods it shows a significant improvement in unfavorable conditions (clouds, coasts, vessels close to each other etc.).

Approximately 90,000 images have been adopted as training set along with some custom HR flagged samples, divided into training and validation sets in the 80-20 proportion and using as test set samples coming from GeoEye, SPOT, DOTA, Black Sky and Sentinel2.

Regarding HR case, a synthetic dataset has been created degrading VHR images and masks by means of Pyramid reducing, an algorithm often used in classic Computer Vision tasks taking resolution pixels from 1.5 m to ~10m, similar to Sentinel2.
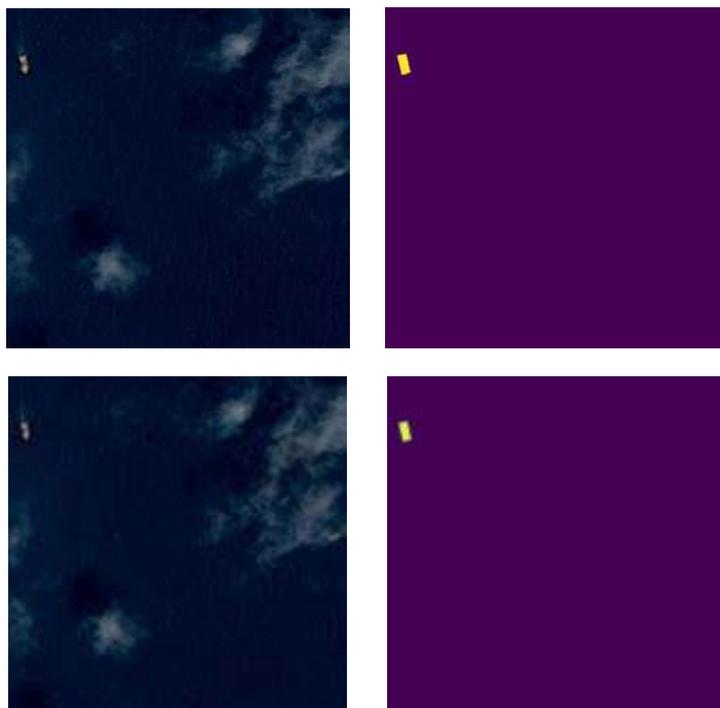
Figure 1. Top-left: VHR image, Top-right: VHR mask, Bottom-left: Synthetic HR image, Bottom-right: Synthetic mask

For Segnet [2] and Unet [3] the function to be minimized is the Binary Cross-entropy that outputs the probability of each pixel being part of class 0 and class 1. The Adam Optimization Algorithm was selected for optimizing the results, instead of using the classic Stochastic Gradient Descent algorithm, and models were trained from scratch with no Transfer Learning Technique

The Mask R-CNN [4] model was initialized with the weights obtained on the MS COCO[6] dataset. For this model, the application of Transfer Learning is used to make the network more efficient in less time. The model minimizes the loss on the class, bounding box and segmentation, and is therefore a linear combination of the 3; The optimizer is the Stochastic Gradient Descent with appropriate values for momentum and learning rate parameters.

Data Augmentation has been an important window of the final model improving and increasing the dataset by several thousands of samples such that the model is customized over the target to segment images coming from any source.

The metrics used to evaluate the accuracy for all three methods for the VHR case is the harmonic-F1-score with the following results:

Table 1. Harmonic-F1-score.

Model	F1 score train	F1 score validation
SegNet	0.7	0.6
Unet	0.79	0.69
Mask R-CNN	0.84	0.75

The confusion matrix is used to have a general idea about the number of ships detected in a validation set of about 15,000 samples as follows:

Table 2. Confusion matrix.

	Ship forecasting	Non-ship forecasting
Ship real	80 %	1%
Non-ship real	1.2%	70%

where the fourth cell is considered as the percentage of cases in which the algorithm predicts just the exact number of vessels.

Once the training is over, the detailed software online workflow is as follows:

- The image is downloaded from its source and is cut into sub-tiles, the size of the sub-titles is a parameter given by the operator based on the size and the source type.
- Each sub-tile is processed by the neural network, previously trained, creating a map of vessels identified within the sub-tile.
- Each of the identified ships is fitted into an ellipse that best characterizes it. Usually, the input images are in the GeoTiff format helping us to shift pixels into real measures like length, width, area in square meters and providing us the coordinates in longitude and latitude.
- Each sub-tile is then merged with the others into a final prediction and a final digital report containing all this information.

The final goal is to have a continuous Maritime Awareness picture with precise and reliable information in near real time. The processing time of the algorithm is less than 1 minute for a Sentinel2 image of 10,000 pixels by 10,000 pixels.
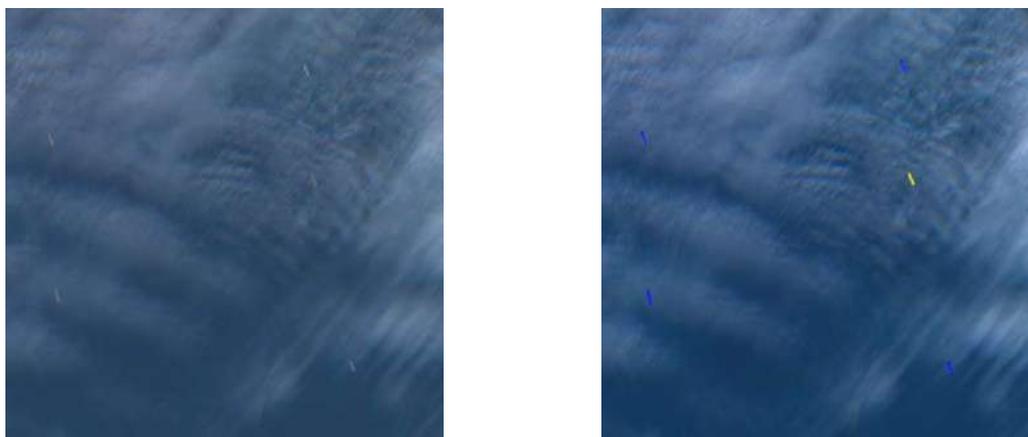
Figure 2. On the left the raw image (Sentinel2), on the right the forecasted image

In Figure 2 above is illustrated an example starting from raw data, a Sentinel2 image in open sea close to the Khark island. The image is cut and processed by means of the Neural Network weights, and each sub-tile has its prediction that then are put together in a final prediction over the entire image as it is shown on the right; an example of the final digital report is given below. The result is stored in a KML file, which opened on the Google Earth portal gives the following information:
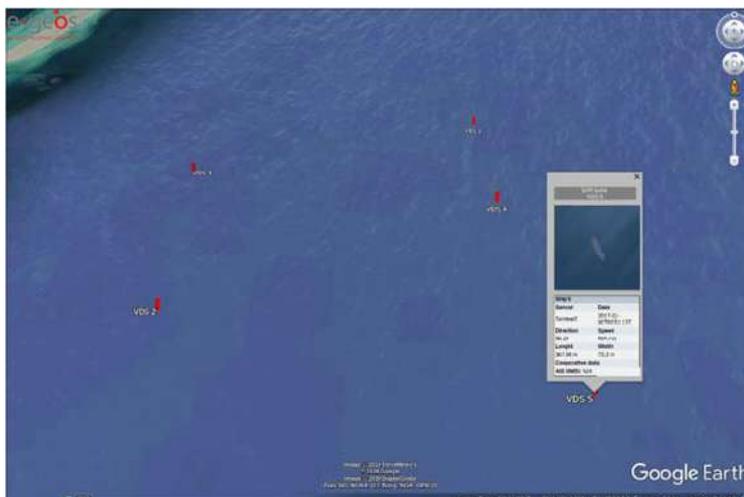
Figure 3. The digital report consists of a KML file that can be displayed in Google Earth with exact vessels position

Figure 4 gives some examples of forecasting over images coming from several different VHR sensors such as SPOT, DOTA, Google Earth.

Figure 4. Example of vessels segmentation on Satellite acquisitions coming from several different VHR Sensors

3. CONCLUSIONS

Results have been very promising and comforting and have been showed in different presentations and workshops including the "Ital-IA" [1] national conference about Artificial Intelligence for Space but the project flow is not over. Preliminary application and operational validation of the selected method will be performed during the next maritime trials in the context of H2020 project MARISA [6]. Our future challenge will be to enlarge the HR dataset with synthetic datasets by means of Generative Adversarial Networks, as well as by degrading VHR images to HR by means of custom filters, to be able to entirely drop the manual work on HR. The work of generating the synthetic dataset has already begun with promising results. Although the Mask R-CNN[4] weights trained on the VHR dataset give fairly acceptable predictions also on the HR dataset, drawing a new network on a flattened HR dataset (not existing today) would give much better results for these datasets. The idea is to follow the concept of particular generative neural networks called CycleGAN [5] that transfer style and ownership from one image dataset to another. The goal is to create datasets of HR, simil-Sentinel2 and Landsat8-like images, starting from VHR datasets and degrading images, leading to labeled HR images. A CycleGAN [5] consists of three networks, two generative and one discriminatory. The first to convert a VHR image into an HR, the second to convert an HR into a VHR and the third to identify whether the image is HR or generated. We want to create a continuous flow of generation and discrimination of satellite image data of all kinds, based on the latest models of artificial intelligence, and to be able to generate, govern, recognize and label every type of image coming from every type of satellite, which is one of the goals of every company involved in the Earth Observation field.

REFERENCES

[1] Bottini, G., Corsi, M., Daffinà, F. and Quattrociocchi, D., "Analysis of Deep Learning models for Ship Detection on VHR and HR optical satellite images", Ital-IA (2019).
[2] Badrinarayanan, V., Handa, A. and Cipolla, R., "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling", Badrinarayanan (2015).
[3] Ronneberger, O., Fischer, P. and Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation", Ronneberger (2015).
[4] Ke, K., Gkioxari, G., Dollàr, P. and Girshick, R.,. "Mask R-CNN", He (2017).
[5] Zhu, J. Y., Park, T., Isola, P. and Efros, A. A., "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", Jun-Yan Zhu (2018).
[6] COCO dataset, "COCO – Common Objects in Context", <http://cocodataset.org/#home> (2019).
[7] H2020 MARISA project website, "MARISA - Maritime Integrated Surveillance Awareness", <http://www.marisaproject.eu> (2019).

Industry is the key for R&D transition to operational end-users - An innovative MSA project example powered by AI, led by CS

M. Degache[1], L. Gardenal[1], J. Charles[1], V. Marié[1], B. Soloch[1], O. Tran[1]

CS, 230 rue Marcellin Berthelot – ZI Toulon Est La Garde –BP 68 83 079 Toulon cedex –France

F. Bouchara[2]

Aix Marseille Univ, Université de Toulon, CNRS, LIS, UMR 7020

ABSTRACT

An industrial company is a key element between R&D laboratories and operational world. Mixing people from both worlds and adding specific competencies will lead to high performance specific products more readily available.

CS Group, with the help of its internal expert network, is involved in the development of innovative solutions for the defence domain. With its partners, CS develops tools and systems to create or improve existing solutions and adapt them to new emerging needs such as, in the maritime field, robust methods for ships detection and recognition. As an example CS investigates artificial intelligence solutions to reduce false alarms rate and support military navy operators in the context of the fight against asymmetric threats. All these efforts aim at developing effective solutions for maritime security supported by operational people and researchers inside and outside the company.

Keywords: Industry, R&D, operational, innovation, artificial intelligence (AI), maritime, detection, recognition.

1. INTRODUCTION

As global threat is increasing, it is more and more necessary to keep a technological advantage against potential enemies. Data processing is speeding up process and our adversaries are able to go faster than us: they have no constraints on funding and regulations.

The best way for NATO Nations to keep their advantage is to rely on fundamental research made in laboratories and universities. But there may be an important time gap between first discoveries and the actual use of newly designed operational systems. Industry is a key in this process: it has to design and produce equipments. Having companies involved both in R&D and in operational fields contributes to a quicker climb of the TRL (technology readiness level) scale.

2. CS INDUSTRIAL VISION: FROM R&D TO OPERATIONAL SOLUTION

2.1 Profile

CS is a mid-cap company employing 2000 persons, with subsidiaries in Europe, North-America and Middle-east. Its four business units (Space, Aeronautic, Energy and Defence) are dedicated to design, build and conduct critical and cyber safe systems. CS solutions are mostly software based and involve topics as various as digital communications, simulation, command & control, big data, artificial intelligence, etc. The company addresses mainly highly innovative projects in its field of expertise where it may have the upper hand given that small companies are too small to conduct them, and big companies not agile enough for fast innovation. Its main customers are government agencies and big industrial companies.

2.2 R&D and operational views

CS dedicates 10% of its revenue to R&D. CS human resources includes people from the academic world with high level degrees of education. The company promotes and relies on their unique abilities to advise and go beyond the state of the art solutions in key strategic sectors. To guaranty the level of excellence a dedicated branch of "experts" has been created to promote their skills and support constant training and qualification. Similarly the staff includes former French Army

---

[1] marianne.degache@c-s.fr; lionel.gardenal@c-s.fr; jehan-christophe.charles@c-s.fr; vincent.marie@c-s.fr; commerce.marine@c-s.fr
[2] bouchara@univ-tln.fr

officers or petty officers, giving insight on the operational point of view to design and operate optimized solutions in the Defence Industry.

2.3 Innovative industrial project: cooperation between main actors and development cycle

CS is engaged in being an active member of selected competitiveness clusters enabling efficient local but also international cooperation. Participating or running projects labelled by these clusters give a good overview of the technological innovation as well as the opportunities of teaming up with non-industrial partners i.e. universities or research centres, but its requires a particular expertise.

Conducting the project by itself is done with classical industrial process: a full V cycle, from conception, specifications, tests and performances verifications. CS is used to conduct these projects with flexibility, diplomacy, persistence. This expertise is at the service of global view for company development, like area surveillance, with maritime surveillance as a branch of it.

3. MSA R&D PROJECT USING AI

3.1 AI use for MSA

3.1.1 History

CS has been involved for more than 15 years in safety and the protection of critical restricted-access areas. Since 2006, CS agency in Toulon has developed solutions such as MWPS (Maritime Warning & Protection System) or SECMAR (SECurity MARitime) using an operational centre and a multi-sensor centralized system to secure maritime area and harbour. The SEMAPORT system, is deployed in particular in the Bay of Toulon, to protect the first French military harbour, the first to use a system of this type; this system detects and automatically monitors suspicious vessels, using radar pictures, day and night optronics systems, AIS... Maritime surveillance and in general, Maritime Situational Awareness (MSA) is indeed a major challenge for the protection of Maritime areas of interests in the framework of the fight against piracy and terrorism.

3.1.2 Roadmap

Image processing based on visible or infrared video streams have been developed by CS for the design of early warning and tracking systems against surface targets. Following both the technological evolutions and operational requirements, major efforts have been dedicated to manage multiple video streams and create graphical user interfaces (GUI) displaying the maritime picture within a Command and Control (C2) system. Besides algorithms have been implemented following classical solutions for the segmentation and tracking of the target against its background. Despite good results, it became obvious that new robust approaches were required to address needs such as the ones of the French Navy (increase of automatic processing).
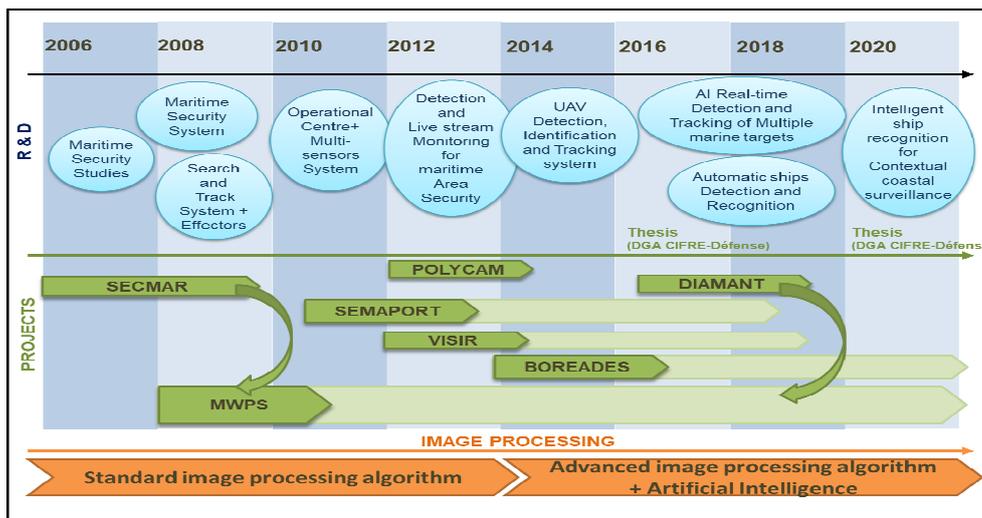

Figure 1: CS Toulon past and current R&D roadmap for maritime security

The above figure shows the innovation push of the past decade and the industrial experience acquired by CS Toulon through its numerous projects[2] in that field. It emphasizes also the step taken in advanced image processing techniques.

Algorithms using artificial intelligence are the current preferred way for these system upgrades. In 2016, CS has invested in a thesis [1] exploring a brand-new temporal and spatial approach for detecting maritime objects. In cooperation with Toulon University, this work has illustrated how image processing on visible wavebands using a rigidity criterion and convolutional deep learning is efficient to separate object detections from the sea background ([2], [3]).

The thesis results have been integrated in the DIAMANT project described in the following section and in turn DIAMANT may contribute to CS global MSA solutions.

3.2 DIAMANT

3.2.1 Initial project Goal

In a partnership with the LIS research centre from Toulon University[3], CS has worked on new methods in order to upgrade the performance of the maritime detection system in the framework of naval combat systems against asymmetric threats. The DIAMANT (Détection Image Adaptée aux Menaces Asymétriques Navales & Terrestres) project was created with the following scientific main objective: develop and consolidate innovative solutions integrating convolutional neural networks for the detection and recognition of boats positioned on the coastal fringe such as asymmetric threats (go-fast, speedboat ...). The operational goal is to improve the performance of threat detection systems for combat systems (military component) and security systems of industrial and port sites (civil component) and unburden the work of the operator with automatic ship classification.

The DIAMANT initial project has been supported by the French governmental agency DGA (Direction Générale de l'Armement) in France.

3.2.2 Scientific work

The proposed approach to ship detection and classification is mainly based on the Fast -RCNN algorithm proposed by Girshick et al. [4]. By factorizing computations between neighbouring region proposals in the RCNN feature space, Fast-RCNN has both reduced the running time and improved its overall detection accuracy. Nevertheless, generation of high-quality region proposals still remained a bottleneck as it is in essence a combinatorial problem. The chosen method takes advantage of the rigidity of real-world objects ([2], [3], [5]), for quickly generating high-quality region proposals for Fast-RCNN.

The proposed approach to rigidity analysis in a high definition maritime video stream begins with the extraction of SIFT [6] key video locations in each new frame. The SIFT descriptor is, first, convolved with a 1D Gaussian profile ($\sigma = 2$), then, transformed into a probability distribution via its mere division by a normalization constant, finally, stored in a discrete histogram of 128 bins. SIFT keypoints are tracked individually across 4K video stream using optical flow and a Kalman filter. The state vector of the Kalman filter denoted by $Y_t = (x_t , y_t , u_t , v_t)^T$ consists of a mere concatenation of the subpixel position (x_t , y_t) of a SIFT keypoint and its 2D velocity vector (u_t , v_t) of which one may only observe a noisy version denoted by Z_t. SIFT keypoint classification as object versus background is based on the estimation of the long-term variation of the dynamic of the SIFT descriptor. To this end, we use the modified version of the Kullback-Leibler pseudo-Distance (KLpD) for deriving a measure of normalized SIFT keypoint textural variation. Suppose two probability distributions P and Q, and define the divergence measure between P and Q as

$$D_{KL}(P,Q) = \sum_i D_i(P,Q)$$

where one has $\forall i = 1, \ldots , 128$:

$$D_i(P,Q) = \inf \left\{ \left\| P(i)\log\left(\frac{P(i)}{Q(i)}\right) \right\|, \left\| Q(i)\log\left(\frac{Q(i)}{P(i)}\right) \right\| \right\}$$

A real-time implementation of above formula is achieved keypoint-wise by simply propagating it across frames via the moving average trick according to the following update scheme:

---

[2] Projects given in Figure 1 do not correspond to an exhaustive list.
[3] http://www.univ-tln.fr/Laboratoire-d-Informatique-et-des-Systemes-LIS.html

$$\hat{D}_{KL}(t) = \begin{cases} D_{KL}(P_1, P_2), & \text{if } t = 1 \\ \alpha D_{KL}(P_t, P_{t-1}) + (1-\alpha)\hat{D}_{KL}(t-1), & \text{if } t > 1 \end{cases}$$

where $\alpha \in (0, 1)$ stands for a real parameter which we have experimentally tuned to *0.1*. The final step consists in classifying every tracked SIFT keypoint either as object or background based on its estimated value of $\hat{D}_{KL}(t)$ using the expectation maximization (EM) algorithm. Once rigid video locations have been extracted, the next step consists in generating in their vicinity a set of high-quality region anchors. This is achieved in the present approach in a similar way to Faster-RCNN by computing three to nine boxes corresponding to different scalings of a square box centered around each rigid keypoint. Fig.2 shows some video-based object detection and recognition results using 4K maritime video. The estimation of the exact class of the maritime object (e.g. sailboat, catamaran, speedboat, etc) is achieved within a 15 ship classes set of useful maritime vessels plus the background class by using a ZF-net [7]. The latter class accepts any object which is not a vessel.
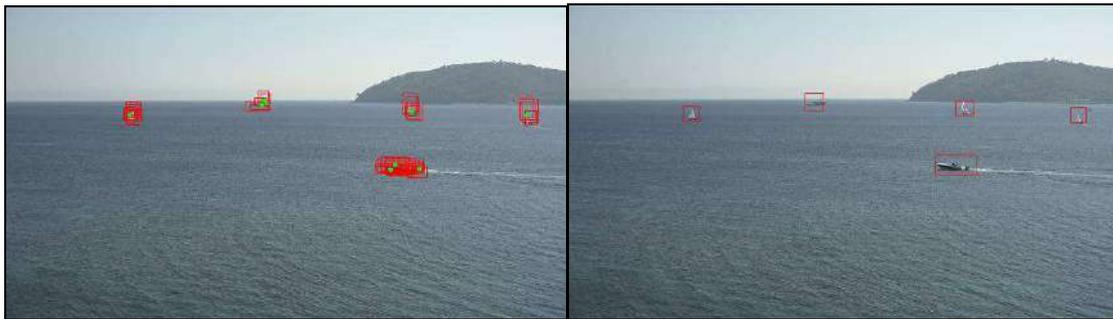
Figure 2: R&D results using DIAMANT 4K acquisition system:
Left: Proposals (in green are shown extracted rigid keypoints);
Right: Maritime object detection.

3.2.3 Workflow

This section (and the corresponding figure below) illustrates the main process workflow in the DIAMANT project. From a real-time video flow or by reading a recorded video sequence from a database (both in 4K format), the main tasks of the DIAMANT project are focused on:
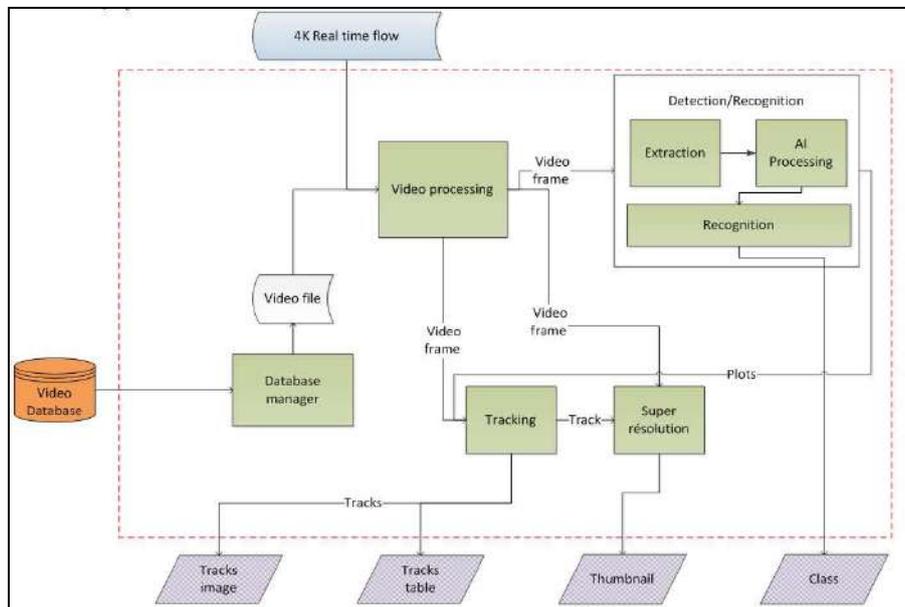
Figure 3: The DIAMANT workflow

- Real time processing: A 4K video flow (chosen to increase detection and recognition) is acquired and processed in order to extract frames. [Video processing module]

- Automatic detection of surface targets: as described in previous part, ships are detected and their recognition class is outputted. [Detection/Recognition module]

- Automatic recognition: DGA symbol is proposed to indicate directly on the screen what class does the ship belong to. [Detection/Recognition module]

- Tracking: conventional tracker is modified by using rigidity analysis. [Tracking module]

- Super-resolution: new approach integrating CNN algorithm is tested to improve image quality. [Super resolution module]

3.2.4 Results

Started in 2016, the DIAMANT project led by CS finishes in 2019, its main results are:
- Design of a 4K image processing chain
- Automatic detection module based on a spatio-temporal processing of the image information and plots selection according to their belonging to an object class
- An automatic recognition processing based on a convolutional neural network approach
- Analysis tools and GUI to select tracks for vignetting, digital zoom and history
- Superesolution method and tool using convolutional neural network, on each thumbnail
- Many real time trials session in the Oursinières harbour located in Toulon area, with cooperative and boats of opportunity.
- "DIAMANT" software integrating the different modules within a synthetic GUI.

The following figure shows an example of results with the DIAMANT tool. Note that the user interface has been implemented in accordance with the ergonomic vision of the DGA, following the goal to present the most synthetic view to the operator. So, the main part of the GUI is a display in real time of 4K video stream recorded by the 4K camera. All ships tracks are added in this display. It is possible to select one to four tracks and display them in thumbnails. DGA classification symbol are displayed for each track. Track trajectory history is added if selected for appearing in a thumbnail.
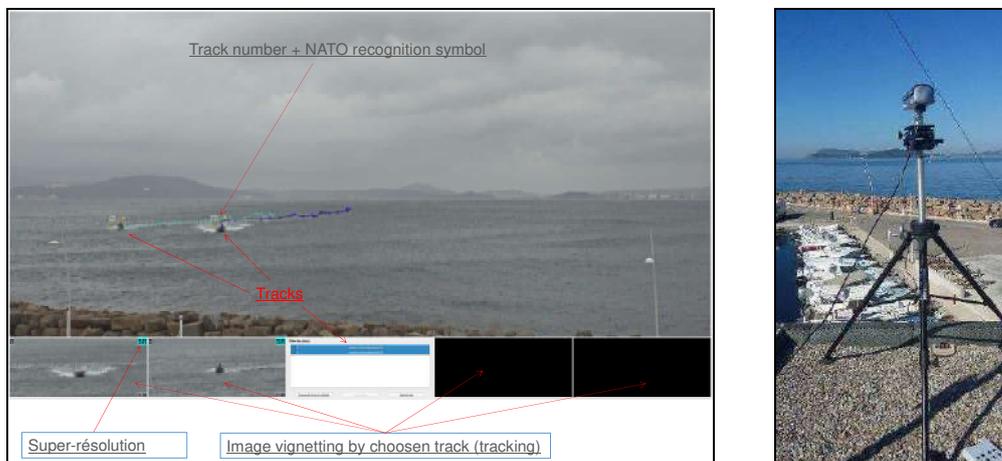
Figure 4: Left : DIAMANT tool GUI, Right : Set up example of the 4K camera during trials.

In this sample result, the ships – two speedboats - come from a 1500 m distance. The 4K camera is positioned at a 15 m altitude above the sea. The next two figures show the capacity of DIAMANT to discriminate two different ships.

Figure 5: Automatic ship track and recognition

Figure 6: Speedboat (track 1) and sloop (track 0) thumbnails with tracks list
(Note: The SR button triggers the super resolution process and the zoom factor - here equal to 1.0)

Next figure show how DIAMANT can operate in non-zero state with a low false alarm rate.

Figure 7: Ship detection in a non-zero sea state

3.3 R&D Follow ups

CS aims at continuing the research with the outcomes of the DIAMANT project. As R&D follow-ups, CS intends to tackle new approaches dealing with infrared images, behavioural analysis supporting an improved classification as well as the development of indicators giving the level of confidence in the processing.

3.3.1 Study on automatic ship recognition in infrared bands

Automatic analysis from imagery is meant to be a great help for operator, even more as the quantity of displayed information grows. The DIAMANT project, using a visible 4K camera, is nevertheless limited to day situations. To address night conditions, CS has already conducted a study to select a new method using infrared imageries for automatic ships recognition.

Convolutional Neural Networks are one of the most practical and commonly used models in the field of deep learning and are highly applicable for analysing imagery, making it the most suitable method for automatic recognition. There is but one major problem coming with this technology: the model needs to use a very large amount of data (thousands of samples) to feed the learning process.

To cover this need, open source data are readily available as far as imagery in the visible band is concerned but it is a different case for infrared images. As a result we cannot rely on an existing good quality infrared images database (containing the specific data we need) to train our network. As a workaround, CS use its own infrared image simulation software (Hesperides [8]) to build a specific synthetic database based on requirements for representativeness (hundreds of "infrared boats" viewed from different angles and distances). Moreover to reduce the volume of data required, we decided to use a more cost-effective deep-learning method and followed the recent method proposed by [9].

The architecture selected as a result of this study should be similar to DIAMANT with a combination of convolutional layers and pooling layers to extract the hidden features, then a fully connected layer to classify the images based on the extracted features. The difference lays in the training method as we use Extreme Learning Machine instead of the usual backpropagation. We believe that this method will allow training each layer fast and individually and producing a model with very good generalization capacities when fed with relatively few data.

3.3.2 Research in contextual coastal surveillance

CS is currently investigating a new research theme making use of contextual coastal surveillance. Plans are thus made to take part of a thesis (between 2020 and 2022) which title is « Intelligent ship recognition for contextual coastal surveillance ». The proposed work looks for solutions involving automatic and tailored ship recognition depending of the contextual situation.

4. MSA FROM INDUSTRIAL POINT OF VIEW

Having a good assessment of a Maritime Situation is known to be a huge challenge. Tens of thousands of various ships and boats, crews, passengers, goods are travelling days and nights all year round. Comparing with a few years ago, the problem is not to acquire data: the problem is to deal with a big mass of information, right or false, accurate or not, aged or not. As it is still the case in a lot of places, classifying and identifying a ship is a matter of a single person using binoculars or relying on an AIS spot on a radar screen. The quality and accuracy of information resulting depends on operator qualification and experience, ship's good will, and flair of people in charge of security.

With new technologies now available, industry is able to propose solutions for managing big data, supporting classification and data fusion as well as presenting a better clearer maritime picture to decision makers. Maritime Situational Awareness key challenge is no more the acquisition of data but the assimilation of such data for a given object and the verification of their quality so as to raise an alert when there are discrepancies. For example, AIS position received by a sensor is to trigger a series of checks, with last know position, draught coherence, picture verification with the closest camera DIAMANT software fitted, etc. In the digital age, technological solutions for MSA may rely on a big network able to automatically exchange information. Administrative data may be stored in a cloud; sensors can integrate Artificial Intelligence (AI) to provide more relevant pieces of information. Tracks histories are to be made available for ship lifetime in order to have comparison elements. The system must be focused on the decision making process: from the large amount of data available, the decision maker must receive the best synthesis, with confidence level and proper operational indicators depending on context. Such indicators can be a picture, a picture comparison, previous track course 5 minutes, 1 day, 1 year ago, etc.

CS has developed some of these technologies, or is able to find partners for that. The company is involved in finding and proposing solutions. Actors' roles are well defined: scientists work on data processing techniques, operational people and politicians have to decide what they want, and industrials have to design and build the system.

5. CONCLUSION

CS has experience in harbour protection, coastal surveillance, and, with its subsidiary Diginext, high sea surveillance (STRADIVARIUS radar). All these elements can contribute to surveillance from EEZ boundaries to pier.

In this paper, we have described the industrial vision of our group and specifically CS Toulon as far as maritime security and related R&D are concerned. A specific focus has been given on research activities making use of AI in a MSA context. DIAMANT project has contributed to our roadmap and R&D follow ups are already identified (in other wavebands, using other techniques and metrics, adding behavioural analysis). Now that a first level of operational and technical requirements is specified, components can be designed for integration in operational systems improving their resulting performance. The easy connexion between lower TRL R&D works and operational feedback, the capability to experiment in realistic environments and the use of industrial processes and methods help the acceleration of the time to market and hence improving faster the security level of end users.

CS drives at contributing further more to Maritime Situational Awareness and go beyond what we have achieved so far, including bid data management, smart mapping, AI use in sensors, in our C2 ground stations or in clouds. Besides being integrator, CS develops also specific technologies when there is a need to fill a gap. Relying on strong partnerships and efficient system architectures enable this approach with reasonable costs and minimum manpower. *In fine* however, human being, with its ability to decide in grey zone, takes the final decision. Digital tools are only there to help, to speed the process, to deliver a better and more complete situation, to close definitely open backdoors to threats, whatever they are.

REFERENCES

[1] V. Marie, "Real-time detection and tracking of multiple marine targets for situational assistance", Toulon University, 2020

[2] V. Marie, I. Bechar, F. Bouchara. Towards Maritime Videosurveillance Using 4K Videos. In: Basu A., Berretti S. (eds) Smart Multimedia. ICSM 2018. Lecture Notes in Computer Science, vol 11010.

[3] V. Marie, I. Bechar, F. Bouchara. Real-time maritime situation awareness based on deep learning with dynamic anchors. 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp 1-6, 2018, DOI: 10.1109/AVSS.2018.8639373

[4] R. Girshick, S. Ren, R. Girshick, and J. Sun. Fast r-cnn. In ICCV, 2015.

[5] I. Bechar, T. Lelore, F. Bouchara, V. Guis, and M. Grimaldi. Object segmentation from a dynamic background using a pixelwise rigidity criterion and application to maritime target recognition. In ICIP, pages 363–367, 2014.

[6] D. G. Lowe. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60(2):91–110, 2004.

[7] M. D. Zeiler and R. Fergus. Visualizing and understand-ing convolutional neural networks segmentation. In ECCV, 2014.

[8] Gardenal L., Tedeschi G., Faijan F., Fauqueux S., Lapierre F., "an optronic sensor performance prediction tool in the stochastic and spatiotemporally variable marine environment", OPTRO, feb. 2018.

[9] Khellal A., Ma H., & Fei Q. (2018). Convolutional Neural Network Based on Extreme Learning Machine for Maritime Ships Recognition in Infrared Images. Sensors, 1490.

Automatic recognition of underwater acoustic signature for naval applications

E. Artusi[*a, b], F. Chaillan[b],

[a]MINES ParisTech PSL Research University, 1 rue Claude Daunesse, 06904 Sophia Antipolis, France; [b]NAVAL Group, NAVAL Research, 83190 Ollioules, France

ABSTRACT

In a military context, where human capabilities are no longer sufficient to process quickly and reliably an ever-increasing amount of data, the implementation of algorithms based on Artificial Intelligence (AI), through the computing power of modern infrastructure, increases the ability to interpret and correlate massive heterogeneous data. This article will present an original automatic underwater acoustic signature recognition technique. The experiments are carried out from public underwater acoustic dataset. Besides, the performance of three architectures based on Mel-frequency cepstral coefficients (MFCCs), Machine Learning techniques and neural networks are compared.

Keywords: military naval application, mission, underwater acoustic signature, signal processing, classification, Machine Learning, Deep Learning

1. INTRODUCTION

The French Navy's surface warships and submarine vessels are designed to carry out various missions as for instance, maritime surveillance, infrastructure protection, participation in interoperability exercises and intervention on a theatre of operation. Thus, the conduct of a naval mission requires the simultaneous integration of a large number of information related to the ship and to its environment, such as sea state, meteorology, AIS data [1], [2], kinematic data and so on.

In addition, a mission in an operational context is obviously subject to the notion of risk and must therefore be reconfigurable in order to being able of managing many contingencies [3]. It is in this context that the ship's commander, in order to make the mission a success [4] must quickly analyze large quantities of massive and heterogeneous data in order to take the right decision depending on the situation. Nowadays, the indicators leading to decision-making are estimated by human beings based on their discernment and expertise. However, human capabilities are no longer sufficient to reliably and quickly process the amount of data collected by the fleet and its respective environment's many sensors. In other words, data volumes are forever increasing while the operational constraint like speed or efficiency used to achieve the mission remains.

For example, during naval missions, the capability to recognize underwater acoustic signatures is essential to be aware in real time of the evolution of the fleet's environment. The operator dedicated to sound analysis identifies and lists underwater acoustics noises of interest among all the detections. Then, as soon as possible he advises the Commander in case of classification of acoustic signature interpreted as a threat. Besides, this is an important function as it allows distinguishing between mechanical noise from a foe's vessel and normal biological activities issued from the underwater landscape; in practice it is difficult to automatically separate similar sounds as well as a human ear.

This fact, associated with the need of rapid classification might drive this operator into a mental overload and a very stressful discomfort. In order to provide him efficient assistance for identifying potential threats despite of fatigue, mental overload and stress it is necessary to design an automatic underwater acoustic signature recognition system.

---

[*] eva.artusi@mines-paristech.fr, eva.artusi@naval-group.com

Despite automatic sound recognition being an important aspect in emerging civilian applications, recognizing efficiently sounds in noisy underwater conditions for military applications and classifying these sound events from a passive device remains a serious difficulty. Actually, the recognition of some everyday real-life sounds such as ice breaking, the barking of a dog, music, or more generally any other acoustic event swamped by noise as well as several sound sources contained within the same recording [5] or bird species based on their song [6], are some of the emerging civilian applications. They ensue from the sound propagation' specificities, the diverse interesting signals' representability like for example noises issued from a submarine. Classification methods inspired by the field of speech recognition using audio characteristics such as MFCCs coefficients [7], [14], which effectiveness for sound classification have been positively shown for civilian applications, should be of interest for military purposes, modulo some upgrades.

This article focuses on the creation of a database, the establishment of classes and on the selection of architectures including machine learning (ML) and deep learning (DL) techniques. It also focuses on a performance assessment to quantify the relevance of our suggested solutions regarding underwater sound recognition designed for military naval applications. This paper is organized as follows: in section 2, the three architectures we will experiment to classify underwater sound recognition are introduced. In section 3, we describe our dataset made up of public sounds, the different configurations we have studied, and the results obtained for each case. Our conclusions and future works are presented in section 4.

2. METHODOLOGY

Historically, acoustic event detection (AED) was processed with characteristics such as MEL Frequency Cepstrum Coefficient (MFCC) combined with classifiers based on Gaussian Mixture Model (GMM), Hidden Markov Model (HMM) or Support Vector Machine (SVM) techniques [8], [9], [10]. More recent approaches use deep neural networks including convolutional networks [11] and recurrent networks [12]. In this article, we will describe the following three architectures:

- MFCCs and SVM,
- VGGish [19],
- An original enhanced version: VGGish and one dense layer.

In the following, after describing the three of them we will evaluate their performance.

2.1. Architectures description

In accordance with the objectives of this study, instead of classically setting up the hyperparameters, which is done, in practice with 5-fold cross validations to improve the reliability of the estimated hyperparameters, here we simply want to show that at least one of the three following architectures would be suitable for military naval applications.

2.2.1. MFCCs features and SVM method

For sound recognition, a feature extraction technique that extracts both linear and non-linear features is required. That is why here the Mel-frequency Cepstral Coefficients (MFCC) are used [23]. The MFCC is a type of frequency representation of the signal, in which any linear frequency f is mapped to the MEL scale according to the non-linear transformation:

$$m(f) = 2595 \log_{10}(1 + f/700) \tag{1}$$

Inversely, each MEL frequency is mapped to the frequency scale according to the following relation:

$$f(m) = 700 \left(10^{m/2595} - 1\right) \tag{2}$$

Consequently, instead of a linear scale we have now a new MEL scale providing rescaled frequencies in accordance with human ear behavior:

- An almost linear scale for frequencies less than 1 kHz, as in this case the transformation can be reduced to:

$$m(f) \approx 3.7f \tag{3}$$

- A log scale for frequencies above 1 kHz.
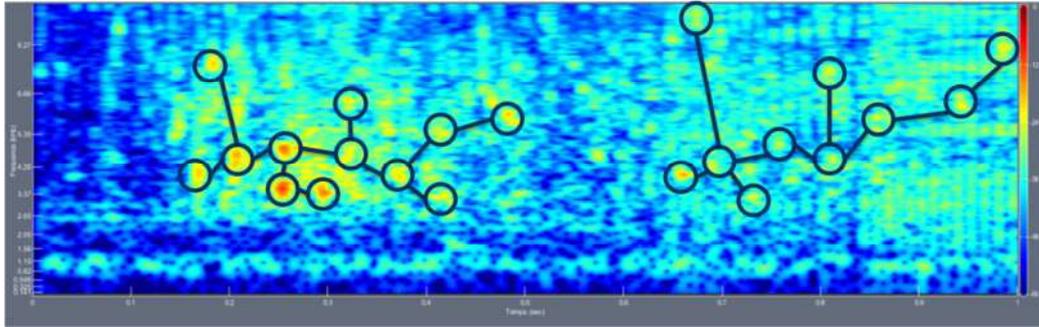
Figure 1: Illustration of the capability of MFCC's coefficient to identify thin details of underwater acoustic signal. X-axis correspond to Time (horizontal) and Y-axis corresponds to Mel frequencies (vertical).

As shown by Figure **1**, the MFCC transform is capable to capturing the important characteristic of audio signals. MFCC is widely used in speech recognition and in sound event recognition [13]. Thus, from any audio signal the MFCC's coefficients are computed according to the processing described by Figure 2.

Figure 2 : MFCC block diagram [22]

In details, the audio signal is segmented into small duration blocks known as frame. Each of the above frames is multiplied with a hamming window. Then, the Power Spectral Density (PSD) of each frame is obtained by taking the squared modulus of its Fast Fourier Transform (FFT). Then, each PSD is multiplied elementwise by a set of triangular band pass filter in order to get filtered magnitude spectrum. It also reduces the size of features involved as there is generally less frequency channel considered than for classical Fourier analysis. Then, the Discrete Cosine Transform (DCT) is applied on the log energy obtained from the triangular band pass filters to finally provide Mel-scale cepstral coefficients. More precisely, $\forall n = 0 \dots N - 1$, $x_l[n]$ is the N-signal containing the samples of the frame l. Its PSD is computed for channels $k = 0 \dots N_{fft}/2 - 1$. Besides, the MEL M-filter bank $H_{MEL}[k, m]$ is defined $\forall m = 0 \dots M - 1$ by

$$H_{MEL}[k, m] = \begin{cases} 0 & \text{if} \quad k < f(m-1) \text{ or } k > f(m+1) \\ \dfrac{k - f(m-1)}{f(m) - f(m-1)} & \text{if} \quad f(m-1) \leq k \leq f(m) \\ \dfrac{f(m+1) - k}{f(m+1) - f(m)} & \text{if} \quad f(m) \leq k \leq f(m+1) \end{cases} \tag{4}$$

This relation shows MEL filters are nothing else than triangular filters centered on each considered MEL frequency m. Finally, according to relation (4) and to the flow graph presented in Figure 2, $\forall m = 0 \dots M - 1$, the $MFCC_l[m]$ coefficients of the lth audio frame $x_l[n]$ sampled at F_s Hz, windowed by $h[n]$, with are given analytically by the following relation:

$$MFCC_l[m] = \sum_{p=0}^{M-1} \log_{10} \left(\sum_{k=0}^{N_{fft}/2-1} \frac{2}{F_s N} \left| \sum_{n=0}^{N-1} x_l[n] h[n] e^{-2i\pi n k} \right|^2 H_{MEL}[k, p] \right) \cos \left(\frac{\pi}{N} (p + \frac{1}{2}) m \right) \tag{5}$$

Hence, let suppose we have L lags, then $\forall l = 0 \ldots L - 1$ if we consider each $M-$vector $MFCC_l[m]$ given by relation (5) as a column of an $(M, L)-$image, it allows us to build the MFCC representation of the audio signal, such as the one illustrated by the Figure 3. MFCC's are commonly used as feature extraction technique in speech recognition system because it approximates the human system response more closely than any other system; nevertheless, it suffers of a lack of robustness in case of low signal-to-noise ratio observation.
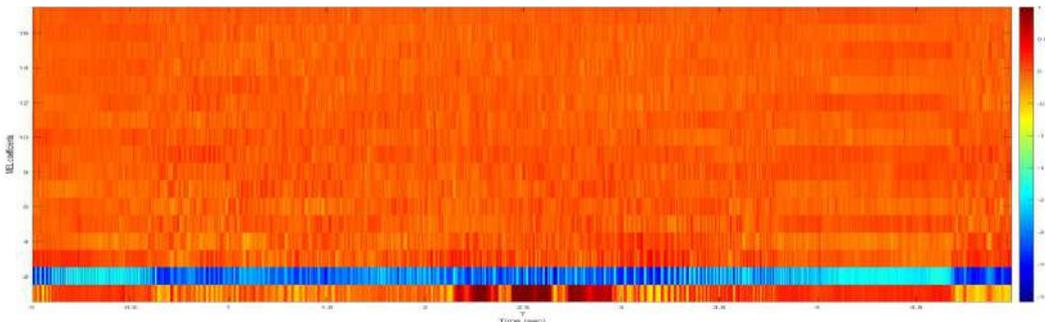
Figure 3 : Mel frequency Cepstrum of a biological signal.

In our case, our automatic underwater acoustic signature recognition must attend the operator dedicated to sound analysis, that's why the use of MFCCs seems to be pertinent. Following the extraction of MFCCs, we use the SVM algorithm as a classifier [15].

2.2.2. The VGGish network

Recently, high-performance Neural Networks for image classification such as AlexNet [16], VGG [17], Inception [18] and others are being tested for audio classification problems. These techniques seem to perform better as the classical one with MFCCs extraction. The architecture we have selected is based on a recent publication [19] trained on millions of YouTube videos. Hence, we take a VGG network slightly modified called VGGish to match with the sound classification technique found in [20].

First, Mel spectrogram is used as input features, computed from the spectrogram of each audio file. Figure 4 shows an example of the Mel spectrogram of a biological signal.
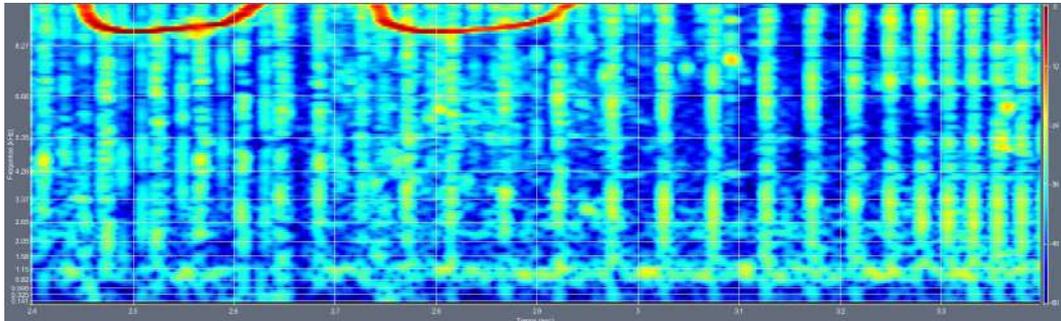
Figure 4: example of a Mel spectrogram of a biological signal

The VGGish we take is a variant of the VGG model described in [17]. In particular, its architecture is modified so that it uses Configuration A with 11 weight layers. Furthermore, the input size was changed to 96x64 for log Mel spectrogram audio inputs. Then, the last group of convolutional and maxpool layers is removed, so there are only four groups of convolution-maxpool layers instead of five, and instead of a 1000-wide fully connected layer at the end, there is a 128- wide fully connected layer. This acts as a compact embedding layer. We use the provided VGGish model that is pre-trained on a large YouTube dataset, which is a preliminary version of what later became YouTube-8M, and then we fine-tune the VGGish model parameters for our application we just add a classifier layer at the end, consisting of parallel logistic for classifier, one per class, which allows multi-class task

Our approach can be considered belonging to Transfer Learning domain [25], as the original VGGish model is modified in order to fit with the new targets we want to recognize, where these new targets are nothing else than classes of underwater acoustics noises.

2.2.3. The VGGish network and one dense layer

Starting from the classical approach [19], we propose an original architecture composed of VGGish followed by one fully connected layer. Here, the VGGish network is considered as a "warm start" for the lower layers of a model that takes audio features as input and adds more layers on top of the VGGish embedding. It is used to fine-tune VGGish because our dataset is different from the typical made up with YouTube video clips.

2.2.4. Synthesis of the three architectures

In summary, performance will be evaluated on:
- A first architecture based on the classification using an SVM classifier of the MFCCs of each audio signal.
- A second architecture using the VGGish network pre-trained that we will re-train on our data.
- A third architecture combining VGGish pre-trained and a dense layer to match our data.

3. EXPERIMENTS AND RESULTS

In this section, the database and the performance metrics we have selected are presented. Then, the three architectures and their performances are described and then discussed.

3.1. Database composition

Public data were collected to train acoustic event recognition architectures. These are from the San Francisco National Park Association [24], which works on maritime history and refers to acoustic events collected during military naval missions of the WWII. Hence, these records constitute a great opportunity to train our models with public but representative data. Nevertheless, despite they are realistic these records are old ones, even if biological noises remain the same, some mechanical noises have slightly changed or are depreciated, so we have to carefully select only up to date representative noises.

Records were divided into 5-second intervals, set up after human observation process as the typical duration of the acoustic phenomena studied. This task is followed by labelling of each 5-second audio frame, manually annotated and checked by professional expertise. Here, a class is dynamically constituted when we have at least 5 examples from the same. In the context of this study, 3 classes belonging to the military field have been retained, Biological, submarine and vessels, respectively labelled as BIO, SM, and SS. The base has initially 478 signals, 80% signals are dedicated to the training set and 20% to the validation set, for a total duration of 39 minutes. Here, the test set is currently only composed of 12 signals including 6 SM, 3 BIO and 3 SS. The training dataset is summarized below:

Classes	BIO	SS	SM
Occurrences	134	158	174

Table 1: Repartition of training dataset

The table 1 allows us to see how classes are balanced, which is important to obtain good performance during learning step and fine-tuning step, avoiding overlearning.

3.2. Performance assessments

The following metrics were selected to estimate performances of the 3 architectures: train accuracy, validation accuracy, confusion matrix, ROC-curve and AUC [21].

Classification accuracy is the ratio number of correct predictions with the total number of input samples. It works well only if there is almost the same number of samples belonging to each class. A confusion matrix is a summary of the results of predictions about a classification problem. The matrix shows for each class the number of correct predictions on its diagonal entries and the number of incorrect predictions on its extra-diagonal entries. In practice, for each example each row of the table corresponds to the actual class, while each column corresponds to the expected class. Even if the confusion matrix is intuitive and so easy to interpret, to take into account the influence of misclassification costs we introduce as an extension of the confusion matrix the Receiver operating characteristic curve. The ROC curve is a probability curve showing the true positive rate in function of the false positive rate. Finally, in order to have a simple scalar metric, easy to interpret, cost-insensitive, summarizing the global performance level of the network, we integrate the ROC curve and we obtain the area under curve (AUC), which represents the measure of separability of one class versus all others. Higher the AUC, better the model is, and conversely. Because of its global cost-insensitive behavior, AUC allows to compare two different classifiers.

3.3. Performances of the three architectures

The train accuracy, the validation accuracy, confusion matrix and AUC-ROC curves are computed from the test set of the 3 architectures and illustrated in Table 2:

Architecture	Train_accuracy	Validation_accuracy
MFCCs and SVM	1.0	0.8696
VGGish	0.9380	0.9130
VGGish and one dense layer	**0.9903**	**0.9203**

Table 2: Accuracy of the train and validation dataset for the 3 architectures

These results allow first to ensure the network well learned and then it generalizes well its knowledges on the validation set. We simply verify overfitting of the networks, this is the case with the VGGish and VGGish and one dense layer architectures, which learned well about the training set and are able to generalize about the validation set. However, the MFCCs and SVM architecture seems to overfit to ensure this, it would be necessary in future work to use cross-validation and add regularization in the event of overfitting.
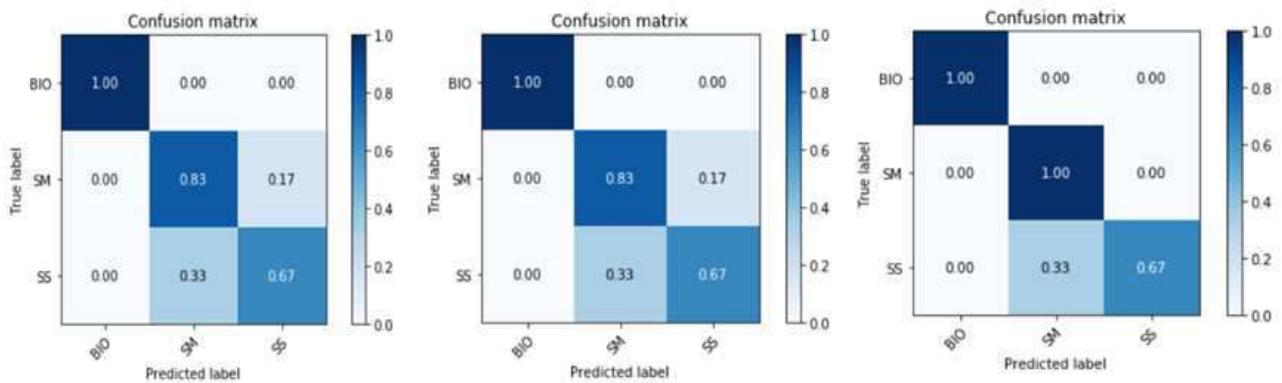
Figure 5: Confusion matrices of the 3 architectures: MFCCs + SVM (left), VGGish (center), VGGish + 1 Dense Layer (right)

Performances obtained on MFCCs+SVM and VGGish architecture are similar, which is quite disappointing in regard of the AI techniques state of the art. However, the Google network pre trained on millions YouTube videos is used; so it therefore seems regular that the network fully recognizes biological signals and has more difficulty to classify SM and SS signals. Nevertheless, the third architecture shows better performances to classify SM and SS by adding one dense layer and by reentering the network to fine-tune VGGish to adapt to the specificity of the input data.

The 3 architectures have more difficulty predicting the SS class which is often confused with the SM class. One possible explanation would be that SM are old and therefore noisy, so their signatures would look like those of an SS.
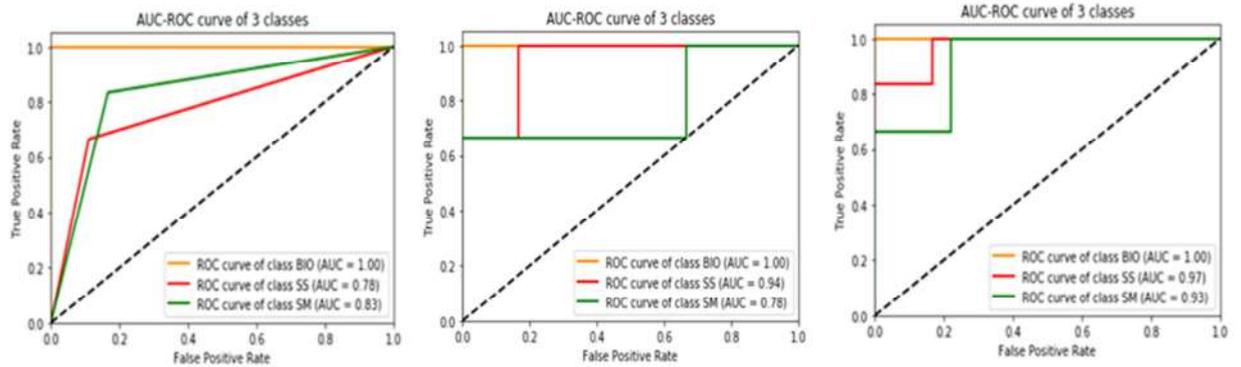
Figure 4 : AUC-ROC curves of the 3 architectures: MFCCs + SVM (left), VGGish (center), VGGish + 1 Dense Layer (right)

Classifiers that give curves closer to the top-left corner indicate a better performance. ROC-curve of class BIO is the same for the 3 architecture, but the ROC-curve of class SM is above that of class SS for the architecture MFCCs + SVM contrary to the 2 others. This may be explained by the fact that VGGish-based architectures use a network pre-trained on YouTube videos with relatively low submarines and surface vessels. We also notice that the performance of VGGish + one dense layer is better than VGGish alone, the fine-tuning of the network worked. The AUC values show that the best performance is obtained on the latest architecture, which is in accordance with the literature. The 3 architectures perform well, but VGGish + one dense layer architecture remains the best taking into account the different metrics.

4. CONCLUSION

Performances obtained via the three architectures are satisfying and it would seem that the last VGGish and one dense layer architecture is the most efficient. Nevertheless, experiments must be carried out with more data, trying to separate them among different classes, in order to take all the benefit deep neural networks which are not much more efficient contrary to the conclusions drawn from the bibliography.

However, the main goal which was the creation of a first database with operational classes, the selection of architectures and performance assessment is achieved, and we can think about future works. In order to see if better performances are possible through deep learning, we consider first the increase and the improvement of the database. Indeed, the current dataset is too old and is not representative anymore, especially for the SS; a new database should be created. Moreover, the recovery of public acoustic data characteristic of naval military field is complicated, that's why an artificial increase in data is being considered. A first step is to create new signals by adding Wenz-type sea noise [19] combined with other techniques from signal processing. Moreover, in order to be as representative as possible of reality, we cannot be satisfied with only 3 classes and will increase them. The neural network architectures that have been used are very deep, it would then be interesting to encourage shallower networks to avoid possible overfitting and observe if performances are better.

REFERENCES

[1] Urakami M., Wakabayashi N. and Watanabe T. "A study on Location Information Screening Method for Ship Application Using AIS Recorded Data,"2018 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom), (2018).

[2] Li Y., Zhang Y. and Zhu F., "The method of detecting AIS isolated information based on clustering and distance, "2016 2nd IEEE International Conference on Computer and Communications (ICCC), (2016).

[3] Nishizaki C., Terayama M. Okazaki T., and R. Shoji, "Development of Navigation Support System to predict New Course of Ship," 2018 World Automation Congress (WAC), (2018).

[4] Shahir H. Y., Glasser U., Shahir A. Y. and When H., "Maritime situation analysis framework: Vessel Interaction classification and anomaly detection," 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), (2017).

[5] Dawei Liang, Edison Thomaz, "Audio-Based Activities of Daily Living (ADL) Recognition with Large-Scale," arXiv: 1810.08691v2 [cs.HC], (2018).

[6] Lukas Müller and Mario Marti," Bird sound classification using a bidirectional LSTM," CEUR-WS.org\vol-2125\paper_134.

[7] Nobuo Sato and Yasunari Obuchi," Emotion recognition using Mel-Frequency Cepstral coefficients," Journal of Natural Language Processing 14(4):835-84, (2007).

[8] Annamaria Mesaros, Toni Heittola, Antti Eronen, Tuomas Virtanen,"ACOUSTIC EVENT DETECTION IN REAL LIFE RECORDINGS,"18th European Signal Processing Conference (EUSIPCO-2010), (2010).

[9] Andrey Temko, Robert Malkin, Christian Zieger, Dusan Macho, Climent Nadeu, and Maurizio Omologo, "CLEAR Evaluation of Acoustic Event Detection and Classification Systems," R. Stiefelhagen and J. Garofolo (Eds.): CLEAR 2006, LNCS 4122, pp. 311 – 322, (2007).

[10] Xiaodan Zhuang, Xi Zhou, Mark A. Hasegawa-Johnson, Thomas S. Huang, "Real-world acoustic event detection," Pattern Recognition Letters 31 1543–1551, (2010).

[11] Sören Becker, Marcel Ackermann, Sebastian Lapuschkin, Klaus-Robert Müller, Wojciech Samek," Interpreting and explaining deep neural networks or classification on audio signals," arXiv:1807.03418v1 [cs.SD], (2018).

[12] G. Keren, B. Schuller," Convolutional RNN: an enhanced model for extraction features form sequential data," arXiv: 1602.05875v3 [stat.ML], (2017).

[13] Satoshi IMAI," CEPSTRAL ANALYSIS SYNTHESIS ON THE MEL FREQUENCY SCALE," IEEE CH1841-6/83/000O0O93, (1983).

[14] Yixiong Pan, Peipei Shen and Liping Shen,"Speech emotion recognition using support vector machine," International Journal of Smart Home, (2012)

[15] Christopher D.Manning, Prabhakar Raghavan and Hinrich Schütze,[Introduction to Information Retrieval], Online edition (c) Cambridge UP, (2009).

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in neural information processing systems, pp. 1097–1105, (2012).

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv: 1409.1556v6 [cs.CV], (2015).

[18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," arXiv: 1512.00567v3 [cs.CV], (2015).

[19] Shawn Hershey, Sourish Chaudhuri, Daniel P. W. Ellis, Jort F. Gemmeke, Aren Jansen, R. Channing Moore, Manoj Plakal, Devin Platt, Rif A. Saurous, Bryan Seybold, Malcolm Slaney, Ron J. Weiss, Kevin Wilson, "CNN ARCHITECTURES FOR LARGE-SCALE AUDIO CLASSIFICATION," arXiv:1609.09430v2 [cs.SD], (2017).

[20] Gordon M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources," The Journal of the Acoustical Society of America 34, 1936, (1962).

[21] N. Japkowicz, M. Shah, [Evaluating Learning Algorithms: A classification Perspective], Cambridge University Press, 1 edition, (2014).

[22] Parwinder Pal Singh, Pushpa Rani, "An Approach to Extract using MFCC," IOSR Journal of Engineering (IOSRJEN), ISSN (e): 2250-3021, ISSN (p): 2278-8719, Vol. 04, Issue 08, ||V1|| PP 21-25, (2014).

[23] Nilu Singh, R.A Khan, Raj Shree," MFCC and Prosodic Feature Extraction Techniques: A Comparative Study," International Journal of Computer Applications (0975 – 8887) Volume 54– No.1, (2012).

[24] https://maritime.org/sound/

[25] Sinno Jialin Pan; Qiang Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, Volume: 22, Issue: 10, Page s: 1345 – 1359, (2010).

7

Big Data architectures and systems

Contents

LIDARs Usage in Maritime Operations and ECO – Autonomous Shipping, for Protection, Safety and Navigation for NATO allies Awareness

NATO (HAF) OF-3 Alexandros Pantazis*[a]

[a]National Technical University of Athens, Laser Remote Sensing Laboratory, Physics Department, 15780 Zografou, Greece, alexpantworld@gmail.com ;

[a]NATO Deployable Air Command and Control Centre,Via Ponte Rosso 1, 44028 Poggio Renatico, Italy, Alexandros.Pantazis@daccc.nato.int ;

ABSTRACT

LIght Detection And Ranging (LIDARs) are cutting edge technological devices and are widely used in atmospheric science, in order to examine the optical and chemical properties of the present atmosphere, in mapping and topography and nowadays their participation in autonomous driving is evolving. In this work a presentation is been made on how LIDARs can be used in the benefit of shipping for navigation and also, autonomous navigation, present weather measurements and detection / self-protection against Unmanned Aerial Systems (UASs), Maritime Unmanned Systems (MUSs), small and radio-undetected boats and also Chemical Biological Radioactive Nuclear (CBRN) agents detection and partially neutralization. The scope of this work is to present the multiple usages that LIDARs can commit in shipping together with a total proposed solution in the all the above, in order for the NATO naval forces to gain their benefit by usage.

Keywords: LIDAR, shipping, autonomous, safety, UAS, MUS, chemical, biological.

1. INTRODUCTION

There are a lot of considerations for military camps, civilian gathering areas as well as for marine environment and shipping, that are not being covered in terms of drones and Unmanned Aerial Systems (UASs) in general and/or Combat UASs (CUASs), Maritime Unmanned Systems (MUSs), Unmanned Ground Systems (UGSs) threats as well as from Chemical and Bio Warfare (CBW) and Radioactive and Nuclear particles. These considerations became a problem especially with the rise of terrorism around the world. In most cases, over 90%, there is no total coverage availability in those areas of safety. In this work, there is a presentation been made with a novel well-established, overall proposal for detection and neutralization - avoidance of all the above, based mainly on LIght Detection And Ranging (LIDAR) and LASER emitting devices, but also working together with other systems forming most effective novel hybrid systems. Another main topic of this work is the well-known efforts been made globally in the Autonomous Shipping (AS) and how could this been provided safely and effectively, also by LIDARs.

2. LIDAR

LIDAR technique have the potential to detect particulate aerosols, incoming CBW agents as well as Radioactive and Nuclear particles after an event like this occurs, with different Lidar techniques and also detect low observable targets like UASs / CUASs, MUSs, UGSs remotely at distances of many kilometers[1 p. vii]. They can provide spatially resolved measurements in 'real-time' at ranges of several kilometers to several tens of kilometers dependent upon several factors such as wavelength, laser power transmittance, ambient conditions and the optical configuration. The 'Laser Radar' is as an effective tool with high temporal and spatial resolution and has been extensively employed to measure the atmospheric and environmental parameters and also applied for military and civilian purposes in many cases such as 'LASER Weapons' or LASER emitting devices.

The detection part depends upon present weather elastic scattering, which is when the scattering frequency is the same as the frequency of the incident light of the LIDAR (like Rayleigh and Mie scattering), or, inelastic scattering which is when there is a change in the frequency[1 pp. 47-58] like Raman scattering. The types of LASERs used by LIDARs are based upon the type of operation like Continues Waved (CW) or pulsed, the mechanism at which population inversion achieved and most important, the state of the active medium it stimulates. The "Free Electron LASERs" which is the most recent type of LASER like behavior, will be using a different kind of procedure for the LASER production and active medium. The application of this particular type of LASER belongs to the upcoming future and could be used mainly as a 'LASER Gun' or 'LASER Weapon'.

The basic elastic LIDAR equation (Eq. 1) is given by the following equation[2, 3]:

$$P(\lambda,R)=P_{OL}\cdot\frac{A_0}{R^2}\cdot\beta(\lambda,R)\cdot\eta(\lambda)\cdot\xi(R)\cdot\Delta R\cdot\exp^{-2\int_0^R[\alpha_{aer}(\lambda,r)+a_{Ray}(\lambda,r)]dr}$$

(1)

where, $P(\lambda,R)$ is the received power in Watts, λ is the laser wavelength, R is distance in m, P_{0L} is the power of the transmitted laser pulse beam in Watts, A_0 is the diameter of the receiver's telescope in m^2, ΔR is the spatial resolution of the LIDAR (depending on the signal sampling frequency), $\xi(R)$ is the geometrical form factor ($0\leq\xi(R)\leq1$) and is a pure number, $\eta(R)$ is the LIDAR opto-electronic efficiency in %, while $\beta(\lambda,R)$ is the backscattering and $\alpha_{aer}(\lambda,R)$, $\alpha_{Ray}(\lambda,R)$ are the extinction coefficients for Mie and Rayleigh scattering, accordingly in m^{-1}.

According to Klett[2, 3], the elastic LIDAR equation can be solved to provide the aerosol extinction coefficient ($\alpha_{aer}(\lambda,R)$) and aerosol backscattering coefficient ($\beta_{aer}(\lambda,R)$), assuming a relationship between these two coefficients, the so-called LIDAR ratio $C(\lambda,R)=\alpha_{aer}(\lambda,R)/\beta_{aer}(\lambda,R)$), So:

$$a_{aer}(\lambda,R)=\frac{\exp[S(R)-S(R_F)]}{\frac{1}{a(\lambda,R_F)}+2\int_R^{R_F}\exp[S(R')-S(R_F)]dR'}-a_{Ray}(\lambda,R)$$

(2)

and:

$$\beta_{aer}(\lambda,R)=\frac{\exp[S'(R)-S'(R_F)]}{\frac{1}{\beta(\lambda,R_F)}+2\int_R^{R_F}\frac{1}{C(R')}\exp[S'(R')-S'(R_F)]dR'}-\beta_{Ray}(\lambda,R)$$

(3)

where,

$$S'(R)-S'(R_F)=S(R)-S(R_F)-\frac{16\pi}{3}\int_R^{R_F}\beta_{Ray}(1-\frac{3}{8\pi\cdot C(R)})dR$$

(4)

$$S(R)\equiv\ln[P'(\lambda,R)\cdot R^2]=\ln[RCS]$$

(5)

RCS is the Range-squared Corrected LIDAR Signal, R_F is a reference altitude for which a molecular atmosphere is assumed, in m, and $P'(\lambda,R)$ is the received power after atmospheric and electronic noise Back-Ground (BG) correction:

$$P'(\lambda,R)=P(\lambda,R)-BG$$

(6)

In case of vertical or near vertical measurements, we can set a reference height at which the atmosphere is purely molecular. In this case, we are able to retrieve the values of $\beta_{aer}(\lambda,R)$ (and/or $\alpha_{aer}(\lambda,R)$) from that height and to the ground. These measurements are, thus, able to provide a clear view of the aerosol load located along the line of sight (LOS) of the LIDAR beam. The optical depth (τ) and total optical depth (τ_{total}) are given by Equations (7, 8) below:

$$\tau_{c_n}(0,R) = \int_0^R a_{c_n}(\lambda, R')dR'$$

(7)

$$\tau_{total}(0,R) = \tau_{C_1} + \tau_{C_2} + ... + \tau_{C_n}$$

(8)

where, R´ is the distance at which $\alpha_{aer}(\lambda, R)$ has been measured with a range resolution of 1 bin (i.e. 7.5 m), τ_{total} is the sum of τ_{Cn}, which is the individual optical depth (n=1,2,3,…) of the (n) prompted atmospheric layers.

The apparently most effective technique for slant pointing LIDAR measurements, for 3D scanning, in an operational environment and under any weather conditions without almost any assumptions, is the "3D stepping" or "LADDER" technique, in order for Klett technique[2, 3] or Raman techniques to be applicable (Fig. 1 - Left)[4, 5, 6, 7, 8].

3. AUTONOMOUS SHIPPING

The shipping and maritime industry is often characterized as being highly conservative and slow to adapt to change, particularly a change as totemic as unmanned shipping. In recent years, the maritime world has been increasingly interested in exploring the benefits of autonomous maritime vessels for freight transportation. This has resulted in a number of exploratory projects including the Advanced Autonomous Waterborne Applications Initiative (AAWA), an AS concept from Rolls-Royce Marine at 2016, AS from Skredderberget at 2018, a Japanese Trans-Pacific test named by Cooper and Matsuda at 2017, the MUNIN research project at 2016, a Chinese AS test location named by Jennings at 2018, an autonomous military ship named by Mizokami at 2018, the DIMECC 'One Sea' Consortium named by Haikkola, 2017 and a start-up company retrofitting old ships to be autonomous named by Constine at 2018.

Many of these efforts focus on the expected benefits of AS, including reduced operational costs, reduced manning, increased operational times, reduced fuel consumption, data analytics[9] and intercommunication also with 5G Networks[10], improved lifestyles for the seafarers and increased maritime shipping capacity named by Kobyliński at 2018 among others. Else have shown more skepticism toward the proposed benefits and have pointed out many challenges that have not yet been solved including legal named by Karlis at 2018, commercial named by Willumsen at 2018 and Technical named by Kobyliński at 2018[9]. One additional benefit that could be brought about by AS is an increased technological and ECO improvement rate in the maritime shipping industry. This higher improvement rate is facilitated by increased potential for upgrading via software named by Greengard at 2015, rapidly improving energy technologies for propulsion named by van Biert et al at 2016 and 'Big Data' driven learning to continuously improve efficiency of transport named by Perera and Mo at 2016.

A ship's ability to monitor its own "health", establish and communicate with what is around it and make decisions based on that information is vital to the development of autonomous operations[6]. The need is to develop a set of electronic senses that informs an "electronic brain" and allow the vessel to navigate safely and avoid collisions. The areas of importance are Sensor Fusion, Sea state, Meteorology, Control Algorithms, Communication and connectivity[6]. Looking at different types of radars, high definition visual cameras, thermal imaging and LIDAR the project AAWA has concluded fusing multiple sensor inputs provides the best results. LIDAR systems (sensors) are high quality state of the art remote sensing systems that are able to monitor environmental parameters, important not only as an environmental baseline, but also for improving the quality and accuracy of measurements which will be mostly helpful for the accurate knowledge of the surrounding weather and sea conditions in real time, in terms of best trajectory calculation, reducing the risk of cruising[6] but also for feeding weather forecasting models for the future transportation planning.

Visibility, Fog and various other atmospheric conditions heading towards the ship, can be detected and measured, along with their calculated time of arrival even with the single wavelength LIDAR[4, 5, 6, 7] making the usage of the latter cost effective. LIDARs already work with the atmospheric conditions detection and measurement for years and accurate algorithms have been created in globally certified eye safe conditions[(1 pp. 203-204), 4, 5, 6, 7]. Ships in this case will be, in parallel, mobile high-performance remote sensing meteorological stations, contributing also to the work of National Meteorological Services or other Meteorological Organizations and through that, to the global Weather Meteorological Organization (WMO), clearly maximizing the accuracy of present weather and sea state measurements and contributing the most, in the weather forecasting models. A novel type of cooperation can be achieved in this way by forwarding the

gathered weather and sea state data of each LIDAR capable, ship, to Meteorological Organizations around the globe in real time, which is also beneficial for the needs of marine units to acquire for future arrangements of cruising in warfare and trade conditions.
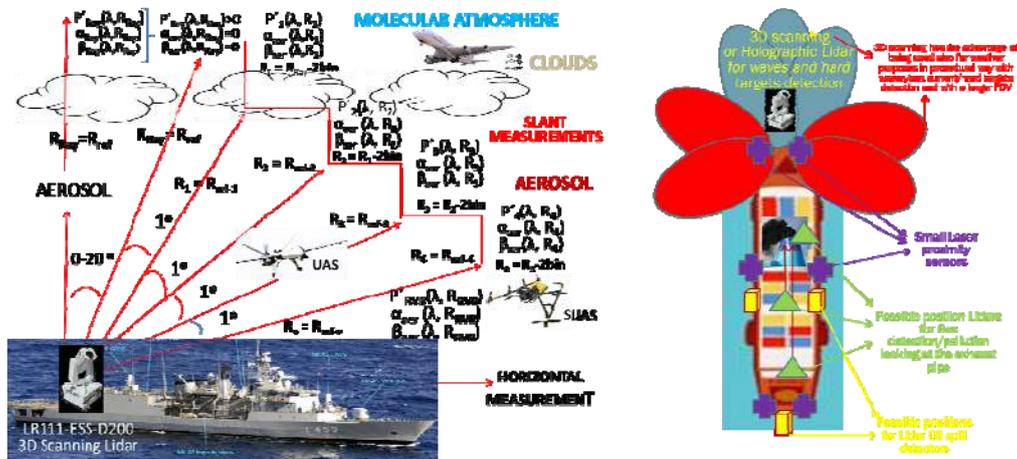

Figure 1. (Left) Visualization of "3D Stepping" or "LADDER" technique[4, 5, 6, 7] with Airplanes and UASs or CUASs detection with 3D LIDAR from a warship. (Right) Proposed LIDARs positioning on ships.

Algorithms developed in[4, 5, 6, 7, 8] can also add to the detection of hard targets and their separation from waves, measuring hard targets and waves velocity and time of arrival and contributing again in ship's safety from crashing – collision, cruising it also away from terroristic attacks in "operational" open or close seas. In particular, specific choice of laser beam wavelengths carries the ability of detecting submarines in some depth and/or surface or Subsurface – Submarine MUSs approaching, subsurface rocks and also measuring, in some depth, the affected for the ship sea current speeds, also through the measurement of the sea current energy spectra, contributing to the cost and fuel consumption effectiveness[9], as well as to safe cruising. LIDARs can also contribute to the accurate detection of oil spills (Fig. 1 – Right) and through that to the detection of earlier presence of ships and submarines in the area as well as their trajectory, even in situations where the 'naked' eye cannot detect.

So, the key measuring parameters of Visibility, Hard Targets detection over and under water, Incoming Fog – weather conditions – big sea waves and Oil spills can be achieved extremely accurately by LIDARs[6, 7, 8]. LIDAR measurements of wind, temperature and humidity 3D profiles as well as sea surface waves directional energy spectra, surface current air velocity and speed of the sea surface, can offer real time situational awareness. Safe alarms triggering can be achieved for approaching severe weather conditions detection and forecasting for optimal route planning, in terms of safe navigation, for fuel consumption and optimization[9] as well as for personnel comfort, providing for all crucial variables for ships in transit. Relatively, marine environment decision making is nowadays solely depended on the reliability and accuracy of weather and sea state prediction as well as personnel experience, but AI or AI – Man interaction with best hybrid sensor system availability and Algorithmic processing[6, 8] is about to "lift" the services provided.

Another area that LIDARs development can leave their fingerprint in shipping is the pollution that the latters create, especially inside harbors and their control by that means. LIDARs can be added near the pipes of the ships along with the stern side where the aerosol by the engines is produced and accurately measure the pollution and the effectiveness of those engines in different conditions and advice the controller of the ship through an AI – Man interaction system, to reduce the consumption when necessary, contributing to the cost effectiveness in terms of fuel consumption[8]. All the above mentioned usages and benefits can be used by autonomous and man handled ships and there will be needed[6]:

a). Fluorence and/or Raman LIDARs for detecting and accurate measurement of oil spill cases, fuel consumption, pollution, BIO and CHEMICAL agents detection of possible terroristic or other attacks and other marine or submarine systems existence by their own oil spills traces in the sea etc (Fig. 1 – Right, Fig. 2 – Left side).

b). 3D scanning LIDAR and/or holographic LIDARs with limited steady angle - position as it performs the on-going monitoring of the present sea status for hard targets detection, big waves, submarines, MUSs and incoming low level weather conditions (Fig 1, Fig. 2 - Right).

c). Smaller laser proximity sensors that will be activated during the mooring phase and during the ship's excursion phase and for safe Navigation inside harbors and narrow areas (Fig. 1 - Right).

d). Importing Navigation Data (Navigation App) and Satellite Communication (GNSS) based, also, in 5G Networks[11] (user dependable).

e). Importing Weather Data from App and from other vessels with corresponding systems.

f). AI Algorithms usage for LIDARs and hybrid system management in total for Data and decision making with an AI system or AI-Man interaction system (user dependable) after evaluation of current Weather / Sea status and other inputs.

g). 3D scanning or Holographic LIDAR, along with other Anti-UAS devices like Microwave radar, for UASs to be protected against possible terrorist or other attacks (Fig 1 - Left).

h). 'Laser Gun' or 'Laser Emitting Device' for self-protection of the ship after Tracking and ID of the incoming possible threat (Figs 2, 3).

i). TDL interoperability acting as Regular Networks with Shared info of all data available / Early warning for ship under attack / under emergency from nearby ships - Distress Call etc. Satellite communication (GNSS) and 5G Networks[11] could be beneficial for the above mentioned functions (Fig. 3 - Right).

j). LIDARs usage as an alternative way of Navigation via Sea Depth scanning (where able) in conjunction with sea depth maps availability and the further development of the latters, especially in narrow and shallow paths, when satellite communication is unavailable because of jamming or interference or any other reason like malfunction (Fig. 2 – Right).

A presentation of LIDARs usage is stated here for the needs of ECO - AS and protection and especially for warships and harbor to harbor or long routed trade ships. This occurred in order for NATO to be informed and perform in most effective way. Under the prism of the new upcoming ECO needs, like pollution control, especially inside harbors but also the threats developed lately like UASs and MUSs, LIDARs and hybrid systems supposed to feed an AI or AI – Man machine interactive navigation and protection device, which eventually will affect maritime patrol and control.
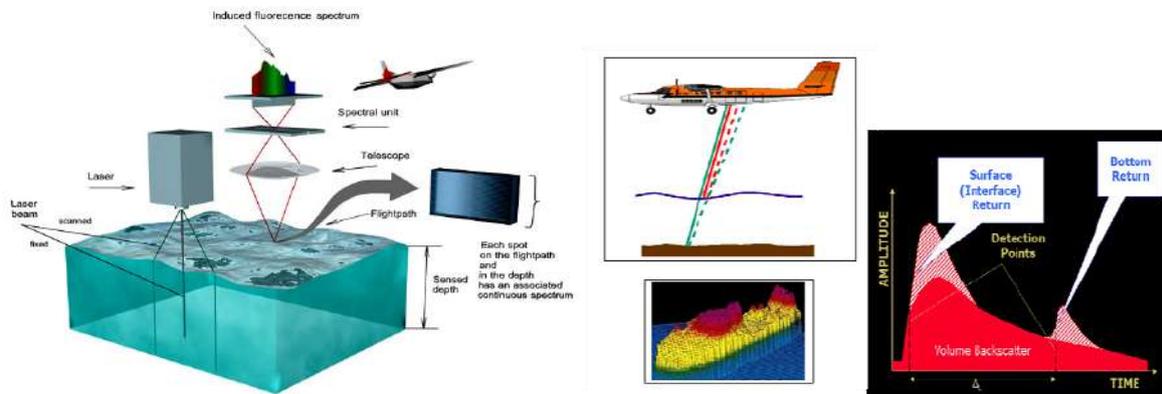
Figure 2. (Left) LIDAR for the detection of sea volumes like plankton, oil spills and through that the detection of early presence of ships – submarines in terms of oil traces in the sea surface. (Right) Aerial and Naval LIDAR και 3D presentation of a shipwreck which can also be used for 3D detection of mines presence Submarines, MUSs, inflatable boats etc in the areas of interest.

4. BIOLOGICAL AGENTS DETECTION

NATO has placed a survey through many kinds of LIDAR techniques and devices and their effectiveness in the detection of early Biological warfare through its RTO – Technical Report – SET 098/RTG-55[12], named "Laser Based Stand-off Detection of Biological Agents", on February of 2010. Biological weapons have become an increasingly important

potential threat in today's military and civilian arenas. They are relatively inexpensive to produce and can yield a significant impact as a terrorist and threaten weapon. Early warning of a biological attack, also with UASs, is essential to establish in a timely defense and to sustain operational tempo and freedom of action.

In addition, the mapping of a biological attack profile is needed to obtain intelligence on affected areas. For these reasons the need to develop methods to remotely detect and discriminate biological aerosols from background aerosols and ultimately, to discriminate biological warfare agents from naturally occurring aerosols, is paramount. The general conclusion was that LIDAR can provide the above mentioned task and also recommends the best kinds of LIDARs that this can be obtained[12]. In order to address these fundamental challenges, several stand-off technologies covering a broad region of the electromagnetic spectrum are being investigated under RTG-055. These technologies include spectrally resolved Ultra-Violet Laser Induced Fluorescence (UV-LIF) at several different excitation wavelengths[12,13], Infrared Depolarization and Long-Wave Infrared (LWIR) Differential Scattering (DISC)[12, 13].

The RTO committee suggested that, active techniques for detection of biological warfare agents may use infrared or ultraviolet Lasers, or both[12]. For Laser induced fluorescence, wavelengths in the ultraviolet are required. The most desired wavelengths are 260 to 300 nm for tryptophan excitation and ~350 nm for NADH excitation. Also excitation of different flavins, primarily riboflavin, is possible. A common excitation wavelength for riboflavin is ~450 nm, but due to the common radiation spectrum as the human visibility (400 to 700 nm) it may carry some limitations in its detection. To obtain good signal to noise ratio, the laser pulses need to be very intense – preferably in pulses of tens of nanoseconds duration with several milli-joule pulse energies. As averaging over many pulses often is required, high pulse repetition rates is also desirable. Continues Waved LIDARs could also do the job without the need of pulses usage[6, 8, 13].

Bio agents but also Chemical ones represent a great threat for marine environment and especially inside harbors in mooring. Opinions are divided as to which is the preferred wavelength: 266 nm UV excites fluorescence primarily from tryptophan, an amino acid present within the bacterial cell wall and tyrosine (also NADH and flavins to a lesser extent) while 355 nm UV excites fluorescence primarily from NADH, a coenzyme found in all living cells and also flavins but not tryptophan. The 266 nm is hence the most appropriate for tryptophan excitation and has a higher fluorescence cross section. In counterpart, 355 nm excitation of NADH may be related to bacterial spore viability. In addition, the attenuation of 266 nm light by atmospheric ozone is approximately 10 times greater than that of 355 nm and so 355 nm LIDAR systems may have a longer detection range. All of the above techniques UV and IR are all up leveling their effectiveness when combined with spectroscopy techniques (Fig. 1 – Right, Fig. 2 – Left, Fig. 3 - Left).

5. CHEMICAL AGENTS DETECTION

CBW (Chemical and Biological Warfare) agents are more commonly delivered close to the ground or sea level, preferably within the altitude region of few hundred meters using the system like rocket shells, missile, UAS, etc. The prevailing atmospheric dynamic processes (wind speed) makes them to spread both vertically and horizontally to form stratified layers / clouds in the atmosphere. Lifetime of these clouds can be from an hour to a day depending upon its nature. In course of time, these clouds settle down and thereby cause health hazard to plants, animals, and human beings. Hence, it is emphasized that the continuous monitoring of the atmosphere nowadays is very important around military areas and marine environments together with civilian population.

LIDAR systems plays a vital role in detecting and identifying these clouds in the atmosphere. A backscattering LIDAR operating at 1064 nm based on Mie scattering principle can search the suspected atmospheric region to look for upcoming clouds, though it cannot identify its composition and concentration. A LIDAR system which contains multiwavelength laser source UV (266 nm and 355 nm), 1064 nm and mid-IR (3-4 μm) wavelengths), common transceiver and scanning gimbal, AI or computer controllers with Graphical User Interface (GUI) can perform the task. A LIDAR master controller or AI triggers IR or UV laser sources to transmit laser radiation into the specific direction for identification of cloud composition if any suspicion arises in the detected cloud signal[14].

In view of the importance of monitoring the atmosphere continuously, a stand-alone backscattering LIDAR system may operate at 1064 nm. Then for the possibly cloud detection and for Chemical agents a DIAL (DIfferential Absorption LIDAR) can be ordered to point the aerosol mass or cloud. Two laser pulses with different wavelengths are emitted into the atmosphere for detection of Chemical Warfare (CW) agents. One wavelength (λ) is tuned exactly to the center of the specific absorption line of the molecule of the interest. The second wavelength (λ_{off}) is detuned to the wing of this

absorption line with no specific absorption. The absorption cross section of the molecule of interest at λ is very large as compared to that at $λ_{off}$. Knowledge of which wavelength has been absorbed (indicated by a highly depleted return signal as compared to that at other wavelengths) gives information about the specific constituent of the atmosphere. Ratio of the return signals at these wavelengths determines the concentration of the molecules of interest due to differential absorption.

A pulsed DIAL laser can always provide the exact range of the detected agents. Selecting the appropriate wavelengths for DIAL measurements involves consideration of factors such as the molecular absorption, interference from other molecular species, atmospheric transmission and scattering, laser transmitter characteristics (for example, gain factor and line width of the chosen wavelength), detector characteristics etc. The emitted laser line widths of these wavelengths should be narrower than the widths of molecular resonant transitions, which, in turn, should be less than the difference between the online wavelengths of two neighboring species. Spectral information on online and offline absorption wavelengths and their cross section values of potential CW agents are very much required for DIAL operation. Many of these agents have distinct absorption bands in the 3-4 μm and 9-11 μm regions[13, 14] and there is relatively less atmospheric attenuation in these spectral regions. No single laser source with the required level of high peak power at each of these wavelengths is used to detect these agents in the atmosphere. CO2 lasers (9 to 11 μm) have commonly been used for the detection of a majority of chemical agents[14]. The generation of 2 to 5 μm wavelengths can be done by various nonlinear techniques, like the Optical Parametric Oscillator (OPO) technique in solid-state lasers.

These LIDAR settings for Chemical agents detection could also be used at the positions of the ship where ECO - pollution measurements are made and for the detection of oil spill traces by other ships, submarines and/or MUSs providing early warning and self-fuel efficiency capabilities at the same time (Fig. 1 – Right, Fig. 2 – Left).

6. DRONE / UAS / CUAS / SUBS / MUS DETECTION WITH LIDAR

The main idea is for a LIDAR system to be made in order for low observable targets (UASs, MUSs) to be detected initially and also for the aerosol detection to be made and for the LIDARs of Fluorescence and DIAL ones to be engaged and detect their potential threats[14, 15]. In particular the idea is:

a). The detection of the target (UAS / CUASs / Drones / Mini Drones / SUBs / MUSs / Bio agents / Chemical agents) by a 3D LIDAR Continuous Waved (CW) or Pulsed with quasi-CW formation in 3 Dimensions (Fig. 3 – Left) and/or a Holographic Lidar. (The contribution of a millimeter Radar for the solid targets with metallic parts - equipment, detection, working in synergy with the LIDAR would be an asset)[15].

b). The Identification (ID) of that target (UAS / CUAS / Drone / Mini Drone / SUBs / MUS) by an Imaging IR (IIR) database (preferably long distance device) or Visual ID (preferably long distance visual camera).

c). The tracking and profile of that target (UAS / CUAS / Drone / Mini Drone / SUBs / MUSs / Bio agents / Chemical agents) by the same or other LIDAR device and also by the IR and Visual camera.

d). The termination - neutralization of that target (UAS / CUASs / Drones / MUSs / Mini Drones) by a 'Laser Gun', preferably, provided by the same LIDAR or other Laser device. It is preferable that the 'Laser Gun' device will be transmitting in a non-visible wavelength like in IR spectrum, avoiding the immediate geolocation of the 'Laser Gun' shooter and visual Early Warning 'proof" of ever shooting towards the target. CO2 LASER could handle this task!

It could also be accompanied by other sensors like:

e). A millimeter Radar for the solid targets with metallic parts - equipment, detection, working in synergy with the LIDAR, would be an asset.

f). Acoustic – Audio sensor tuned to the acoustic frequency made by a number of UAS / CUAS / Drone / Mini Drone / MUS for the alert of nearby targets presence, but avoiding the sound of the "whispering" wind blowing at the same time.

g). An InfraRed sensor (IR or Imaging IR) for "extra" (Thermal) detection of nearby targets and preferably from a respected distance like from few km away.

h). A visual camera especially for detecting moving objects scanning the perimeter of the ship, preferably from a respected distance like from few km away.

i). RF and WiFi seekers and jammers for the detection of radio guided UAS / CUASs / Drones / Mini Drones / MUS to detect and be able to intervene to their guidance.
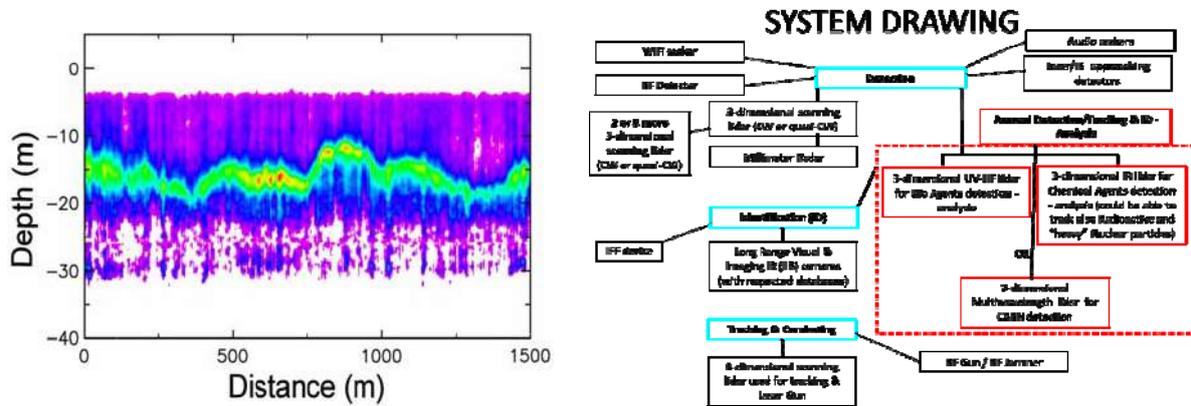

Figure 3. (Left) Vertical slice of thin plankton layer. Processed NOAA LIDAR return as a function of depth and distance along the track is encoded as color from violet -least return to red-highest return. (Right) Draw of the proposed system connected between different sensors[6].

REFERENCES

[1] Measures, R.M., "Laser Remote Sensing – Fundamentals and Applications", Kreiger Publishing Company, Krieger Drive, Malabar, Florida, 32950, 1992.

[2] J. Klett, "Stable analytical inversion for processing LIDAR returns," Appl. Opt. 20, 211-220 (1981).

[3] J. Klett, "LIDAR inversion with variable backscatter extinction ratios," Appl. Opt. 24, 1638-1643 (1985).

[4] A. Pantazis, A. Papayannis, G. Georgousis, "Lidar algorithms for atmospheric slant range visibility, meteorological conditions detection and atmospheric layering measurements", Appl. Opt., 56, 6440-6449 (2017).

[5] A. Pantazis, A. Papayannis, G. Georgousis, "Lidar algorithms in 3D scanning for atmospheric layering and planetary boundary layer height retrieval: comparison with other techniques", Appl. Opt., 56, 8199-8211 (2018).

[6] primordialphoton.com

[7] A. Pantazis, A. Papayannis, "Lidar algorithms and technique in 3D scanning for planetary boundary layer height and single-beam–single-pointing wind speed retrieval", Appl. Opt., 58, 2284-2293, (2019).

[8] PCT/GR2018/000037 by A. Pantazis.

[9] C. C. Spandonidis, N. Themelis, G. Christopoulos, C. Giordamlis, "Evaluation of Ship Energy Efficiency Predictive and Optimization Models Based on Noon Reports and Condition Monitoring Datasets", Data Analytics, 103-108 (2018).

[10] C. L. Benson, P. D. Sumanth, A P Colling, "A Quantitative Analysis of Possible Futures of Autonomous Transport", INEC (2018).

[11] C. C. Spandonidis, C. Giordamlis, "Data-centric Operations in Oil & Gas Industry by the Use of 5G Mobile Networks and Industrial Internet of Things (IIoT)", ICDT, 1-5 (2018).

[12] NATO RTO RT SET-098/RTG-55, "Laser Based Stand-off Detection of Biological Agents", February 2010.

[13] Vladimir A. Kovalev, William E. Eichinger, "Elastic LIDAR", Wiley Interscience, 2004.

[14] S. Veerabuthiran, A.K. Razdan, "LIDAR for Detection of Chemical and Biological Warfare Agents", Defence Science Journal, Vol. 61, pp.241-250 (Delhi, 2011).

[15] G. J. Kunz, H. P. Th. Bekman, K. W. Benoist, L. H. Cohen, J. C. van den Heuvel and Frank J.M. van Putten, "Detection of small targets in a marine environment using laser radar", TNO Defense, Security and Safety, Remote Sensing of the Coastal Oceanic Environment (2005).

Big Data Infrastructure for heterogeneous sources and Data fusion services applied to Maritime Surveillance

Giuseppe Vella[a], Giovanni Barone[a], Viviana Latino[b], Domenico Messina[a], Vito Morreale[a]

[a]Engineering Ingegneria Informatica, P.le dell'Agricoltura 24, 00144 – Rome (Italy); [b]Eka Srl Via Monteroni s.n., C/O Edificio Dhitech - Ecotekne, 73100 Lecce

ABSTRACT

The large amount of data coming from different sensors involved in Maritime Surveillance context that are available today are not usable by maritime security systems since they are not accessible at the same time and, often, they are not interoperable and their usage is often regulated by different national and transnational policies that makes difficult the sharing of information between public authorities. The main challenge described in this paper is to overcome such difficulties and provide a Big Data Infrastructure (BDI) that will be the information hub used by all the Data fusion Services of the partners involved in the MARISA (Maritime Integrated Surveillance Awareness) project. The BDI will allow the Data fusion services to store and retrieve different kind of data coming from different sensors like Automatic Identification System (AIS) or Synthetic Aperture Radar. Once the data are stored according to a specific project data model that extends the Common Information Sharing Environment (CISE) model, all the information fused at different levels are represented by a unified multi layered Maritime Situational Awareness application.

Keywords: Big Data Infrastructure, heterogeneous source integration, Data Fusion Services, CISE data model

1. INTRODUCTION

Europe has a coastline almost 68000 km and the maritime area under the jurisdiction of European Union (EU) Member States is larger than the total land area of the EU. This means that from one side sea is vital for European economy for the development of commercial activities like tourism, fishery, transports or even mineral extraction or wind farms, while from the other side it can also be used by criminals and terrorists to commit crimes that unavoidably raise threats like: piracy, trafficking of drugs, irregular immigration, smuggling, illegal fishing, environmental crimes and maritime accidents/disasters. Indeed Europe needs to enhance cross-border and cross-sectoral cooperation to deliver maritime security, to optimize the information exchange between Legacy systems used by the National Coordination Centers, to secure maritime ecosystem, to reduce maritime traffic and human lives, to prevent and react to maritime crimes committed in the sea or near the costs. Countries need a common operating framework, a common format for the information to be shared, an interactive environment that can provide not only a Maritime Situational Awareness for a real operational environment but a toolkit that will support end users to observe the current situation at the sea, comprehend which can be potential threats, determine if vessels on the sea are fulfilling pre defined criminal behaviors, react whenever a crime occurs. The Big Data Infrastructure implemented and described in this paper aims at integrating data coming from heterogeneous sources and giving the chance to the different machine learning algorithms identified with the MARISA Toolkit as Data Fusion Services (nearly 30) and developed in the MARISA project to identify suspicious and even recurring criminal behaviors of vessels and reduce risks at the sea.

2. BACKGROUND

Large amounts of "raw" data are being collected nowadays, at unprecedented scale, coming from different sources, from different sorts of assets from different EU Member States, from the Internet and social networks, and gathered for different security purposes, in a variety of formats, are available but not necessarily exploitable because they are not accessible at the same time nor interoperable, until they are "fused" and made "understandable" to all systems supporting information exchange, situational awareness, and decision-making and reaction capability at the EU external maritime borders. In [1] the improper usage of the same MMSI (or IMO, call sign, destination extracted by AIS or Long Range Identification and Tracking) by more than one vessel is described as an operation managed by a data fusion approach

that tackles the recognition and the maintenance of each ship-track at different stages. (1) Gating: a selection of the track candidates is accomplished by using geometric and kinematic considerations. (2) Data Association: an association between the incoming message and existing tracks is made through a nearest neighbor approach. (3) Track Management: Track Management: track initiation, confirmation and deletion are implemented [2][3]. [4] explains how different sensors can be used to monitor and to track different targets to fuse them into one single target utilizing optical and microwave sensors on platforms such as satellites and airplanes, thus avoiding the limitations of the sensors, but this introduces limitations in the platform. The most limiting factor is the interrupted data availability, since no airplane is able to stay in the air constantly during the whole year and during all weather conditions. Meanwhile, satellites, which are orbiting around the earth, will be over the zone of interest for a limited time only. The combination of rule-based architecture for the mining of AIS data stream and statistical models to BDI and frameworks largely remains to be investigated [5]. Some progresses have been made in the extraction of maritime Pattern of Life through an unsupervised approach that has been tested using extensive datasets[6]. We propose to use a Big Data Infrastructure that orchestrates the flow related to data coming from the CISE network, the sensors (AIS, Radar) to reconstruct maritime Patterns of Life and data related to the Meteo Oceanographic Conditions, provides access to the different data fusion services that elaborates data through several steps and stores newly produced information in the proper type of storage allowing all the rule-based behavior analysis Data Fusion services to detect anomalies and risks that are displayed in a multi-layered Maritime Situational Picture.

3. METHOD

3.1 Overview

The MARISA Toolkit architecture is composed of different layers that allow all the modules to exchange data (Vessels detected, Alerts, Routes, Documents and Open Source INTelligence data) starting from the data ingestion till to the visualization phase. In particular we will focus on the anatomy of the BDI that is the knowledge hub for the whole toolkit and allows storing all the information coming from the external data sources.
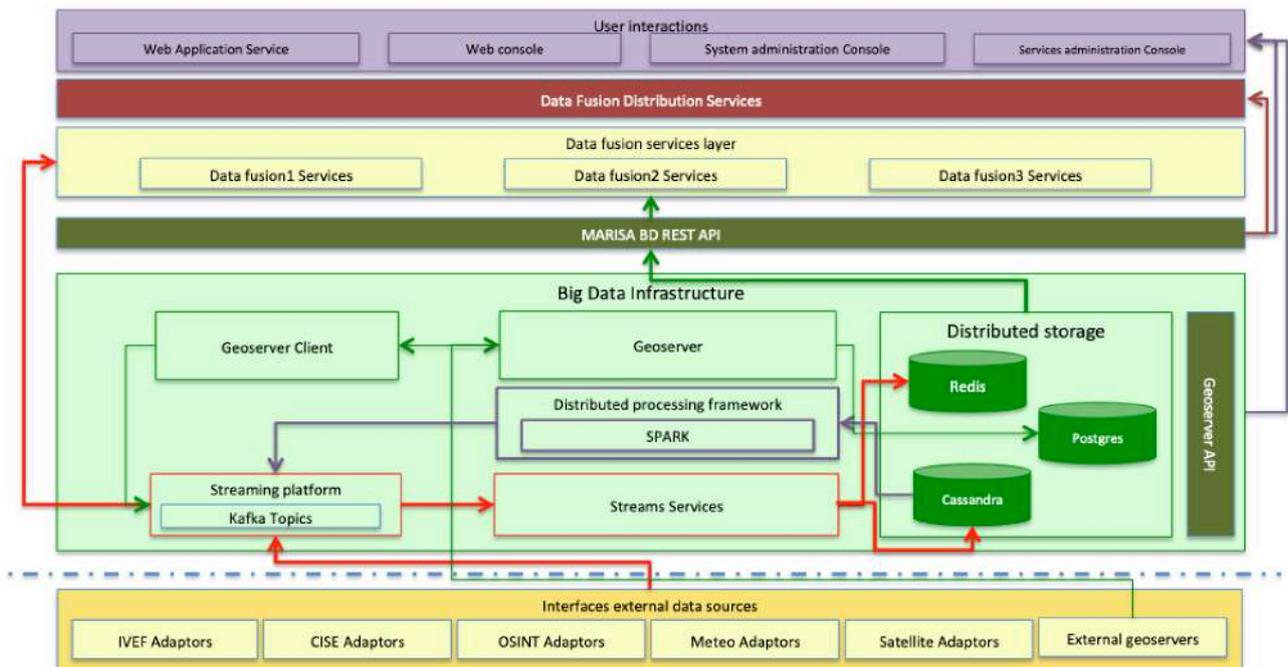

Figure 1. Overall MARISA Architecture.

The BDI allows to Data Fusion Services to retrieve data, elaborate them and create data products, and to the Representational State Transfer (REST) hub layer, the Data Fusion Distribution Services (DFDS), to distribute data to any other external system all the information produced by the Data Fusion Services (DFS). Last but not least through the

BDI rest Application Programming Interface (API) and the Geoserver API the BDI allows to present the information to the end users. In the following paragraphs the different layers that interact with the BDI will be introduced.

Interfaces for external data sources - this layer allows the different adapters to ingest data from external sources: Interchange VTS Exchange Format (IVEF), CISE, Satellite, OSINT, Weather data and External GIS servers that in particular store Web Feature Service (WFS) that will be fused by the different Data fusion services and Web map Service (WMS) that will be represented in the multi layered Maritime Situation Picture. All the data coming from these sources will be adapted to the MARISA data model and stored through the streaming platform based on Apache Kafka [7] and Spring Cloud Data Flow [8]. The BDI allows to manage different processes: real time data that are ingested by means of streaming services and stored temporarily on Apache Kafka and on Redis [9] for further queries and offline ingestion of data stored into big table databases like Cassandra [10] or on Geoserver [11] in case of data related to Satellite, AIS or Meteo Oceanographic conditions.

Data Fusion Service layer - after the ingestion is finished, all the data fusion services of the different levels (observation, comprehension and projection of future states) can consume either in one step data using Kafka API and producing information that will be stored and available as data fusion products or can produce further new data that will be elaborated in several steps by other data fusion services that in the end will produce information that will be ready to be visualized.

User interactions layer - the user interaction layer has the objective to present the Data Products of the Data Fusion Services according to the configurations that have been set for each service. This layer accesses the BDI via Kafka consumers. Moreover, the UI layer is able to retrieve textual data produced by some DFS by means of the REST API and maps with geographical features directly accessing to the BDI geoserver that aggregates data coming from proprietary partners geoservers (e.g. WMS and WFS layers).

Data Fusion Distribution Services - The Data Fusion Distribution Services layer interacts with the MARISA REST BDI API, the Geoserver API and the Data Fusion Services API in order to retrieve information for each data product provided by each Data Fusion Service, allowing queries to external systems.

3.2 Specific approach

The core of the BDI is the streaming platform that is based on Apache Kafka and Spring Cloud Data Flow. The processes for the streaming platform are modeled using Domain Specific Language by creating a stream that contains: the name of the topic (the MARISA entity) that contains data coming from sensors or processed information as output of the Data Fusion Services; the source of data with a direct access to external geoservers as data providers; the sink or the component that will allow to store information both in real time on Redis or on Cassandra. Moreover External Data Fusion services can store previously analyzed information directly on Postgres. In case of external Geoservers that produce daily amount through the processing of AIS or Synthetic Aperture Radar data and inject them into the centralized geoserver, the BDI retrieves those information from the geoserver by means of an authenticated HTTP Client Source component that stores information through the streaming platform on Kafka with a sink on Redis and Cassandra for Business Intelligence purposes.
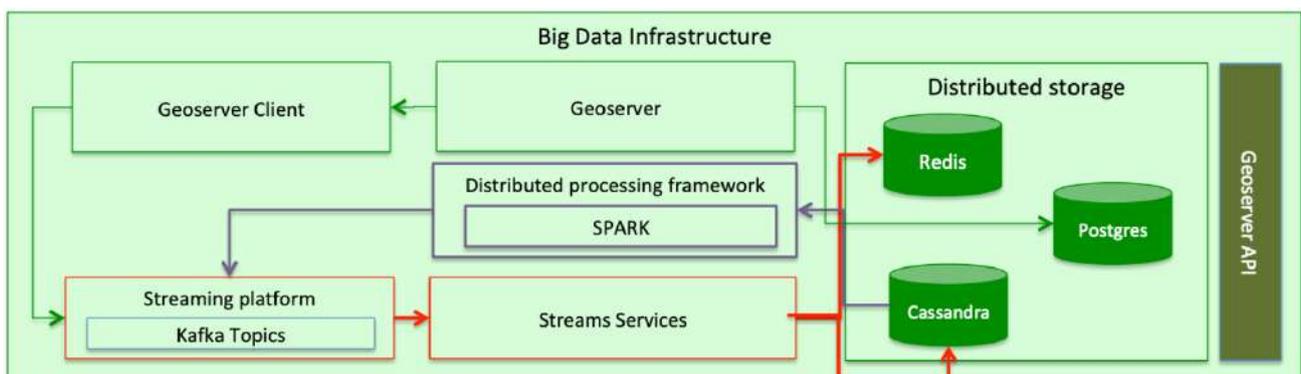

Figure 2. The Big Data Infrastructure

3.3 Data Fusion results

Once information has been stored according to the processes described in the previous section, they are displayed in different layers of the MARISA Situational picture enabling a further analysis of the situation thanks to the rule-based Behavioral Analysis services that detect anomalies starting from the detected vessels fused by DFS. Meteo Oceanographic Condition service output will be not only displayed to the end users of the MARISA project (Marina Militare, Hellenic Minister of Defence, Guardia Civil, Portuguese Navy, Dutch Coast Guard) but they will be used by other prediction services to forecast routes or arrange missions to solve critical situations.
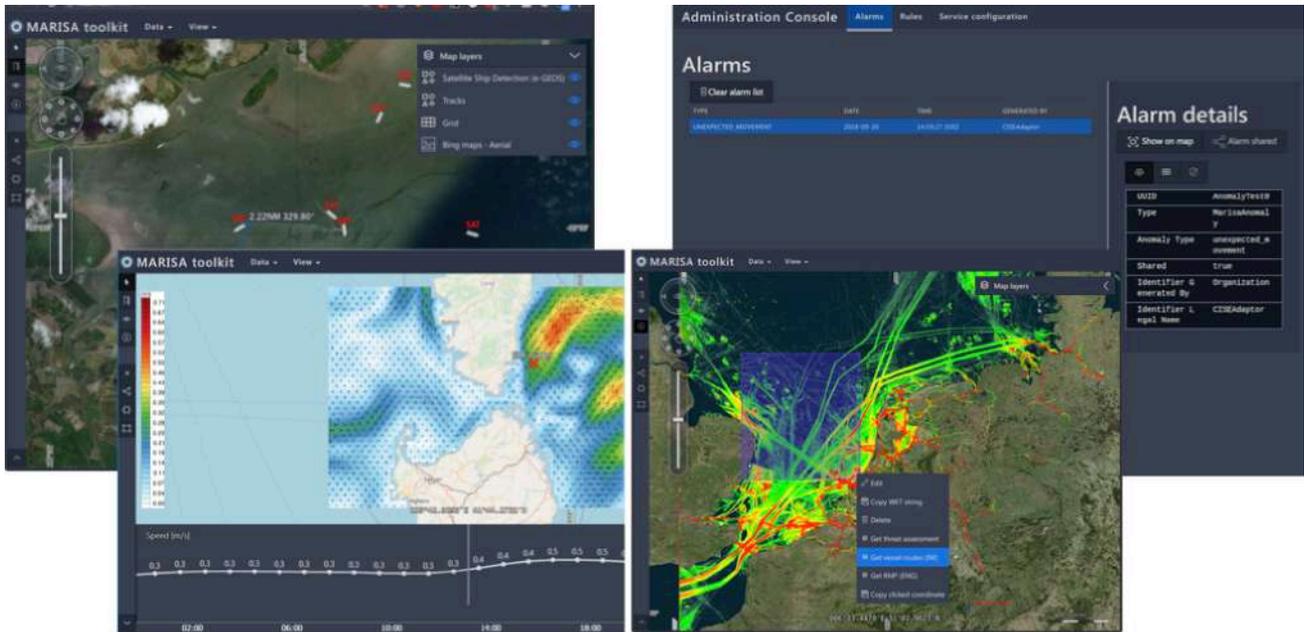

Figure 3. The multi layered Maritime Situational Picture and the real time Alarms console.

4. BIG DATA INFRASTRUCTURE RESULTS

The BDI is an infrastructure that distributed on a cluster composed by 4 nodes. The operating system of every node is based on CentOS Linux 7.5.1804 and each node is composed by 4 high frequency single core Intel Xeon Gold, 47 GiB of RAM and a storage capacity of 200 GiB. The orchestration of the components of the infrastructure is performed by Kubernetes. Kubernetes master is on the master node while the other three nodes are used to distribute the workload.

Data stored in the BDI is managed by three main technological components: Apache Kafka, Redis and Apache Cassandra. Kafka is composed by a unique instance of Apache Zookeeper and Apache Kafka that stores between an average of 20 and 50 messages per second with an average size of 1500 byte per message. The daily volume of data managed by Kafka is nearly 1.7 GiB taking into account traffic data in the network. For what concerns Redis, it allows to manage up to 8 GiB of data. It represents the live streaming situation and it allows to store up to 3 days. Apache Cassandra is like a data center composed by three nodes with a Simple Strategy as a replication strategy and a replication factor is equal to 3. This configuration guarantees high reliability and fault tolerance. By means of Cassandra is possible to store up to 2 months data. The size of data managed by the previously described components will increase in a more powerful infrastructure with the increasing of the cluster nodes as explained in the next section.

4.1 Historical data elaboration on Cassandra via Spark.

In order to test the scalability and the performances of the BDI we have performed some tests on data stored in Cassandra. Data are elaborated and extracted through the usage of Apache Spark [12] distributed process engine. In particular for the MARISA project purposes we have tested Spark using a driver (master node) and two workers (executors) executing it in cluster via Kubernetes.

In the table below there is an example of the response time after a request for elaboration of the data stored on Cassandra after Spark is executed to elaborate and perform queries on data using 1 executor (worker) or 2 executors.

Table 1. Overview of the messages managed by Cassandra and the related response time of Apache Spark analysis

# of records on Cassandra	Response time (in seconds) – 1 worker	Response time (in seconds) – 2 workers
1.000	28''	29''
10.000	28''	39''
100.000	78''	62''
1.000.000	92''	64''

After the performing of the tests illustrated above, we can conclude that Apache Spark scales horizontally with the increasing of nodes and an increasing dataset. Response time on a dataset of 1000 records is almost the same by using one or two workers. With the increasing of the size of the dataset, response time improves with the increasing of workers considering that the analysis service run with Apache Spark includes the overhead of data traffic on the network and the time needed to deploy drivers and workers.

ACKNOWLEDGEMENTS

REFERENCES

[1] Fabio Mazzarella, Alfredo Alessandrini, Harm Greidanus, Marlene Alvarez, Pietro Argentieri, Domenico Nappo, Lukasz Ziemba, "Data Fusion for Wide-Area Maritime Surveillance", MOVE Workshop on Moving Objects at Sea, Brest, France (2003).

[2] Monica Posada, Harm Greidanus, Marlene Alvarez, Michele Vespe, Tulay Cokacar, Silvia Falchetti, "Maritime awareness for counter-piracy in the Gulf of Aden", IGARSS, 249-252 (2011).

[3] Silvia Falchetti, Marlene Alvarez Alvarez, Tulay Cokacar, Harm Greidanus, Michele Vespe, "Improving Co-operative Vessel Tracking for a Maritime Integrated Surveillance Platform", Proc. NATO RTO SCI-247 Symposium on "Port and Regional Maritime Security", Lerici, Italy, (21-23 May 2012)

[4] Dejan Nikolic, Nikola Stojkovic, Zdravko Popovic, Nikola Tosic, Nikola Lekic, Zoran Stankovic, Nebojsa Doncov: Maritime Over the Horizon Sensor Integration: HFSWR Data Fusion Algorithm. Remote Sensing 11(7): 852 (2019)

[5] Ronan Fablet, Nicolas Bellec, Laetitia Chapel, Chloé Friguet, René Garello, Pierre Gloaguen, Guillaume Hajduch, Sébastien Lefèvre, François Merciol, Pascal Morillon, et al., "Next Step for Big Data Infrastructure and Analytics for the Surveillance of the Maritime Traffic from AIS Sentinel Satellite Data Streams", Conference: Big Data for Space Conference, BiDS'2017, Toulouse, France, 1-4 (November 2017)

[6] Nicola Forti, Leonardo Millefiori, Paolo Braca, "Unsupervised Extraction of Maritime Patterns of Life from Automatic Identification System Data" in Proc. of the OCEANS 2019 MTS/IEEE Conference, 2019

[7] https://kafka.apache.org/

[8] https://spring.io/projects/spring-cloud-dataflow

[9] https://redis.io/

[10] http://cassandra.apache.org/

[11] http://geoserver.org/

[12] https://spark.apache.org/

A Cloud Architecture for big data analytics and real-time anomaly detection in global maritime applications

Mariano Alfonso Biscardi[a], Marcello Cinque[b], Marco Corsi[a], Filippo Daffinà[a], Raffaele Della Corte[b,c], Alfonso Di Martino[c], Claudio Perrotta[c], Salvatore Recano[c], Dino Quattrociocchi[c]
[a]e-Geos S.p.a., via Tiburtina 965, Roma, Italy ; [b]Dept. of Electrical Engineering and Information Technology, Univ. of Naples, via Claudio 21, Naples, Italy; [c]Critiware S.r.l., via Carlo Poerio 89A, Naples, Italy

ABSTRACT

The availability of global and live vessel data streams, improved in time and space resolution, acquired both trough automatic tracking systems (AIS) and satellite target detection, enables a large set of applications, ranging from real-time monitoring of vessel positions to the detection of anomalous behaviours, such as, entering forbidden areas, meeting at sea or sudden change of heading or speed. However, providing both real-time and historical analytics, on significant datasets (e.g., years of data acquisitions) poses severe technical challenges, in terms of performance and scalability. To tackle these issues, in this paper we present a software architecture, underlying a software product named SEonSE, designed to manage large amount of historical vessel data and enabling real-time big data analytics worldwide, with the aim to serve different stakeholders for different purposes. The architecture has been implemented and deployed in the cloud and its functioning has been tested on real live datasets, including global AIS data and satellite vessel detection.

Keywords: Cloud, big data analytics, anomaly detection, maritime, vessel detection

1. INTRODUCTION

The use of the sea as main mode of goods transportation worldwide is progressively increasing over the years. Suffice to think that, according to Eurostat, in 2016 the 80.8% of European exported goods have been transported over the sea, with an increasing trend of about 8% in ten years over other modes of transportation [1]. At the same time, the availability of global vessel identification information coming from "cooperative systems" (e.g. AIS, VMS or LRIT), coupled with vessel information extracted from satellite data (that allow to uncover non-cooperative, or so-called "dark" vessels), is enabling a wide set of applications in the fields of global maritime situational awareness, including the detection of anomalous behaviours [2]. The increase of global sea use also corresponds to an increase in size and density of vessel data available worldwide. For this reason, the use of big data solutions in this domain is gaining momentum [3].

As a matter of fact, the data source that current and future maritime intelligence systems have to handle possesses all the three basic "V"s characterizing a "big data" data set [4], i.e.: (i) *volume*, as gigabytes of AIS and satellite derived vessel data are acquired at global level and on a daily basis; these data need to be properly stored and further analyzed to perform anomaly detection; (ii) *velocity*, as the new data are updated with a pace that can reach a worldwide update at most every 5 minutes; and (iii) *variety*, as multiple data sources are encompassed to build the intelligence required to perform the intended analysis (e.g., to compute an "index risk" associated to each vessel).

In this paper, we present our experience in this direction, presenting the recent evolution of the SEonSE platform towards a cloud-based architecture for big data analytics over worldwide, real-time vessel data. SEonSE (standing for Smart Eyes on SEas) is an e-Geos software product conceived to provide web services, and a related map based web interface, for real-time data and intelligence on vessel information acquired through a variety of sources, such as, the automatic tracking system (AIS), encompassing both collaborative and dark vessels, the latter determined by the on-line analysis of multi-sensor satellite imagery. The first versions of the platform, based on traditional approach and tools (e.g. relational database), were conceived to be used on limited areas of interest over a limited period of time. The need of real-time processing and analysis of worldwide vessel information for global anomaly detection called for a complete re-visitation of the underlying software architecture, towards a cloud-based scalable solution for big data analytics. According to the historical data in our possess, global collaborative vessels data increased of 895% in 5 years, ranging from about 480

GBs in 2013 up to 4.3 TBs in 2018, encompassing about 4 billions vessel data updates per year. On average, about 12 GBs of data are produced everyday, accounting for about 13 millions vessel data updates per day, with frequent updates (in the order of few minutes).

Among the many technical issues arising from the management of such massive amount of real-time data, we specifically focus on the following two: 1) sustained ingestion, i.e., the ability to sustain the pace of updates, being capable to ingest and analyze all new data on time, before the start of the subsequent update; 2) fast response to anomaly queries, i.e., to provide answers to anomaly detection requests, over a given spatial area and in a given time interval, in a reasonable time, without impacting the real-time ingestion; 3) capability to online extract maritime patterns of life by aggregating and fusing all the available information at run-time . To face these issues, the SEonSE cloud architecture is designed to perform scalable on-line data gathering and processing, using in-memory optimizations, and to execute batch pre-processing jobs in the background. This pre-processing approach enables to elaborate and store the outputs of analytics queries in advance, as the ingestion progresses, so to provide almost immediate answers to analysts, if not for the time needed to transfer the result of queries over the network.

In the rest of the paper, we present some details of the cloud-based architecture for big data processing underlying SEonSE (section 2), and provide an initial view of results, focusing especially on the processing time of the global live AIS data stream (section 3). Final remarks and outlook for future work is presented in section 4.

2.   SEONSE CLOUD ARCHITECURE

Figure 1 presents the main elements of the SEonSE cloud architecture for big vessel data analytics. In its current implementation, it is based on (and hosted as) Amazon Web Services (AWS) [5], and, as such, it re-uses some key elements, such as S3 for storage, the NAT Gateway for internal re-routing, and the Application Load Balancer.
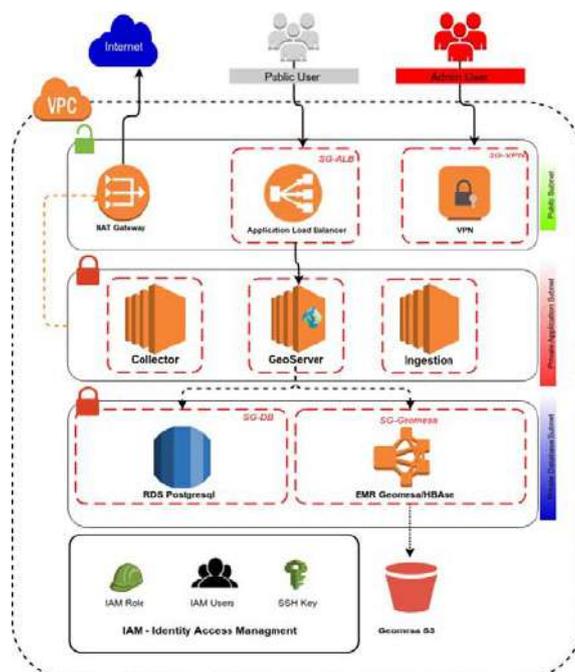
Figure 1. SEonSE Cloud Architecture for big vessel data analytics

The architecture presents 4 main layers, described in the following, starting from the lowest:
- Storage and IAM: at this layer we find the storage of data S3 buckets [5] (used to store raw data and database files, both for Postgres and HBase) and the management of the users along with their roles and privileges to access the platform.

- Private Database Subnet: at this layer we deploy the database components on Elastic Compute Cloud nodes (EC2) [5] in a private subnet, namely a postgresql relational database and a Geomesa/Hbase instance [6] (encompassing a master and more slave nodes). The postgresql relational database, extended with postgis, is used to store configuration metadata and the latest snapshot of worldwide vessel data, particularly useful for the on-line use of the platform. Geomesa, instead, is used to store and retrieve historical vessel data and pre-processed analytics. Geomesa is an open source tool for large-scale geospatial storage, querying and analytics on distributed computing systems. It provides spatio-temporal indexing on top of several nosql database solutions, among which Hbase [7], which we chose due to its inherent scalability, its automatic table sharding capabilities, and the native support of the Hadoop Distributed File System (HDFS). We deployed Geomesa on top of the ElasticMapReduce AWS service to take advance of its built-in scalability support and the native HBase support.
- Private Application Subnet: the private subnet implemented at this layer hosts application level components, among which:
  - o The Collector, which gathers vessel data in real-time from external providers (e.g., AIS vendors) and stores them in raw format (also for backup purposes) on S3;
  - o The Ingestion, that, starting from the last update provided by the Collector, performs the core operations of the platform, namely, producing the live global vessel snapshot on postgresql, updating the historical information on Geomesa, and computing (and storing on Geomesa) anomaly detection events;
  - o GeoServer, a de-facto standard open-source solution to expose collected and processed data (both coming from postgresql or Geomesa) as OGC standard web services, such as WMS and WFS.
- Public Subnet: at this layer we find the components used to expose the managed data to final users, using GeoServer in a replicated fashion (to improve both the performance and the availability) through the Amazon Application Load Balancer service. At this layer, it is also deployed a Virtual Private Network to let admin users access and configure the system.

The architecture share similarities with other initiatives in the maritime domain [8][9], confirming once more the need of the technological shift on the SEonSE platform. The focus of this paper is on the opportunity to perform pre-processing and anomaly computation already during the ingestion, as presented in next subsection, in order to have data about anomaly ready to be queried without further on-line computation.

2.1 The Ingestion and pre-processing chain

Among the architectural components presented, the Ingestion plays a central role. It includes both a processor task to produce the HBase table of historical vessel data and a set of analytic tasks that produce tables of pre-processed data items containing detected anomalies. The implemented chain is exemplified in Figure 2.


Figure 2. The ingestion chain
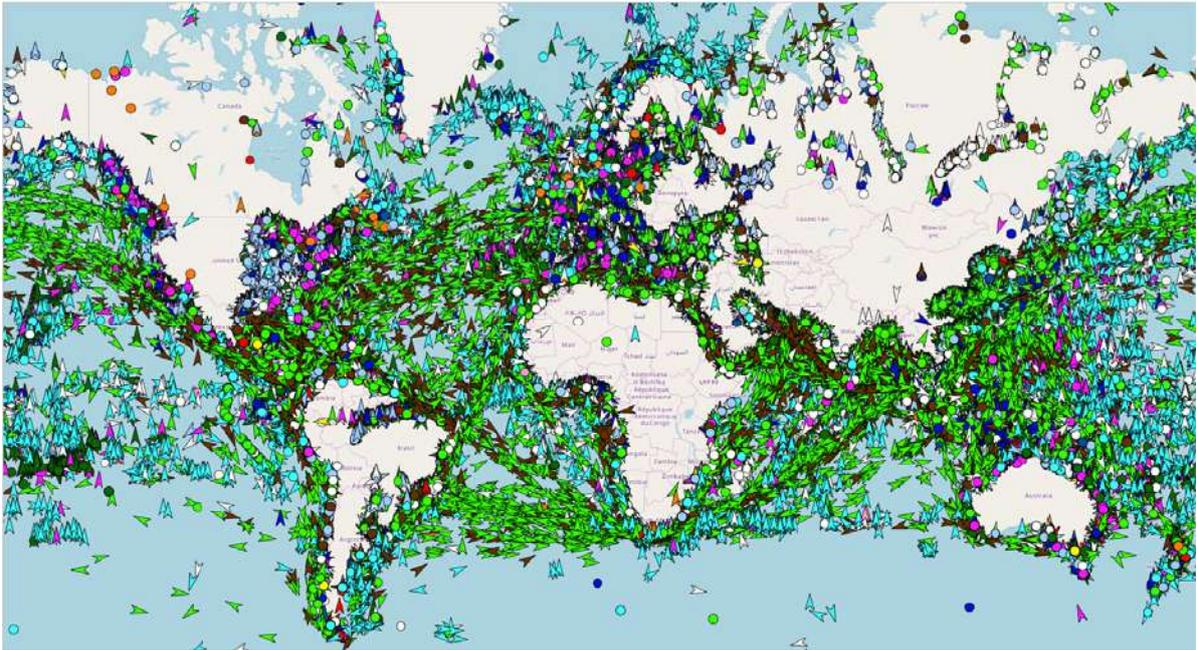
Figure 3. The global AIS vessel live data provided by SEonSE

Worldwide vessel data (encompassing about 300.000 vessels, on average) are gathered on-line by the collector every at most 5 minutes, and saved on S3, which acts as a local backup storage and store for further processing, if not covered already by pre-processing. As soon as a new data instance is saved, the ingestion processes it (through the Live Processor, that continuously waits on newly published contents in S3) and (i) produces the Live vessels snapshot in PostgreSQL, containing always the newest real-time information about vessels, and (ii) updates the vessels history with the newly acquired data. The history accumulates the data of vessel during the overall monitoring period and can be used to perform historical queries, in a given time interval, on peculiar vessels or on a given geographical area of interest (AoI).

The vessel history is then processed by Analytic Processors, that implements several anomaly detection algorithms and produce different types of anomaly events in Geomesa/Hbase. Similarly to the history, the anomalies table can be queried over a given event type, time interval and AoI to retrieve current or past information about vessels' anomalous behavior, such as, entering/exiting an AoI, standing in a AoI, performing a change of name or a sudden change of speed or heading, a rendezvous between vessels, etc. Details on algorithms are omitted here due to company restrictions.

The key advance of the solution is that anomalies are pre-processed during the ingestion. Hence, queries on anomalies do not require to actually elaborate the analysis algorithm, but are read operations from Geomesa, taking advance of its optimized indexes. To keep-up with the high pace and volume of acquired data, the processors are equipped with in-memory databases of vessels information and status, in order to quickly perform the required elaboration (without accessing the databases on the permanent storage) and to store only the results (e.g, anomalies found) on HBase, at the end of a transaction.

3. LIVE INGESTION TEST RESULTS

We implemented, configured, and launched the ingestion chain for global, live AIS data on two Amazon EC2 instances with 2 vCPUs and 64GB RAM. A new snapshot is produced by the Collector on S3 every 3 minutes, each containing about 30000 AIS vessels updates. The execution of the whole ingestion and collection chain completes in about 70 seconds, on average, including the reading of the collected snapshot from S3, the reordering of messages, the update of the global AIS vessel live data (shown in Figure 3), the update of the vessel history table on Geomesa/HBase, and the

pre-processing and storage of anomalies. Hence, in the current settings, the implemented chain is able to keep-up with the pace of updates, producing meanwhile pre-processed data about anomalies (with insert operations of new data that do not block concurrent read operations) that can be subsequently and efficiently queried from Geomesa, without requiring any further processing.

(a)   (b)
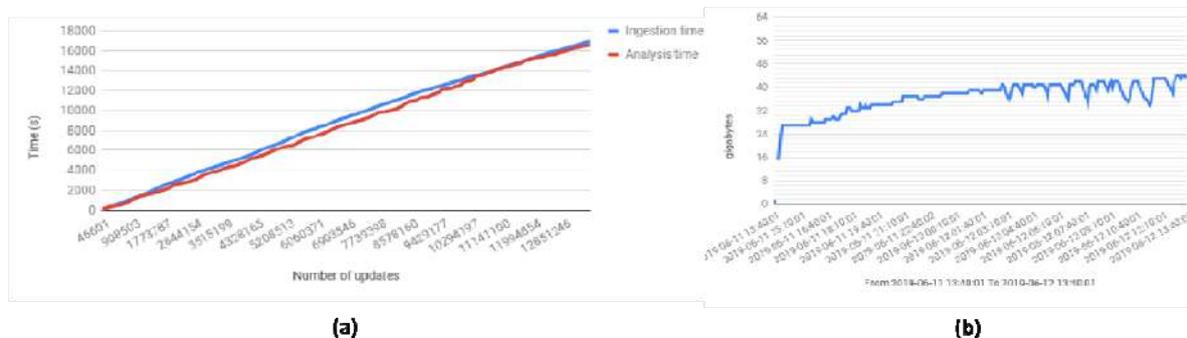
Figure 4. (a) Ingestion and analysis times evolution, and (b) memory usage, over 1 day of collection and analysis

Figure 4.a depicts the evolution of ingestion and analysis times over 1 day of collection. As it can be seen from the figure, times increase linearly with the number of AIS vessel updates, with no observable degradation phenomena over 1 day of operation.

We observe a corresponding increase of the usage of the memory during the process (Figure 4.b), due to our in-memory database, which however stabilizes around 40 GB. Small increases of memory usage are still observable after the stabilization, suggesting us to perform periodic memory maintenance operations (e.g., forcing application reload) once or twice a week. Over 1 day of operation (with data acquired in June 2019), about 13 millions updates are managed, with a total processing time (including both the ingestion and analysis) of about 9.5 hours.

4. CONCLUSIONS

In this paper we presented the architectural evolution of SEonSE, an e-GEOS product for maritime situational awareness and intelligence applications, needed to perform real-time maritime big data ingestion and analytics of vessel information. The proposed solution takes advantage of both state-of-art open-source solutions for the management of large amounts of geospatial information, such as Geomesa/Hbase on top of an HDFS cluster, and in-memory optimizations, to keep-up with the pace of updates. The solution has been implemented and deployed in the cloud, using AWS. Preliminary test results show the practical feasibility of the solution and demonstrate that, with a reasonable amount of computing resources, it is possible to sustain the pace of worldwide live vessel updates, while producing vessel histories and anomalies.

Future work will be devoted to the ingestion and pre-processing of historical data, to produce a ready-to-use permanent store of vessel information for maritime intelligence. This is particularly useful to analyze past behaviors of vessels, and hence evaluate an index risk to be associated to each vessel.

REFERENCES

[1] Eurostat Report DS-022469, "International trade in goods by mode of transport", data extracted in May-June 2017.
[2] R. O. Lane, D. A. Nevell, S. D. Hayward and T. W. Beaney, "Maritime anomaly detection and threat assessment," 2010 13th International Conference on Information Fusion, Edinburgh, 2010, pp. 1-8.
[3] Filipiak, Dominik & Stróżyna, Milena & Węcel, Krzysztof & Abramowicz, Witold. (2018). Anomaly Detection in the Maritime Domain: Comparison of Traditional and Big Data Approach. 10.14339/STO-MP-IST-160.

[4] Paul Zikopoulos, Chris Eaton. 2011. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data (1st ed.). McGraw-Hill Osborne Media.

[5] James Murty. "Programming Amazon Web Services: S3, EC2, SQS, FPS, and SimpleDB". O'Really, 2008.

[6] James N. Hughes, Andrew Annex, Christopher N. Eichelberger, Anthony Fox, Andrew Hulbert, Michael Ronquest, "GeoMesa: a distributed architecture for spatio-temporal fusion," Proc. SPIE 9473, Geospatial Informatics, Fusion, and Motion Video Analytics V, 94730F (21 May 2015).

[7] Mehul Nalin Vora, "Hadoop-HBase for large-scale data," Proceedings of 2011 International Conference on Computer Science and Network Technology, Harbin, 2011, pp. 601-605.

[8] I. Lytra, M. Vidal, F. Orlandi and J. Attard, "A big data architecture for managing oceans of data and maritime applications," 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), Funchal, 2017, pp. 1216-1226.

[9] L. Cazzanti, A. Davoli and L. M. Millefiori, "Automated port traffic statistics: From raw data to visualisation," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 1569-1573.

Development and Integration of
Coastal Surveillance Radar System over four seas

MERT Abdullah Aykut, ÜZÜMCÜ Serdar, ATAY Fatih, HAVELSAN A.Ş,
Mustafa Kemal Mahallesi 2120 Cad. No: 39 P.K.: 06510 Çankaya- ANKARA +90312 219 57 87
Email: aamert@havelsan.com.tr, suzumcu@ havelsan.com.tr, fatay@havelsan.com.tr

ABSTRACT

In the last decade surveillance implementations get more interest due to the increase of vessel traffic and at same time increase of terrorism, smuggling activities, and illegal immigration activities on waterways. The EU regulations and combat against criminalism is on the uppermost challenges at the Turkish maritime borders which leads to a national initiatives in maritime surveillance with the goal of national security in four seas in relation to the geographic location of the country. The Coastal Surveillance Radar System (CSRS) provides situational awareness with adequate radar coverage of the coasts, territorial waters and exclusive economic zones of Turkey and create a defined maritime picture supported by data received from radars, electro-optical sensors and some other military and ministerial institutions related to coastal security and sea traffic. The CSRS aims to increase cooperation level of military and ministerial institutions and organizations in charge of coastal areas of Turkey, and enhance efficiency of reconnaissance, patrol, and search and rescue activities.

Keywords: coastal surveillance radar system, maritime border, patrol, and search and rescue

1. INTRODUCTION

The Coastal Surveillance Radar System (CSRS) Project aims at securing a maximum possible sensor coverage in the maritime jurisdiction at the times of peace and war, and form a near real time like surface picture to ensure more efficient patrol and search-rescue operations in four seas, so that it would be possible to enhance the level of collaboration with the other public entities. The project goal is make it possible to perfectly and efficiently monitor all the territorial waters and adjacent zones, and struggle with such crimes as trafficking, illegal immigration, fishery and maritime pollution.

The CSRS Project is expected to form an recognized Surface Picture by means of the fusion and identification of the system trace data through the Automatic Identification System (AIS) as received from the sensors installed in the Coastal Surveillance Stations in the Turkish Maritime Jurisdiction and the radars and systems of the public entities/organizations integrated with the Maritime Information System. The CSRS Project is an integrated project of wide coverage, on which the information produced by the Maritime Information System is mutually shared with many public entities/ organizations. The following Figure 1 shows The Coastal Surveillance Radar System (CSRS) Project.

The system comprises of two networks, the Coastal Surveillance Data System (CSDS) and the Maritime Information System (MIS). For the information exchange between the Ankara Operation Headquarters, Identification and Monitoring Centres and the Coastal Surveillance Stations, CSDS will be used, while the identified Common Data Pattern defined on MIS will be essential for the Service Based Architecture to secure the information exchange among the civil/military public entities in line with the individual entity requirements and within the framework of the sharing rules. Following are the entities covered by MIS: the Coast Guard Command, the Ministry of Transportation, Maritime Industry and Communication, the Navy Command, the Ministry of Environment and Urbanization, the Ministry of Food, Agriculture and Husbandry, the Ministry of Customs and Commerce and the General Directorate of Meteorology.

Figure 1. The Coastal Surveillance Radar System (CSRS) Project

During the project, one operation centre, four identification and tracking centres and 11 coastal surveillance stations are equipped with radars, Electro-Optical (E/O) and other sensors for integration with the public entities through the CGRS software. The coastal surveillance stations act as local control stations, the tracking centres work as zonal control station and the operation centre behave as national control station. The hierarchical and distributed data processing are performed over these mentioned control stations [2].

1.1 The Purpose of the Project

The purpose of the project is to establish a maritime tactical picture with integration of enough radar coverage over the whole coastal line and utilizing the electro-optical and integrated state agencies information. It has the goal of increasing the efficiency of law enforcement and search & rescue operation agencies. The gathered information and fused data are publish with state agencies.

1.2 System Components

The operation center is connected with the identification & tracking centers (KİM) and the coastal surveillance station (CSS) dedicated network with sufficient security measures. The CSS are equipped with radar systems, E/Os (IR and TV) and communication equipment [6]. Figure 2 shows an equipment located on coastal surveillance station CSS.
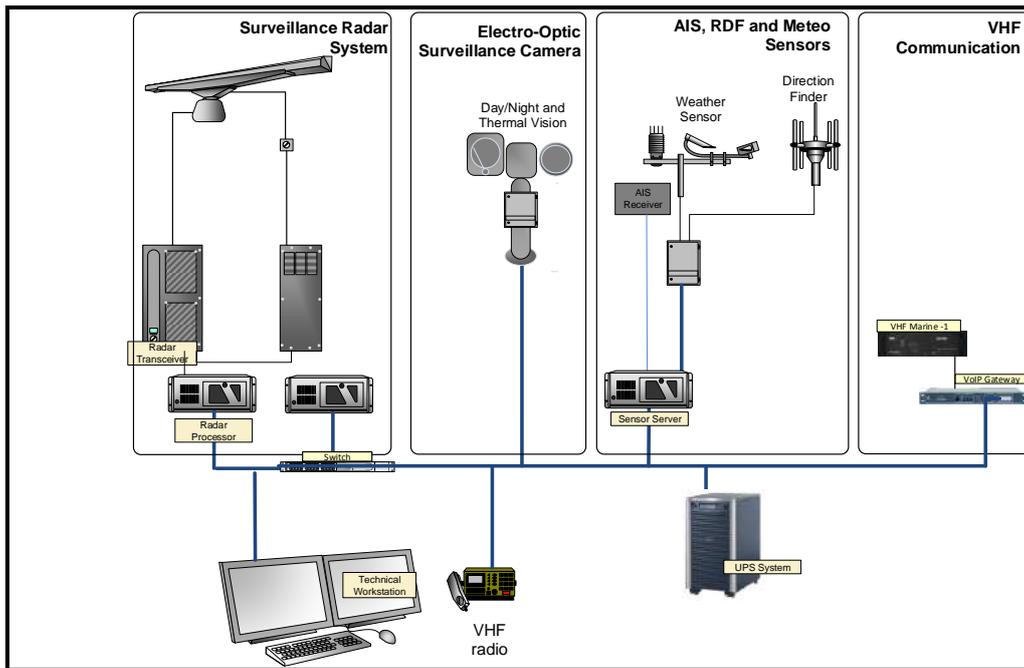
Figure 2. Equipment located on coastal surveillance station CSS

2. MAIN FUNCTIONS OF CSRS

The Coastal Surveillance Radar Systems software has the following functionalities.

2.1 Sensor Management Detection/Control

The Sensor Management Detection function is targeted to monitor the surface mobility in the coverage area of the sensors (Radar, AIS, E/O) located in the Coast Guard (CG) Stations and Traffic Surveillance (TS) Stations, which are designed as to be remotely controlled unmanned stations.

2.2 Data Fusion/Identification

The information received from the CGRS sensors and through MIS are fused to form the tactical seascape. The identification data received from the AIS's are used to automatically produce the identified seascape by insertion of identification data from E/O. The VTS software has a data fusion functionality for position calculation. The AIS and radar data are fused together to get the closest real position of ship, some track information must be visualized inside the system. Fusion process can be done two different ways, such as distributed (Location based) and centered (Sensor based) fusion. For distributed fusion 1st level fusion will be done at station between radar and AIS tracks. 2nd level fusion will be done at the center between station tracks. For centered fusion All AIS and Radar tracks will be processed at the center. Since the AIS source is the ship itself. The same data can be received from different sensors (station). The system drops if the same data comes another sensor. Fusion process will be done between only one AIS track and other radar tracks. This method increases center process, and center hardware capability.

2.3 Decision Support

The decision support function checks the element status information (location, local weather forecast, etc.) and the Coast Guard resource information for determination of the most appropriate resources, and provides the commissioning suggestions to the responsible advisor at the operation center.

2.4 Event Management

The event management function gives the capability to ensure the management of research/ rescue and other duties, and the determination of such planning as routing, element number, etc. It determines the most appropriate Coast Guard units that can intervene any event, and draw up and implement the operational plans.

2.5 Communication (Voice, VoIP) Control

The communication control function allows to manage VHF radio communication to ensure necessary coordination with the elements / targets in the field of interest and an interface to public telephone network.

2.6 Record and Replay

It provides a synchronized data store for the operators to make a post-event analysis of the sensor data and system information (including voice and video). The recorded data can be optionally archived on the external media through the operator software.

2.7 System Monitoring and Management

The monitoring of system equipment facilitates the system management, and increases the operational availability. It provides any necessary inputs for transfer of the resources of disabled centres into the standby centre. The status data of the equipment in the headquarters and stations are collected and monitored.

2.8 Stakeholder Entity Integrations

The currently available systems of the public entities to be integrated through MIS are included in the CGRS by means of the adapter software components.

3. CAPABILITIES & SPECIFICATIONS OF CSRS

Figure 3 shows a simple scenario of mission execution based on a reporting received. Upon receipt of a reporting about a sea vessel by phone call, e-mail or fax, or emergency calls from VHF and AIS of a sea vessel, the Operator examines the track created in the system for the maritime vessel if it is in the coverage area of radar and EO. The Operator uses EO sensor cameras that have the necessary coverage for the vessel to contribute to the identification of the track and to verify the reporting / call received. He can query the system related to the track and see the event history log about the vessel using the previously registered information in the system. For example, if the vessel is in the list of Suspicious Vessels in the system, it is more likely that the notification received is correct. If the vehicle is out of the range of the system sensors, the Operator initiates a track at the location reported in the system, and makes manual identification for the track using the EO sensors and notification information received through the call. The operator-initiated track is converted into a system track via fusion and correlation capability of the system in ITC and then it is sent to the MOC for global fusion and correlation by using additional information provided by other ministerial systems [1]. Finally, the track is included into the Recognized Maritime Picture. According to the type of the event information received via the call, the Operator may initiate a Task Planning or Search / Rescue Planning. The operator sends the plan to the patrol ships of the Coast Guard Command, which act as law enforcement agencies, through the CSRS communication infrastructure. When the CGC patrol received the task plan / order, it starts to execute the task plan and goes to the specified location and carries out the duties in the scope of the plan. Following the completion of the task, a report about the completion of the task plan/order, including necessary information, is prepared by the CGC patrol ship and sent to the ITM. The task result report is transferred to the CSRS database and the task is finalized.
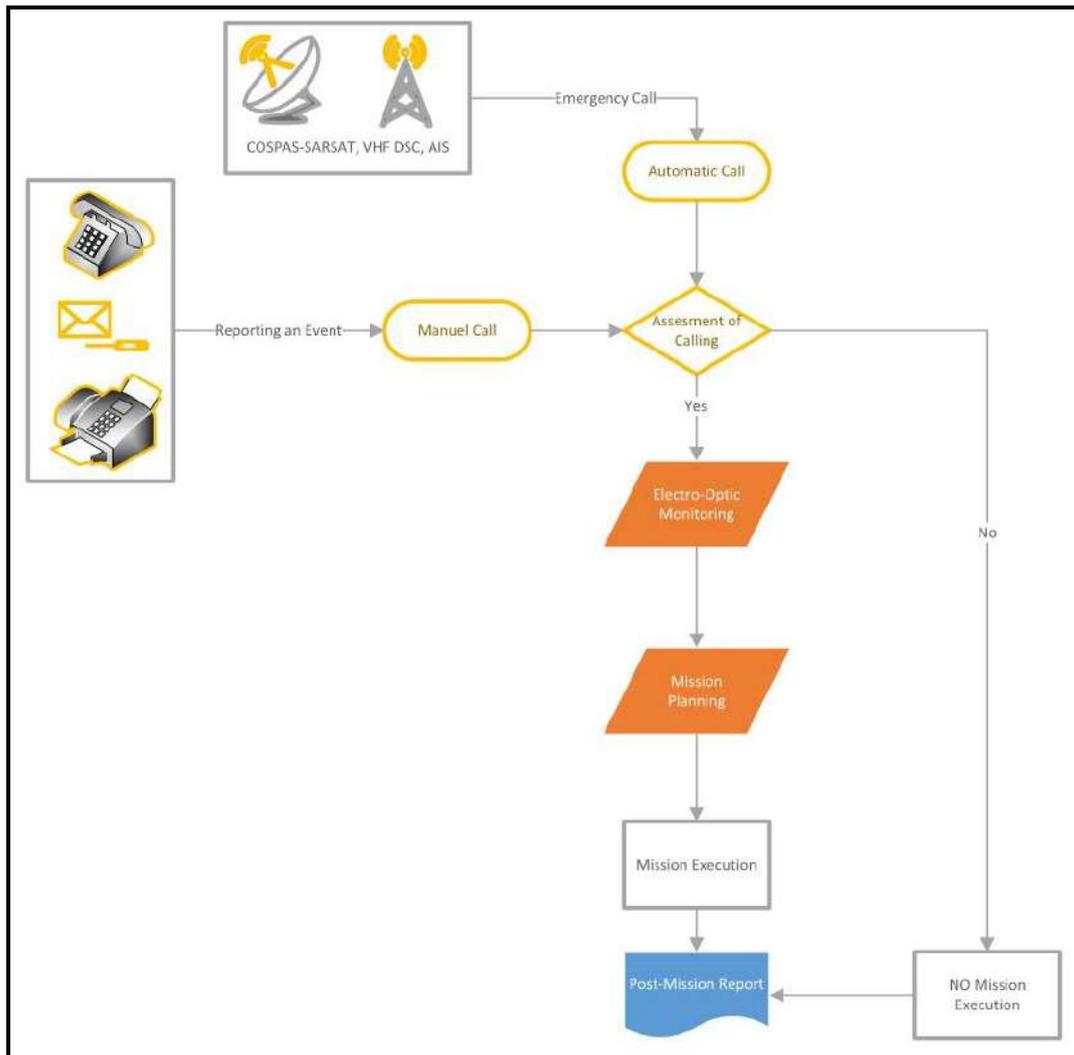
Figure 3. A simple scenario of mission execution based on a reporting received

Figure 4 shows the control system of the radars installed in the network and locations of vessel. All radars installed in this network use the same hardware configuration and software so that the appearance of the radar display will be the same. Some detected targets will be shown as small dots, large dots, and lines, which are depend on the shape and size of the targets similar to sea target recognition [3]. The Automatic Identification System (AIS) monitors the information sent by the ships such as ID, destination, type of cargo. This AIS is a supplementary to the radars. Figure 4 shows the display of this AIS with the Marmara Sea region in which the Turkish Straits Vessel Traffic Services radars are also integrated [4][5].
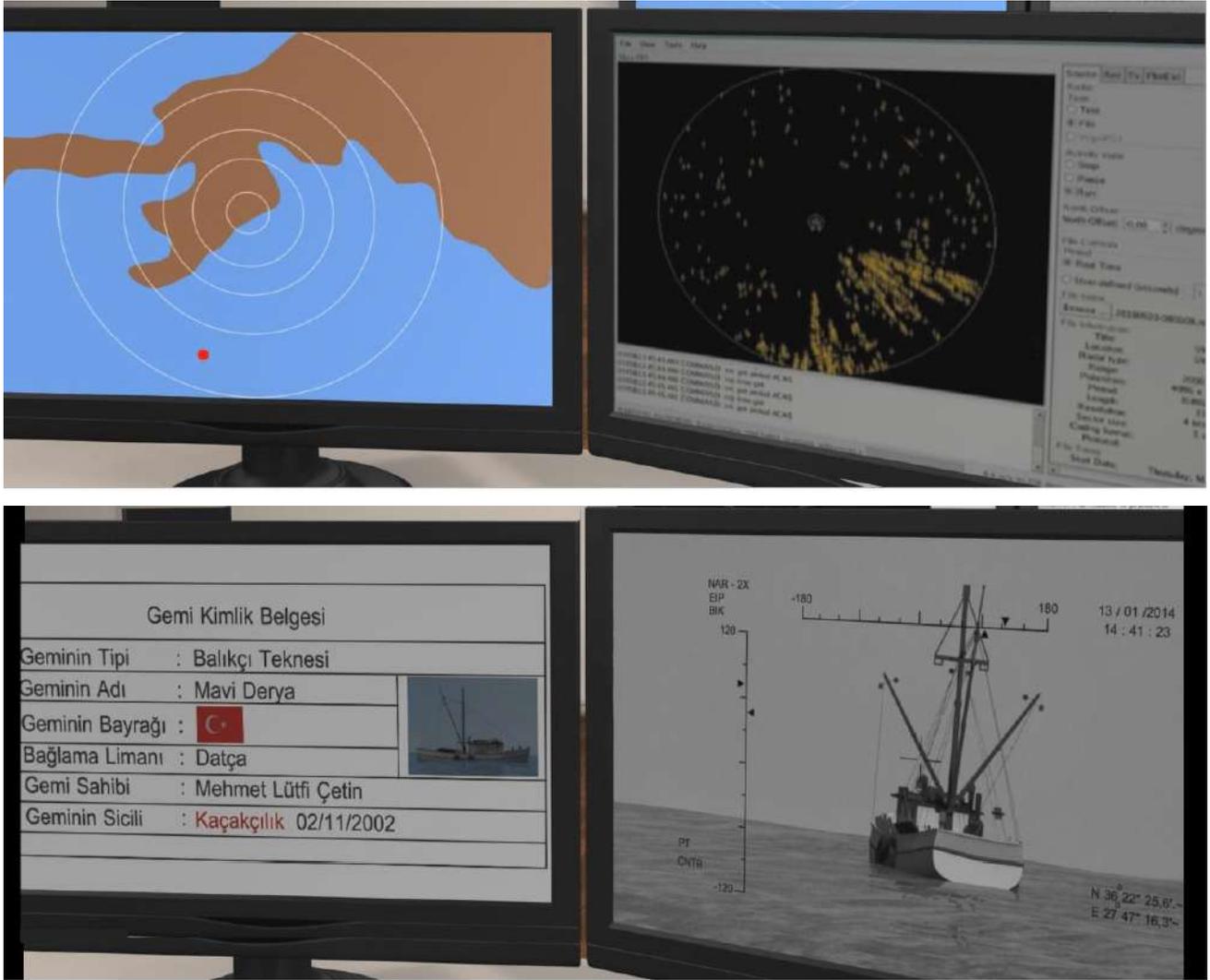
Figure 4. Displays of CSRS Application Software

4. CONCLUSION

Systems of Systems projects as in this paper a brief layout given CSRS technology should be used. The Coastal Surveillance Radar System (CSRS) provides situational awareness with adequate radar coverage of the coasts, territorial waters and exclusive economic zones of Turkey. The CSRS aims to increase cooperation level of public institutions and organizations in charge of coastal areas of Turkey, and enhance efficiency of reconnaissance, patrol, and search and rescue activities. With CSRS, Turkish Coast Guard will be able to deal much more decisively with smuggling, illegal immigration, fishing and sea pollution.

REFERENCES

[1] Carthel, C., Coraluppi, S. and Gignan, P., [Multisensor Tracking and Fusion for Maritime Surveillance], NATO Undersea Research Centre (NURC) Viale S. Bartolomeo 400, 19126 La Spezia, Italy (2007).

[2] Singh, A. K. and Aggarwal, V. "Coastal Surveillance in Multi Sensor Environment: A Design Approach," Pankaj Tiwari Centre for Artificial Intelligence and Robotics, Bangalore, Karnataka, India-560093 (2016).

[3] Van den Broek, A. C., Van den Broek, S. P., Van den Heuvel, J. C., Schwering P.B.W., and van Heijningen, A.W.P., "A multi-sensor scenario for coastal surveillance," TNO Defence, Security and Safety, P.O. Box 96864, 2509 JG, The Hague, The Netherlands (2007)

[4] Turhangil, H. C., Döner, M., "Turkish Straits Vessel Traffic Services (TSVTS)," HAVELSAN, ISSN 2667-8187, Feb. 2019.

[5] Üzümcü, S., Mert, A., "Usage of Digital Twin Technologies during System Modeling and Testing in Vessel Traffic Services System Project", Incose Symposium 2019.

8

Maritime Situational Awareness and ethics

Contents

Redefining Situation Awareness for the Maritime Information Warfare Domain

Anna-Liesa S. Lapinski

Defence Research and Development Canada, 9 Grove Street, Dartmouth, NS, Canada B3A 3C5

ABSTRACT

Maritime Situational Awareness (MSA) is generally understood in the physical world. The challenge is achieving it, not understanding the concept. However, what does situation awareness (SA) mean in the Maritime Information Warfare (MIW) domain? An ongoing project at Defence Research and Development Canada is addressing that question. Stepping back from the maritime domain, it can be observed that Information Warfare (IW) overlaps with all other warfare areas, but as a warfare area in itself, the concept of SA should still exist. In an effort to ultimately understand what SA in MIW means, this paper first breaks down the information battlespace into four key components: content, content generator, conduit, and repository. These are posed as sub-battlespaces in the information battlespace, analogous to above water and underwater sub-battlespaces. Then, the types of SA elements are proposed: physical, digital, and human. This is important because while one warfare activity might be heavily human centric, another might be technology centric. However, to adequately conceptualize SA for any warfare activity, there should be cognizance that all three types could exist. The paper concludes with a proposed definition of SA in IW that could be translated to the MIW domain.

Keywords: Information Warfare, Maritime Information Warfare, Situation Awareness, Situational Awareness, IW, MIW, SA

1. INTRODUCTION

On April 1, 2018, Defence Research and Development Canada (DRDC) started a new project in the Maritime Information Warfare (MIW) program: Maritime Information eXploitation (MIX). Part of the MIX project is to recognize the elements of the information environment that must be catalogued to achieve the required situation awareness (SA) for activities within the MIW domain.

To conduct activities in any warfare environment, awareness of the elements surrounding and affecting the activities are required. The accumulation of the necessary awareness is a key step in achieving SA. One of the questions MIX is meant to answer is: what does the term "situation awareness" mean in the MIW domain? However, before that can be answered, it seems logical to take a step back and answer the question: what does the term "situation awareness" mean in the Information Warfare (IW) domain? Once a general understanding is attained of the meaning of SA in IW, it can be translated to answer the original question. After SA in MIW is understood, the next step will be to isolate the elements that need to be catalogued to achieve SA in MIW.

To understand what SA in IW means, this paper begins by breaking down the information battlespace into parts. It then identifies the common categories of elements that contribute to SA in the IW domain. A generic definition of SA in IW is then proposed.

1.1 Information Warfare

In an article in which a commander within Directorate of Naval Information Warfare (DNIW) in Ottawa, Canada is interviewed on IW [1], IW is defined as "the provision, assured use and protection of information, processes, systems and networks, and limiting, degrading and denying that of adversaries to achieve operational advantage across the battle space." The IW domain, therefore, is seen as touching many aspects of warfare; e.g., communications, cyber operations, electronic warfare, information operations, and intelligence. MIW is therefore information warfare as it relates to the maritime domain.

1.2 Situation Awareness

A popular definition of SA is by Endsley [2] which states that "situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." This definition was actually developed within the context of pilot cognition. The question is then, how does that definition translate into the information battlespace?

To begin to answer that question, it is useful to divide Endsley's definition into three parts: cataloguing (i.e., perceiving), comprehension, and predicting. Arguably, in Endsley's definition of situation awareness, comprehension and anticipating the future cannot be achieved until "the elements in the environment within a volume of time and space" begin to be catalogued; therefore, the first part of the definition is critical to attaining the second and third parts of the definition. This paper will focus on that first step: the cataloguing of the required elements as this is a key step to fully attain SA.

In IW, to catalogue the elements it is first required to acknowledge that at the most general level "space," in the Endsley definition, is the information battlespace. In the information battlespace, there may be "inform and influence" operations (within the psychological realm) as well as technical operations (within the technical realm), as Porche et al. [3] proposes. For example, spreading propaganda is an "inform and influence" operation, while disabling an adversary's communication infrastructure would be a technical operation. Porche et al. [3] provide examples of a large variety of specific types of activities under these two operation types.

It should also be pointed out that the IW SA needed to carry out a specific activity will be smaller in scope than the all-encompassing SA for the entire IW domain. Whether for the smaller scoped IW SA or the all-encompassing SA aimed at the entire IW domain, it is useful to understand the information battlespace prior to building SA. The next section attempts to do this by breaking down the information battlespace into parts.

2. BREAKING DOWN THE INFORMATION BATTLESPACE

The following presents an argument that the information battlespace is made up of four smaller battlespaces: content battlespace, content generator battlespace, conduit battlespace, and repository battlespace (these battlespaces are a revision of what was proposed in Hazen et al. [4]). Dividing the information battlespace into smaller battlespaces is analogous to dividing the ocean battlespace into above water and underwater. Within each of these smaller battlespaces, both offensive and defensive activities can take place. In the same way ocean activities can span both above water and underwater battlespaces, activities in the information battlespace can span more than one of these smaller battlespaces. The following explains the smaller battlespaces, with examples.

Figure 1. The information battlespace divided into parts; i.e. smaller battlespaces.

Content is an all-encompassing term that represents the information in the IW domain. Content could be a thought, a memory, sensor output, algorithm output, text, an image, a document, a song, etc. Content would also include the medium that contains the content, such as paper, binary file, sound waves, etc., as it can be challenging to separate the content from the medium. Activities in this battlespace might include obtaining the adversary's content, obtaining content about the adversary, protecting our own content, controlling the publicly released content so as to control the narrative, etc.

A **content generator** produces the content described above. It could be a person's brain, something felt, tasted, heard, seen, or smelt by a person (i.e., the human acting as a sensor), a sensor, an algorithm, etc. Both humans and technology can be content generators. Activities in this battlespace might include acting against an adversary's content generator, protecting our own content generator, purposely having a generator produce false content, etc.

The **conduit** is the mechanism for getting content from one entity to another. Conduit is used as an all-encompassing term to represent the manner in which content is moved. This could be speaking, sign language, a computer network or

broadcast, an application used to move/distribute content (e.g., a text messaging application, Facebook, Twitter), the handing of a hard or soft copy of a document to someone, etc. Even though Figure 1 shows only one conduit, there could be many unique conduits placed in series or parallel carrying the same content. For example, the process of converting a thought into spoken words that travel through the air as soundwaves can be thought of as several conduits in series; e.g., from neurons firing, to manipulating mouth muscles, to pressure variation in the air. Also, content going into a conduit may end at a content generator; or content leaving a repository may end in another repository. Activities in this battlespace might be to disrupt, infiltrate, or protect a conduit, set up a conduit to spread false content, etc.

The **repository** is a place for content storage. It could be a person's brain, a server, a filing cabinet, etc. Activities in this battlespace might be to infiltrate, destroy, or protect a repository, etc.

3. SA ELEMENTS OF IW

Under these four smaller battlespaces, it is proposed that there are three types of elements to be catalogued so as to achieve SA: physical elements, digital elements, and human elements (includes psychological elements). The following gives examples of elements under the smaller battlespaces for illustrative purposes.

Physical elements used to help build SA would include an awareness of the physical world that has a direct bearing on what is happening in the information battlespace. For example, in the content generator battlespace, an SA-building element could be the knowledge of the existence and the status of the adversary's sensors. In the content battlespace, an SA-building element could be the physical status of a collection of hard-copy-only paper documents, for example, ones that are in transit. In the conduit battlespace, an SA-building element could be the status of the physical security of critical telecommunication fibres. In the repository battlespace, an SA-building element could be the status of physical infrastructures that house important libraries or servers. It would also be the physical components of the server; such as drives, the casing, etc.

Digital elements used to help build SA would include awareness of the digital world that has a bearing on what is happening in the information battlespace. For example, in the content generator battlespace, an SA-building element could be monitoring critical algorithms for evidence of tampering. In the content battlespace, an SA-building element could be monitoring internal electronic documents for tampering or corruption. In the conduit battlespace, an SA-building element could be status reports on attempted intrusions into a protected computer network. In the repository battlespace, an SA-building element could be monitoring the structure (i.e., not the data) of critical databases, looking for corruption.

Human elements used to help build SA would include facts that are human related, that have a bearing on what is happening in the information battlespace. For example, in the content generator battlespace, an SA-building element might be tracking the location of human assets who provide eye-witness information. In the content battlespace, an SA-building element might be monitoring how the public or adversary perceives a situation. In the conduit battlespace, an SA-building element might be monitoring a network of people that is in place to relay critical information. In the repository battlespace, an SA-building element could be monitoring the cognitive abilities of a workforce [6] and how their cognitive function is impacting short term and long term memory.

Dividing up the information battlespace and SA elements as described above is not critical for achieving SA in IW, but given how large the information battlespace is, it is proposed to be useful. Such a division could aid in identifying gaps and needs, as well as helping to combat the enormity of the SA problem. It is also an explicit reminder that there are physical, digital, and human elements to SA in IW for those who might be biased to one type of element.

It is promising that the battlespace division and the SA elements encompass cyber-physical-social systems (CPSS). A cyber-physical system has both physical and software components interwoven and interacting with each other, such that the interactions alter the system; e.g., autonomous vehicles. A CPSS has humans also woven into the system. For example, Zhong et al. [5] proposes that a command and control organization could be treated as a CPSS. A CPSS has content generators, content, and conduits as part of the system. Even the terminology, cyber-physical-social, generally aligns with the terminology used here: digital, physical, and human elements. Given that monitoring a CPSS might be a task in maintaining SA, it is encouraging that the model can account for all the aspects of a CPSS.
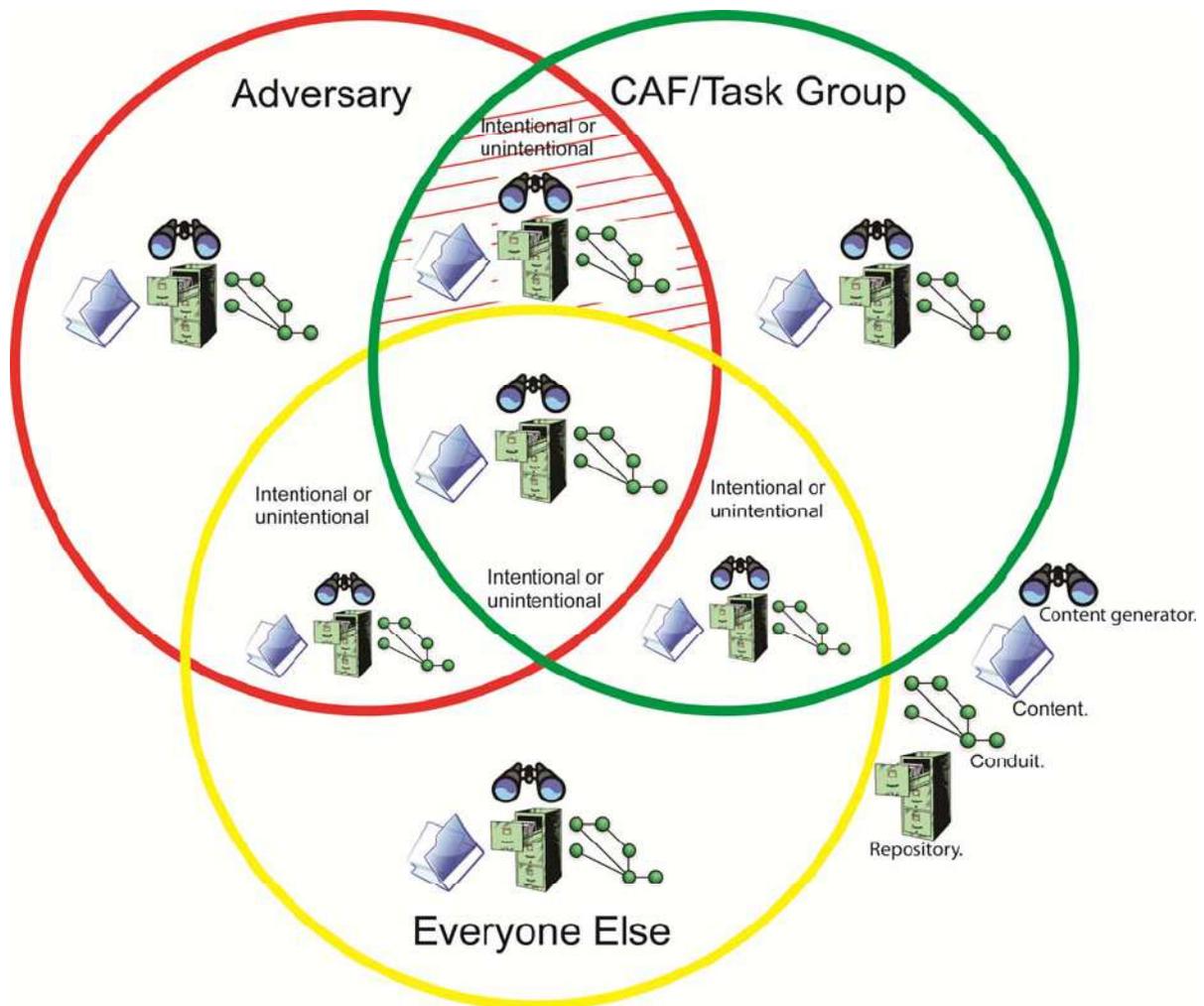
Figure 2. A Venn diagram of access to content generators, content, conduits, and repositories. The green circle indicates what we actually access. The red circle indicates what the adversary has access to. The yellow circle indicates what everyone else (possibly including other government departments) has access to. When the circles overlap, it means both parties have access to certain content generators, content, conduits, and repositories. The segment with the red lines indicates the overlap that the example focusses on.

4. DEFINITION OF SA IN IW

So, what does the term "situation awareness" mean in the Information Warfare domain? To modify Endsley's definition, it is postulated that it means:

> *The perception of the physical, digital, and human elements within the content generator, content, conduit, and/or repository battlespaces of the IW environment, during a volume of time; followed by the comprehension of the meaning of the elements, and the projection of their status in the near future.*

In MIW, all the SA elements would be those specifically supporting maritime activities and operations.

The changes to the Endsley definition may impact the design of future decision support systems by bringing awareness to the needs of MIW. For example, those who generate a Recognized Maritime Picture (RMP) may decide to include information aspects in the picture that have not been included previously or might want to generate a separate Recognized Maritime Information Picture. This would impact the design of future systems supporting RMP generation. However, the level of importance MIW SA has and whether the design of a decision support system needs to be altered, depend on the needs of the mission and mandate that the decision support systems are supporting.

5. AN EXAMPLE

Imagine someone is tasked with providing situation awareness specifically regarding access within the four MIW battlespaces for a specific maritime mission. The general Venn diagram regarding access might look something like Figure 2. We consider here the part of the Venn diagram where the Canadian Armed Forces (CAF) and the adversary both have access, or are suspected to both have access, to the same things. Table 1 gives fictitious examples of information that, in the context of the example, is being used to build MIW SA. The information is organized based on the battlespaces and SA elements discussed in this paper. For clarity, only the Digital-Content Generator box has more than one SA-building report, but in reality this might not be the case.

Table 1. The information battlespace versus the types of situation awareness elements found in IW. A blank box implies there is nothing to report.

	Content generator	Content	Conduit	Repository
Physical SA elements	Adversary known to gain physical access to coastal AIS receiver network. Monitoring via commissionaire in progress. No suspicious activity in last 24 hours.		Adversary known to gain access to farm under which digital fibre is laid. Monitoring via satellite in progress. No suspicious activity in last 24 hours.	Suspicious access to room containing database servers at 03:07. Records indicate they left the room at 03:20. Investigation in progress.
Digital SA elements	Pattern of life algorithm used to guide surveillance tasks was confirmed to be altered by a hostile outside entity. Monitoring of algorithm code in progress. No suspicious activity in the last 24 hours. Data provider known to also sell to adversary has stopped providing data since 15:08.	Modified documents placed on a lightly protected CAF network have not yet been downloaded.	Unclassified network was infiltrated by unknown entity at 04:22 and then kicked off the network at 04:23. Investigation in progress, but initial findings suggests an adversary sympathizer actor.	
Human SA elements	Data analyst reports being approached by suspicious individual at a bar Tuesday PM. Background check of the individual in progress but initial findings indicates they might belong to an adversary cell.	See Content generator note.		

While the example is simply for illustrative purposes, it should be stated that if the investigation mentioned in the Physical-Repository box results in discovering it was a trusted CAF member who entered the room, that report might be moved to the CAF/Task Group access list to maintain awareness. Other investigative findings might also move information to another place in the Venn diagram in Figure 2.

However, the task and the Venn diagram are merely being used to show how information can be categorized. All the captured information in Table 1 helps form the MIW SA regarding access and can contribute to a Recognized Maritime Information Picture.

6. CONCLUDING REMARKS

Further research is required to validate that the model outlined in this paper is correct and useful. One test of its accuracy will be to identify an actual maritime mission/operation and identify all the elements of the information environment that must be catalogued to achieve the required SA for activities within that mission/operation. The mandate to maintain a recognized maritime picture and common operating picture appear to be adequate initial test subjects. In this situation, monitoring the content generators, content, conduits, and repositories, to ensure continuous and accurate information would be a large part of maintaining MIW SA for that mandate. The SA elements would likely be dominated by physical and digital elements, with minimal human elements.

REFERENCES

[1]   Blakeley, D., "Information as War," Royal Canadian Navy, Crowsnest, 6-7 (2017).
[2]   Endsley, M. R., "Situation awareness global assessment technique (SAGAT)," Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, 3, 789-795 (1988).
[3]   Porche, I., Paul, C., York, M., Serena, C. C. and Sollinger, J. M., [Redefining information warfare boundaries for an army in a wireless world], Rand Corporation, (2013).
[4]   Hazen, M. G., Isenor, A., Desharnais, F. and Randall, T., "Characteristics of Information Warfare: The Battle for the Narrative," 22nd International Command and Control Research and Technology Symposium, 1-14 (2017).
[5]   Liu, Z., Yang, D.-s., Wen, D., Zhang, W.-m. and Mao, W., "Cyber-physical-social systems for command and control," IEEE Intelligent Systems, 26(4), 92-96 (2011).
[6]   Yan, C., Fu, K. and Xu, W., "On Cuba, diplomats, ultrasound, and intermodulation distortion," Computers in biology and medicine, 104, 250-266 (2019).

Collaborative Space-based Maritime Situational Awareness (CSMSA)
- Pathway to Global Maritime Cooperation for
Security, Safety, Environmental Protection, and Resource Conservation

Guy Thomas[a]

[a]Multinational Maritime Security Center of Excellence, 2100 Mt Royal Ter,
Baltimore, MD 21217-4848, United States of America

ABSTRACT

Collaborative Space-based Maritime Situational Awareness (CSMSA), also known as C-SIGMA, "Collaboration in Space for International Global Maritime Awareness " envisions a global collaboration of all space-faring nations linking together existing and planned unclassified space system capabilities to form a worldwide collaborative network. This effort would be coordinated via regional centers to create international global maritime awareness, available to all nations of good will. It would be a huge step toward global maritime security but it would also provide substantially improved safety, environmental protection, resource conservation, as well as disaster mitigation and recovery. This paper also discusses both the need for C-SIGMA and what its components should include. These components include the software to know which sensing satellite to task when, the utility of five different satellite types, and the dynamic data analysis tools to build knowledge and understanding from the derived information and other maritime related data sources. The questions of governance are not specifically addressed as those answers are in the political realm, but suggestions are provided.

Keywords: Satellites, maritime security, maritime situational awareness, maritime domain awareness, space systems, S-AIS, SARsats, C-SIGMA

1. INTRODUCTION

"Maritime Security has different dimensions, including but not limited to Maritime Situational Awareness (MSA), Law enforcement, maritime safety, maritime environment, maritime science & technology, maritime trade & economy, maritime law and public health. Therefore, in national terms, Maritime Security can only be achieved by a "whole of government" approach. If we succeed in applying this approach together with like-minded countries in a multi-national environment, we can attain our common Maritime Security objectives"

"There can be no Maritime Security without Maritime Situational Awareness."

Figure 1. Satellite AIS (S-AIS) Tracks now define the Maritime Domain

2. BACKGROUND

In the 21st century it is well known that the cyber world has expanded exponentially but unnoticed by many, since 2004 and increasing steadily since then, there has also been an on-going revolution in space-based Earth observation systems and, led by space-based AIS, their utility over the world's waterways has increased dramatically. These capabilities not only support safety and security at sea but can also significantly assist in economic and environmental stewardship and resource protection, as well as disaster mitigation and recovery. This is especially true of the remote areas of the world such as the Arctic, and the resource rich areas in the underdeveloped world such as the Gulf of Guinea, the South China Sea, Micronesia, and the Indian Ocean. The potential contributions of space-based Earth observation systems to maritime awareness is of growing interest to the world's naval and law enforcement forces, as well as to environmental preservationists, governmental transport, commerce, maritime, environmental protection, and disaster preparedness ministries, in addition to ship brokers, and others with an interest in the marine domain, its environment, and the protection of its resources, but coordination to maximize these capabilities is lacking.

Ongoing research started just after 9/11 (September 2001) shows that no single country or international organization has the ability and resources to fully support the safe, secure and efficient use of the maritime domain as well as the conservation and protection of the marine environment with its finite resources of fish, minerals and oil, as well as to substantially assist oceanic commerce. In that no one country has sufficient resources, including space craft, substantial international collaboration is essential to achieving these objectives in a balanced manner. Indeed, this effort may need to be managed by an agency of the United Nations. Among the greatest needs, as well as the greatest opportunities for international collaboration are presented by the multiple national and regional efforts to develop the doctrine and concept of operations to coordinate the use of the space technologies now available for detecting, identifying and tracking vessels well offshore, on a global scale. These systems are especially suited in areas with shared international interests such as the Arctic, or in pirate infested waters, or in areas known to support smuggling or resource theft of all types such as the Gulf of Guinea.

As of mid-2019 there is a virtual tidal wave of new space-systems with earth observations capabilities being planned and built. This includes a robust number of satellites (175+) equipped to collect the United Nations' International Maritime Organization (IMO) mandated Automatic Identification Systems shipboard anti-collision beacon. There are also several dozen of synthetic aperture radar satellites now being built. These are the two most critical part of space-based maritime awareness. C-SIGMA would not be possible without S-AIS as it reports the position of every legal ship many times a day, often with less than a two-minute delay between position reports, and SARsats are very useful because of their all-weather, day/night imaging capabilities.

Additionally, as indicated above, the cyber world is also enjoying a similar expansion, with artificial intelligence and machine learning leading the way. The time to take advantage of these significant opportunities that are on the horizon and rapidly bearing down on the world is now, as this tide of technology rises. Catch the wave! Seize the moment!

3. SYSTEM DESCRIPTION

There are eight elements that must be integrated for effective results. (Five in space, two on the ground.)

Two of the six different satellites types employ passive sensors:

1. Automatic Identification System (AIS), an automated short message system designed for collision avoidance and traffic control in congested waters but is now also used globally as a primary ship identification and tracking system. It is the key component of this concept.

2. Unclassified Signals Intelligence satellites (SIGINTSats)   (New Capability as of 2019)

Another three of the six different satellites types employ imaging sensors:

3. Synthetic aperture radar satellites (SARsats)

4. Electro-optical (EO) imaging satellites

5. Video optical satellites

The sixth element are the communications satellites

6. M2M communication satellites. Individual transponders sending short formatted status reports to communications satellites

The last two elements deal with the tasking of the satellites and analysis of their collected data plus information from other sources.

7. The ground infrastructure, (terminals), software tools and licenses to allow users of the system to determine which spacecraft to task when to obtain the desired results, and to so act.

8. The software tools to correlate, fuse and analyze the information generated by the space systems, including S-AIS track data, the basis for all analysis, along with all other pertinent data resident in all reachable data sources such as port, financial, shipping and broker records.
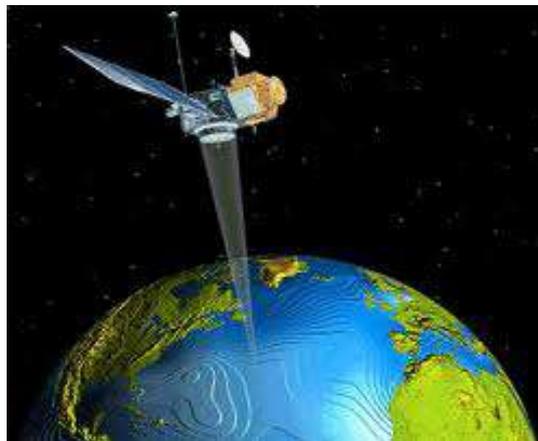

Figure 2: Many Satellites now Watch the Seas

4. OPERATIONAL CONCEPT

Tracking the reports from S-AIS is now revealing the "Pattern of Life" of the maritime world on a global scale. This pattern has been recorded for more than the past 10 years and offers great insights into normal and abnormal behavior in the maritime domain. Optical, video and especially SAR radar images of many sea areas have also greatly increased our understanding of what is abnormal. Unclassified signals intelligence exploitation of the emissions in the maritime domain are offering yet another source of highly useful data. Ships also use both text messages and machine-to-machine data communications, the technological heart of the "internet of things" (IoT), to report the state of many things onboard the ships and in the waiting ports. Those data items include the status of engines, the temperature and pressure inside containers, fuel levels, oil pressure in the engines and critical bearings, and a myriad of other things. Information technologists have also been gathering shipping and ship broker records as well and have been analyzing all of the above data using artificial intelligence and machine learning to the problem of dynamic data analysis for at least the last 15 years. (Yes, five years before the first S-AIS receivers reached space but the job became much easier when they did start collecting AIS from all over the globe.) The lessons learned have been very fruitful, and it will only get better. This is especially true as more and more people share more and more data.

The IT capacity now exits to analyze all of this myriad data. ICG Solutions alone reports it ingests and analysis several billion discreet data points per day. The collaboration of all of this data on a worldwide scale would allow watch centers in each region of the world to gain a much better understanding of what is happening in their area of responsibility and

where it fits in the global picture. To be sure, much of the information will be regional specific, but a great deal of it will be directly affected by what is happening elsewhere in the world.

Space systems will not replace terrestrial systems, but they will make the terrestrial systems much more effective. Instead of sending a maritime patrol aircraft or a law enforcement cutter on an area patrol, they will be sent to search where abnormal activity is known to be.

There are already as least two organizations employing unclassified space-based systems for enhanced maritime situational awareness. They are both based and focused on Europe. They are the European Maritime Safety Agency and FRONTEX, the Border Guard and Coast Guard of the European Union. France and India have recently announced they are going to jointly build a space-based maritime security system for the Indian Ocean. There are several other organizations which are now also using Space as means to increase their maritime security. The proposal here is not that we replace theses entities, but rather that we bring them all together to enhance the sharing of resources and allow for their use in a more effective and efficient manner. It would allow for the more well-developed nations of the world to assist those that are not so developed, for the betterment and benefit of all.

5. BONUS ROUND

While C-SIGMA/CSMSA would go a long way toward satisfying many of the world's varied needs for maritime situational awareness, it would also have the added benefit of providing a focal point for the creation of the global maritime security system envisioned in "A Cooperative Strategy for 21st Century Seapower," the current major policy statement of the US military maritime services. The coordination needed to implement C-SIGMA would provide a focus for the efforts to achieve a common goal of protecting the maritime environment as stated in that document and would go far in bringing the lawlessness of such places as the Gulf of Guinea under control. Space-based earth observation does not replace terrestrial systems but does make them substantially more effective and is a significant start to fulfilling the core need of knowing who is where on the world's waterways.

C-SIGMA/CSMSA also directly supports of the US Presidential Policy Directive Four (PPD-4), US National Space Policy (NSP), 28 June 2010, and could be the international mechanism to satisfy its Implementation Task #1, quoted below. PPD-4 emphasizes U.S. leadership in space and directs international collaboration on mutually beneficial space activities for the purpose of broadening and extending the benefits of space to all mankind. PPD-4, which is unclassified, has a classified Implementation Directive. However, Task Directive #1 is unclassified and states:

> "(U) Working through the National Maritime Domain Awareness Coordination structure, the Secretaries of Defense, Homeland Security, Transportation, State and Commerce, will develop an unclassified, international available program to foster international collaboration using civil and commercial space systems to enhance global maritime domain awareness to provide: enhanced safety of life at sea; increased mutual security of all users of the maritime domain, improved protection of the maritime environment and the resources of the sea; improved flow of commerce; and better monitoring of the condition and performance of the Marine Transportation System."[iii]

The implementation of that directive has been held in abeyance for some unknown reason but implementing C-SIGMA/CSMSA could well be the key to building the envisioned, truly global, maritime security system, and would substantially assist in many tasks including detection of illegal smuggling of all types, of goods, arms, drugs, and people; much improved maritime pollution control and resource protection, such as illegal fishing and stealing of oil, as well as dramatically assisting humanitarian assistance and disaster recovery operations. Remote ocean surveillance in such areas as the Arctic, the southern oceans, and the western Pacific would clearly immediately benefit many people both in and out of those regions.

6. WHERE NEXT?

Implementing C-SIGMA/CSMSA in the Arctic and a few other locales such as the Gulf of Guinea and the western Pacific could be the needed stepping-stone to the implementation of Global Maritime Awareness for the betterment of the entire world.

Since April 2005 C-SIGMA has held 11 highly successful international conferences at locals all over the world including two in Italy, one in Japan, Ireland, England, And Portugal plus Canada and the US. These conferences have been attended by all the major Earth observation and AIS satellite builders and operators, and most if not all, of the builders of dynamic data analysis software focused on Earth observation. Many users of this data from all corners of the globe have also participated. Many, if not all attendees, agree we must do something like what is described above, but the question remains "How?" Both NATO and the Multinational Maritime Security Center of Excellence have expressed an interest, not nothing specific has materialized yet. Money and manpower seem to be the stumbling blocks. Given the high potential payoff, it should not be.

REFERENCES

i Multinational Maritime Security Centre of Excellence Concept of Operations, pg 1
ii Ibid
iii Classified Annex, US Presidential Policy Directive 4, 28 June 2010
iv The first four conferences were called TEXAS (Technical EXchange on Ais via Satellite)

An Ethical Framework for Maritime Surveillance Technology Projects

Sari Sarlio-Siintola[a*], Tuomas Tammilehto[a], Saara Siintola[a,]

[a]Laurea University of Applied Sciences, Ratatie 22, 02200 Vantaa, Finland

ABSTRACT

The ethics of Maritime Surveillance is a topic of increasing importance in both academia and other forums. This development owes partially to new legal obligations, such as those set out in EUs new data protection legislation. Also the funders of innovation programs are increasingly expecting projects to pay attention to and address various ethical issues. The ethical challenges involved in the development and piloting of technology-based maritime surveillance solutions are multifaceted from both the research and development perspective, and from the viewpoint of the final solution to be created. The purpose of this paper is to present a framework for a) the identification of ethical, legal and societal aspects in technology innovation projects, and b) the operationalisation of these aspects as concrete requirements. Furthermore, in order to concretise the proposed framework, we discuss the outcomes of ethical analyses of two Horizon2020 maritime surveillance projects, MARISA and RANGER.

Keywords: Maritime Surveillance Technology, Ethical, Legal and Societal Aspects (ELSA), Responsible Research and Innovation (RRI)

1. INTRODUCTION

The ethics of Maritime Surveillance has been a topic for vivid discussions in both academia and various other forums, reports and statements. Especially concerns related to the tension between privacy and security on the other, has been a central focus in the debate.[1]Two centrepieces of EU law in the area of data protection, the General Data Protection Regulation ('GDPR'), and the Data Protection Law Enforcement Directive ('LED'), compel the carrying out of a specific Data Protection Impact assessment (PIA) prior to any processing of personal data that is likely to result in a high risk to the rights and freedoms of natural persons – including development work in maritime surveillance projects.

In addition to privacy-related concerns, the implications of new surveillance technologies for asylum seekers and refugees have been deliberated by several scholars.[2] [3] Due to the fact that both EU law and various international conventions regarding e.g. human rights, the rights of refugees, and Search and Rescue activities impose obligations on states to help and protect those in need, the increased situational awareness enabled by the new technologies will also lead to an increased responsibility to act.

The paper is organised as follows. In section two, we give a general introduction our approach to ethics work that covers both research and development processes, and the solution to be created during those processes. In section three, we present a methodology for identifying ethical, legal and societal aspects of technology projects aiming to produce innovations for the market. Finally, in section four, we discuss the operationalisation of the identified aspects as ethical requirements and guidelines for both technology and organisational arrangements.

---

[*]sari.sarlio@laurea.fi; phone +358 (0)40 513 9118; www.laurea.fi

2. ETHICAL DIMENSIONS OF TECHNOLOGY-BASED PROJECTS

Technology research and development projects are multidimensional from the point of view of ethics, legislation and societal impacts. In addition to traditional research integrity, it is essential to ensure also the comprehensive ethical and social sustainability of the solution being developed. Ultimately, sufficient ethical sustainability is a prerequisite for the social and political approval and market potential of any solution. Both research integrity, validation of the ethical features of the solution, and the use of beta versions in real time settings need to be addressed, also already during trials or other such tests (see Figure 1).

The ethical guidelines of Horizon2020 focus heavily on traditional research integrity-related issues and as such offer limited guidance in terms of the ethical and societal sustainability of real time setting piloting and trials, or the final product/solution itself. This can pose challenges for developers, for instance in projects that develop and pilot together with end users information systems (solutions) processing personal data; the implications of data protection related requirements can namely look very different for different aspects of the project.

From the research integrity perspective, the rights of end-users (and other natural persons) taking part in the projects must of course be secured. This may influence e.g. the collection of personal data or the dissemination of photos in which individuals can be identified. In addition to this, the solution itself must comply with numerous requirements set out in the GDPR and/or the LED, such as embodying the principles of Data Protection by Design and by Default. In other words, data protection must be integrated into the architecture by utilising privacy enhancing technologies or similar. Lastly, data protection compliance is required already during development and trials/pilots – the phase during which privacy and data protection features themselves are also being validated. In some cases the problem can be solved for instance by using fake data, but meaningful piloting often requires a real data. The take-home message is that ensuring data protection compliance is essential regardless of whether data protection issues are a central focus or purpose of the project.[1]
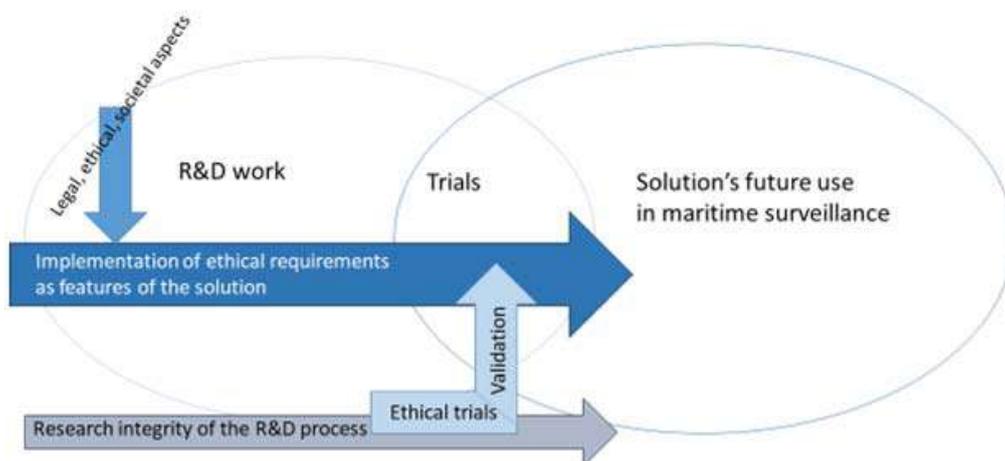

Figure 1 Ethical Dimensions in Technology-Based Projects

---

[1] An important thing to understand is that all data relating to an identified or identifiable person, directly or indirectly, constitutes personal data under the GDPR and the LED, and all processing of such data needs an explicit legal basis, such as a legal obligation or consent (Article 6). This means that for instance AIS data constitutes personal data, the processing of which falls under the scope of data protection legislation. As the installing and use of AIS could be interpreted as an expression of consent for the typical processing of such data this is normally not a problem.

In identifying the ethical challenges and opportunities that relate to the solution being created itself, a distinction should be made between the layers of technology, user processes, and business/governance models. This is essential, as the implications of ethical, legal and societal requirements often look different for different layers. Ethical requirements which can be implemented as technical features of the solution can be handled in the technical planning, implementation, and validation in a way that is analogous to end-user requirements. On the user process level, the implementation of ethical requirements concern for instance user manuals or administrative arrangements such as the training of users. On the business/governance models level, the relevant considerations could concern for example the division of responsibilities between different actors or various kinds of preparations and feasibility considerations to be done before implementing the solution into a specific environment.

Important is also to remember that the features of the developed solutions may vary according to the environment in which they will be implemented, which may have implications on ethical requirements on all layers of the solution. Both MARISA and RANGER, for instance, can be implemented as either stand-alone solutions or as part of the Common Information Sharing Environment (CISE). See Figure 2.
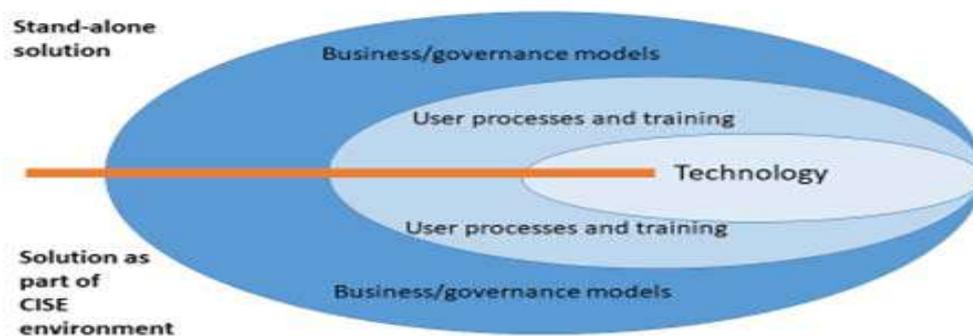

Figure 2: Ethical Layers of Technology-Based Solutions

3. THE METHOD OF ETHICAL ANALYSIS

Leeze et al. argue that ethics in security research must be seen as a way of putting critiques to work, not as a mere legitimising function of 'ethics approval'.[4] The same argument is widely discussed in the context of ELSA and RRI research. The approach we have developed during the MARISA and RANGER projects aims to provide a model/method/framework for doing just that. The purpose was to maximise the benefits of both projects while preventing or minimising any ethical risks.

In both MARISA and RANGER, we divided the analysis work into the following components: 1) a critical ethical analysis of the technology and its use in the relevant context (border control, customs, search and rescue, environment and general law enforcement), 2) a legal framework for the project (including development, the solution itself, and its future use), and 3) a Social Impact Assessment (SIA) and a Data Protection and Privacy Impact Assessment (PIA). The results of this work where then encapsulated into a set of concrete ethical requirements for the project, as well as a Code of Conduct containing ethical principles to be embedded both in training material and Business Model documentation.

We conducted the ethical analyses and the legal frameworks as desk-top studies where, among other things, we have analysed the content of various regulations, guidelines and policy papers. The SIA of each project was has carried out by having various stakeholders and experts do brainstorming work, the results of which were integrated in the project planning and risk management to mitigate potential problems and to promote positive impacts across the lifecycle of developments. The practice is participatory and it increases understanding of change and capacities to respond to change.[5] Central to the SIA approach is that ethical issues (concerning both positive and negative societal impacts) are taken into account already in the design-phase of innovation. Ethics are thus understood not only as legal and moral

constraints for innovation, but also as active catalyst of innovation from which value can be derived. The PIA work in MARISA and RANGER was organised in collaboration with project partners, utilising a PIA tool provided by CNIL (Commission Nationale de l'Informatique et des Libertés).

Key ethical challenges identified in the MARISA and RANGER projects are disclosed below, in Table 1.

Table 1: Main Ethical Challenges in the MARISA and RANGER Projects

Main Ethical Challenges in the MARISA and RANGER projects	
Challenges	Layers of the solution
Tensions between different rights and values, such as freedom and security, which are likely to become more pronounced as a result of the new security technologies.	- Business & governance models - User processes
Ethical and legal issues relating to privacy and data protection in both current and future configurations of RANGER and MARISA, including both technical and organizational arrangements	- Business & governance models - User processes - Technology
RANGER's impact on wildlife and humans in the region where the radars are installed. Regardless of whether the risks are real or only fears, it is ethically and societally important to address the issue.	- Business & governance models - Technology
Ethical and legal issues relating to OSINT, Big Data and AI in MARISA. These include the need for human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness (including awareness of and strategies to control subconscious biases), environmental and societal well-being, and accountability.	- Business & governance models - User processes - Technology

In the next step, the ethical, legal, and societal framework built as a result of the analysis work was refined into smaller, more detailed ethical requirements that were then categorised into the classes 'ethical awareness', 'ethical analysis' or '(any) activity'. At this stage the requirements should be specific and concrete enough to be associated with the relevant phase or layer of the project: pilots and trials, technology, user processes, business and governance model, or generally on the solution.

Finally, a specific Code of Conduct was formulated for each solution, based on the results of the analyses. The codes are designed for end-users, decision makers, and developers of the solution; the idea is that they shall be embedded both in training material and Business Model documentation..

4. THE RESULTS OF THE ANALYSIS AND THEIR IMPLMENTATION

The value bases of the ethical requirements and the Code of Conduct (for MARISA and RANGER) are derived from both fundamental human rights values and norms established in international and EU law, and various ethical issues raised by the end-users and other stakeholders.

The Code of Conduct establishes principles that should be taken into consideration when deploying, using and developing a solution, and concern the totality of ethical and societal considerations: the technology itself, how the technology will be used, as well as the whole business model/procurement as part of the European Maritime Surveillance ecosystem. A Code of Conduct should be subject to reviews and updates; when implementing a solution in a specific user community contexts, the principles are to be further specified and integrated into other existing codes of conduct. In the table 2, we list the main sections of the current versions of MARISA Code of Conduct and RANGER Code of Conduct.

Table 2: Contents of the RANGER and MARISA Code of Conducts

Code of Conduct (titles)
1 The Justification of MARISA/RANGER is Based on Ethical Grounds
2 The Humanitarian Imperative and the Rights of the People at Sea
3 Transparency, Liability and Human Decision Making
4 Privacy and Data Protection
5 Value for End-users Involvement
6 Moral Division of Labour in Maritime Surveillance and SAR
7 Robustness, Accountability and Learning

The ethical requirements defined must be taken into account in the technology development and organizational arrangements related to user process descriptions and training, as well as in governance- and business modelling. In the Table 3, here below, there are presented some ethical requirements of the MARISA project.

Table 3: Examples of MARISA's ethical requirements

(MARISA-G21) Recognize third countries in the sea as both end-users of MARISA, and as partners in solving shared problems with the help of new technology.	*Desirable/ Essential Activity* *MNGMT UC, AM,BM WP1, WP2*	*The MARISA Advisory Board include a representative from a third country. The point will be addressed during the MARISA workshops and the Advisory boards.* *This issue is relevant also in the various future User Communities and Business/Adoption Models of MARISA.*
(MARISA-T1) Provide transparency and proper functionalities to help estimate the quality, reliability and validity of various data to be used. Code this information for the end-user to help her in the decision making.	*Essential Activity* *TECH WP3-WP5*	*This requirement is translated into several requirements in the technical baseline. Specific KPIs have been defined to monitor the fulfilment of the functionalities during the validation. Rules can be configured by the users. Refers to technical documentation (D3.x, D4.x, D5.x)* *MARISA_UR_GEN_55, MARISA_UR_GEN_60 various MARISA_UR_DF1 requirements* *The AI-checklist will also be used in order to investigate the transparency issues in MARISA.*
(MARISA-U2) Operational decisions shall never be made by a computer, not even the most efficient one: it must always be a human who makes the final decisions. MARISA can only assist in operational decision making, by providing information to the end-user/decision makers. The end-users must be informed regarding these liability issues in the training material.	*Essential Awareness* *TRAIN WP8*	*The users will be always in the loop, the toolkit will support decision making and planning being the final decision lies on the end-users. This is clearly explained in the training and user manuals*

(MARISA-B5) Organizational activities concerning Data Protection must be applied as part of the governance model for each new implementation of MARISA. Conducting a light PIA before the implementation is essential.	*Essential Activity* *BM/GM WP8*	*The final ethics deliverable D2.13 provides basic guidelines the organizational activities. These are to be embedded in MARISA exploitation/business modelling and in training material. See also MARISA code of conduct in D2.13.*

CONCLUSIONS

It should be evident that ensuring the proper implementation of ethical requirements is essential for any project. In spite of this, ethical compliance has long been near synonymous with proper research ethics, other important dimensions having been left with more or less an anecdotal status. The problematic nature of such a narrow perspective is often particularly accentuated in cases where a project's subject matter falls under a security topic. The RANGER project provides an illustrative example of this: when technological advancements lead to an increased surveillance capacity for authorities (in this case in the form of novel over-the-horizon radars), so do the moral and legal duties to act against ill will and to help those in distress; with great power comes great responsibility. Furthermore, the developed technology can fundamentally change practice and customs: the moral division of labour can be altered, a change that calls for holistic ethical considerations.

To answer to these challenges, we have attempted to develop a systematic framework for identifying ethical aspects from a more comprehensive viewpoint than that provided by the traditional science and research integrity perspective. One goal is to help developers and practitioners of technological innovations to turn these aspects into tangible sets of ethical requirements to be addressed during all phases of the project and on all layers of the solution being created.

The critical thing to realise is that ethics is not about declaring principles. Rather, it is intertwined in every aspect of a project and beyond, from the proper development of products and services, to their use, and all the way up to business and governance processes.

ACKNOWLEDGMENTS

REFERENCES

[1] Wright, D. and Raab, C. D., "Constructing a surveillance impact assessment," Computer Law & Security Review 28(6), 613- 626 (2012).
[2] Jeandesboz, J., "Beyond the Tartar Steppe: EUROSUR and the Ethics of European Border Control Practices," In Burgess, J. P. and Gutwirth, S. A. (eds) Threat Against Europe? Security, Migration and Integration, VUBPress, Bruxelles, 111-132 (2012).
[3] Crépeau X., Report of the Special Rapporteur on the human rights of migrants, François Crépeau - Regional study: management of the external borders of the European Union and its impact on the human rights of migrants, UN Human Rights Council 24 April 2013, A/HRC/23/46.
[4] Leese, M., Lidén, K., & Nikolova, B., "Putting critique to work: Ethics in EU security research," Security Dialogue, 50(1), 59–76 (2019).
[5] Zwart, H., Landeweerd, L. & van Rooij, A., "Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'," Life Sci Soc Policy (2014) 10: 11.
[6] Esteves, A. M., Franks, D. and Vanclay, F., "Social impact assessment: the state of the art," Impact Assessment and Project Appraisal, 30(1), 34-42 (2012).

9
Maritime Unmanned Systems

Contents

Maritime Cyber Situational Awareness Elaboration for Unmanned Vehicles

Olivier Jacq[a], Pedro Merino Laso[b], David Brosset[a,c], Jacques Simonin[a,d], Yvon Kermarrec[a,d]
and Marie-Annick Giraud[e]

[a]Chair of Naval Cyber Defense
Funded and supported by École navale, ENSTA Bretagne, IMT Atlantique, Thales and Naval
Group, CC 600 F-29240 Brest Cedex 9, France
[b]French Maritime Academy (ENSM) - 38 Rue Gabriel Péri BP 90303 F-44103 Nantes Cedex
04, France
[c]Naval Academy Research Institute, Ecole navale - CC 600 F29240 Brest Cedex 9, France
École navale - CC 600 F-29240 Brest Cedex 9, France
[d]IMT Atlantique - Lab-STICC, F-29238 Brest Cedex 3;
[e]SOFRESUD, 777 Avenue de Bruxelles F-83500 La Seyne sur Mer, France

ABSTRACT

In history, concurrence for dominance at sea has always been a major concern for most sea-facing countries. As the vast majority of worldwide goods exchanges are made by sea, an answer to the growth of the worldwide commerce over the last twenty years has been to increase the size and on board value of ships sailing between eastern and western countries. Digitalization has also been an answer to ease the routing, management, conning and overall return on investment of such ships. But as the cyber surface attacks consequently widened, the maritime sector became a real interest for state-sponsored and criminal hackers. The concern of a major cyber-attack targeting ships at sea or naval shore infrastructures is no more science fiction, and late incidents do confirm the urgency to act. "Safety is paramount" is a known motto on board ships, the crew being drilled for fire and water floods. However, the awareness of the sector on cyber remains low and if modern ships have a reduced crew, autonomous ships are arriving with just nobody on board. This is the reason why, on top of the usual cyber-security measures, the implementation of intrusion detection systems on unmanned vehicles is essential to detect, react and stop any incoming threat. In this paper, we outline the main challenges for the years to come on Maritime Cyber Situational Awareness (MCSA) for autonomous and remotely controlled vehicles.

Keywords: Cybersecurity, maritime, situational awareness, autonomous, unmanned vehicles

1. INTRODUCTION

The use of digital systems is now essential for any civilian or military maritime activity. As a consequence, the availability, integrity and confidentiality of those systems are of the uttermost importance. Newest generations of ships as well as Unmanned Surface Vehicles (USVs), Unmanned Air Vehicles (UAVs) and Unmanned Underwater Vehicles (UUVs) are built around computers and digital sensors to ensure all their missions, from communications, positioning and conning to situation awareness, combat systems, intelligence and, of course, engine and platform management. Another characteristic of the digitalization of ships is a high use of network interconnection to help automation and ease the overall situation awareness, giving the commanding officer and crew a better understanding of what's going on in their own ship and in surrounding maritime activities. The number of sensors available over the network is increasing and heterogeneous systems historically isolated are now interconnected, widening the surface attack of a ship and underlining the urgency to protect them properly.

Further author information: (Send correspondence to O.J.)
O.J.: E-mail: olivier.jacq@ecole-navale.fr
P.M.L.: E-mail: pedro.merino-laso@supmaritime.fr
D.B.: E-mail: david.brosset@ecole-navale.fr

Advanced research, testing and operational deployments are under way on autonomous and unmanned vehicles for surface, aerial and underwater use. As embarked crew is replaced by automation and remote control operations, cyber-security is no option but an important challenge. Research works address this problem from multiple axes such as risk assessment,[1] secured architectures[2] and novel role-based access definition.[3] Adding autonomous features on already vulnerable protocols and systems like Industrial Control Technologies (ICT) draws new attack vectors and an underlined incapacity of facing cyber-attack from within an unmanned vehicle. Detecting and hunting cyber-threats will become a more complex task of that equipment. Specifically, their reaction in case of an emergency such as a cyber-attack has to be anticipated and formalised, down to the software and the hardware, to avoid any physical damage to self, to peers, or to the environment. But, before being able to properly react to a cyber-attack, those vehicles first have to be designed with a very high level of security and safety to be protected from most advanced threats all over their life cycle. Proper detection and reaction schemes must then be designed to detect, alert, and safely react, as one but also as a whole when a coordinated attack against an autonomous vehicles fleet is concerned. As an answer, the Sea4M project[2] aims to define a dedicated security architecture for managing and controlling USV fleets. This project has proposed a novel Human-Computer Interaction (HCI) development that adapts itself to particular roles to address Maritime Situational Awareness (MSA) for specific user's needs. This work is a first step to include MCSA in Sea4M's architecture to provide a full overview of the system state for each particular user.

2. RELATED WORK

In this section, we will underline that Situational Awareness (SA) is a challenging goal for tele-operated systems and that MCSA should be seen a subset of it. The elaboration of MCSA depends on three prerequisites: an efficient knowledge of digital assets, a dedicated maritime CSA and performing visualisation schemes.

2.1 Maritime digital assets cartography

The "knowledge of us" is an essential knowledge both to SA and cyber-security. To correctly assess risk and protect systems, one needs to understand the way systems are designed, set up and run. However, applied to maritime, the commanding officer (CO) of a ship has not designed the boat he's commanding. If he has an excellent overview of her reaction at sea, facing winds and waves, the usual knowledge he has of his onboard Information Technology (IT) and Operational Technology (OT) is usually the multiple function displays set at the bridge. This fact can be explained because the design of a ship is based on the integration of multiple Commercial Off-The-Shelf (COTS) systems, for instance Electronic Chart Display Information System (ECDIS), Automatic Identification System (AIS), Global Navigation Satellite System (GNSS) as well as Programmable Logical Controller (PLC) and Industrial Control Systems (ICS) for platform management. Each of those systems is designed by subcontractors. The shipbuilder usually only has an integration role to make the systems communicate together. If alarms are, however, frequently generated in a logging window per system, there is usually no central or global awareness on how the ship systems are performing.

2.2 Cyber Situational Awareness

Situational Awareness has been described by Mica Endlsey as a "state of knowledge".[4] Her work, based on situation perception, comprehension and projection, has been applied to meet the Air Force needs. One can think of the wide use of Head Up Displays (HUD) in aviation as a very good example of help for the pilot to achieve SA. Applied to maritime, MSA can be seen as a fusion of the "knowledge of them" (the maritime environment such as weather conditions, position of other ships, situation intelligence) and the "knowledge of us" (own situation concerning security, safety, conning, weapons and ship's status) in perspective with mission assignment[5] to help decision makers. Therefore, MSA depends on internal and external-facing sensors, intelligence, and mission planning and knowledge.[6] MSA can have different levels of abstraction, at a ship's, fleet and HQ or shipowners' levels. However, sensors and computers are, most of the case, domain-focused (engine situation, platform management, ECDIS, etc.). The existence of interconnection protocols such as National Marine Electronics Association's NMEA0183 and following greatly helps data sharing, mostly within the bridge. However, other domains from engine and platform management still use dedicated visualisation, platforms and protocols. As a consequence, most commanding officers still have to switch from a display to another to achieve situational awareness.

Franke and Brynielsson[7] remind us that Cyber Situational Awareness (CSA) should not be seen as a goal by itself, but as a subset of global SA, underlining an important effort still has to be drawn for the fusion of CSA to SA, and therefore for the fusion of Maritime CSA (MCSA) to MSA. Current efforts on the MCSA research cover the intrusion detection and architecture[8] as well as an overview of MCSA to help decision makers for the maritime sector.[9] Challenges for elaboration of the MCSA remain multiple. Maritime systems encounter a number of physical constraints which complexify MCSA elaboration, this complexity getting just higher when it comes to unmanned and autonomous vehicles. A first difficulty is isolation at sea, which sets a high level of dependency on satellite links. IT and OT support at sea also remains complicated due to the usual lack of IT/OT experts on board. Satellite coverage and environmental conditions may also lead to temporary link failures and limited bandwidth. The very long life-cycle and design of the ship imply the onboard integration of heterogeneous COTS systems coming from various builders. Their onboard integration is often minimal, mostly just set to enable communication, without much - if any - security features such as content control and filtering or network isolation. Subcontractors, as well as ship integrators, are reluctant to authorize patch management on board to mitigate vulnerabilities. Most software, such as ECDIS, have to undergo new tests to get fully requalified and avoid regression. The same also applies to drones and autonomous vehicles: for instance an UAV would have to be re-qualified after patching and follow intense sea trials to give a high assurance on the absence of any regression due to the patch. When intrusion detection systems are deployed on board, their ability to detect attacks remains low. This is due to their limited possibility to process maritime and industrial specific protocols: even if correctly set up and maintained with up-to-date signatures, such sensors also have to be monitored on a permanent basis by experienced shore-based cyber-security engineers to detect threats. Finally, while MCSA should just be a subset of MSA, cyber-security experts usually lack maritime experience and therefore often fail to measure the impact on the ship's mission. On the contrary, commanding and watch officers' awareness on cyber remains low and lack knowledge, time and procedures to search for causality assessment and precisely measure the possible impact and spread of an attack on the ships' IT and OT systems.

2.3 Visualisation

The fusion and visualisation of all available information for the crew are important challenges. While the amount of data coming from sensors is increasing, the risk that crucial information is not taken into account or that only partial information is given is real. Several studies are working on augmented reality and HUDs for the maritime sector[10] and draw the design for the bridges of the future. But, due to the lack of connexion and common protocols between cyber Intrusion Detection Systems (IDS) and common bridge tools, MCSA and cyber alarms are not part of it. Finally, when talking of a task force commander or headquarters staff, MCSA, when achieved, has to be displayed using a proper abstraction level. For instance, the usual high level of false positives encountered on traditional IDS sensors underlines that only qualified incidents have to be shared with mission commanders. Causality and impact assessments are also key values commanders are looking for, to confirm or modify their action course.

3. CYBER SITUATION AWARENESS FOR SURFACE UNMANNED VEHICLES

The progress in digitalization and autonomy will cross another step forward with unmanned vehicles. In a few years (if not now), a modern fleet will combine such vehicles, either as a whole or as a subset of traditional maritime assets. Remote surveillance of a vehicle's fleet will be paramount to accomplish the mission safely. SA for those assets has to take into parameter the cyber state of the vehicles. However, their small form factor, limited capacities in energy and communication as well as environmental constraints will challenge intrusion detection. Unmanned vehicles add complexity, heterogeneity and widen the attack surface of traditional maritime assets.[1] Their "autonomous" way of acting should not lessen the attention on the fact that they add new and strong dependencies on satellite communications, position systems and external facing sensors (radars, cameras). Whether acting alone or as a coordinated fleet, those unmanned digital systems of systems could be lured, locally or remotely, just by flawing the sensors or by reverse engineering their software to exploit design flaws.

3.1 Situation Perception

On an intrusion detection point of view, unmanned vehicles will embark a high number of new sensors, flowing real time or differed-time data to shore or ashore back-end platforms. This volume of data also means a possibility

to corrupt data being sensed and transmitted to change the way back-end platforms behave. Detecting such intrusions on a high volume of data is a hard task. While the programmed reaction to an unknown behavioral or value on a sensor is highly critical and has to be formally approved and tested, the real behaviour, which can be hard-coded in the software, is usually only known by the vehicle designer.

On a whole fleet of unmanned vehicles, the multiplicity of intrusion detection system agents (whether network or host oriented) on multiple sources adds complexity: detecting a coordinated cyber-attack gets even more complicated. Finally, perception relies on detection schemes: anomaly-based detection for those vehicles is a solution, using techniques such as machine learning to learn of the normal behaviour in various conditions. However, this approach needs training datasets, implying a very high level of training of the model, meaning a test in every possible condition is necessary to later flag any deviation. If this condition is not met, the false positive rate can be very high. On the other hand, a behaviour-based detection also seems complex: looking for Indicators of Compromise (IoC) on such systems requires having a very high level of knowledge on its way of working (operating systems, network communications, processes). On specific chipsets, real-time operating systems and proprietary networking protocols, using such sensors might reveal not to be pertinent.

Finally, this multiple-agent and multiple-source detection scheme on unmanned vehicles will need data to be sent to a dedicated Security Operations Center (SOC) for analysis. Choosing between real-time and past-mission analysis is a crucial choice. Real-time means a permanent, dedicated and secure link will have to be set between vehicles and the SOC, which is nearly impossible to achieve - or at a very high price on discretion and bandwidth for UUVs, and which would have an important impact on available bandwidth for maritime UAVs too. Long haul USVs and autonomous commercial ships are probably best targets to meet this requirement. The quality features for situation perception (freshness, truthfulness, completeness) and hence the overall quality of MCSA will highly depend on the choices made at this step.

3.2 Situation Comprehension

Situation comprehension is met when causality assessment, impact assessment and attacker perception are met. Assessing causality means the data gathered during the recognition phase is complete enough to determine where the attack came from. However, forensic experts need a lot of time and data to qualify a complex attack. For instance, snapshots of memory, full low-level copies of hard drives and live captures of network traffic and processes usually needed. This amount of data is impossible to send over a network on a live unmanned system and has to be done afterwards (if the vehicle is not lost). This causality assessment also means all passwords, architecture information and software are accessible, understandable and known by forensic experts. The lower price of storage can give hope to the fact that a lot of logging and debugging data will be held on board such vehicles.

Understanding the attacker's perception is based on causality assessment but has to be cross-checked with generic (such as MITRE's) or dedicated[11] attack frameworks, but also with impact assessment. However, assessing the impact once again implies the overall architecture and working of the vehicle are known, which is rarely the case. This is due to the fact that such vehicles are often COTS by themselves, and that manufacturers set a high level of intellectual property on both software and hardware, making it a "black box" to the final client. Analyzing data to separate the good from the evil will need the manufacturer to work with forensics experts.

A dedicated maritime SOC including autonomous vehicles is not always available and the operators of the system will always be the first line of defence. Unmanned fleet remote operation is realised by a particular role hierarchy that defines, for each user, specific duties and responsibilities. MCSA would gain in being abstracted for these roles with adapted pertinent information and take into account crisis management times.

3.3 Situation Projection

In such conditions, predicting the next step of the attacker is also a real challenge. With the low real-time data available to SOC analysts, setting a priority level on an event may lead to advanced attacks not being detected due to the lack of data and knowledge. On the contrary, a basic detection or a saturation of low-level detections may trigger unwanted reaction from the vehicle or from the SOC and decision makers, which could mainly be caused by stress or fear of a major impact on the fleet and act in prevention to avoid any further impact.

4. CONCLUSION

Digitalisation of ships and shore infrastructures has widened the attack surface of the maritime domain. While the current awareness of the maritime domain on cyber remains low, and as important incidents are unveiled, the fore-coming and increasing development and use of unmanned surface, aerial and underwater vehicles have to trigger a new way of thinking, designing and using digital assets in the maritime world. The evolution and correction of long-term processes for regulation authorities, subcontractors and ship builders such as *secure by design* and dynamic patching to cope with new threats will take a very long time. In this article, we have shown that intrusion detection for maritime assets, and specifically the integration of MCSA as a part of SA was an efficient way to go, to detect and react accordingly in case of a cyber-attack. Now is the time to embark highly efficient prevention, intrusion detection and resiliency features in unmanned vehicles. If this challenge is not properly taken into account, in case of targeted cyber-attack, it could lead to major cyber-physical issues once the vehicles are at sea.

ACKNOWLEDGMENTS

REFERENCES

[1] Tam, K. and Jones, K., "Cyber-risk assessment for autonomous ships," in [*2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*], 1–8, IEEE (2018).

[2] Merino Laso, P., Brosset, D., and Giraud, M.-A., "Secured architecture for unmanned surface vehicle fleets management and control," in [*2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*], 373–375, IEEE (2018).

[3] Merino Laso, P., Brosset, D., and Giraud, M.-A., "Defining role-based access control for a secure platform of unmanned surface vehicle fleets," in [*IEEE 2019 Oceans Conference*], IEEE (2019).

[4] Endsley, M. R., "Toward a theory of situation awareness in dynamic systems," *Human factors* **37**(1), 32–64 (1995).

[5] Riveiro, M., Falkman, G., and Ziemke, T., "Improving maritime anomaly detection and situation awareness through interactive visualization," in [*2008 11th International Conference on Information Fusion*], 1–8, IEEE (2008).

[6] Gad, A. and Farooq, M., "Data fusion architecture for maritime surveillance," in [*Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002.(IEEE Cat. No. 02EX5997)*], **1**, 448–455, IEEE (2002).

[7] Franke, U. and Brynielsson, J., "Cyber situational awareness–a systematic review of the literature," *Computers & Security* **46**, 18–31 (2014).

[8] Jacq, O., Boudvin, X., Brosset, D., Kermarrec, Y., and Simonin, J., "Detecting and hunting cyberthreats in a maritime environment: Specification and experimentation of a maritime cybersecurity operations centre," in [*2018 2nd Cyber Security in Networking Conference (CSNet)*], 1–8, IEEE (2018).

[9] Jacq, O., Brosset, D., Kermarrec, Y., and Simonin, J., "Cyber attacks real time detection: towards a cyber situational awareness for naval systems," in [*2019 CyberScience*], 1–2, IEEE (2019).

[10] Wahlström, M., Karvonen, H., Kaasinen, E., and Mannonen, P., "Designing user-oriented future ship bridges–an approach for radical concept design," *Ergonomics in design: Methods and techniques* **1**, 217–231 (2016).

[11] Jones, K. D., Tam, K., and Papadaki, M., "Threats and impacts in maritime cyber security," *Engineering and Technology Reference* **1**, 1–12 (2016).

Smoothing optimal RRT* trajectories for recovery of an AUV by a moving surface vessel.

Simon Williams[a], Xuezhi Wang[b], Daniel Angeley[a], Christopher Gilliam[b],
Bill Moran[a], Trevor Jackson[c], Richard Ellem[c], Amanda Bessell[c]
[a] University of Melbourne, [b] RMIT University, [c] DST Group

ABSTRACT

In this paper we will show how the use of a kinematic constraint (a Dubins Path) can smooth optimal paths generated by a probabilistic algorithm (RRT*) for generating optimal paths. We have extended the original algorithm to take account of moving objectives, extended the cost functions to plan trajectories, improve the accuracy of sensor measurements, and incorporate position estimates and covariances into the sampling mechanism to focus the trajectories.

The main drawback of these methods is that the optimality is only probabilistic. Consequently, any finite implementation produces variable and disjointed trajectories that are not suitable in reality. Several smoothing techniques were tried and the Dubin's path produced the best results.

Keywords: sonar, trajectory optimsation, AUV

1. INTRODUCTION

Autonomous Underwater Vehicles (AUV) are expected to play an increasingly important role in a wide range of undersea applications such as conducting surveys to map the seabed and locate bottomed objects, sensing and characterising the undersea environment, and monitoring the movements and behaviours of surface vessels, underwater vehicles, other sub-surface objects and/or marine life. Vehicles operating in the underwater environment typically rely on sonar (i.e. underwater acoustics) as the primary method for sensing and communication over extended ranges due to the significant attenuation of electromagnetic signals. Here we consider using a passive sensor on an AUV operating in cooperation with a surface vessel that supports deployment and recovery of the AUV.

The problem of path planning for a mobile surface vessel that maintains supervisory contact (using an acoustic communications link) with an AUV executing a pre-planned mission was considered by Best.[1] Here we consider an AUV that conducts survey missions in a more autonomous manner, and on completion of the mission or when signalled by the support vessel via the communication link, uses its on-board sensors and optimal trajectory path planning to navigate safely back towards the support vessel to rendezvous and surface within a safe standoff distance for subsequent recovery. The support vessel is assumed to be mobile and is expected to transit along a straight line trajectory at a constant relatively slow speed towards a safe region (away from other vessels) to prepare for recovery of the AUV. Signals transmitted via the acoustic communications link can be used to assist the AUV with the detection and identification of the support vessel but are not used to control or program the trajectory or destination of the AUV. The final recovery location will not be known to the AUV or the support vessel in advance and will depend on how long it takes for the AUV to navigate to a safe recovery location.

The AUV is equipped with a forward facing high-frequency sonar sensor array that is used for sensing and navigation, which operates primarily as a passive acoustic sensor. The sensor array receives the acoustic energy emitted from all nearby vessels or objects located within the sensor's field of view, defined by the angular sector containing bearing and elevation angles within $\pm60°$, with a finite angular beam resolution of $\pm3.5°$. An on-board sonar processing system determines the direction of arrival of the received acoustic signals and uses this to estimate and track the positions and velocities of all observable signal sources over time. The estimated state information associated with each track is then used for optimal trajectory path planning to enable the AUV to

safely navigate towards the recovery vessel in the shortest amount of time while avoiding any other vessels or objects in the area.

The presence of other vessels or objects will impose sensor measurement and physical space constraints within the optimal trajectory path planning algorithms. Sensor measurement constraints result from reduced detectability of the support vessel caused by interference from other noise sources within the same angular sector or increases in overall received noise levels from noise sources in close proximity to the sensor. Physical space constraints result from maintaining minimum safe separation from other vessels or objects, and, from specified non-navigable areas such as shallow water areas, shorelines, islands, physical structures in the water, shipping channels or other restricted areas. These latter 'non-navigable' physical constraints are implemented as pre-programmed geographic regions within the on-board processing system of the AUV.
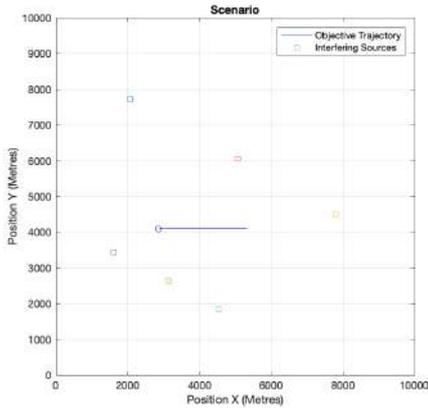

Figure 1: Example scenario for an AUV to navigate. The path of the moving objective is indicated by the blue line and the squares indicate the position of the interfering sources. Note that the AUV is initialised at the origin.

```
function RRT*(o, g, η, N)
    t ← {o}
    n ← 0
    while n < N do
        x ←SAMPLE
        x' ←STEER(t, x)
        t ← t + {x'}
        t ←REWIRE(t)
        n ← n + 1
    end while
    return t
end function
```

Figure 2: Original RRT* algorithm

1.1 Randomly Exploring Random Trees

Figure 1 shows a scenario where an AUV needs to navigate a fleet of interfering sources using its passive sonar to rendezvous with a moving recovery vessel.

Previous work[2,3] was based on $n-$step lookahead on the tracker processing. The new algorithm is based on the Randomly Exploring Random Trees (RRT*) algorithm developed in the robotics literature[4,5] for motion planning. The probabilistic nature allows soft constraints to be implemented as required.

The RRT* algorithm is provably optimal for a fixed objective and arbitrary allowable manoeuvres at each time step, which does not hold in our case. Figure 2 shows the original RRT* algorithm.

It begins with picking a point at random in the region of interest (SAMPLE). Then you move from the nearest point in the tree towards that point as far as is feasible (STEER) usually by giving a maximum allowable distance travelled η. Limits on feasibility are placed by speed restrictions and obstructions in the region. The cost of this step is then added to the cost recorded at the originating node. But, before this cost is encoded in the tree, a search is conducted locally to see if there is an alternative route to the new point that is lower cost (REWIRE). One of the keys to the RRT* algorithm is that this search need only proceed *locally*, within a distance r given by

$$r = \min \left\{ \left(\frac{\gamma \log n}{\xi_d n} \right)^{1/d}, \eta \right\} \tag{1}$$

where d is the dimension of the space, n is the number of sample points, γ is a constant and ξ_d is the volume of the unit ball in $\mathbb{R}^d$ Figure 4 shows how, if a new lower cost link is possible then, the original link is deleted and
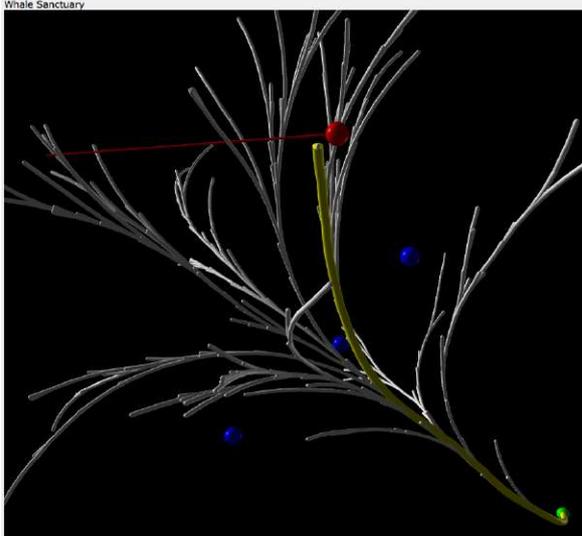
Figure 3: An example RRT tree with the ship and its motion marked in red, interfering sources in blue and the AUV's initial position in green and optimal path in yellow.
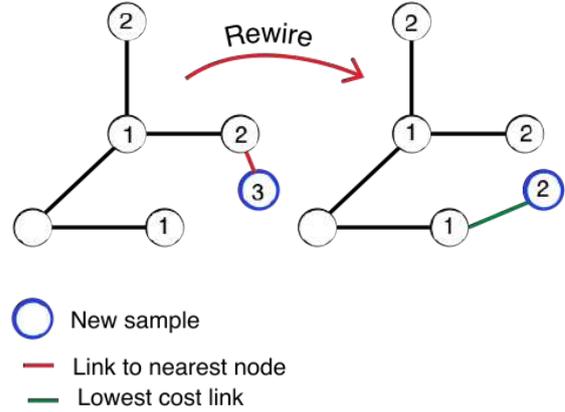

Figure 4: Diagram of rewire step in RRT* algorithm. The new sample (blue) is linked in to the tree at its nearest neigbour (red). A search, that only extends a maximum of one link length around the sample, reveals a lower cost link. The old edge is deleted and the new edge created (green).

the alternative instantiated and the new (lower) cost encoded at the new node.

This algorithm is probabilistically complete and asymptotically optimal for finding a path to a goal through an unobstructed region. Specifically the probability of the algorithm converging to the optimal path increases to 1 as the number of samples N increases.

2. MOTION PLANNING FOR TRAJECTORY OPTIMISATION

In this section we shall describe our extensions to the RRT* algorithm[6,7] to plan optimal trajectories that address the use of mobile objectives, dynamic cost functions, kinematic constarints, focussed sampling, and goal-oriented stopping criteria

2.1 Implementing RRT*-based Trajectory Optimisation

The tree shown in Figure 7 is one step in our trajectory optimisation algorithm. Ultimately only the first step of the yellow path will be used to initialise another run of the RRT* motion planner. The full trajectory optimisation algorithm is described in Figure 6. This involves generating an RRT* tree at each step and using the first edge of the tree that ends closest to the target, where the FIRSTSTEP(t, i) returns the first step in the branch of tree t at depth i which is closest to the goal.

2.2 Mobile Objective and Dynamic Costs

In order to incorporate the motion of the objective into the algorithm we take advantage of the fact that the AUV uses a tracker to keep an estimate of the position and velocities of its recovery vessel and the other vessels in its field of view. Using this information and the fact that each step in the tree corresponds to an elapsed time, we can use an updated position of the objective when calculating in the tree.

For example, if our new point x is attached to the tree at a depth k from the origin, then the cost function can be evaluated using the new objective location $l_i = l_0 + kv$ where l_0 is the original location and v is the velocity estimate in units of tree steps η.

This enables the implementation of dynamic cost functions where we penalise the alignment of an interfering source and the objective within the same received beamwidth ($\pm 3.5°$), and maintain the objective within the sensor's field of view ($\pm 60°$). This avoids the loss of detections that would result from the recovery vessel being

```
function Cost(x_i, l_i, s)
    c ← η
    if Angle(x_i, s) - Angle(x_i, l_i) ≤ 3.5° then
        c ← c + 50
    else if Angle(x_i − x_{i−1}, l_i) ≥ 60° then
        c ← c + 25
    end if
    return c
end function
```

Figure 5: Dynamic Cost function

```
function OptimiseTrajectory(l, s)
    path ← [InitialNode]
    for i in 1..l do
        tree ←RRT*(s)
        path ← path + FirstStep(tree, l − i)
    end for
    return path
end function
```

Figure 6: RRT*-based Trajectory Optimisation

obscured by interfering noise sources associated with other nearby vessels, or located outside the sensor's field of view, which would impact the ability of the tracker to provide accurate target state estimates. The form of these dynamic cost functions is shown in Figure 5, where x_i is the location of the new node which is i steps from the root, l_i is the expected location of the objective after i steps and s are the locations of the interfering sources.

2.3 Kinematic Constraints

The RRT* algorithm[4] has been modified to incorporate motion constraints for the AUV by limiting the maximum change in direction at each time step to 10°. This alters the STEER procedure so that the vector joining the new point x' to the tree always lies within 10°of the previous leg of the tree.

Figure 7 is the result of our modified RRT* algorithm. 1000 samples were taken from a uniform distribution across the region which is a 6x6x6 cube centred at (0,0,0). The AUV starts at (-3,-3,-3) pointing directly at the recovery vessel and has a maximum speed of $\eta = 0.2$ per unit time. The recovery vessel starts at (2,2,-2) moving in direction (0,0,1) at a constant speed of 0.09 per unit time.
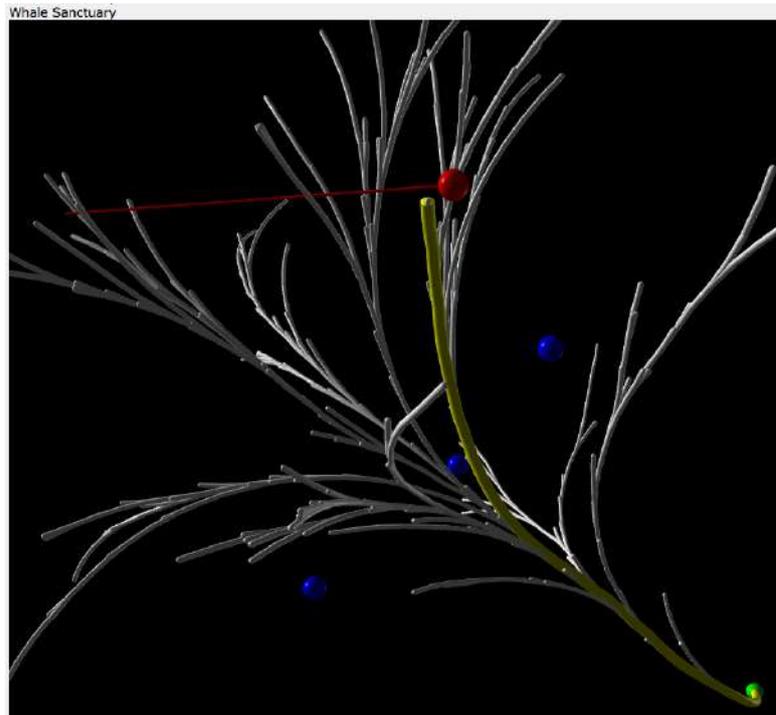

Figure 7: An example RRT tree with the ship and its motion marked in red, interfering sources in blue and the AUV's initial position in green and optimal path in yellow.

2.4 Trackers and Focussed Sampling

We model the measurements made by the AUV's sensor using a simplified sonar equation with a simplified uniform propagation model and a discrete beam pattern for each of the sensor received beam components.

The AUV's tracker is a Sequential Monte-Carlo Multi-Hypothesis Tracker (SMC-MHT) described in.[2,8] It is a multiple hypothesis tracker which detects and estimates the underlying target state with unknown measurement origin by enumerating and evaluating all possible measurement origin hypotheses. The sequential Monte Carlo sampling technique is used to approximate the posterior probabilities of the association hypotheses and make the approach computationally tractable.

It is important to be able to separate out the effect of the tracker on the trajectory optimisation algorithm. We shall see one approach by processing the data in the next section. Our approach is using what we have termed a 'Ground Truth' tracker, which outputs the actual position of the target with a simulated diagonal covariance an order of magnitude smaller than those predicted by the actual trackers. This enables repeatable runs with out interference from the random nature of the tracking algorithms.

Usually the SAMPLE procedure uses a uniform distribution over the area to be covered. Here we use the position and covariance estimates from the tracker and the parameters of a multivariate normal distribution and sample the space focussed around the objective location.

2.5 Goal oriented stopping criteria

In the original RRT* implementation the cost minimised the distance to the objective. However, a more realistic criteria for a moving observer is to use its passive sensor to approach the target, and then to terminate the approach or switch to other close-range sensing methods to avoid complex nonlinear near-field estimation effects associated with angles-only tracking problems. Accordingly the stopping criteria used here was changed to closing within 300m of the target and the cost function adjusted to minimise distance from the target rather than distance travelled.

This is a fundamental distinction between our approach and the other greedy approaches, which are limited in the number of steps ahead they look by the exponential growth of the number of possiblities to consider. Instead, we can set goal-oriented stopping criteria for the trajectory optimisation, like the one detailed above: Stop when you are within 300m of the objective.

3. OPTIMAL PATHS FOR THE RELAXED DUBINS SHORTEST-PATH PROBLEM

A Dubins path is the shortest curve that connects two points in the two-dimensional Euclidean plane (i.e. x-y plane) with a constraint on the turn and with prescribed initial and terminal tangents to the path, and an assumption that the vehicle traveling the path can only travel forward.

In our case we have no constraints on the direction of the AUV at the end of each step. This generalised version of a Dubins path is called the relaxed Dubins shortest-path problem. The standard Dubins shortest path problem specifies an initial and final configuration (x, y, θ), and results in six possible path types: RSR, LSL, RSL, LSR, RLR and LRL, where R is a right turn at minimum turning radius, L is a left turn at minimum turning radius and S is straight ahead. In contrast, the relaxed Dubins problem only has four possible path types: RS, LS, RL and LR. Intuitively, this is because it doesn't need the final turn to fix up the orientation at the final point.

Bui and Boissonnat[9] show the regions where each path type is optimal. If you canonicalise the problem by translating and rotating until the initial configuration is (0, 0, 0), the partition of the plane shown in Figure 8 arises i.e. LS and RS are the optimal paths for all points except those within the minimum turning radius circles to the left and right of the initial configuration.
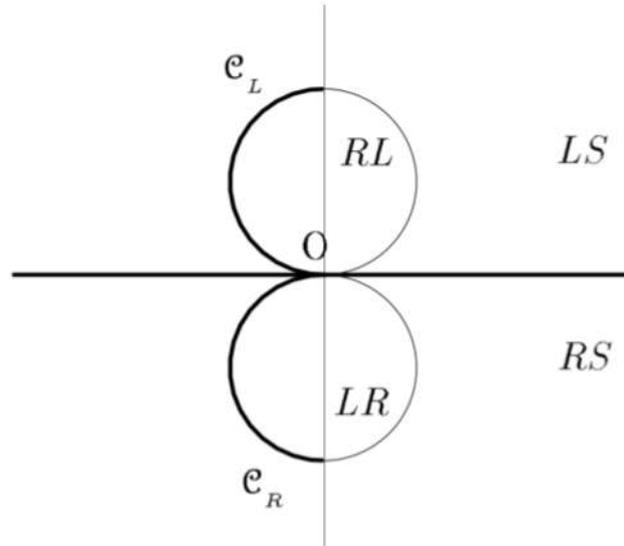
Figure 8: Shows the partition of the plane into Dubins' paths starting from an initial configuration origin moving along the x-axis

4. RESULTS AND CONCLUSIONS

We ran simulations using the length of the Dubins path as the criterion for the optimal trajectory for both the Sequential Monte Carlo Multiple Hypothesis Tracker (SMC-MHT) and Ground-Truth trackers. The results of the comparison with the simulations used in the previous work[6] are shown in Figures 9–12, showing much smoother paths. Figures 9 and 10 shows the tracks in black and the optimal trajectories in red for the straight line and Dubins' path for 30 simulations using the SMC-MHT tracker and Ground truth tracker respectively. Figures 11 and 12 show a histogram of the relative trajectories between the AUV and the estimated position of the recovery ship over the same 30 simulations.

It is clear from both the raw and relative trajectories that using the Dubins model has smoothed the paths generated compared to the straight line algorithm. As an indication of the narrowing of the relative histogram the smallest eigenvalue of the covariance of the data respresented in Figure 11 is 2.65×10^4 for (a) the straight line but only 1.52×10^4 for the relaxed Dubins paths.

REFERENCES

[1] Best, G. and Anstee, S., "Motion planning for autonomous underwater vehicle supervision," in [*Proceedings of Australasian Conference on Robotics and Automation*], (2014).

[2] Morelande, M. and Ellem, R., "Trajectory optimisation for 3D angle-only tracking (Milestone A)," tech. rep., University of Melbourne (5 2014).

[3] Morelande, M. and Ellem, R., "Trajectory optimisation for 3D angle-only tracking (Milestone B)," tech. rep., University of Melbourne (9 2014).

[4] Karaman, S. and Frazzoli, E., "Incremental sampling-based algorithms for optimal motion planning," *Robotics Science and Systems VI* **104**, 2 (2010).

[5] Karaman, S. and Frazzoli, E., "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research* **30**(7), 846–894 (2011).

[6] Williams, S., Wang, X., Angeley, D., Gilliam, C., Moran, B., Jackson, T., and Ellem, R., "Dynamic target driven trajectory planning using RRT*," in *22nd International Conference on Information Fusion*.[10]

[7] Wang, X., Williams, S., Angeley, D., Gilliam, C., Moran, B., Jackson, T., and Ellem, R., "RRT* trajectory scheduling using angles-only measurements for AUV recovery," in *22nd International Conference on Information Fusion*.[10]
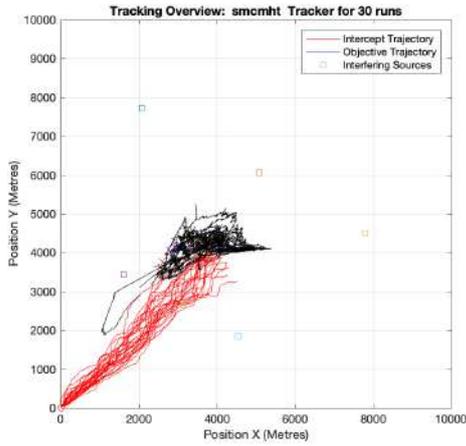
(a) Straight Line
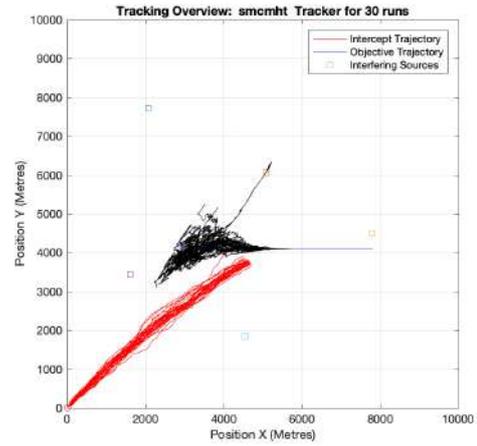
(b) Dubins path

Figure 9: Target track (black) and optimal trajectories (red) for 30 runs using RRT* trajectory planning for the scenario in Figure 1 using an SMC-MHT tracker
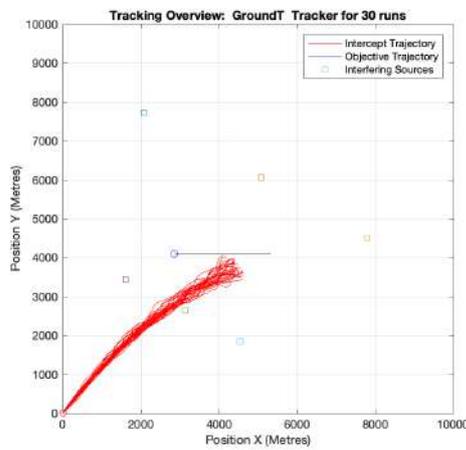

(a) Straight Line
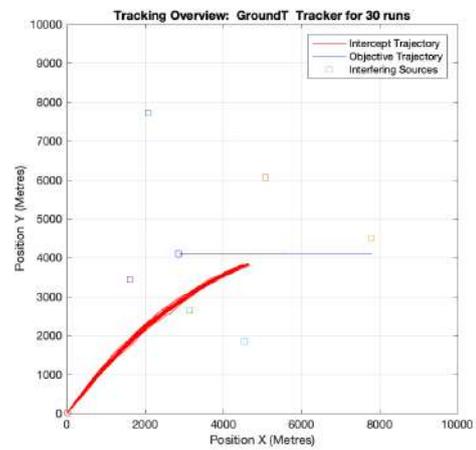
(b) Dubins path

Figure 10: Target track (black) and optimal trajectories (red) for 30 runs using RRT* trajectory planning for the scenario in Figure 1 using the Ground Truth tracker

[8] Reid, D. B., "An algorithm for tracking multiple targets.," *IEEE Transactions on Automatic Control* **24**(6), 843–854 (1979).

[9] Boissonnat, J.-D. and Bui, X.-N., "Accessibility region for a car that only moves forwards along optimal paths," Tech. Rep. 2181, INRIA, BP93, 06902 SOPHIA-ANTIPOLIS (1994).

[10] [*22nd International Conference on Information Fusion*] (2019).
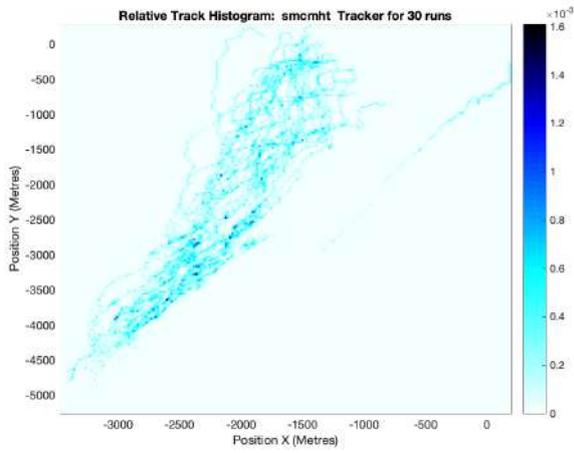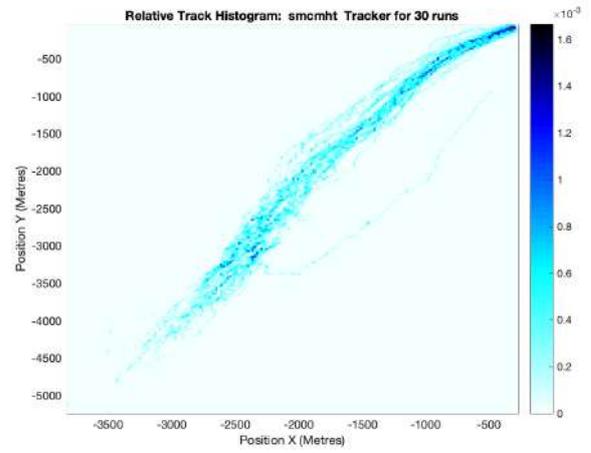
(a) Straight Line   (b) Dubins path

Figure 11: Probabilty of relative observer-target track using RRT* trajectory planning for the scenario in Figure 1 using an SMC-MHT tracker
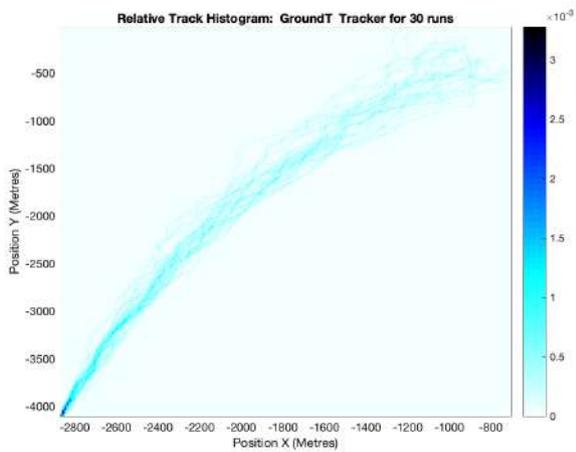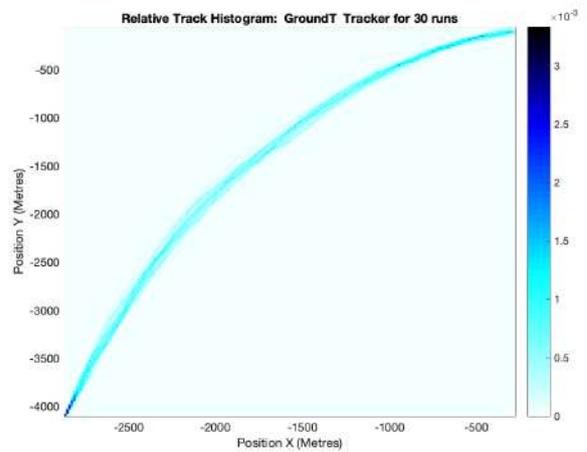

(a) Straight Line   (b) Dubins path

Figure 12: Probabilty of relative observer-target track using RRT* trajectory planning for the scenario in Figure 1 using the Ground Truth tracker

How to measure distance between the elements of an underwater robotic swarm by power Leds in unknown sea water conditions

R. dell'Erba*[a]

[a]ENEA Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Via Anguillarese 301,00123 Rome

ABSTRACT

One of the most important tasks of an underwater swarm is to know its geometric configuration; this can be achieved from the knowledge of the relative distances between the swarm elements as geometry distance problem. The method we propose here can be a possible support to compute them. Aim of this paper is therefore to measure distance between underwater AUVs swarm using cheap power Leds as light source and photodiode as receiver in unknown light water adsorption conditions. The method is based on light signal exchanged between the machines and the distances are calculated by the unknown water adsorption coefficient; the receiving photodiode produces a current we can correlate with distance and water adsorption coefficient, resolving the unknown parameters by moving the robots and stressing the emission conditions of the diode. In a previous paper we stressed work conditions of a power Led, we are using as optical modem in shallow water, to vary its emission characteristics. On these results we can perform a set of measurements leading to the knowledge of distances and adsorption coefficient $a(\lambda)$.

Keywords: Power Leds, Underwater robotic swarm, distance measurement.

*ramiro.dellerba@enea.it; phone +390630486128; fax +390630486038; http://robotica.casaccia.enea.it/index.php?lang=en

1. INTRODUCTION

It is well known as a robotic swarm presents advantages with respect of a single Underwater Autonomous Vehicle (AUV). These are in the speed-up coming by the parallelism and in the increase of reliability by redundancy [1], provided that the lack of one member can be easily managed by redistributing the job among the others like, in natural environmental, as used by the bees [2]. A swarm can interact with a human operator as a single object, without the problem of controlling a large number of individuals; job sharing between the elements is an internal swarm task. Moreover the human operator has the possibility to examine an object concurrently from more than one point of view, leading to a better perception of the surrounding environmental. A swarm is able to perform tasks in a more fast and robust way with respect of a single machine; but the most important feature, we are working, is its capability to span communication from the surface to the basement. This can allows a quasi-real time communication and therefore to interact with the underwater system also using a remote console that can be locate on the coast. Practically we can realize a multi-hop network with variable geometry. The swarm can adapt its configuration depending on the exploration mission and communication task. This could limit the use of the expensive surface ships to the deployment phase, taking advantage of the parallel exploration to shorten times and have many other advantages [3], [4]. ENEA is working in robotics since a long time (1961) and underwater robotics is a key topic of our laboratory [5]–[10]. Some years ago, (2006) we have moved our studies from a single autonomous underwater vehicle (AUV) to a swarm of very low cost cooperating robots. One of the most important tasks for an underwater vehicle is its localization into the sea; for a swarm there is also the need to know its geometric configuration. To get absolute localization at least one element of the swarm must have a precise position (typically on element emerges and fix position using GPS); localization of the whole swarm follow. The knowledge of the configuration is a very important issue for many applications often depending by the assigned task that can also vary during the mission [5], [11], [12]. Configuration can be computed from distance measurements [5] that is the purpose of this paper. Usually is done by ultrasonic measurements but our intention is to use absorption light coefficient as support to traditional measurements.

1.1 State of art

Acoustic communications are the standards in submarine environmental [13]. Unfortunately reflections, fading and other phenomena make difficult and sometimes not reliable their measurements. The need to continuously exchange data among the nodes (the swarm elements) to get the distances represents a considerable burden for the network operation to calculate the network configuration, also using suitable algorithms, because it force to frequent short messages that deoptimize the exploitation of the communication channel, mainly for the long times needed to switch from a message to another one [14]. A possible solution, which improves both the time allocation in the acoustical protocols and unloads the acoustical channel burden, is to couple the acoustic protocol with optical device [15], [16]with the intention to collect distance measurements between the robots more precise, using sensor data fusion. Optical methods are very powerful but their performances are affected by many strongly variable parameters like salinity, turbidity, the presence of dissolved substances that change the color and the transparency of the water in different optical bands [17]. Moreover the amount of solar radiation, in shallow water, heavily affects the signal to noise ratio. Our current approach uses a mixed strategy based on the variable exploitation of the optical channel depending on the environmental conditions. In favorable conditions the transmission protocol will freely decide which channel to adopt depending on the priority, i.e. distance-to-cover and dimension of the message itself. In less favorable conditions the optical channel will be limited to the fundamental synchronization task, generating a light lamp that will optimize the message passing through the acoustic channel. So far we have realized an optical modem with cheap power Leds that it is working together acoustic system. It must be outlined that only dense swarms can take advantage of such an approach because only in these situations, with internal distances ranging from few meters to a maximum of 20 meters, there are the conditions suitable to use light signals for sync and measurements. Moreover, backup solutions based on the "all acoustical" approach, must remain available because it is always possible to find dirty waters with no way to use light signals. In this paper we work about the problem to calculate distance using optical signals. In a previous paper we proposed the use of a cheap power led system to support acoustic devices in localization and configuration computation of an underwater robotics swarm [18]. The system was based on light signal exchanged between the machines and used power led, of different wavelength, to calculate distances between them. The unknown water conditions, affecting the light propagation, required a local measure of the absorption function $a(\lambda)$ using a known distance. Moreover we investigated as power supply and flash duration of the Leds can be stressed to vary light emission spectrum. By this experience now we are able to measure distance without the local measure of the absorption function. In this paper we show as, modifying spectrum, we shall be able to measure $a(\lambda)$ and, consequently, the distance d.

1.2 Our prototype

In Figure 1 and Figure 1Figure 2 the swarm element, named Venus and realized in our laboratory, and the optical modem prototype are shown. Its characteristics are the following: Max depth 100m; Max speed 4 Km/hr; Weight about 20 Kg; Autonomy 3hrs; Dimensions are 1.20mX0.20m diameter. Standard sensors include a stereoscopic camera, sonar, accelerometer, compass, depth meter, hydrophones side-scan sonar. In the optical modem the Leds' chips and photodiode are visible; the system is repeated on three faces to cover 360°.

Figure 1 Robot prototype during test in Bracciano's lake

Figure 2 Optical modem

Note we are dealing with a system thought as to be a component of a swarm of about 20 objects. The distances between robots are between 3 and 50 m. Therefore, the maximum distance possible between two robots is about 1000 m, as a very particular alignment case; the average value of the distances was considered about 10 m. Technology of dense swarms is an answer to these problems need larger acoustical bands and allowing other physical channels to be exploited.

2. RESULTS
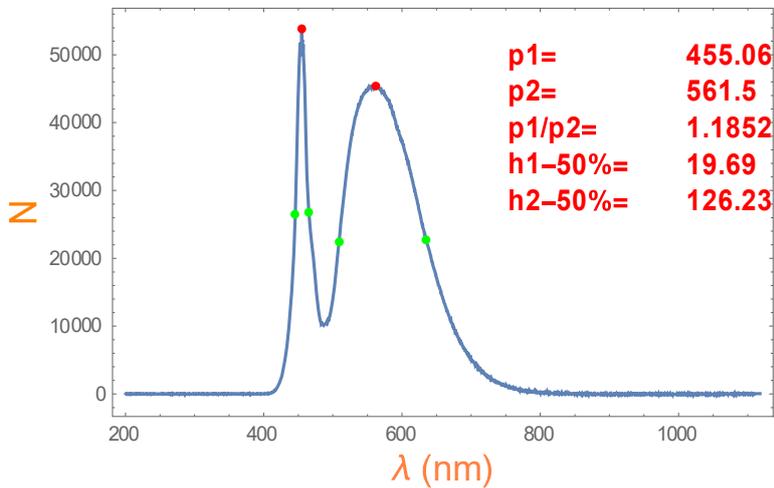
Figure 4 White Led emission

In a previous paper [18] we have considered cheap power Leds that we are using to build an optical modem and tried to stress power supply and flash duration to vary the characteristic of light emission; typical emission of the white Led is shown in Figure 4 while one experimental result, concerning the peak emission value vs. power supply, can be seen in Figure 3. We used white Led ENSW10-1010-EB1 by EDISTAR together with a photodiode OSD100-E by Centronic. The measurements have been performed by HR4000 instrument, by Ocean Optics, with a resolution of 0.5 nm wavelength. The reason to attempt to vary the Led's emission characteristics is economic: owing to the large number of swarm elements the use

of complex variable frequency system is not sustainable. So we modify the peak shape emission by the power tension and perform distance measurements based on water light adsorption. The water adsorption is generally unknown so, in the previous work, we performed a local measurement of the attenuation coefficient by a measure on the known distance head-tale of the robot. Unfortunately this implies we have Leds emission on the head of the torpedo-like robot and photodiode on the tale. This determines constructive problems of the optical modem so far we want got rid of this architecture looking for another solution. So far stressing Led emission we get information useful to model water attenuation coefficient and, consequently, the distance. Practically we vary the spectrum emission of the Led, varying power supply, and

Figure 3 White Led first peak wave length shift value vs tension

perform an attempt to model absorption coefficient. In this paper we propose a theoretical way to compute it with not yet experimental data but simulating how could be the collected current from the photodiode using a simple equation. Only the data concerning the emission of the Led are experimental. Later we try to match the current using some water categories, or linear combination of them, from the literature.

Figure 5 Light absorption in sea water in different condition [14]

Light absorption by a medium, as a function of distance from the source is a well known phenomenon. In a first approximation, for a spherical wave, we can assume an exponential law for the decay of the signal intensity. So we can write

$$I = f(\theta, \varphi) \frac{I_0 \, e^{-a(\lambda)d}}{d^2} \qquad (1)$$

Where I is the measured intensity signal, I_0 the emitted intensity, $a(\lambda)$ the absorption function, $f(\theta,\varphi)$ a shape factor and d the distance. We do not consider here the shape factor, known for the characteristics of the Led and of the photodiode leading to a modify of the spherical shape. Typically you should consider an emission diagram of 60°x60° therefore the energy is spread on this solid angle. The $a(\lambda)$, describing how the signal is attenuated as function of the wave length λ, is strongly affected by the water conditions as can be seen in Fig. 5 due to the courtesy of [19]. We are assuming that water characteristics do not vary in the volume containing the robots. There is no way to obtain, $a(\lambda)$ and d separately by a simple measure if the intensity I. In a previous work [20] we used monochromatic sharp peak blue Led to stress so we can measure the derivative of Eq. 1 respect to λ that give us the product $a'(\lambda)*d$. Computing it for two, unknown, distance d_1, d_2 we can obtain the ratio d_1/d_2 that is what we need. So far, considering the ratio of Eq.1 for the two distances, we can obtain the product $a(\lambda)*d_1$. Now, considering the Log of Eq.1 we have $a(\lambda)*d_1 + 2Log(d1)$ and consequently d_1, without any model of $a(\lambda)$.

This is very nice but we are now using real white Leds so we use, in this case, a semi-experimental procedure. We use
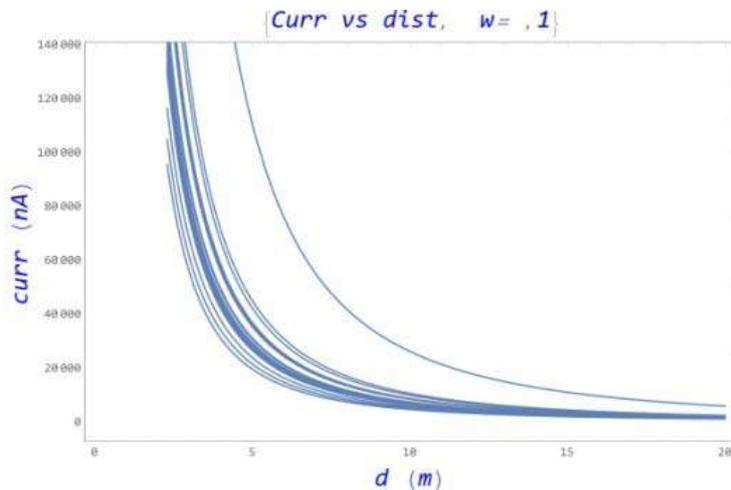
Figure 6 Collected current by photodiode for different Led stress vs distance in clear water (w=1)

experimental stressed emission curves of the white Led together with the, furnished by the factory, responsivity of the photodiode to calculate the collected current at different, unknown, distance. The collected current is computed by the integration of the product of the experimental emission curves, the responsivity of the photodiode and Eq.1. We use, as $a(\lambda)$, the eight curves of Fig.5 to build a database. If we stress our Led emission on different unknown distances we obtain a set of curves like Fig.6 and Fig. 7, depending on water quality. We now use the eight curves of Fig.5 as preliminary database, so we have 8 plots like Fig.6. Now we perform our measurements in unknown water and distances, stressing the Led in 40 different ways, obtain a 40 points set. These points are a vertical line on one of the 8 plots like Figure 6 posed at unknown

distance; the points must coincide on the curves. We repeat the operation for some other unknown distances. We have to match our 40 points set into one of the 8 plots at a certain distances. One measurement could be present ambiguities leading to multiple choices so we repeat the measure at other distances as well as at the same one. At the end will be only one possible solution to match all. As "match" we mean a procedure concerning the minimization of the distances of our set of points from the curves, through optimization procedure. We are assuming the water behavior can be described by Figure 5 or by a combination of its curves; this could be reasonable because the figure is considering a very large water conditions cases. If, as usual, our water is not comprise into the eight curves we generate it from linear combination of them loosing in precision. Practically the steps are the following. 1) Simulate the data. We choose one of the eight kind of water (2 in our example or linear combination of them) and a distance d (7.1 in our example). The collected current C form the photodiode is:
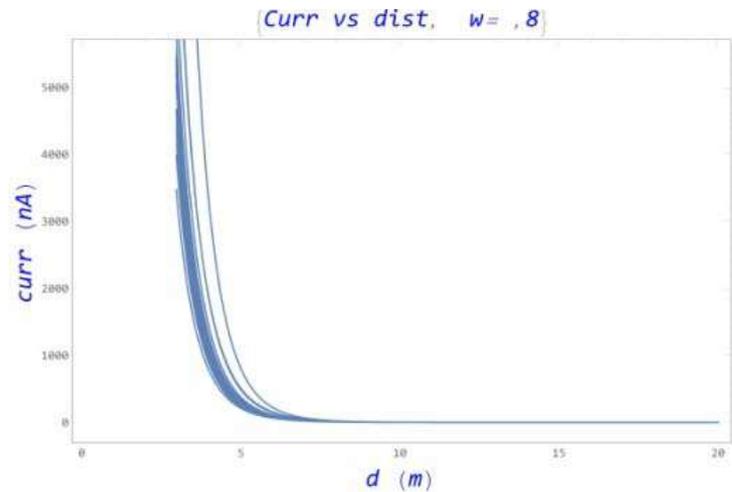
Figura 7 Collected current by photodiode for different Led stress vs distance in turbid water (w=8)

$$C = \int_{390}^{700} R * \text{EmLed}(\lambda) * \text{Eff}(\lambda) * f(\theta, \varphi) \frac{I_0\, e^{-a(\lambda)d}}{d^2}\, d\lambda$$

Where EmLed is the emission curve of the Led (Figure 1) in 40 different stress conditions and Eff is the efficiency response of the photodiode; we have also introduced a random noise R of the 10%. 2) With this data we try to match, for each picture like Figure 6 and Figura 7 (or linear combination of them) the set of 40 points until we find the distance d as can be seen in Figure 8 and Figure 9. In the figures our measurements are the vertical points and there is just one distance, in a particular kind of water, to minimize the distances of the points from the curves; we move the vertical lines to find the distance for the best match between the curves and the points.
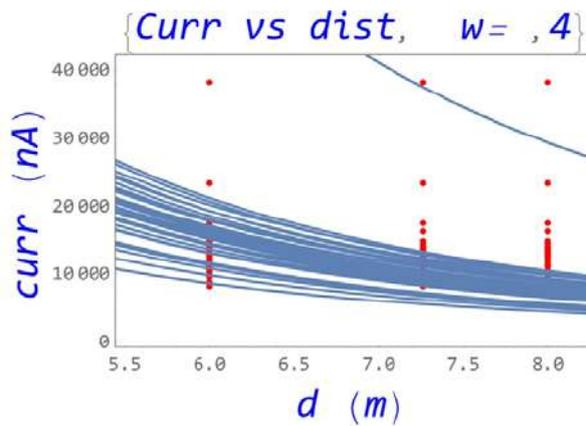
Figure 8 Collected current by photodiode for different Led stress vs distance in water (w=4) and distance=7.1m. The red points are our, simulated, measurements. We try to match the curves without success
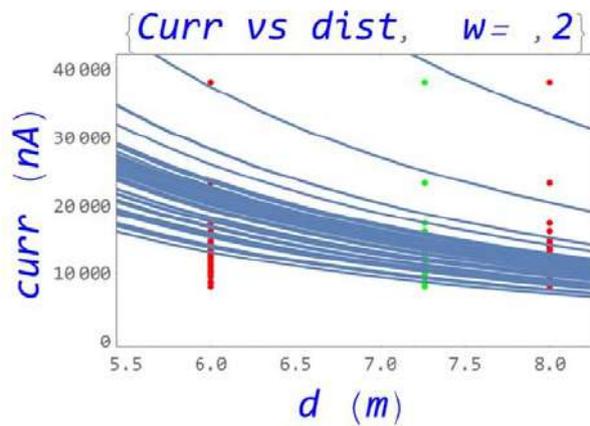
Figure 9 Collected current by photodiode for different Led stress vs distance in water (w=2) and distance=7.1m. The red points are our, simulated, measurements. We try to match the curves: best result is for d=7.26 m

Simulation computing can be seen in Table 1 where the simulation for ten different distances is shown. Therefore we perform 40 flashes for each distance and 50 measurements with random errors, stressing the Led in different conditions, for ten unknown distances. The results are quite good. Take in account that each simulated measurements is affected by a random 10% noise generate to make more realistic the computing. In Table 2 we report the results of mixed water conditions (Linear combination of water of kind 2 and 4). The results are less accuracy, especially at low distances, but still useful.

Real distance	Estimated distance	Standard deviation
2.8	2.6	0.5
3.2	3.1	0.1
3.5	3.5	0.19
4.2	3.7	0.9
7.2	6.9	1.6
10.1	9.8	0.8
12.1	12.1	0.7
12.5	12.4	1.5
17.3	16.5	2.6
19.1	15.8	4.8

Table 1 Estimation of the unknown distance *d* using one of the eight water of Figure 5

Real distance	Estimated distance	Standard deviation
0.2	2.0	0.5
0.8	3.1	0.1
3.7	3.3	0.1
5.2	3.7	0.9
8.3	6.9	1.6
10.8	9.9	0.8
12.6	12.1	0.7
13.4	12.4	1.5
16.5	16.5	2.6
19.1	15.8	4.8

Table 2 Estimation of the unknown distance *d* using mix of the eight water of Figure 5

3. CONCLUSION

We have proposed a method to measure distances and extinction coefficient in unknown underwater conditions by using cheap power Led whose emission is stressed by power source. The method could be a support to standard acoustic measurements and can be used to calculate configuration of an underwater robotic swarm that we are developing in our laboratory. We have built an optical modem by using cheap commercial Leds and modifying power tension we are able to modify emission peak. In this paper we proposed semi- experimental results where the emission peak is experimental but the collected current form the photodiode is computed by simple equation. Simulating these collected current we use a database of water condition, or its linear combination, to match the data. Varying the unknown distances we can build a set of compatibility to identify the extinction coefficient of the water.

Experimental measurements are in progress to verify the concrete possibility to estimate extinction coefficient a(λ) by this method. Errors sources are many. This method must be considered as an iterative method whose precision is increasing with many measurements. These measurements must be integrated with some other source like acoustic, as usual in robot science. The use of more than one Led, with different frequency emission, suddenly increase precision [18]. This because working on the ratio intensity at the different frequencies we can enhance our measurements owing to the increasing stability of the photodiode current, with respect of that of a single source. Practically we increase the sensitivity of the signal to the measured distance, by enlarging the responsivity dynamic. The work is in progress in our laboratory with experimental campaign into the Bracciano Lake, close to our laboratories.

REFERENCES

[1] C. Xian-yi, L. Shu-qin, e X. De-shen, «Study of self-organization model of multiple mobile robot», *Arxiv preprint cs/0601062*, vol. International Journal of Advanced Robotic Systems, Volume 2, Number 3 (2005), ISSN 1729-8806, 2006.

[2] D. Karaboga, «An idea based on honey bee swarm for numerical optimization», Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, 2005 [Online]. Available at: http://www-lia.deis.unibo.it/Courses/SistInt/articoli/bee-colony1.pdf. [Consultato: 07-nov-2014]

[3] B. Anderson e J. Crowell, «Workhorse AUV-a cost-sensible new autonomous underwater vehicle for surveys/soundings, search & rescue, and research», *Proceedings of MTS/IEEE OCEANS, 2005*, pagg. 1228–1233, 2005.

[4] V. P. Shah, «Design considerations for engineering Autonomous Underwater Vehicles (2007). Design considerations for engineering AUVs».

[5] R. dell'Erba, «Determination of spatial configuration of an underwater swarm with minimum data», *International Journal of Advanced Robotic Systems*, vol. 12, n. 7, pag. 97, 2015.

[6] R. dell'Erba e C. Moriconi, «Harness: A robotic swarm for environmental surveillance», in *6th IARP Workshop on Risky Interventions and Environmental Surveillance (RISE), Warsaw, Poland*, 2012.

[7] R. dell'Erba e C. Moriconi, «HARNESS: A ROBOTIC SWARM TO EXPLORE AND PROTECT UNDERWATER MARINE COAST», presentato al Quarto Simposio Internazionale: "Il monitoraggio costiero mediterraneo: problematiche e tecniche di misura, Livorno, 2014.

[8] R. dell'Erba e C. Moriconi, «HARNESS: A Robotic Swarm for Harbour Security"», presentato al International workshop Port and Regional Maritime Security Symposium, Lerici (Sp) Italy.

[9] R. Dell'Erba e C. Moriconi, «Bio-inspired Robotics — it». [Online]. Available at: http://www.enea.it/it/produzione-scientifica/edizioni-enea/2014/bio-inspirede-robotics-proceedings. [Consultato: 15-dic-2014]

[10] R. Dell'Erba, C. Moriconi, e A. Trocciola, «CAN ROBOTIC TECHNOLOGY SAVE CULTURAL ASSETS SUBMERGED?», *ARCHEOMATICA-TECNOLOGIE PER I BENI CULTURALI*, vol. 9, n. 3, pagg. 22–26, 2018.

[11] R. Dell'Erba, «The Localization problem for an underwater swarm», *Technical report ENEA*, 2012.

[12] C. Moriconi e R. dell'Erba, «The Localization Problem for Harness: A Multipurpose Robotic Swarm», in *SENSORCOMM 2012, The Sixth International Conference on Sensor Technologies and Applications*, 2012, pagg. 327–333 [Online]. Available at: http://www.thinkmind.org/index.php?view=article&articleid=sensorcomm_2012_14_20_10138. [Consultato: 04-apr-2014]

[13] «Underwater acoustic localization for small submersibles - Kottege - 2010 - Journal of Field Robotics - Wiley Online Library». [Online]. Available at: http://onlinelibrary.wiley.com/doi/10.1002/rob.20378/full. [Consultato: 03-apr-2012]

[14] F. S. Schill, «Distributed communication in swarms of autonomous underwater vehicles», The Australian National University, 2007.

[15] S. Han, Y. Noh, R. Liang, R. Chen, Y.-J. Cheng, e M. Gerla, «Evaluation of underwater optical-acoustic hybrid network», *Communications, China*, vol. 11, n. 5, pagg. 49–59, 2014.

[16] N. Farr, A. Bowen, J. Ware, C. Pontbriand, e M. Tivey, «An integrated, underwater optical/acoustic communications system», in *OCEANS 2010 IEEE-Sydney*, 2010, pagg. 1–6.

[17] «IEEE Xplore Abstract - Optical wireless underwater communication for AUV: Preliminary simulation and experimental results». [Online]. Available at: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=6003598&abstractAccess=no&userType=inst. [Consultato: 12-gen-2015]

[18] R. dell'Erba e C. Moriconi, «High power leds in localization of underwater robotics swarms», *IFAC-PapersOnLine*, vol. 48, n. 10, pagg. 117–122, 2015.

[19] W. Bogdan e D. Jerzi, *Light Absorption in Sea Water*, Springer. 2007 [Online]. Available at: http://download.springer.com/static/pdf/107/bok%253A978-0-387-49560-6.pdf?auth66=1416993589_ea5bcc47d8e3ca441bbb50259745bb4b&ext=.pdf. [Consultato: 26-nov-2014]

[1]

[20] R.dell'Erba, submitted to Journal of Applied Mathematics and Mechanics

CopterCurrents
High Resolution Surface Current Fields
Measured by a Small UAV

Jochen Horstmann[*], Ruben Carrasco and Michael Streßer
Helmholtz-Zentrum Geesthacht, Max-Planck-Str. 1, 21502 Geesthacht, Germany

ABSTRACT

Within this paper, we describe a methodology called CopterCurrents to retrieve high-resolution ocean surface currents utilizing video sequences of ocean surface gravity waves. The video sequences were acquired by an off the shelf quadcopter utilizing a small, nadir looking 4k video camera. For stabilization of the video, the camera is mounted on an actively controlled gimbal. For geocoding of the video data, the aircrafts data logger records all flight data collected by the drone. Within the first processing step, the data are corrected for lens distortion and geocoded to a rectilinear coordinate system at water level. In the next step, the directions, lengths as well as phase velocities of the imaged surface waves are extracted from the geocoded video sequence. Finally, the properties of the surface waves are used to estimate the surface current vector, by retrieving the difference of the observed phase velocities to that given by the linear dispersion relation of surface gravity waves. To demonstrate the capability as well as applicability of CopterCurrents, we acquired video sequences in different environments and compared the resulting current fields to measurements obtained by other instruments such as an acoustic Doppler current profiler.

Keywords: Surface currents, ocean waves, video processing, UAV, drone, remote sensing, dispersion filter

1. INTRODUCTION

High-resolution surface current fields in rivers and coastal areas are of major interest for engineers and researchers. In particular as the currents are one of the most important drivers of morphological changes and often a major source of complications for engineering projects. Therefore, 2D current maps with high spatial and temporal resolution are an important source of information for the design and planning of engineering projects. In general, the measurement of 2D surface current fields with high spatial and temporal resolution is challenging and very costly to obtain. Typically, surface currents are measured by acoustic Doppler current profilers (ADCP), which measure the speed and direction of a fluid at a single point (Eulerian measurements) along a vertical profile, or by surface drifters, which describe the path of a water body at the surface (Lagrangian measurement). To obtain a surface current map in space and time, land based high frequency (HF) radars are often used [1]. Unfortunately, HF-radars measure surface current fields with spatial resolutions > 500 m and temporal resolutions > 30 minutes. In order to overcome these limitations, numerical circulation models are utilized, which are complex, require a lot of additional input such as bathymetry and tidal levels, and are still not an ideal solution for high-resolution current fields. In the last decade, microwave radars (marine radars) have been utilized to obtain surface currents [2]. However, the methods utilized for marine radars are limited to the smallest waves the radars can resolve, which are typically > 15 m and therefore often a drawback in inlets, harbors and rivers as well as under low wind and wave situations.

Within this paper a very cost effective method is being introduced and applied to retrieve ocean surface currents utilizing a video sequence of surface waves that are acquired in the range of visible light by an off the shelf drone. The methodology is based on the dispersion relation of ocean surface gravity waves [3], which has been successfully applied to marine radar image sequences [4, 5] as well as to airborne video sequences [6]. In contrast to the previous airborne application [6], today's off the shelf quadcopters allow to geocode video imagery without the need of ground control points within the image utilizing a truly low cost instrument [7, 8]. The methodology can also be extended to estimate the vertical current shear near to the surface by considering waves of different wave lengths.

---

[*] Jochen.Horstmann@hzg.de; www.hzg.de

2. ACQUISITION SYSTEM

All video data utilized within this study were acquired with the Phantom 3 Professional, an off the shelf quadcopter from DJI (Figure left hand side). The Phantom 3 carries an Ultra-High Definition (4069 x 2160 pixel) video camera with acquires video imagery with a frame rate of up to 25 Hz. The camera is equipped with a 1/2.3 CMOS chip and a 20 mm (35 mm format equivalent) rectilinear lens with f/2.8 at ∞. The lens covers a 94° field of view with near to no barrel distortion as well as negligible chromatic aberration. At a flight height of 100 m the pixel resolution is approximately 0.04 m. For stabilization of the camera, the quadcopter is equipped with an active 3-axis gimbal, which very successfully accommodates for pitch, roll and yaw movements of the aircraft (Figure 1 right hand side). The pitch can be controlled within -90° to +30° at a maximum angular speed of 90°/s and an angular control accuracy of ±0.02°. For all video acquisitions within this paper, the camera was operated at nadir (down looking) and set to Ultra-High Definition (4069x2160 pixel) with a frame rate of 25 Hz. All videos were stored on the cameras internal Micro SD card.

Figure 1: On the left hand side the Phantom3 Professional from DJI on the right hand side the Brushless gimbal with the 4K video camera mounted below the aircraft.

The automatic hovering system of the Phantom 3 Professional allows an accuracy of the selected position within ± 0.5 m in the vertical and ± 1.5 m in the horizontal. The hovering accuracy is maintained at wind speeds of up to 8 m/s. During the entire flights all flight information, such as geographical position, height, heading as well as the tilt, roll and yaw angle of the aircraft and gimbal are recorded by the aircrafts data logger. The positioning and gimbal accuracy as well as the geocoding has been investigated by Holman et al. [9] showing an overall excellent performance for observation of the ocean surface.

3. METHODOLOGY

In a first step all video data are corrected for the lens distortion and then geocoded to a rectilinear coordinate system at the level of the water surface (UTM 32). For this purpose the quadcopters recorded telemetric data, e.g. latitude, longitude and flight height, are utilized. To show the overall performance of the system a video image sequence of the Elbe River was acquired near the city of Lauenburg, Germany (Figure 2). For this sequence, the quadcopter was hovering over the river acquiring several minutes of video data. Within the video sequences (typically 30 to 60 s) the propagation of surface waves can be observed in space and time. These videos allow to measure surface wave properties such as the wave length and phase velocity, which in turn enable to retrieve the surface current vector. The surface current results from the difference of the observed phase velocity to that given by the linear dispersion relation of surface gravity waves,

$$\omega = \varpi(\boldsymbol{k}) = \sqrt{g\,k\,tanh(k\,d)} + \boldsymbol{k} \cdot \boldsymbol{U} = \varpi_0(k) + \boldsymbol{k} \cdot \boldsymbol{U}, \tag{1}$$

were $\boldsymbol{k}$ is the wave number vector, g the gravitational force, d the water depth and $\boldsymbol{U}$ the two-dimensional near-surface current, commonly known as velocity of encounter. This physical property of surface waves is also being utilized very successfully for retrieving ocean surface currents from HF-radar [1] as well as from marine radars [4, 5].

At an altitude of 100 m the ground resolution is approximately 0.04 m, which allows resolving wave lengths greater than 0.08 m. In the following example of the Elbe River the investigation has been limited to surface waves within 0.3 m and 3 m. Due to the high frame rate of 25 Hz aliasing [3, 8] is a minor issue when analyzing these data sets and has been neglected. In a next step, the geocoded video data are transformed from the space-time domain into the wavenumber-frequency domain using a 3D Fast Fourier Transformation (FFT). In the wave number frequency domain the energy of the

Figure 2: Single image from the video sequences of the Elbe River collected in front of Lauenburg, Germany.

surface gravity waves is located on the so-called dispersion shell, which is defined in (1). Figure 3 shows a wavenumber frequency plain with the dispersion shell (curve) fitted to the data. The shape of the dispersion shell is changed due to the velocity of encounter, which is the sum of the drone's velocity and of the near surface currents. Within the final step, the velocity of encounter is estimated by a simple least squares fitting technique [3, 8]. Finally, the surface current is retrieved
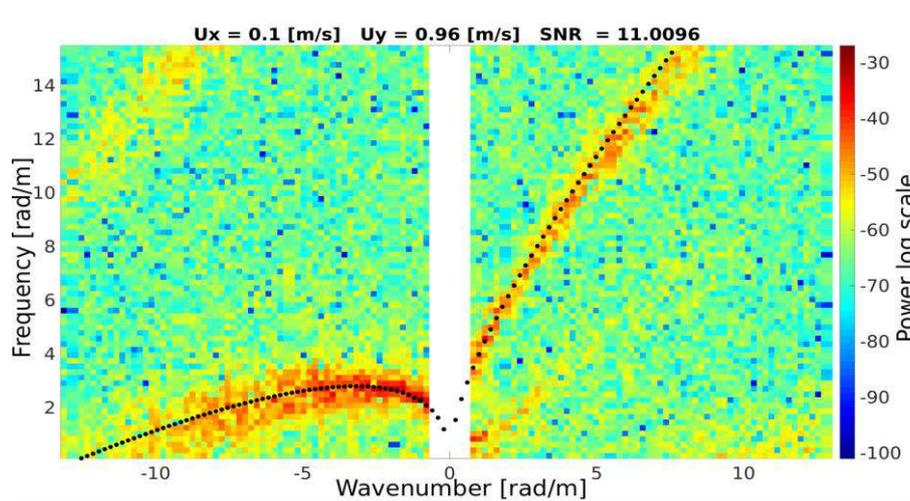

Figure 3: Wave number frequency plain along the main flow direction of the Elbe River retrieved from a box of 20 m x 20 m in the main flow of the Elbe River (yellow square indicated in Figure 2).

by removing the contribution of the drone's movement from the velocity of encounter. This entire method is named "CopterCurrents" and the source code, designed in MATLAB, has recently been made available at the GitHub depository (https://github.com/RubenCarrascoAlvarez/CopterCurrents).

Within this study, the methodology has been extended to retrieving surface current shear by extracting the currents for different wave number intervals [10]. As the retrieved currents are a weighted average of the current at the surface down to the penetration depth of the wave (approximately half the ocean wavelength), the current shear can be retrieved by estimating the currents at different wavenumber intervals. The contribution to the surface currents are largest at the surface and decrease exponentially with depth. In case of a linear current profile, the video-derived current corresponds to an effective depth of 7.8% of the wavelength.

4. RESULTS

To test the entire methodology several flights were performed over the Elbe River at the city of Geesthacht and Lauenburg, Germany. In a first attempt video data were collected from the weir across the Elbe River at Geesthacht, were the current direction and flow pattern in front of the open gates of the weir are nicely depicted by the current field retrieved via CopterCurrents (for details, refer to Horstmann et al. [7]). An initial validation of the methodology was performed at the Elbe River in front of Lauenburg (Figure 2). The drone-retrieved currents (Figure 4) were compared to ADCP data collected by a small boat over the entire area imaged by the drone. These results showed an overall root mean square error of 0.09 m/s with a negligible bias [8].
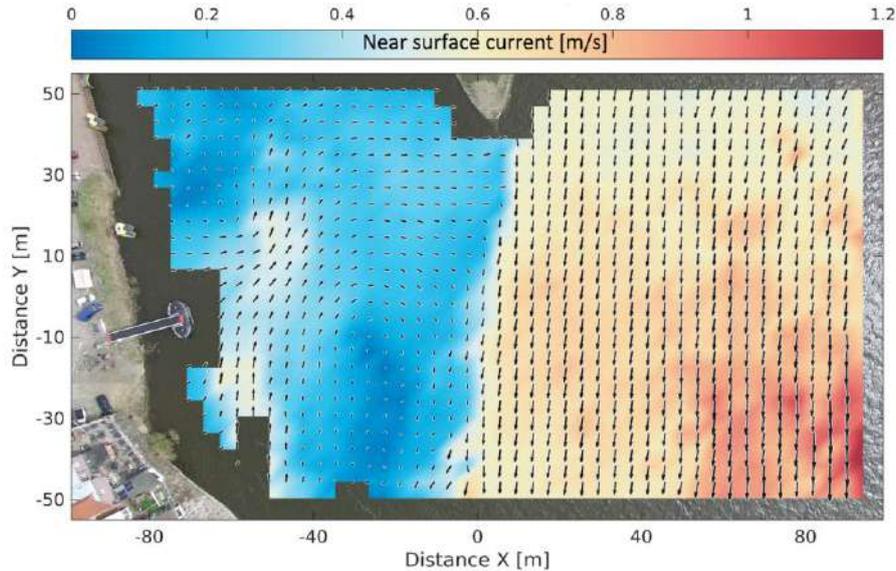

Figure 4: Surface current field resulting from the image sequence of the Elbe River (Figure 2) in front of Lauenburg.

Within this study, the CopterCurrents method is extended to retrieve the upper surface current shear by estimating the currents from different intervals of wavelengths. For testing the method two areas were selected in the Elbe River in front of Lauenburg as indicated in Figure 5 and 6. The first area is located in a channel to the port of Lauenburg with very low
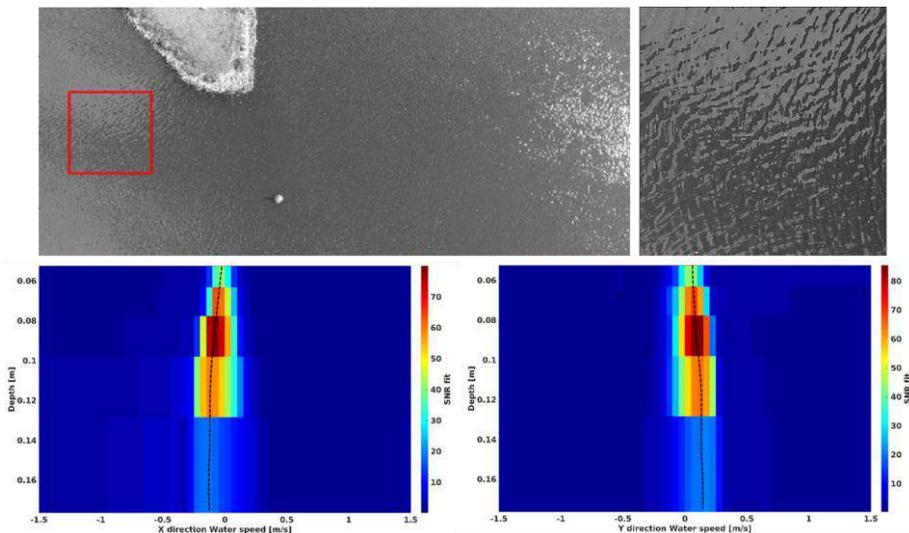

Figure 5: Surface current shear in the channel to the port of Lauenburg with low current speeds. The shear in X-direction (lower left) and Y-direction (lower right) as retrieved from the video data.
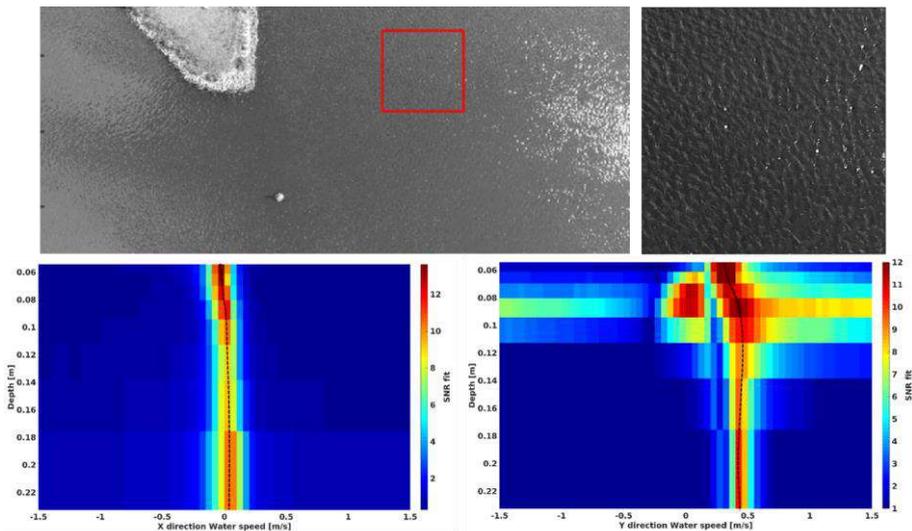
Figure 6: Surface current shear in the main flow of the Elbe River with current speeds of 0.5 m/s. The shear in X-direction (lower left) and Y-direction (lower right) as retrieved from the video data.

currents (Figure 5) and the second within the main flow of the Elbe River (Figure 6). During the acquisition, the winds were blowing upstream with approximately 7 m/s. Within the channel to the port of Lauenburg (Figure 5) there is nearly no vertical shear in X- and Y-direction, respectively. Note, that in this case the shear can only be estimated in the upper 0.12 m due to the lack of longer waves. The vertical shear in the main flow of the Elbe River is depicted in Figure 6. A significant shear can be observed in the upper 0.1 m in particular along the main flow direction.

5. CONCLUSION AND OUTLOOK

A methodology, namely CopterCurrents, was developed to extract surface current fields with a spatial resolution of < 5 m and a temporal resolution of up 10 s utilizing video recordings of an off the shelf small drone. A preliminary validation was performed by comparing the current fields to ADCP measurements, which resulted in an overall error of < 0.1 m/s [8]. Within this study, several flights were conducted under different environmental conditions to further test the capability of the method. The video data have been converted to surface current fields and the SNR has been utilized as a first kind of quality parameter for the measurements. Due to the small size of the quadcopter, measurements can only be performed at wind speeds below 8 m/s and are limited to flight durations of approximately 20 min within a range of 1 to 2 km. The overall performance of CopterCurrents has shown very promising results under various conditions e.g. current fronts, bathymetric jumps, small scale eddies and river flow fields. Therefore, the entire source code of the CopterCurrents toolbox has been made available at the GitHub depository (https://github.com/RubenCarrascoAlvarez/CopterCurrents).

CopterCurrents has been extended to estimate the vertical current shear. The method was applied to video data acquired over the Elbe River showing significant surface shear in the upper 0.1 m. To validate the method an experiment has been carried out in the Bear Cut Inlet, in Miami, USA, in 2018. Several flights were undertaken with co-located drifter and ADCP measurements as well as by deployment of various objects in the water (e.g. bamboo plates and oranges) to retrieve the current velocity via particle tracking velocimetry (PTV) from the drones video data. The main goal of the Bear Cut Inlet experiment was to investigate the upper surface shear observed with the extended version of CopterCurrents in more detail. Preliminary results showed a clear dependence of the copter retrieved surface currents on the local wind conditions in particular in the upper 0.1 to 0.2 m. An in depth analysis of the data is ongoing. The results will be presented during the Maritime Situational Awareness Workshop in Lerici, Italy and will be published shortly.

ACKNOWLEDGMENT

REFERENCES

[1] A. Rubi, J. Mader, L. Corgnati, C. Mantovani, A. Griffa, A. Novellino, C. Quentin, L. Wyatt, J. Schulz-Stellenfleth, J. Horstmann, P. Lorente, E. Zambianchi, M. Hartnett, C. Fernandes, V. Zervakis, P. Gorringe, A. Melet and I. Puillat, HF Radar Activity in European Coastal Seas: Next Steps toward a Pan-European HF Radar Network, Front. Mar. Sci., Vol. 4(8), doi: 10.3389/fmars.2017.00008, 2017.

[2] J. Horstmann, J.C. Nieto Borge, J. Seemann, R. Carrasco and B. Lund, Wind, Wave and Current retrieval utilizing X-Band Marine Radars, Chapter 16 in Coastal Ocean Observing Systems, Elsevier, p. 281-304, 2015.

[3] C. Senet, J. Seemann, and F. Ziemer, The Near-Surface Current Velocity Determined from 5 Image Sequences of the Sea Surface, IEEE Trans. Geosci. Remote Sens., Vol. 39(3), p. 492-505, 2001.

[4] W. Huang, R. Carrasco, C. Shen, E. W. Gill and J. Horstmann, Surface Current Measurements Using X-Band Marine Radar With Vertical Polarization, IEEE Transactions on Geoscience and Remote Sensing, Vol. 54(5), p. 2988-2997, May 2016. doi: 10.1109/TGRS.2015.2509781, 2016.

[5] B. Lund, B.K. Haus, J. Horstmann, H.C. Graber, R. Carrasco, N.J.M. Laxague, G. Novelli, C.M. Guigand, and T. Özgökmen, Near-Surface Current Mapping by Shipboard Marine X-band Radar: A Validation, Journal of Atmospheric and Oceanic Technology, Vol. 35(5), p. 1077–1090, doi:10.1175/JTECH-D-17-0154.1, 2018.

[6] J.P. Dugan, C.C. Piotrowski, and J.Z. Williams, Water Depth and Surface Current Retrievals from Airborne Optical Measurements of Surface Gravity Wave Dispersion, J. Geophys. Res., Vol. 106(C8), p. 16903-16915, 2001.

[7] J. Horstmann, M. Streßer, and R. Carrasco, Surface Currents Retrieved from Airborne Video, in Proc. OCEANS-Aberdeen, doi: 10.1109/OCEANSE.2017.8084957, 2017.

[8] M. Streßer, R. Carrasco and J. Horstmann, Video-Based Estimation of Surface Currents Using a Low-Cost Quadcopter, in IEEE Geoscience and Remote Sensing Letters, Vol. 14(11), pp. 2027-2031, doi: 10.1109/LGRS.2017.2749120, 2017.

[9] R.A. Holman, K.L. Brodie, and N.J. Spore, Surf Zone Characterization using a Small Quadcopter: Technical Issues and Procedures, IEEE Trans. Geosci. Remote Sens., vol. 55(4), p. 2017–2027, Apr. 2017.

[10] B. Lund, H.C. Graber, H. Tamura, C.O. Collins III, and S.M. Varlamov, A New Technique for the Retrieval of Near-Surface Vertical Current Shear from Marine X-band Radar Images, J. Geophys. Res., Vol. 120(12), p. 2169-9291, 2015.

10
Persistent surveillance

Contents

Spire / ICEYE – Arctic Maritime Awareness Platform

Fabien Guillaume[a*], **Quentin Gollier**[b**], **Christoffer Winquist**[c***], **Simon Andersson**[c]

[a]Spire Maritime, 33 Rue Sainte-Zithe L-2763, Luxembourg, [b]Spire Global Luxembourg, 33 Rue Sainte-Zithe L-2763, Luxembourg, [c]ICEYE Oy, Maarintie 6, FI-02150 Espoo, Finland

ABSTRACT

Spire Global and ICEYE are two key actors in the Maritime geospatial intelligence ecosystem. With more than 70 Low-Earth Orbit (LEO) satellites constantly collecting Automatic Identification System (AIS) information on vessels across the world's seaways and weather data enabling advanced local weather forecasting capabilities, Spire Global is providing complex awareness products across the maritime industry. ICEYE has recently launched several Synthetic Aperture Radar (SAR)-capable satellites, enabling an all-weather, 3 x 3-meter resolution radar solution already tuned to the needs of global maritime actors. We jointly believe there is a strong opportunity to pool these capabilities and expertise to generate a combined Arctic Maritime Awareness Platform (AMAP) that can gather, in a single integrated solution, AIS-based ship monitoring, SAR-based ship monitoring, sea-ice data, iceberg spotting, ice mapping and weather forecasts to create an advanced and wide-area situational awareness platform over the Arctic region. This solution would notably leverage Spire's very high satellite density over the northern latitudes, and the expertise of both actors in generating complex data visualization tools and interfaces.

Keywords: Arctic, Maritime, Awareness, Polar, Satellite, Defense, SAR, AIS

1. INTRODUCTION

Satellites have emerged from the Cold War as key purveyors of geospatial intelligence (GEOINT), combining the ability to generate highly precise data on a global scale, irrespective of national or regional boundaries (Airbus, 2019). However, it is only recently that commercial providers such as Spire Global, Planet Labs and ICEYE have emerged among larger satellite manufacturers and operators as credible and sophisticated actors on this stage, powered by rapidly decreasing manufacturing and launch costs. Due to their ability to generate high amounts of data at a fraction of previous costs, the high revisit rates enabled by their growing constellations, frequent launches, and quick technical iteration cycles, these new commercial GEOINT providers have started to deliver new value to military, security and intelligence organizations across the world (Erwin Sandra, 2018).

Spire Global currently provides key data and capabilities to the US Air Force and prime U.S. defense contractors; and provides cutting-edge geospatial data to global monitoring organizations (such as Global Fishing Watch) and intelligence agencies which leverage its ability to supply continuous data streams of vessel and aircraft locations, and weather data at all-altitudes. ICEYE is also active in this market, proving dedicated capabilities in monitoring vessels across key waterways through its all-weather SAR imaging (Amos Jonathan, 2019). As an example, ICEYE has demonstrated the ability to monitor dark tankers in the Persian Gulf trying to hide when exporting Iranian oil despite the US embargo.

*fabien.guillaume@spire.com, +3522855031, *quentin.gollier@spire.com, +33689558305, ***chris.winquist@iceye.fi, +358503878670
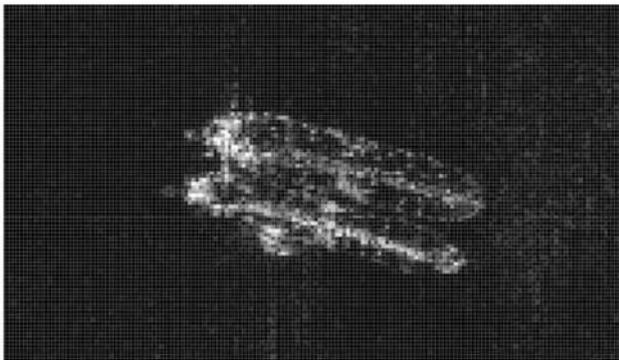
OIL TRANSSHIPMENT, ICEYE-X2

ICEYE

Oil flattens the sea surface, reflecting radar energy away from the satellite. Oil therefore appears dark

Figure 1. Radar imagery of possible oil transhipment; and radar intelligence on a resulting oil spill from a Saudi tanker

Since their inception, the two companies have had a key focus on the maritime sector where a combination of security, logistics and environmental challenges have made the provision of global datasets uniquely valuable for a variety of actors across the value chain (ESRIN, 2011). This has resulted in a joint expertise with uniquely complementary datasets able to provide a deep, dense and continuously refreshed information for governmental and defense actors, covering regions where terrestrial sources are inadequate or unavailable.

2. ARCTIC CHALLENGES

The Arctic has recently come to the forefront of the global security landscape. As a vast and largely inhospitable region, the North Pole has until now remained on the sidelines. A combination of oil discoveries, environmental changes and shifting geopolitical interests have contributed to its rise as a strategic domain for multiple actors (including Russia and, more recently, China) (American Security Project, 2013) (European Parliament, 2018).

Despite important transformations brought about by climate change, the Arctic remains a challenging area for traditional intelligence and signal collection, in part due to its vastness and its extreme climate, as well as the changing patterns of ice formation and movements, which heavily constrain the ability of military and security organizations to monitor this space. Furthermore, the rise in importance of the Arctic, largely due to the discovery of new natural resources exploitable by the creation of new waterways and the melting of the ice cap, has led to an increase in the traffic of military and commercial vessels operating in the area. These combined trends put pressure on the organizations involved in the control and protection of the northern latitudes, as the increased need for monitoring and surveillance can hardly be bridged with terrestrial data collection methods. The Arctic is the ideal playground for satellite-based monitoring systems.

As an illustration, the U.S. Coast Guard recently launched its first satellites as part of the 'Polar Scout' mission in collaboration with the Department of Homeland Security (Schreiber, 2018). Also this year, Canada awarded two contracts for three satellites aiming to "*leverage innovative science & technology expertise* […] *to identify, assess and validate technologies in support of air and maritime surveillance, particularly in the North*" (Boucher, 2019). These recent programs have highlighted the exceptional challenge of collecting ground-based data in the region as one of the key drivers for their mission.

3. THE ARCTIC MARITIME AWARENESS PLATFORM

What truly enables a wide-ranging Arctic Maritime Awareness Platform (AMAP) is the fusion of complementary data from new global constellations of AIS and SAR satellites that have been coming online in the last five years, and are growing rapidly. The very-high density of sensors and frequent revisit rates open up the possibility for affordable, near-real-time monitoring.

The AMAP concept presented herein, would draw on the existing capabilities of both actors (Spire and ICEYE) to create an integrated, fast-refreshing solution that would assist NATO and its Member States to monitor the vast Arctic space. It would notably focus on deploying the following capabilities:

- AIS monitoring
- SAR observation of sea ice
- SAR observation of vessels
- SAR observation of facilities and vehicles
- Weather monitoring and forecast
- ADS-B monitoring

These data sets would be integrated into a single software platform available through an API or a dedicated UX accessible to authorized platforms and users.
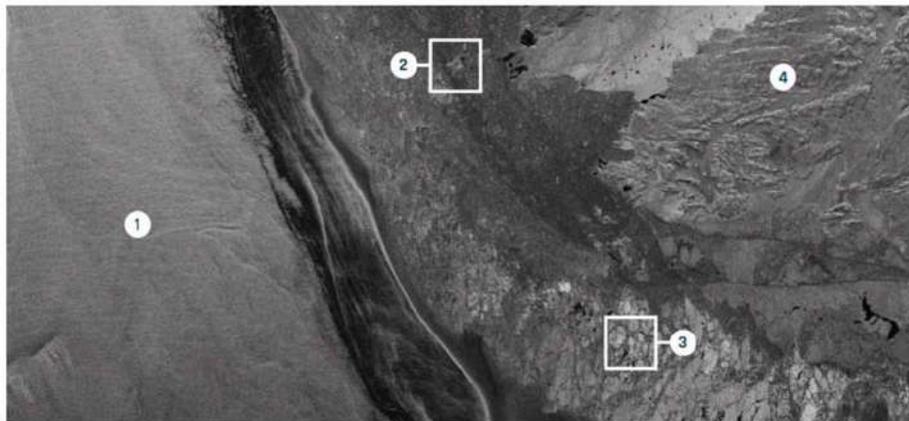
Figure 2. ICEYE Example imagery of the Kara Strait within the Arctic Circle, with 1- Open sea with a variety of sea states, 2- a small rocky island, 3- ice on top of water, and 4- Vaygach island

ICEYE currently provides dark vessel detection based on convolutional neural network machine learning in combination with AIS data to find vessels that disengage their AIS transponder. ICEYE has also experience in providing operational sea ice monitoring services. Spire's existing algorithms on vessel location and predicted direction could easily be complemented with the suite of analytics developed by ICEYE, in addition to raw imaging data. The combined solution would therefore provide a wide envelope of tailored services plugging the gaps remaining in the overall situational awareness of the Arctic region. It would enable decision makers to task NATO-specific assets in a directed and efficient manner.

Spire already commercializes its "Data Fusion" product to maritime clients, leveraging software platforms currently developed for large U.S. security-focused organizations (public and private). This package is notably valuable for fleet operational efficiency and trade monitoring, as this customer segment relies on highly accurate data at a global scale. We believe that this solution would be of high interest for defense organizations with a remit to monitor complex and vast areas where our satellites have unique coverage and revisit rate advantages.
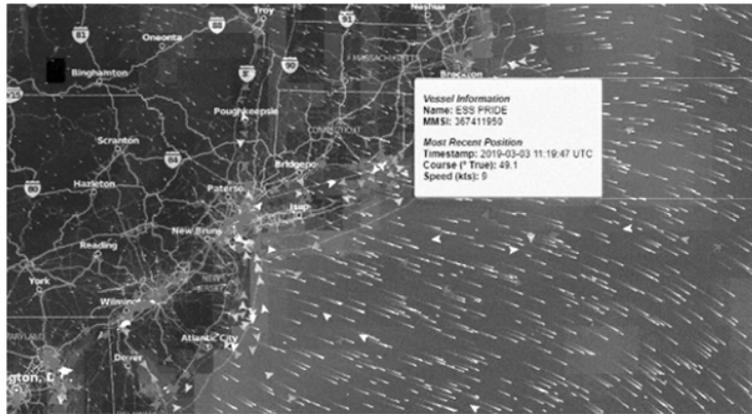

Figure 3. Spire's data fusion software, integrating AIS vessel location with proprietary weather forecast

This proposal, if selected and completed to the satisfaction of interested parties, could then be extended to additional regions (e.g. Strait of Hormuz, Taiwan Strait, etc.) where fast-refresh, actionable information is integral to a mission-critical concept of operations. Historical datasets could also be of assistance in generating models of typical ice patterns or patrol routes of non-NATO military vessels, bringing additional information to the table.

Figure 4. Visualization of Spire's maritime data set over the Persian Gulf

In addition to Spire and ICEYE's hardware and software capabilities, the AMAP would leverage the high density of Spire's arctic coverage, with a large number of satellites (20+) currently in a polar orbit. This would enable the product to sport a very high revisit rate for some data sets, thus ensuring that the information provided remains as up-to-date and accurate as possible.

Figure 5. Visualization of Spire's current orbit density above the Arctic

4. ADDITIONAL ELEMENTS

Spire/ICEYE data fusion for AMAP can also be enhanced with the monitoring and prediction of the evolution of the ice and weather in the arctic region to ensure safe maritime operations. ICEYE SAR data can be fused with Spire AIS and weather prediction models to create an automated alerting system.
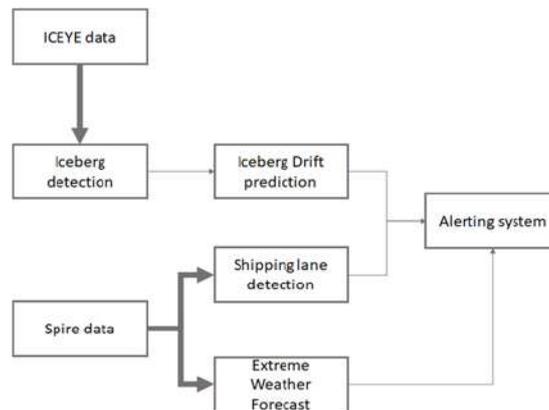
Figure 6. Proposed workflow for the Spire/ICEYE Arctic data fusion platform

SAR data can be used to implement an object-based method for automatic iceberg detection (A.K. Mazur, A.K. Wåhlin, A. Krężel. (2016, November 25). This technique combines edge-detection, segmentation, feature extraction and classification algorithms to efficiently extract objects of relevance from SAR data. An up-to-date iceberg database can be implemented and integrated with other data streams, such as weather forecasts, AIS-derived shipping lane information, etc.
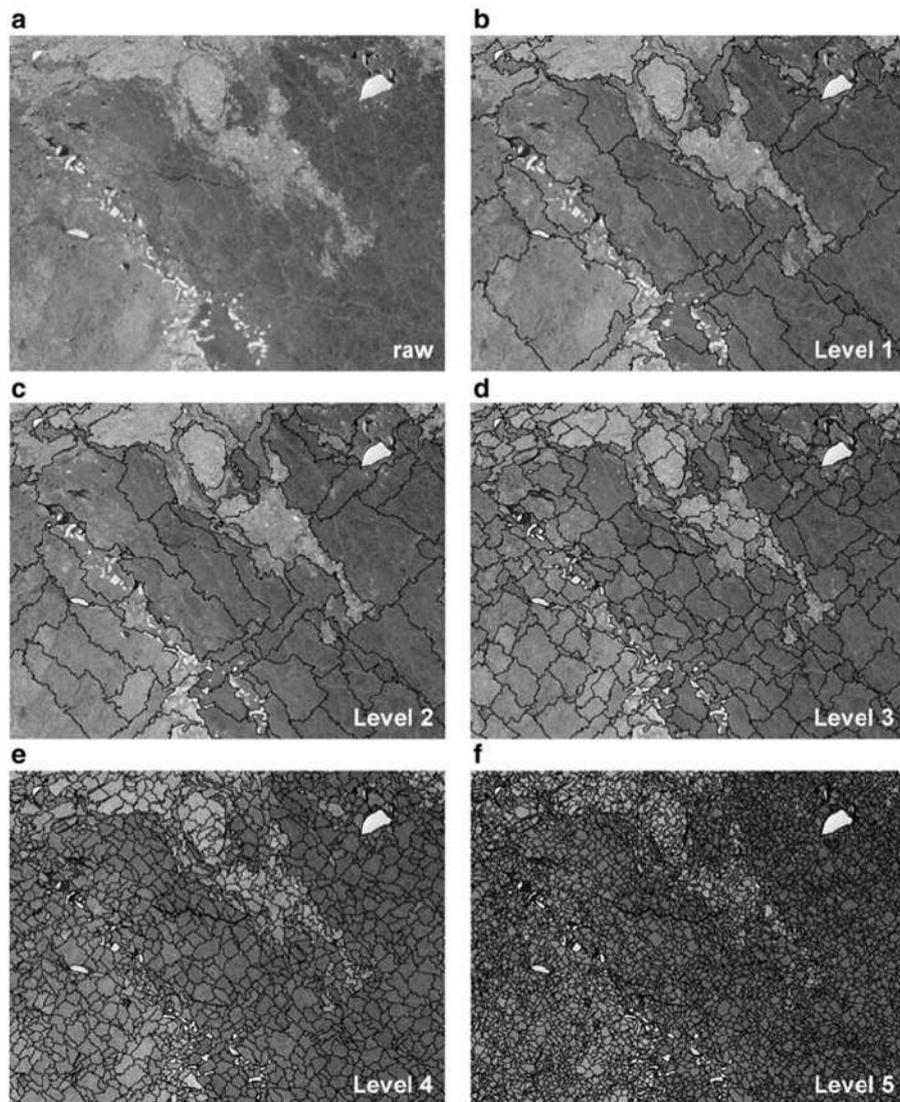
Figure 7. Iceberg detection on LANDSAT ETM+ images in the Amundsen sea
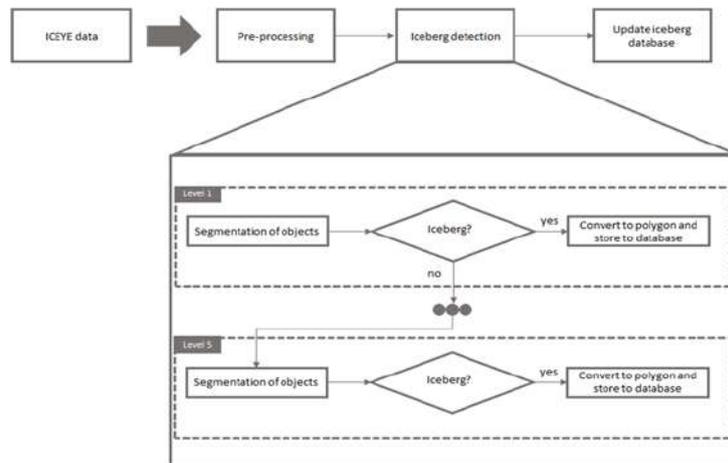
Figure 8. Proposed workflow for the Spire/ICEYE Iceberg Database

Building upon the Iceberg database, a drifting model can be used to predict the future trajectory of icebergs. The drift model can be based on (T. J W. Wagner, I. Eisenman, Rebecca W. Dell. (2017, April)). It implements a Lagrangian iceberg model that requires 3 inputs: sea surface current and wind, and sea surface temperature.
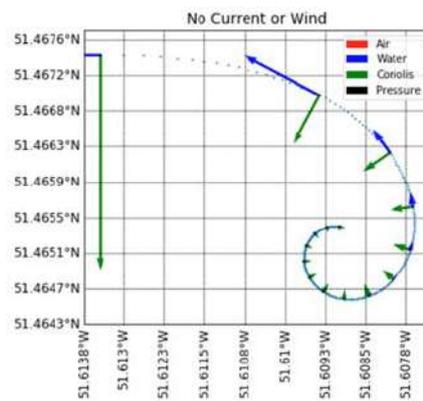
Figure 9. Example of a drift without current or wind.

The predicted trajectory can serve as input to an alerting system to estimate if and when an iceberg will intersect a vessel's trajectory. Real-time shipping lane monitoring can be achieved using the Traffic Route Extraction and Anomaly Detection (TREAD) algorithm (Giuliana Pallotta, Michele Vespe and Karna Bryan. (2013, June 4)). TREAD identifies vessel objects from raw AIS data streams for a given Region of Interest. After labelling the entry, exit and stationary points in the feed, routes are identified and the maritime traffic network is recreated to predict vessel trajectories.
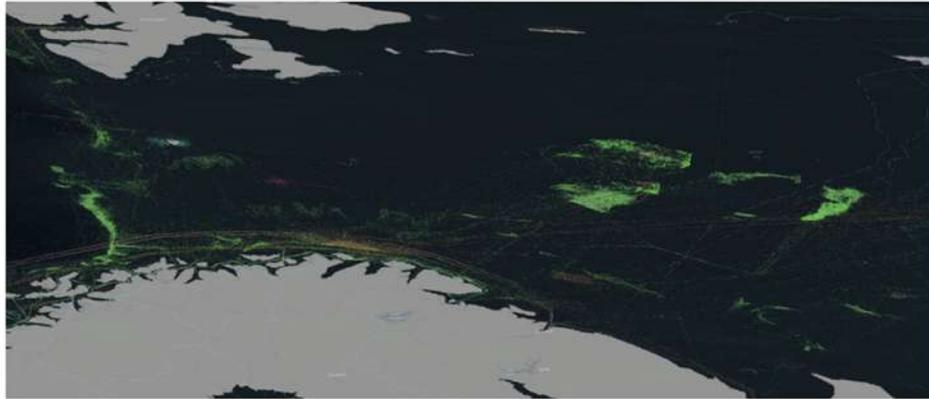
Figure10. Spire vessel density map of the Barren sea

5. CONCLUSION

The Arctic Maritime Awareness Platform (AMAP) proposed in this paper could offer NATO and other participating organizations a decisive advantage in the monitoring and management of the Arctic region, using a set of high-density, fast-refreshing SAR and RF space-based sensors already deployed and in operation. Both ICEYE and Spire bring decisive, actionable data that can complement and support existing NATO assets for a more efficient decision making process. The wide depth of mutually interacting and mutually reinforcing data sets combined in the AMAP could be a valuable asset to NATO in the strategic vantage point of the Arctic.

REFERENCES

Airbus. (2019). Satellite Imagery Intelligence for Defence Operations. Retrieved from Intelligence Airbus Defence and Space: https://www.intelligence-airbusds.com/en/8309-satellite-imagery-intelligence-for-defence-operations

American Security Project. (2013). PERSPECTIVE: The Arctic – Five Critical Security Challenges. 19: September.

Amos Jonathan. (2019, May 24). 'Sabotaged' tanker in Gulf of Oman leaked oil. Retrieved from BBC: https://www.bbc.com/news/science-environment-48387033

Boucher, M. (2019, February 6). Canadian Arctic Security Increasing with Acquisition of Three Satellites and Polar Radars. SpaceQ, p. 1.

Erwin Sandra. (2018, April 23). To keep foothold in government market, satellite imaging companies have to gain trust. Retrieved from SpaceNews: https://spacenews.com/to-keep-foothold-in-government-market-satellite-imaging-companies-have-to-gain-trust/

ESRIN. (2011, July 18). Pooled Satellite Data for Maritime Surveillance on the Horizon. Retrieved from European Space Agency: https://www.esa.int/About_Us/ESRIN/Pooled_satellite_data_for_maritime_surveillance_on_the_horizon

European Parliament. (2018). China's Arctic Policy. Brussels: European Parliament.

Schreiber, M. (2018, December 10). U.S. Coast Guard launches its first satellites, with an eye on Arctic security. ArcticToday, p. 1.

A.K. Mazur, A.K. Wåhlin, A. Krężel. (2016, November 25). An object-based SAR image iceberg detection algorithm applied to the Amundsen Sea

T. J W. Wagner, I. Eisenman, Rebecca W. Dell. (2017, April). An analytical model of iceberg drift

Giuliana Pallotta, Michele Vespe and Karna Bryan. (2013, June 4). Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anomaly Detection and Route Prediction

BalSAR: A stratospheric balloon-borne SAR system and its use for maritime surveillance

Marco Martorella[a,b] and Elias Aboutanios[c]

[a]Dept. of Information Engineering, University of Pisa, Via Caruso 16, 56122 Pisa, Italy
[b]Radar and Surveillance Systems National Lab., CNIT, Galleria Gerace 18, 56124 Pisa, Italy
[c]School of Electrical Engineering and Telecommunications, UNSW Sydney, Sydney, 2052, Australia

ABSTRACT

Surveillance systems are continuously employed for both military and civilian applications, including homeland security and border protection, which are two main concerns to NATO and in particular to the Science for Peace and Security (SPS) programme. Several platforms and systems, developed in past years, have turned into surveillance systems that are currently used in such scenarios. This paper describes a high-altitude balloon-borne synthetic aperture radar (BALSAR) system, which is currently under development as part of a NATO funded project within the SPS programme. Such a system will be able to performance surveillance tasks by acquiring radar data, forming SAR images and using them in order to extract valuable information.

Keywords: High-altitude platform, stratospheric balloon, airborne surveillance, synthetic aperture radar

1. INTRODUCTION

Military and civilian information gathering is an essential part of maintaining security and significant effort and money are spent on systems to enable these functions. Current technologies, which mainly employ satellites, aircraft (both manned and unmanned - UAVs), and drones suffer from a number of shortcomings. Space-borne systems operate from a large distance and, provided careful constellation design, are able to cover almost all areas on the surface of the earth. However, space missions are expensive and their use depends on the presence of a satellite over the designated area which can only happen at particular times that are determined by the orbit. Also, they do not offer a rapid and timely response as the ability to repeat measurements over a given area is constrained by the satellite orbit, which may impose intervals of several hours or even days between revisits. Airborne systems operate at much lower altitudes and piloted aircraft missions put the lives of pilots at risk as they are inherently vulnerable to attacks. Their size, and therefore radar cross section, and flight altitude limits make them easier to detect and shoot. Moreover, the cost of such aircraft is quite high justifying the launch of an intelligent missile against them. While UAVs take the pilot out of the equation, they still suffer from all of the other problems associated with piloted aircraft including cost and vulnerability. In fact, all of these systems are not expendable and must be protected.

High altitude platforms (HAPs)[1] have the potential to complement the above two systems and address many of their shortcomings. HAPs have been proposed in military applications for the gathering of surveillance data and are poised to play a key role in the area of national security. The USA, China, Japan and European Union have had military HAPs projects. These platforms enjoy a number of unique advantages with respect to both aircraft and satellites. HAPs operate at altitudes exceeding 20km, and include certain aircraft, airships and balloons.[1] Their high altitude gives them a higher degree of immunity against attack as compared to aircraft while providing them a wider field of view.[2] On the other hand, they are a cheaper alternative to traditional satellite systems as their development and deployment costs are much lower than those of spacecraft. Their comparatively low altitude, with respect to spacecraft, makes them more versatile and recoverable meaning that

---

Further author information: (Send correspondence to Marco Martorella)
Marco Martorella: E-mail: marco.martorella@iet.unipi.it, Telephone: +39 050 2217 624
Elias Aboutanios: E-mail: elias@ieee.org, Telephone: +61 (0) 2 9385 5010

they can be maintained and even upgraded. Also for remote sensing applications, they do not suffer from long revisit times that are a drawback of satellite systems.

While some systems employing HAPs exist[1–3] or have been proposed, most tend to be either airships or UAVs, and are large and very expensive. The Zephyr[4] is a solar powered UAV that is developed by Airbus. It is described as a High Altitude Pseudo-satellite (HAPS) as it is designed to hover for an extended period of time over a designated region. Among the intended uses of the Zephyr, Airbus lists maritime and border surveillance, environmental surveillance, missile detection, navigation, and continuous imagery. Lockheed Martins ISIS[1] and Raytheons Radar Aerostat[5] are two US projects developing high altitude airships. All three systems are manoeuvrable and are intended to provide long mission durations (on the order of days, months or even years at a time). Consequently, they are quite expensive and require significant infrastructure for deployment. As an example, the UK Ministry of Defence was reported to have an order for two Zephyr worth USD18 million.[1]

Stratospheric balloons have been used for decades for scientific experiments and remote sensing and both NASA[6] and JAXA[7] conduct high altitude balloon missions. These free floating balloons are very large and carry payloads weighing many hundreds of kilograms or even a few tonnes. Stratospheric balloon constellations with trajectory control have also been proposed for communications,[8,9] and scientific missions.[10] Small (sounding) balloons are extremely cheap even in comparison to the large HAPs discussed above. These are mainly used for weather sensing,[11] education[12,13] and amateur activities.

We propose a new high-altitude (in excess of 20km) balloon-borne Synthetic Aperture Radar (SAR) system that overcomes the shortcomings of existing systems without compromising performance, such as resolution, signal-to-noise ratio and hence target detection and recognition. The new system has a number of significant advantages over the aforementioned solutions. It is very low cost and hence expendable, rapidly deployable, has low probability of intercept (LPI) characteristics and is practically immune to attacks, requiring a very expensive guided missile to intercept it. There are a number of scenarios where airborne/space-borne radar surveillance is needed. Such scenarios include border protection, battlefield surveillance and, in more general terms, critical area surveillance where ground systems cannot be deployed either because the area is far from any logistic support or because of the unacceptably high risks that would be involved in deploying the system. In this paper we describe the proposed architecture and detail various subsystems.

2. PROPOSED SOLUTION

The proposed solution employs a combination of a very light SAR system and a high altitude balloon platform to provide a balloon-borne SAR (BalSAR) system. The BalSAR system would provide the means to fly at a very high altitude, 20km and above, and therefore operate at a safe distance from the surveillance area. In addition to the BALSAR system, the project will also produce the support systems including the flight prediction and mission planning software, launch (balloon inflation and release) system, telemetry tracking and command (TTC) system, and HAP recovery system. A high altitude balloon borne SAR has in fact been recently reported in Ref. 14, which demonstrates the feasibility of the concept. Our system, however, aims to achieve a number of specific aims including low cost, rapid deployment and standardization of the bus.

The BALSAR system is shown in Fig. 1. It comprises a balloon capable of lifting a High-altitude Platform (HAP) weighing around 20kg, a flight termination system, and a parachute to facilitate the HAP recovery. Capitalizing on the experience in building the UNSW cubesat UNSW-EC0,[15,16] which formed part of the European QB50 constellation,[17] the HAP itself is modeled on the cubesat standard.[18,19] Platform standardization will enable the use of commercial-off-the-shelf (COTS) components, directly leading to greater flexibility, as well as lower costs and reduced risk. Additionally, as COTS components improve being driven by market competition and pressures, improvements in the HAP will result. Finally a standard bus will greatly facilitate the opportunities of payloads other than the SAR system to be carried by the HAP. This enhances the utility of the proposed HAP system. Therefore, the standardisation of the HAP is one of the core goals of this project.

The project involves significant challenges both in the SAR and HAP systems. In order to deliver a low-cost and rapidly deployable system, the overall size and complexity must be kept low. The proposed BALSAR system rides on the stratospheric winds and does not include any trajectory control. Therefore, the mission goals are ensured through an innovative architecture that includes the subsystems and support systems described in what follows.

Figure 1. Illustration of the BALSAR system. Note that the radome is not drawn in order to show the antennas.

3. THE FLIGHT SUBSYSTEM

The flight subsystem consists of the balloon, termination mechanism and parachute. The balloon is helium filled and is rated to carry a payload weighing 20kg up to an altitude of 30km. As the balloon's trajectory is not controlled, a flight termination subsystem is included in order to terminate the mission in case it diverges significantly from the set flight plan. The termination subsystem is positioned between the parachute and balloon and receives the termination commands from ground control. Flight termination is achieved by burning a nichrome wire to separate the parachute and HAP from the balloon. The termination subsystem can also be used if a desired landing site is specified. The mission is then terminated at the point that results in the HAP landing at the designated site. Also note that automatic flight monitoring may be programmed into the on-board computer (OBC) which can then command the termination mechanism to activate if a pre-specified mission envelope is exceeded.

The parachute ensures that the HAP descends at a rate that prevents damage if it is to be recovered. Thus the parachute is located below the termination device and is sized to provide the desired descent rate, which is typically between 5 and 10 m/s. During the descent phase of the flight, the parachute automatically opens when the atmospheric density becomes high enough (typically at altitudes higher than 15km).

4. THE HAP SUBSYSTEM

The proposed HAP architecture is shown in Fig. 2. As alluded to earlier, it is modelled on a spacecraft bus, and in particular on a cubesat design. Therefore, the HAP comprises an OBC, an electrical power subsystem (EPS) and associated batteries, a communications module (comms), inertial measurements units (IMUs), and various positioning units to provide tracking of the platform. The HAP subsystem also houses the payload, that is the SAR subsystem. Micro-controller Units (MCUs) are employed as intermediaries between the OBC and a number of other subsystems in order to perform specific tasks associated with these subsystems. This philosophy ensures that critical tasks are handled by their decidated MCUs which ensure the reliable operation of the HAP. The main subsystems of the proposed HAP architecture are described below:

- OBC: the OBC manages the flight, executes the schedule, and monitors the various subsystems. Upon powering the bus, the OBC will first initialise the other subsystems and load the flight parameters to the SAR and stabilization microcontroller units (SAR-MCU and S-MCU respectively). The OBC will then periodically query the other subsystems to check the health of the HAP. Deviations from the nominal flight envelope will lead to termination of the flight. Finally, if requested by ground control, the OBC will gather relevant information on the HAP subsystems and communicate it to the ground via the comms subsystem.

- EPS: the HAP carries two electrical power subsystems, one for the bus (EPS-BUS in Fig. 2) and another for the SAR payload (EPS-SAR in Fig. 2). Each EPS includes its own battery pack and is fully controllable allowing various rails to be switched on and off. The use of a separate EPS for the payload is dictated by the power requirement of the radar.

- COMMS: the COMMS module operates at UHF in the amateur band. It provides communications with ground control in order to monitor the HAP health and track it. It also permits critical commands, such as flight termination to be uplinked to the HAP. Note that the SAR data is stored onboard and is not downlinked to the ground during the flight.

- SAR-MCU: This microcontroller interfaces the SAR payload to the HAP bus and performs two primary functions: firstly, it controls the operation of the radar by first turning it on at the right point of the flight, instructing it to start the acquisition and then stopping and turning it off. Secondly, it stores the position tags for the radar snapshots. To this end, the SAR-MCU receives an interrupt from the radar every time the latter acquires a snapshot and then fetches the position data, tags it and stores it in the positioning file.

- PNT-MCU: the position, navigation and timing MCU logs the data from the PNT subsystem and then services requests for tracking data from the OBC, SAR-MCU, and ACS-MCU (which drives the antenna stablization system). This configuration allows the PNT-MCU to sample the high precision navigation unit at the maximum rate and then accommodate the different rates at which the various requests are made by each subsystem. The PNT system includes a GPS receiver as well as IMUs (inertial measurements units) comprising accelerometers and gyroscopes.

- ACS-MCU: In order to minimize the weight of the HAP, only the antenna arrays of the radar are stablized. The attitude control and stabilization MCU provides the interface to the stabilization subsystem. The ACS-MCU is present with the required relative pointing direction of the antennas with respect to the direction of motion of the platform. During the acquisition phase, the ACS-MCU will then use the PNT information that it receives from the PNT-MCU to calculate the absolution pointing direction which it will then relay to the controller of the stabilization subsystem.

- SEP-MECH: the separation mechanisms are included primarily to ensure the safety of the flight and control the risk of the mission. The OBC will continually monitor the flight parameters and verify that they are within the acceptable mission envelope. Should the flight move outside this envelope, the separation mechanisms are activated to terminate the flight. Two independent mechanisms are included to provide redundancy. One of these mechanisms, SEP-MECH2, is an independent system that was mentioned in the previous section. The other, namely SEP-MECH1, is connected to the OBC and forms part of the HAP bus. In addition to the PNT system described above, multiple tracking subsystems, denoted as T1, T2 and T3 in Fig. 2, will be used ensure adequate tracking system redundancy at all stages of the flight.

- SEN-MCU: the sensor suite MCU provides the functionality for various HAP environmental sensors, such as temperature and pressure, and other flight information sensors, such as flight dynamics sensors, to be . These are logged by the SEN-MCU and stored on a dedicated SD card. The MCU can also be queried by the OBC if the relevant data is required.

In addition to the subsystems detailed above, the HAP provides environmental protection to the bus and payload. During the flight, the atmospheric temperature can drop as low as -70° and the pressure decreases to approximately 1% of its value at ground level. Foam insulation, combined with the heat generated by the various subsystems, will ensure that the internal temperature of the HAP remains above -20°. Furthermore, thermal and vacuum testing will be employed prior to the flight to verify the system performance under the expected environmental conditions.
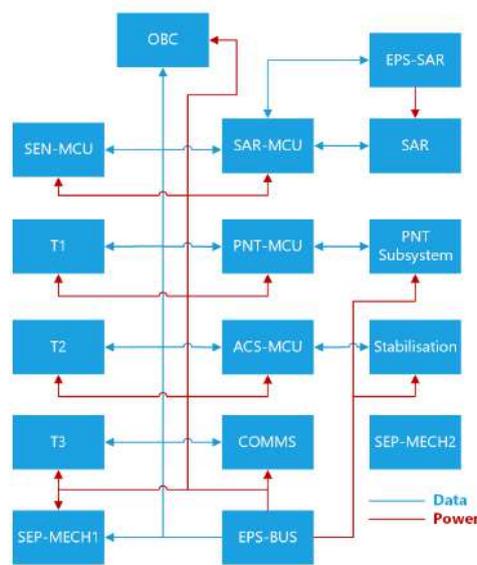
Figure 2. Architecture of the HAP bus showing the data and power interconnections.

5. THE SAR PAYLOAD

As the HAP is limited to around 20kg in total weight, the project requires a miniaturized SAR system. The SAR system is restricted to a total of 10kg, with the electronics weighing around 7kg and the antennas 3kg. Furthermore the SAR draws its power from the HAP, which then places a power consumption requirement on it. The limits on weight and power consumption are particular challenges for the SAR design.

5.1 Radar Front-End

The radar front-end is made up of a single board in standard ITX format ($17 \times 17cm$) that is stacked and interconnected according to the requirements to other boards (Embedded Digital processor and power supply subsystem) and an external power amplifier. The front end implements an X-band direct-conversion Linear-FMCW radar transceiver architecture. The transmitted waveform is generated by a PLL-based frequency synthesizer. This approach focuses on a compact, low cost and low power consumption solution that allows for the generation of large bandwidth and high chirp rate Linear-FMCW waveforms. The front-end mainly consists of the following sub-sections:

- Waveform generation and transmitter: this consists of a PLL based, X-band, programmable signal generator (phased detector, VCO and loop filter), a RF pre-amplifier, a splitter for the generation of the OL signal, a RF medium power amplifier (MPA) and a digital control interface. The generator is locked to the same low noise reference source that feeds the ADC stages, in order to perform coherent data processing.

- RF receiver: the receiver includes an input limiter, a RF band-pass filter, a low-noise amplifier and a quadrature demodulator. The demodulator is fed by OL signals generated by the aforementioned synthesizer.

- Base-band signal conditioning: down-converted signals are sent to a base-band signal conditioning stage that employs a programmable attenuator and an active band-pass filter. This allows to tune the overall gain and adapt the signal to the input dynamics of the AD converters.

- Power amplifier: the output of the transmitter section is sent to a solid state GaN based linear high power amplifier (HPA), with a maximum output power of 40 dBm.

Parameter	Value
Waveform Type	Linear-FMCW
Frequency range	9.3 to 9.9 GHz
Chirp rate	up to 1 THz/s
Output power	40 dBm (max)
TX attenuator	$0 - 30$ dB (1 dB step)
Minimum detectable signal	-130 dBm
Noise figure	6 dB
IF bandwidth	from 10 to 40 MHz
RX attenuators	$0 - 40$ dB (0.5 dB step)

Table 1. Radar front-end specifications
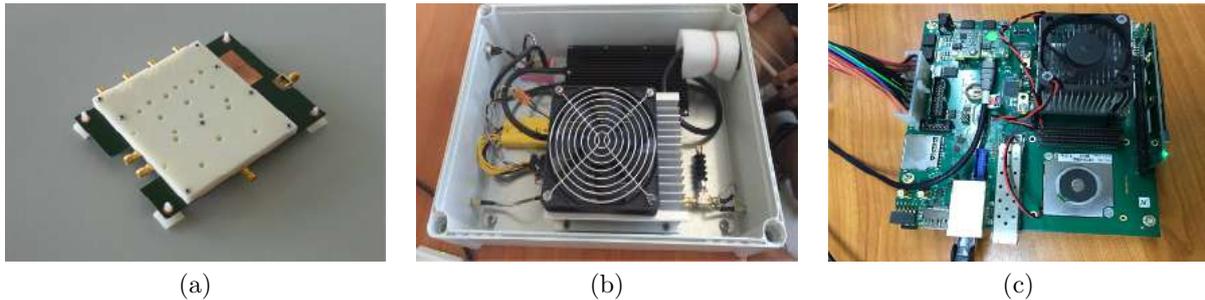
Figure 3. SAR Payload - (a) front-end, (b) power amplifier, (c) embedded digital radar processor.

- Power down-conversion subsystem: this is a multiple output down-conversion subsystem that employs mixed linear and switch-mode topologies in order to fulfill the power requirements of all the previous described sections. Such subsystem down-converts the voltage(s) provided from the HAP and converts them to the rails required by the various SAR subsystems.

The radar front-end specifications are summarized in Table 1, and picture of a test board of the radar front-end is shown Figure 3(a).

5.2 Power Amplifier

The power amplifier is a COTS device by Keylink Microwave. This model is a GaN based high power amplifier operating between 9.1 GHz and 10.1 GHz and offers a wide dynamic Range with 10 W of output power. It has long term reliability and high efficiency and it is ideal for X-Band linear applications. A picture of the same component used for another system is shown in Figure 3(b).

5.3 Embedded Digital Radar Processor

The Embedded Digital Radar Processor has been designed around the Trenz Electronic TEBF0808 carrier board which is a baseboard for the Xilinx Zynq Ultrascale+ MPSoC modules TE0808 and TE0803. A picture of the developed system is shown in Figure 3(c).

The main sub-systems are:

1. Carrier board

2. System on a Module

3. Acquisition board

4. Mass memory storage

6. MISSION SUPPORT SYSTEMS

In order to enable the mission execution, a number of supporting subsystems are being developed. These include the flight prediction and planning, inflation and release system, and telemetry tracking and command system. These are described in what follows.

6.1 Flight Prediction and Planning

Flight prediction and mission planning are important aspects of the system. Given a launch location, date and time, the flight prediction software uses weather data to determine the expected flight trajectory. This allows the ground track as well as the landing site to be determined. Mission planning, however, can require that a number of waypoints be observed by the balloon. Therefore, when provided with this waypoints, the mission optimization software uses the mission planning program to iteratively determine the balloon inflation and flight parameters in order to give the optimal trajectory that is as close to the desired waypoints as possible.

6.2 Inflation and Release System

The balloon will have a pre-launch diameter of approximately 10m and will therefore display a large area to any light breeze. Therefore, it needs to be anchored to the ground and and handled properly during inflation. This requires an inflation and release rig that is being developed as part of the project. The rig is modular to facilitate its transport to the launch site. It is also easy to set up and is able to provide measurement of the neck lift of the balloon to ensure the correct inflation is achieved.

6.3 Telemetry Tracking and Command System

Although the SAR data will be stored onboard the HAP and will not be downlinked to the ground during the flight, a is being developed to continually track the HAP and monitor its state. The TT&C system, which operates at UHF in the amateur band, comprises a mobile ground station that will enable simple commands to be uplinked to the HAP and health check data to be received from it. Additionally, the TT&C subsystem receives the position data from the HAP and updates the estimated flight path and landing position. This permits the mission to be monitored and decisions to be made on the termination of the flight if required.

7. CONCLUSIONS

This paper describes a novel balloon-borne synthetic aperture radar that is under development as part of a NATO funded project. The full system realisation is predicted to be completed by 2020 and results in terms of SAR imagery should appear soon after the system completion. The BALSAR system is intended to be low-cost and rapidly deployable in order to provide enhanced surveillance capability in hostile environments. Furthermore, the BALSAR system will find applications in remote sensing and monitoring applications, such as border protection and disaster monitoring.

REFERENCES

[1] d'Oliveira, F. A., Melo, F. C. A. L. A. d., and Devezas, T. C., "High-Altitude Platforms Present Situation and Technology Trends," *Journal of Aerospace Technology and Management* **8**, 249 – 262 (09 2016).

[2] Wang, W. and Shao, H., "High altitude platform multichannel sar for wide-area and staring imaging," *IEEE Aerospace and Electronic Systems Magazine* **29**, 12–17 (May 2014).

[3] Yang, H., Li, Z., Wu, J., Huang, Y., Yang, J., and Yang, X., "Near-space slow sar high-resolution and wide-swath imaging concepts," in [*2013 IEEE Radar Conference (RadarCon13)*], 1–5 (April 2013).

[4] "Airbus zephyr high altitude pseudo-satellite (HAPS)." https://www.airbus.com/defence/uav/zephyr.html. [Online], last accessed June 2019.

[5] "Raytheon aerostat." http://www.raytheon.com/capabilities/products/jlens/. [Online], last accessed June 2019.

[6] "NASAs Scientific Balloon Program." http://asd.gsfc.nasa.gov/balloon/. [Online], last accessed June 2019.

[7] "JAXA high altitude balloon projects." http://global.jaxa.jp/projects/sas/balloon/topics.html. [Online], last accessed June 2019.

[8] "Google loon project." https://loon.com/. [Online], last accessed June 2019.

[9] "Facebook high altitude platform project." https://code.fb.com/connectivity/high-altitude-connectivity-the-next-chapter/. [Online], last accessed June 2019.

[10] Nock, K., Heun, M., and Aaron, K., "Global stratospheric balloon constellations," *Advances in Space Research* **30**(5), 1233 – 1238 (2002).

[11] "NOAA NWS Radiosonde Observations." https://www.weather.gov/upperair/factsheet. [Online], last accessed June 2019.

[12] "Stratoflights classroom on the edge of space." https://www.stratoflights.com/en/education/info/. [Online], last accessed June 2019.

[13] "UNSW's BLUEsat High Altitude Balloon Team." https://bluesat.com.au/projects/high-altitude-balloons/. [Online], last accessed June 2019.

[14] Zaugg, E. C., Margulis, A., Bradley, J. P., Kozak, A. H., and Roehrich, W. K., "SAR imaging from stratospheric balloons: First results," in [*IEEE Radar Conference*], (2019).

[15] Osborne, B., Aboutanios, E., Dempster, A., Cetin, E., Heiser, G., and Glennon, E., "UNSW EC0 CubeSat Design: Experiments in Radiation Tolerance Critical Systems, GNSS Remote Observation and 3-D Printed Satellite Structures," in [*5th European Cubesat Symposium*], 41 (2013).

[16] Cheong, J. W., Southwell, B., Lam, C., Bultitude, J., Andrew, W., Green, S., Osborne, B., Dempster, A. G., Aboutanios, E., and Crowe, W., "Design and Development of the UNSW QB50 Cubesat - EC0," in [*International Astronautical Congress*], (2016).

[17] Muylaert, J., Reinhard, R., Asma, C., Buchlin, J., Rambaud, P., and Vetrano, M., "QB50: An international network of 50 cubesats for multi-point, in-situ measurements in the lower thermosphere and for re-entry research," in [*ESA Atmospheric Science Conference, Barcelona, Spain*], 7–11 (2009).

[18] Puig-Suari, J., Turner, C., and Ahlgren, W., "Development of the standard cubesat deployer and a cubesat class picosatellite," in [*2001 IEEE Aerospace Conference Proceedings (Cat. No.01TH8542)*], **1**, 1/347–1/353 vol.1 (March 2001).

[19] Spangelo, S. C., Kaslow, D., Delp, C., Cole, B., Anderson, L., Fosse, E., Gilbert, B. S., Hartman, L., Kahn, T., and Cutler, J., "Applying model based systems engineering (mbse) to a standard cubesat," in [*2012 IEEE Aerospace Conference*], 1–20 (March 2012).

Persistent Maritime Surveillance Against Underwater Contacts using a Wave Gliders: Fleet Composition and Effectiveness

Ronald Kessel*[a], Craig Hamm[b], Martin Taillefer[c]

[a]Seamount Analytics, 1501-445 Laurier Ave W., Ottawa, Ont. Canada; [b]Sigma Ocean Research, 9 Ancroft Court, Ottawa, Ont., Canada; [c]Maritime Way Scientific, 5-1420 Youville Drive, Ottawa, Ont. Canada.

ABSTRACT

Autonomous platforms hold obvious promise for maritime operations. That promise faces reality when costs, overall effectiveness, fleet composition, infrastructure, and METOC limitations are assessed. Operational dimensions can be explored using multi-disciplinary modeling and simulation (M&S) combining platform physics (power, propulsion, sensors, communications, sea keeping) with oceanographic influences (wind, waves, currents, ice, solar energy, sound propagation and ambient noise) and economics (costs, benefits, business model)—all in light of top-down objectives. The resulting virtual demonstration can be critiqued and improved by application-domain experts. Successful demonstrations can garner the consensus that large capital projects require for advancing capability. The Innovation for *Defence Excellence and Security (IDEaS)* programme of the Canadian Dept. of National Defence (CAN DND) has therefore funded a project on wave gliders equipped with thin-line sonar arrays for autonomous persistent (multi-month) maritime surveillance against underwater contacts. This paper reviews selected key results from that analysis.

Keywords: Maritime surveillance, autonomous systems, wave gliders, thin-line passive sonar arrays, operational analysis

1. INTRODUCTION

Wave gliders are unmanned surface vessels (USVs), nominally 3 m in length, that take propulsion energy from ocean waves and battery recharge from the sun (see Fig. (1)). They travel at 1.5 to 2.25 knots in sea state 3 and higher [1, 2], they can tow payloads like sonar arrays [3], and they have been proven in multi-month deployments, mainly for oceanographic purposes thus far [4]. **Thin-line passive sonar arrays** are a new generation of sonar technology that are light-weight and low-power which can be towed by wave gliders. They have progressed alongside signal analysis and **auto-detection algorithms** for detecting underwater contacts.
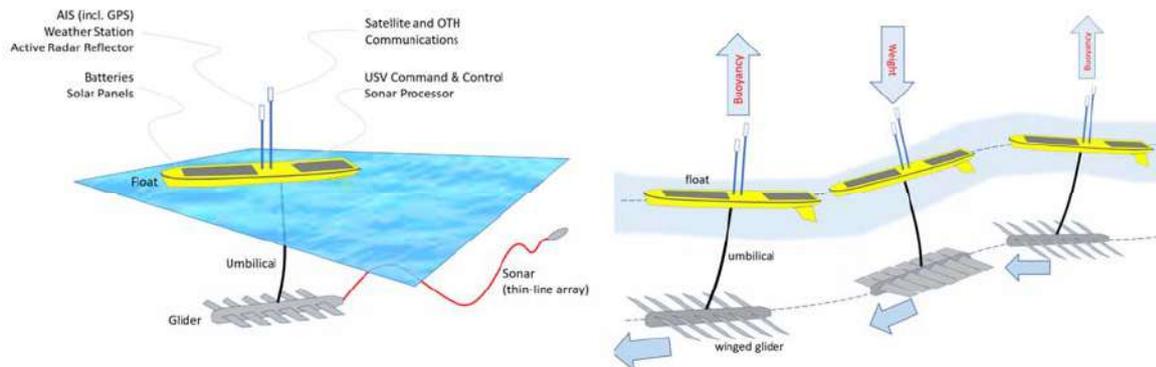

Figure 1 Schematic of a sonar-equipped wave glider (left) and its propulsion by wave action (right). Graphics based loosely on wave gliders by Liquid Robotics (USA).

The use of sonar-equipped wave gliders has often featured as a bullet item in marketing brochures and slides, and their use has been demonstrated to a limited extent at sea using a single platform [5, 6], apparently unchallenged by

underwater contacts. But a larger picture of capability has so far never been addressed. Selected parts of that picture are presented here, including: (1) *Requirement Reconciliation*—the match with requirements for autonomous capability given by the CAN DND IDEaS innovation program; (2) *Surveillance Coverage*—the relationship between deployed fleet size and coverage effectiveness amid seasonally changing propagation conditions (a key cost driver); and (3) *Exogenous Limits*—the impediments of harsh METOC conditions in selected Canadian waters. Space does not allow for the impact of current and waves on wave glider speed, the expected variability of sonar coverage, communications, ship carriage requirements, capability costing and business model, which are all part of the larger project. A capability demonstration video of wave glider fleets will be shown in the presentation.

2. REQUIREMENT RECONCILIATION

A brief summary of the requirements ("wish-list") for highly autonomous persistent maritime surveillance as given by the CAN DND IDEaS programme, and their answer by sonar-equipped wave gliders, are summarized in Table (1). The main discriminators between wave gliders and unmanned underwater vehicles (UUVs) and other unmanned surface vehicles (USVs) are summarized in Table (2). Much more could be said both pro and con [7], but the tables make a good case for taking wave gliders seriously for this surveillance capability.

Table 1 Requirement Reconciliation of sonar-equipped wave gliders for Persistent Maritime Surveillance against underwater contacts in Canadian maritime approaches.

Autonomy	**The capability must be highly autonomous with on-board automatic signal analysis**—Wave gliders self-navigate long distances between waypoints with a passive, rugged, low-maintenance persistent propulsion mechanism (waves) and payload battery recharge (solar).
Persistence	**Multi-month operation on station**—Wave gliders have proven capable of three-month continuous operation on station. Deployments are limited mainly by biologic fouling. Navigation and communications are maintained through to sea state 7, with proven survival through much higher sea states (hurricane).
Rapidly Deployable	**Must be deployable from military ships without ship engineering modifications, and deployable from shore, and operate without cable or tether**—The small size and weight of wave gliders and thin-line arrays make deployment and recovery from military ships possible with a small crane. Autonomous long-distance transits to and from station (without ship transport) do not impair time on station.
Covert	**Sensors must be passive for covertness**—Thin-line sonar arrays are passive and covert. The wave glider underwater profile is very low. They are deemed to be generally undetectable by unalerted underwater contacts.
Effective	**Must have high probability of detection with acceptable rate of false alarm**—As with passive sonar generally, **e**ffectiveness depends on ambient noise levels (shipping and wind-generated), local sound-speed profile, length of the sonar array, sonar depth in the water column, and the noise radiated by the underwater contact. Effectiveness is treated below.
Real Time Communications	**Must include high-speed communications for _continuous near real time_ monitoring and tasking**—Wave gliders can carry near real time satellite communications with less than 2-minute latency and permitting two-way communications (Iridium and RUDEX), enabling remote status and position updates, re-tasking, and transfer of underwater signal information for analysis at headquarters.
Scalable	**Coverage area and duration must increase uniformly without change of hardware**—Sonar-equipped wave gliders can be added to increase the extent and effectiveness of coverage.
Affordable	**Costs and value must be objectively substantiated**—Costs, including infrastructure for 24/7 C2 monitoring are treated in a forthcoming full report.

Table 2 Selected technology-class discriminators between wave gliders, other unmanned surface vehicles (USVs), and unmanned underwater vehicles (UUVs)

Real-Time Comms	Wave gliders support continuous near real time communications through satellite or radio. UUV communications require either surfacing or underwater communication systems and nodes.
Power	The unlimited propulsion and electrical power of wave gliders from waves and sun, though at low levels, provide robust simplicity (few parts), long untended endurance, relentless persistence, and low liability of fuel spills, fire, collision—all of which are key autonomy enablers. UUVs face the overhead of docking stations and behaviours, or relatively frequent recovery for recharge.
Transit to	Wave gliders can make long autonomous transits to/from station with no operational penalty apart from transit

Station	time. UUVs (prop-propelled) generally avoid long transits to station owing to power constraints.
Sea Keeping	UUVs naturally avoid extreme sea states (except during recovery, deployment, and surface comms), where USVs must face or sail to avoid them. Adverse sea conditions are therefore considered below.
Sonar Depth	UUVs can easily change the sonar operating depth to capitalize on propagation conditions, where arrays towed by USVs are generally limited to shallower depths (10 to 100 m).

3. SURVEILLANCE COVERAGE

A key cost driver of surveillance is the number of platforms required for effectiveness. The wave glider coverage will follow that of passive sonar coverage generally, which determines the spacing and number of wave gliders required for each deployment. The results generalize to other towing platforms, autonomous or manned.

Sonar modeling was carried out using three-years of monthly propagation conditions at locations within each operating area in Fig. (2), all in waters of roughly 200 m deep. The propagation model was a ray model (*Bellhop* extended to NX2D). Wave-gliders were arranged in a picket-fence surveillance pattern of variable spacings S between 1 and 30 km as in Fig. (3); sonar depths varied between 10 and 100 m; sonar capability was varied through a set of directivity indices 6 to 18 dB; and targets varied between depths 20 to 200 m, with 3 radiated tones varied between 100 and 1000 Hz, with strengths corresponding to *slow-quiet* targets (90 to 100 dB SPL re 1 m narrowband tones for target speed 2 to 6 knots) to *loud-fast* targets (110 to 120 dB SPL re 1 m tones for target speeds 14 to 20 knots).

The sonar integration time was 10 seconds, the probability of false alarm was 0.0001, and the receiver operating characteristic (ROC) curves were Swerling II for narrowband processing. The sonar array towed by each wave glider was horizontal (HLA), with beams ± 30° to broadside. The ambient noise level was the classic Wenz shipping and wind noise, where the shipping level was varied uniformly between light and heavy, and wind speeds were the 3-hourly now-casts from historic wind data over 3 years at locations in each operating site in Fig. (2).
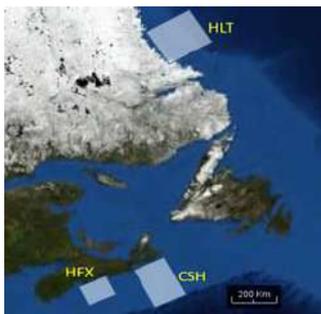

Figure 2 Operating areas on the east coast of Canada selected by the project team.
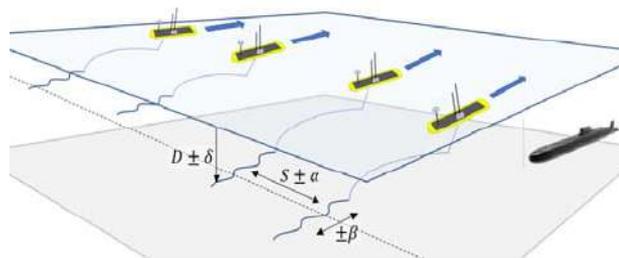
Figure 3 Picket-fence surveillance pattern assumed for persistent maritime surveillance. The separation distance is S in the figure. The modeling included positioning uncertainties that are beyond the scope of this paper.

Monte Carlo simulation was used to explore coverage versus separation distance using over 90000 randomized picket-fence/sonar/operating-area/target combinations. Coverage effectiveness was based on the *cumulative probability of detection CPD* for each target sailing through the picket fence, computed from the cumulative combined single-look probabilities of detection for all wave gliders in the picket fence for each target sail-through. The best sonar depth was found for each season in each operating area against slow-quiet targets and used in all ensuing coverage analyses. No optimization against target depth was used because the target depth was presumed unknown.

Surveillance *Effectiveness* was defined as the portion of simulated target encounters with *CPD* > 80 %, where *CPD* is the probability of detecting a target at some time during its transit through the picket-fence surveillance.

Effectiveness > 80 % was considered *highly effective* surveillance capability (i.e., over 80 % of target encounters across the full range of realistic operating conditions have an overall probability of detection of over 80 %).

Effectiveness > 50 % was considered *moderately effective* surveillance capability (i.e., over 50 % of target encounters across the full range of realistic operating conditions have an overall probability of detection of over 80 %).

Fig. (4) shows the surveillance *Effectiveness* as a function of wave-glider separation distance *S* for the entire the set of randomized target encounters, at all operating areas (sites), across all seasons, with seasonally optimized sonar depth, for 32-element sonar arrays with directivity index 15 dB. The shaded region of the graph shows the large effectiveness variation with site, season, and target. Large variation of effectiveness is typical for passive sonar across the set of realistic conditions. Each surveillance deployment must be tailored for the seasonal conditions, as one generally must for passive sonar, adjusting search operations for the conditions on hand.

Surveillance effectiveness is _high_ above the dashed line "HIGH", and moderate above the dashed line "MEDIUM" in Fig. (4).

The mean in Fig. (4) indicates that a 2000 m nominal wave glider spacing *S* spacing can assumed for high effectiveness. A 20 km picket-fence of high effectiveness therefore nominally requires a fleet of about 10 wave gliders. Considering the shaded region in Fig. (4), however, one expects that the same fleet will provide a high effectiveness in some operating areas, and for some seasons, with spacing over 10 km, across a picket line of length of over 100 km.

The mean in Fig. (4) similarly indicates that a 25 km nominal wave glider spacing *S* spacing can assumed for moderate effectiveness. A 20 km picket-fence of moderate effectiveness therefore nominally requires a fleet of about 1 wave glider. The variation in passive sonar coverage results in the large variation on the deployed fleet size and coverage performance. The variation is a property of long-range passive sonar across the full range of realistic operating conditions. It true of passive sonar generally and is not unique to surveillance using sonar-equipped wave gliders.
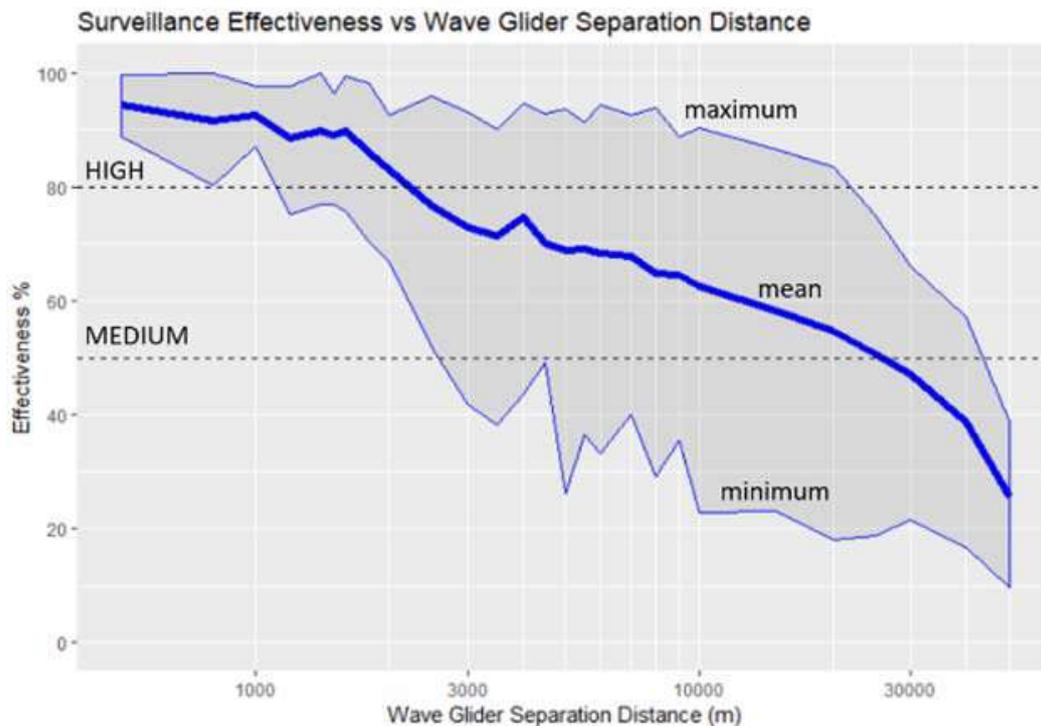

Figure 4 Surveillance effectiveness versus wave glider separation *S* across all simulations of operating area, season, and target saliency. The passive sonar properties assumed are described in the text.

The impact of changing conditions such as operating area, season, and target saliency can be explored using the Monte Carlo simulation. Table (3) summarizes some of the variation assessed using the mean of the surveillance effectiveness for subsets of the Monte Carlo simulations, grouped by operating area and target saliency. The variation with operating area and target saliency is large, as expected with passive sonar generally.

Table 3 The nominal (mean) wave glider separation distance by operating area and target saliency.

	Separation S (km)							
	FULL Target Set (0 to 120 dB)		QUIET Targets (90 to 100 dB)		MEDIUM Targets (100 to 110 dB)		LOUD Targets (110 to 120 dB)	
Operating Area	HIGH Effectiveness	MEDIUM Effectiveness	HIGH Effectiveness	MEDIUM Effectiveness	HIGH Effectiveness	MEDIUM Effectiveness	HIGH Effectiveness	MEDIUM Effectiveness
HFX Outer Halifax Harbour	15	40	5	12	20	40	42	> 50
CSH Coastal Shelf	2	20	1.5	4	2.6	30	9	50
HLT Labrador Sea	2	10	0.6	2.1	2.6	10	7	45

4. EXOGENOUS LIMITS

Wave gliders survive sea state 6 and higher, but the performance of passive sonar arrays is expected from experience to be limited to sea state 6 and less owing to noise and motion. As with small craft generally, sea states above 8 should be avoided using METOC forecasts. Table (4) summarizes the frequency of sea states greater than 6 and 8. Similar limitations are faced in sonar operations by surface platforms, manned and autonomous.

Table 4 Relative frequency in % of sea state greater than sea state 6 and 8 in 3-hour intervals over three years (2005-07) in the operating areas in Fig. (2).

	% Sea State > 6—Surveillance Suspended				% Sea State > 8—Area Avoidance			
Area	Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter
HFX	6.9	0.2	7.7	15.8	0.00	0.11	0.23	0.48
CSH	7.2	1.1	6.8	20.2	0.26	0.19	0.17	0.44
HLT	18.0	2.7	18.1	32.6	0.74	0.22	1.43	1.59

Battery recharge by solar panels can limit operations if a platform's power load exceeds recharge. Exceedance is more likely in winter and at higher latitudes. Fig. (5) plots the solar power captured based on actual historic now-cast solar data in the operating areas for one year (2014) assuming a 1 W average hotel and communication load, a 3.5 W sonar 32-element-array load, a 10 W processing load, and a 2 m^2 sonar panel with 80 % efficiency. Low battery power and possible recharge down time may occur when sonar recharge approaches or falls below that level (dashed line in Fig. (5)). The resulting recharge down time depends on battery size and the low-power strategy of the wave gliders (not included here).
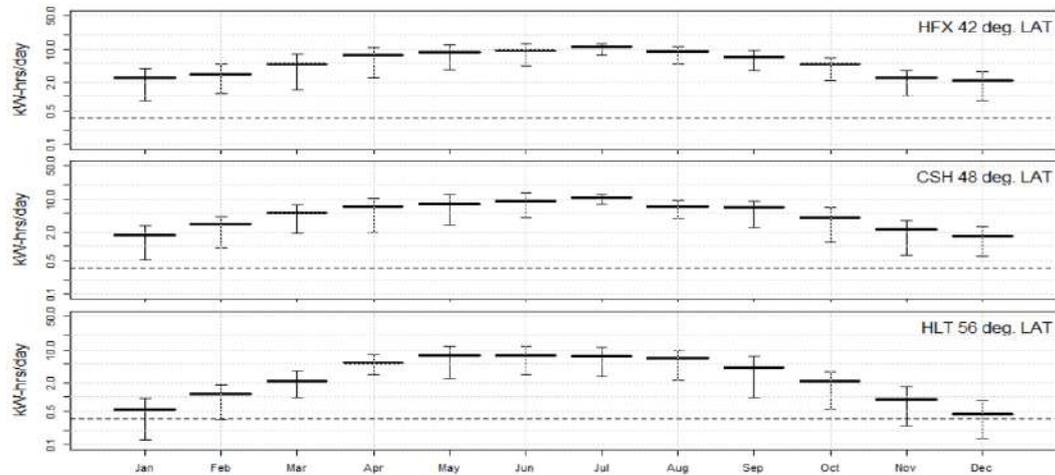

Figure 5 Mean daily solar power acquired by each wave glider. The 80 % confidence interval is indicated. The dashed line is the minimum required to maintain battery charge level.

5. CONCLUSIONS

Overall effectiveness requires a good match with operational requirements, a high probability of detection, and acceptable down-time due to exogenous conditions, which were briefly reviewed here for persistent maritime surveillance using sonar-equipped wave gliders.

It is believed that no platform lends itself so well to persistent surveillance with high autonomy as sonar-equipped wave gliders. None meets capability requirements so well—for multi-month persistence, ease of deployment and recovery from port or ship, for near real time communications (for monitoring and re-tasking), and for absence of anchor and cabling, to cite a few key requirements.

Most importantly for extended use in home waters, no other class of autonomous maritime platforms afford equally high autonomy and independence with such low liabilities of operation owing to the wave glider's combination of light weight, slow relentless speed, its few moving parts, and the absence of combustible, spillable fuels. Low liability is paramount for the distant operation of autonomous systems outside of combat.

The Technology Readiness Levels (TRLs) of the key technologies for persistent maritime surveillance were estimated to be TRL7-8 for wave gliders, TRL 6 for thin-line acoustic arrays, and TRL 5 for highly autonomous autodetection for passive sonar.

A plausible fleet size in Canadian waters is on the order of 10 wave gliders, giving *moderate* surveillance effectiveness for a picket-fence line on the order of 300 km under favorable sonar conditions, and 30 km under challenging conditions, while giving *high* surveillance effectiveness on the order of 200 km under favorable conditions, and 10 km under challenging. In practice, the wave glider spacing, overall picket-fence line of coverage, and deployed fleet size would be determined for a given operating area and season ahead of each deployment using modeling of the kind used in this study.

The operational footprint of the sonar-equipped wave glider fleet includes procurement costs, maintenance, deployment, recovery, intervention in the event of failed platforms, satellite communications, and 24/7 command center watch (excluding response measures prosecuting detections). This project also estimated the operational costs ($/month/km-picket-fence) for a given fleet size, and the incremental costs of adding one wave glider. A business model in which surveillance services are contracted from a commercial enterprise was assumed, as opposed to government/military ownership of the capability. Contracted services capitalize faster and more efficiently on capability updates and developments—expected in satellite communications, autonomous behavior sets, and autodetection with accumulated data, wave glider platforms, and thin line arrays, and so forth [8]. Persistent surveillance lends itself to contracted services insofar as surveillance operations remain bounded, stopping short of legal enforcement and the use of force, which remain the task of government agencies who act on the information provided by the surveillance provider. It is a natural division of labour. Autonomous surveillance is allocated to commercial enterprise(s), while law enforcement, sovereignty assertion, and use of force are allocated to government agencies acting on the information provided by the contracted surveillance. A precedent for the commercial services model exists in manned aerial maritime surveillance, contracted from PAL Aerospace by the Canadian Government (Maritime Security, Fisheries and Oceans) [9]. PAL Aerospace provides airframes, pilots, flight crew, sensors, and data management to the Canadian government for large-area aerial surveillance, mostly routine but some special tasking as well. They provide a complete aerial ISR service. A similar business model is proposed here for persistent underwater maritime surveillance.

This project provides realistic qualitative and quantitative reference points in the vision for the new capability. The proposed way ahead [8] aims to support continual rapid advance of high autonomy for Canadian maritime surveillance, and to maximize the larger benefit to the emerging maritime robotic sector in Canada. It is therefore based on standardized open architectures for autonomous systems, and on an inclusive business model for contracted surveillance services, which together:

- avoid stifling progress with exclusive proprietary systems

- provide avenues for contributions to capability from diverse players, including Canadian industry, universities and allies (USA, AUS, NATO)

- streamline at-sea collaboration between industry and university partners by ensuring compatibility between their systems and the fielded system

- motivate innovation and speeds their uptake into at-sea surveillance capability

- support the export of maritime surveillance services to clients in addition to the Canadian government

- support the wider use of subsystem components (broaden their market) in applications other than maritime surveillance

REFERENCES

[1] Smith, R.N, Dasy, J., Hinez, G., Andersonz, W. and Sukhatme, G.S., "Predicting Wave Glider Speed from Environmental Measurements," Proc. OCEANS'11 MTS/IEEE KONA, Waikoloa, HI, USA, 19-22 Sept. 2011

[2] Kessel, R., "How Sea-Wave Spectra Naturally Limit the Speed of Wave Gliders", Technical Note, Seamount Analytics, Ottawa, 31-Mar-2019

[3] Manley, J.E., "USVs as Tow Platforms Wave Glider Experience and Results," Proc. OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 19-23 Sept. 2016

[4] Manley, J.E., Carlon, R. and Hine, G., "Ten Years of Wave Glider Operations: A Persistent Effort", Proc. OCEANS 2017 – Anchorage, 18-21 Sept. 2017

[5] GeoSpectrum Technologies Inc., "Tower Direction Sensor", https://geospectrum.ca/commercial-products/directional-sensors/towed-directional/, last viewed 12-June-2019

[6] Moldoveanu, N., Combee, L., Caprioli, P., Muyzert, E. and Pai, S., "Marine seismic acquisition with 3D sensor arrays towed by wave gliders," Proc. SEG International Exposition and 86th Annual Meeting, pp. 168-172, (2016)

[7] Kessel, R., Hamm., C., Stephanson, D., and Taillefer, M., "Persistent Maritime Surveillance Using Wave Glider Technology: Interim Project Report", Full Report, Maritime Way Scientific, Ottawa, 31-Mar-2019

[8] Kessel, R., Hamm., C., Stephanson, D., and Taillefer, M., "Persistent Maritime Surveillance Using Wave Glider Technology: Final Project Report", Full Report, Maritime Way Scientific, Ottawa, 30-Jun-2019

[9] See for instance PAL Aerospace, "PAL Aerospace Awarded Aerial Surveillance Contract by the Government of Canada", CISION, News, 04-Mar-2019, https://www.newswire.ca/news-releases/pal-aerospace-awarded-aerial-surveillance-contract-by-the-government-of-canada-865862016.html ; and "Provincial Aerospace: Navigating a Niche Industry", The Canadian Business Journal, Jan 2008, http://www.cbj.ca/biz_action/JAN/provincial.html

Acoustic clutter removal, Hong Kong Ship Channel

W. Jobst and L. Whited
Applied Marine Physics, LLC

D. Smith

CyberSmiths, Inc.

ABSTRACT

With the widespread availability of recreational and military underwater vehicles, maintaining underwater security in ship channels and harbors is more challenging than ever before. Ship traffic, high background noise and widespread acoustic clutter make detection of potential threats technically difficult and costly. The objective of this paper is to demonstrate how maximum length sequences, m-sequences, can remove both stationary and moving clutter in a ship channel or harbor environment, considerably improving opportunities for threat detection. As an example of the potential for clutter removal, the performance of a notional traffic monitoring system with underwater coverage is simulated for a segment of the Hong Kong Ship Channel.

Keywords: m-sequences, clutter, harbor defense, ship channel defense, coherent subtraction

Introduction

Maximum length sequences have been used to probe the dynamics of our oceans since Steinberg and Birdsall's 1964 pioneering experiment in the Florida Straits.[1] In the following 50 years, oceanographic studies using m-sequences with near-stationary acoustic sources and receivers were conducted at longer range, eventually leading to tomography, the Heard Island experiment, the Greenland Sea experiment, ATOC and many others. The list of current and past investigators is far too lengthy to properly credit all but a few.[2-14]

Recently, m-sequence studies and experiments have focused on harbor and near-shore defense, rather than oceanography.[15-17] In these areas the challenge has been to find small acoustic targets in the presence of clutter from underway commercial shipping, bottom features and moored vessels. The objective of this paper is to demonstrate, through simulation, how m-sequences can remove both stationary and moving clutter in a ship channel or harbor environment, considerably improving opportunities for threat detection in near-shore environments.

A Notional System for Ship Channel Monitoring

The Hong Kong Ship Channel is one of many possible locations for our notional system. As shown in Figure 1, a promontory provides a physical acoustic barrier between the acoustic projector and the horizontal line array receiver. *Separating the source and receiver eliminates the direct acoustic path* which, because of its high energy, can limit the performance of bistatic systems. (Unlike historic "ping-the-listen" sonars, m-sequence transmissions are low power and long duration. "Listening" continues throughout the signal transmission.) The geometry also places moving clutter

sources near-perpendicular to the source and receiver paths, *minimizing the Doppler channels and array beams that need to be processed.*
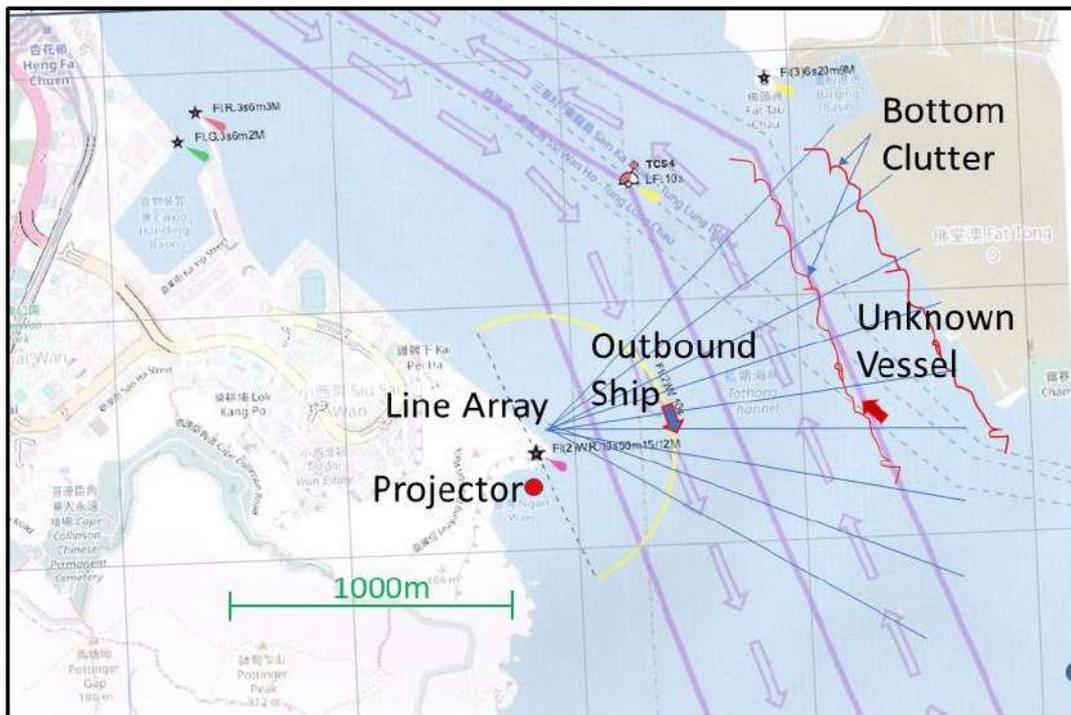

Figure 1: A notional system for ship channel monitoring

In Figure 1 the red lines near the far shore and on the shipping channel boundary indicate regions where acoustic clutter from the channel bottom is expected. Ten clutter points with random amplitudes are generated at both the far shore and the dredged shipping channel boundary within the 5-degree receiving array beam. In addition to bottom clutter from the channel boundary and far shore, reverberation is modeled by 33 clutter points distributed between the intersection of the source and receiver beam patterns and the far shore. (The projector beam pattern, typically 60 degrees, is not shown.) Reverberation energy is determined by bottom and surface backscattering coefficients associated with a rocky bottom and by the scattering area.[18-19] Considering the channel boundary, far shore clutter and reverberation, 53 clutter sources are included in the receive beam.

Two moving vessels are modeled. The outbound ship is traveling at 20 knots with a +15dB target strength. Doppler-shifted, direct and surface-reflected acoustic paths are included, to and from the ship, resulting in four closely spaced acoustic arrivals at the receiver. Similar modeling represents the unknown inbound submerged vessel traveling just beyond the ship channel. The speed of the unknown vessel is 6 knots, target strength -15dB and depth 6m. The objective is to detect the submerged vessel in the presence of reverberation, bottom clutter and the much larger +15dB moving ship.

"Simulation" is a word that has been used to describe a variety of systems with widely varying complexity. As shown in Figure 2, the simulation in this work separates the signal processing and environmental algorithms in such a way that the environmental portion of code can be replaced by the "real" ocean, thereby simplifying the transition from simulation to experiment.

The m-sequence is designed, usually with some trial-and-error, in the first block. Sequence duration is chosen to exceed the out-and-back bistatic distance. For the chosen environment, the carrier frequency is 5kHz, with 8 samples per cycle, 6 carrier cycles per sequence digit, 2047 digits per sequence and 5 sequences per transmission. The resulting duration of a single sequence is over 2 seconds, long enough for acoustic propagation across the ship channel and back. The carrier frequency was chosen based on the availability of reasonably priced, directional, linear acoustic projectors.
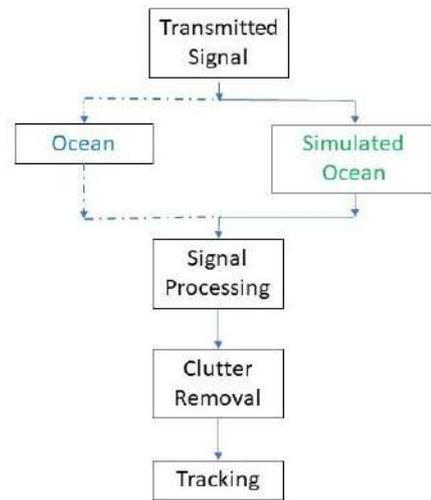
Figure 2: Simulation

In the environmental simulation, sound speed is isovelocity. Doppler spreading of surface-reflected paths is included for shipping and for the unknown vessel. Ambient noise spectrum level is 70dB, representative of harbors.[19] The line array beam pattern is modeled as a pie-shaped segment with coherent gain for clutter and incoherent gain for noise.

The received, demodulated signal, representing clutter and target, is the superposition of 61 time delayed and Doppler-shifted m-sequences. Filtering, demodulation and correlation are similar to what has been published in the past[15] with the exception that the received signal is demultiplexed into 48 sequences, each with 2047 complex samples. Each sequence is then processed individually with time and frequency corrections before correlation with a zero Doppler reference. In addition, *the received signal is bandpass filtered to smooth the m-sequence phase transition and to avoid computational noise from the interpolation algorithm*. With these changes the zero Doppler autocorrelation noise floor, shown in Figure 3(a), is more than 250dB below the autocorrelation peak. The peak width at 3dB points is 56 samples, one sample for
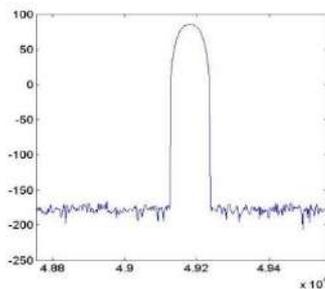

Figure 3(a) Autocorrelation
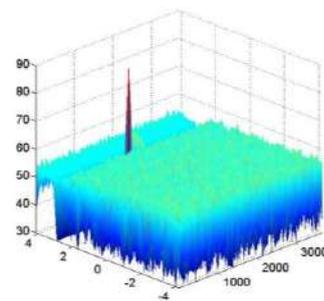
Figure 3(b) Ambiguity Function

each of the 56 time-multiplexed m-sequences. The autocorrelation noise floor for Doppler-shifted signals (not shown) is limited by filtering and is approximately 80dB below the peak. The ambiguity function floor for Doppler shifted signals, Figure 3(b), is approximately 26dB below the function peak and has near-uniform variability.

Simulations with noise, clutter and shipping are shown in Figure 4. Axes are source-object-receiver Doppler (+4 to -4 knots), source-object receiver distance (0 to 3600 meters) and received energy after beamforming and m-sequence processing ($10\log|A|^2$). Noise alone is shown in Figure 4(a). Reverberation, beginning at the intersection of source and receiver area coverage, clutter from the dredged shipping channel and clutter from the far shore have been added to noise in Figure 4(b). In Figure 4(c), clutter from a +15dB moving ship and a -15dB underwater vehicle (hidden by the ambiguity function floor) have been added.
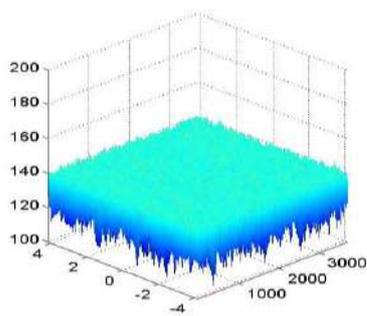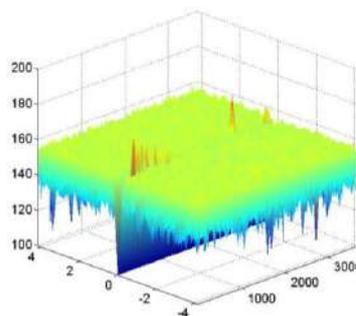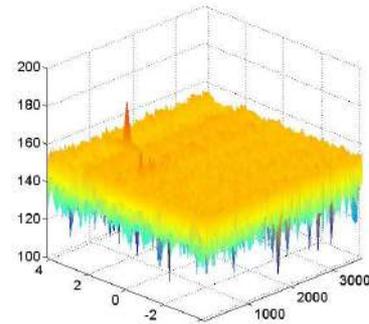

| Figure 4(a) | Figure 4(b) | Figure 4(c) |

Figure 4: Ambiguity function with (a) ambient noise, (b) noise, reverberation and boundary clutter, (c) noise, reverberation, boundary clutter +15dB shipping and -15dB underwater vehicle.

We define "coherent subtraction" as the operations that transform clutter from the ambiguity function to the time-domain where it can be subtracted from the original received signal. The coherent subtraction steps are: (1) identify areas containing clutter in the ambiguity function range-Doppler plane, (2) time and frequency correct clutter from the complex ambiguity function to its original Doppler frequency, (3) inverse transform the clutter segment from the ambiguity function domain to the time domain, (4) coherently subtract the complex time-domain clutter from the demodulated, complex received signal, and (5) reprocess the "cleaned" signal. The process is illustrated in Figure 5.

The ambiguity function before coherent subtraction is shown in Figure 5(a). Although reverberation, bottom clutter and the submerged -15dB inbound vessel are present, they are masked by the ambiguity function floor generated by the outbound ship. In Figure 5(b), the +15dB outbound vessel has been coherently subtracted from the received signal, and the ambiguity function reprocessed. Correlations from the unidentified vessel, bottom clutter and reverberation remain. After the third iteration, Figure 5(c), boundary clutter residuals have been coherently subtracted, and only the -15dB, 6 knot target remains.
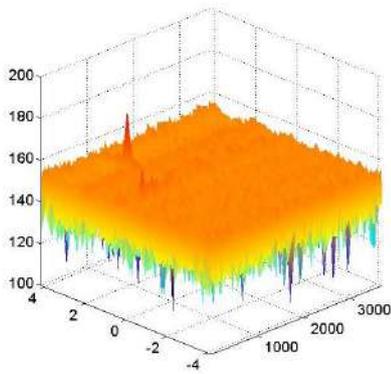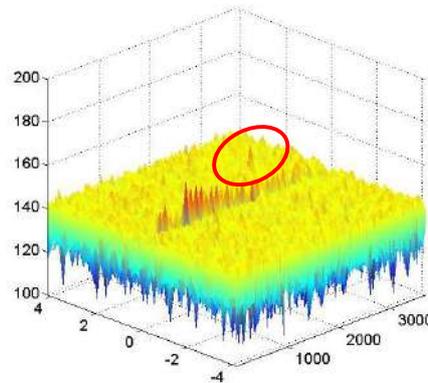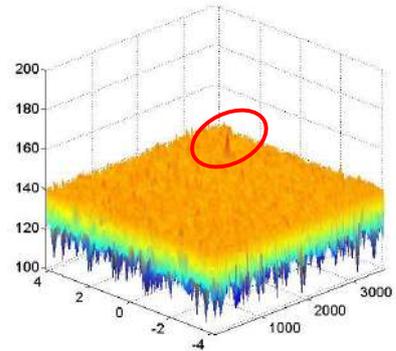
| Figure 5(a) | Figure 5(b) | Figure 5(c) |

Figure 5: Ambiguity function (a) before coherent subtraction, (b) after one iteration removes clutter from the moving ship, (c) after a second iteration removes reverberation and boundary clutter leaving the submerged target clearly visible.

Summary

With the widespread availability of recreational and military underwater vehicles, maintaining underwater security in ship channels and harbors is more challenging than ever before. This work demonstrates that, with attention to computational noise, coherent subtraction may provide a practical method for reducing clutter in a complex ship channel or harbor environment. Careful placement of the acoustic source and line array receiver can reduce signal processing complexity, minimize equipment cost and minimize the personnel required for system operation and maintenance.

References

[1] Steinberg, John C. and Birdsall, T. G. "Underwater Sound Propagation in the Straits of Florida" J. Acoust. Soc. Am., 39, (1966)

[2] Steinberg, J., Clark, J., DeFerrari, H., Kronengold M. and Yacoub, K. "Fixed system studies of underwater acoustic propagation" J. Acous. Soc. Am., 52, (1972)

[3] Jobst, W. and Dominijanni, L. "Measurements of the temporal, spatial and frequency stability of an underwater acoustic channel" J. Acoust. Soc. Am., 65, (1979)

[4] Munk, W. and Wunch, C. "Ocean acoustic tomography: A scheme for large scale monitoring" Deep-Sea Research, 26, (1979)

[5] Spiesberger, John L., Spindel, Robert C. and Metzger, Kurt "Stability and identification of ocean acoustic multipaths" J. Acous. Soc. Am., 67, (1980)

[6] Metzger, K. "Signal processing equipment and techniques for use in measuring ocean acoustic multipath structures" Ph.D. Dissertation, University of Michigan, Ann Arbor (1983)

[7] Birdsall, T. G. and Metzger Jr., K. "Factor inverse matched filtering" J. Acoust. Soc. Am., 79, pp. 91-99, (1986)

[8] Munk, Walter, Baggeroer, Arthur et al, "The Heard Island papers" (A collection of contributed papers by many authors), J. Acoust. Soc. Am., 96, (1994)

[9] Munk, Walter H. "Acoustic thermometry of ocean climate" J. Acous. Soc. Am., 100, (1998)

[10] Spiesberger, J. L. "Comparison of measured and modeled temporal coherence of sound near 75 Hz and 3683 km in the Pacific Ocean" J. Acous. Soc. Am., 124, pp. 2805-2811, (2008)

[11] Chang, Henry Siang-Ih "Detection of weak, broadband signals under doppler-scaled, multipath propagation" Ph.D. Dissertation, University of Michigan, (1992)

[12] DeFerrari, Harry A. "The application of m-sequences to bi-static active sonar" J. Acous. Soc. Am., 114(4), 2399-2400, (2003)

[13] DeFerrari, Harry "Eliminating clutter by coordinate zeroing" J. Acous. Soc. Am., 117, 149th meeting (2005)

[14] DeFerrari, Harry and Wylie, Jennifer "Ideal signals and processing for continuous active sonar" ICA Session 2pSP1, Montreal, CA, 2-7 June, 2013

[15] Jobst, W., Smith, D. and Whited, L. "Multistatic detection and tracking using linear maximal sequences" J. Acoust. Soc. Am., October 2010

[16] Jobst, W., Smith, D. and Whited, L. "Multistatic detection and tracking using linear maximal sequences" Proceedings of the 20th International Conference on Acoustics, ICA 2010, Sydney, Australia, 23-27 August 2010

[17] Jobst, W., Smith, D. and Whited, L. "Target detection and tracking in shallow water using maximal length sequences" 2010 International Waterside Security Conference, Marina di Carrara, Italy, 3-5 November 2010

[18] Dahl, Peter H. "On bistatic sea surface scattering: Field measurements and modeling" J. Acoust Soc. Am., April 1999

[19] Urick, Robert J. "Principles of Underwater Sound" 3rd edition" McGraw-Hill, 1983

Underwater sound speed Netcdf calculator

Petros Bitsikokos[a], Ioannis Bitsikokos [b]

[a] Military oceanographer (OF-1), Hellenic Navy Hydrographic Service, 229 Av. Mesogeion TGN 1040, Cholargos, pmpitsikokos@gmail.com; [b] Software Engineer, independent researcher, 14 Agnoston Martiron st. Nea Smirni, Athens Greece, jmpitsikokos@gmail.com

NETCDF UNDERWATER SOUND VELOCITY CALCULATION TOOL

The paper describes and presents preliminary results of a Netcdf Transformer tool, developed under java environment, for calculating and producing underwater Sound Speed Netcdf files. This study has been conducted using E.U. Copernicus Marine Service Information (CMEMS). By combining distinct Salinity and Temperature Netcdf data output from CMEMS, a compatible with low end machines Netcdf calculator is created, suitable for "at sea" calculations of the underwater sound speed. The calculations performed are based on Mackenzie (1981) 9 term formula for underwater sound speed. Products of the calculator are depicted in results. The soundscape / acoustic environment of the area bounded by 36N to 42N and Prime Meridian to 009E is calculated, with the distinct halo-clinic fronts of the area in mind. As input data, monthly mean hind casted average salinity and temperature data are selected for a period of 12 months. The calculator input data are the salinity and temperature Netcdf data outputs of Mediterranean Sea physical reanalysis component (product MEDSEA_REANALYSIS_PHYS_006_004), as made available by CMEMS. The output is a single Netcdf sound speed data file, plotted using NASA's Panoply software. The use of the calculator, permits the production of Military Oceanography (MILOC) related products, thus allowing the comprehensive and robust identification of the soundscape, therefore directly enhances the Environmental Knowledge and Operational Effectiveness (EKOE) principle for all level commands.

Keywords: Underwater sound speed, acoustic environment, soundscape, Netcdf underwater sound speed analysis tool, Operational Oceanography, MILOC support

INTRODUCTION

One of the fundamental aspects of providing Military Oceanography (MILOC) related products, is the underwater acoustical description of each area of operations. Therefore a tool for the calculation of the underwater sound speed and the production of the relevant thematic maps is needed to support all levels of command. In order to achieve this, a Sound Speed Calculator is created with the aim to produce Netcdf format files containing sound speed values for the selected area. The input files are Netcdf data from CMEMS MFS model MEDSEA_REANALYSIS_PHYS_006_004, as made available by CMEMS online catalogue[1-2] set. The output of the program is a new Netcdf file, containing sound velocity values and has the exact same space-time dimensions as the original source files. A secondary input option for the calculator is the use of a single source Netcdf data, as available from the global reanalysis model of CMEMS [3] (not included in the current paper). The full analysis of the input data used, goes beyond the scope of this paper, but for a more detailed description and analysis, the reader is encouraged to visit references 1 to 3.

METHODOLOGY (CALCULATOR DOCUMENTATION)

The NetcdfSoundVelocityTransfomer.jar is a Java Application which uses the *netcdfAll-4.5* and *javax.swing* java libraries to read, construct and manipulate .nc files through a user friendly desktop graphic interface. This client side application is customizable to utilize .nc files, with salinity and temperature data of the common chronological period with the purpose of creating a new .nc file with same timestamps and sound velocity data derived/transformed from the previous two files.

The input files for the program to process are netCDF (network common data form) files, with array-oriented scientific data, which pertain to unidata' s Common Data Model (CDM) and their exact file type is **NetCDF-3/CDM**. One primary and one secondary implementation of source data are developed.
This paper presents only the primary mode. By designating the input data as 2 distinct .nc files[3] with one file containing salinity data (variable name: *vosaline*) and the second file containing temperature data (variable name: *votemper* or *thetao*)
Alternatively it works with one global analysis .nc file [4] containing both salinity and temperature variables.

The algorithm for calculation of the new variable is based on the Mackenzie formula for underwater sound speed [5]. The use of this formula, allows the calculator to be used on low end machines, usually carried "at sea", by avoiding complex calculations which are resourceful demanding.

Salinity file description

The following description is an example of a salinity file named: sv04-med-ingv-sal-an-fc-m_1555267048618.nc downloaded from CMEMS online catalogue plotted by Panoply, a cross-platform application which plots geo-referenced and other array datasets like netCDF, HDF (Hierarchical Data Format) and GRIB (General Regularly-distributed Information in Binary form).

The values of interest in this file are *so* variable which is the salinity float type value, possible names to parse are either *vosaline* or *so*, both acceptable by the application.


Figure 1. . Caption of variable description for sv04-med-ingv-sal-an-fc-m_1555267048618.nc salinity file from Panoply

Temperature file description

The following description is an example of a temperature file named: sv04-med-ingv-tem-an-fc-m_1555266829350.nc .

The critical variable for this file is the *thetao* variable which is the temperature float type value, possible names to parse are either *votemper* or *thetao*, both acceptable by the application.

Figure 2. Caption of variable description for sv04-med-ingv-tem-an-fc-m_1555266829350.nc temperature file from Panoply

Variable definition

- **FillValue**: The _FillValue attribute specifies the fill value used to pre-fill disk space allocated to the variable. The fill value is returned when reading values that were never written. If _FillValue is defined then it should be scalar and of the same type as the variable. If the variable is packed using scale_factor and add_offset attributes (see below), the _FillValue attribute should have the data type of the packed data. This application accepts 1.0E20f as fill value so this value is shown as NaN(empty) in Panoply.

- **missing_value:** This attribute is not treated in any special way by the library or conforming generic applications, but is often useful documentation and may be used by specific applications. The missing_value

attribute can be a scalar or vector containing values indicating missing data. These values should all be outside the valid range so that generic applications will treat them as missing.

- **units:** A character string that specifies the units used for the variable's data.

- **coordinates:** The value of the coordinates attribute is a blank separated list of names of auxiliary coordinate variables and (optionally) coordinate variables. There is no restriction on the order in which the variable names appear in the coordinates attribute string.

- **standard_name:** The physical description of a variable

- **long_name:** A long descriptive name. This could be used for labeling plots, for example. If a variable has no long_name attribute assigned, the variable name should be used as a default.

- **ChunkSizes:** Array with the 4 numbers/sizes of chunks for each dimension of the variable

DATA INPUT

As a case study, the monthly mean hindcasted averaged data of salinity and temperature for the area bounded by 36N to 42N and Prime Meridian to 009E was selected. The input Netcdf files are downloaded from CMEMS interactive catalogue for the period from Jan 1992 to Dec 1992. Depth range selected is from the surface down to 313m (partial use of the available z levels of the model). The area chosen represents an area with distinct halo clinic variations, as the surface water masses originating from the Atlantic Ocean, after entering the Mediterranean Sea through the Gibraltar strait, meet with the saltier waters masses of the Mediterranean Sea. As an indicative representation of the input data 2 thematic maps are created using NASA's Netcdf data viewer software Panoply [6].
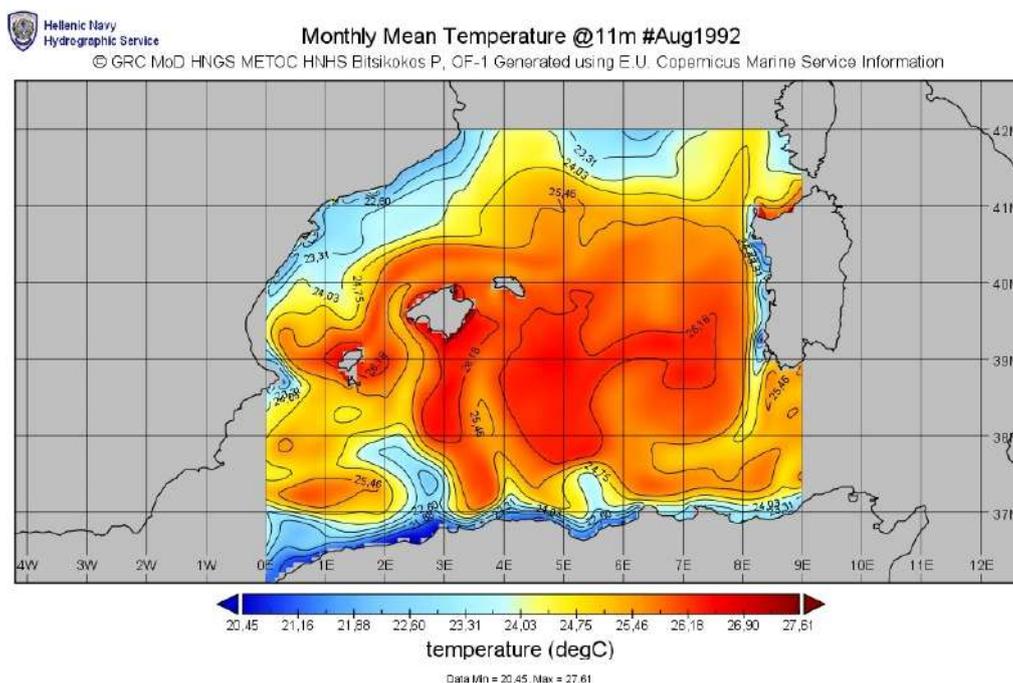

Figure 3. Input Netcdf file plotted. It presents the monthly average Sea Temperature at a depth of 11m for Aug 1992.
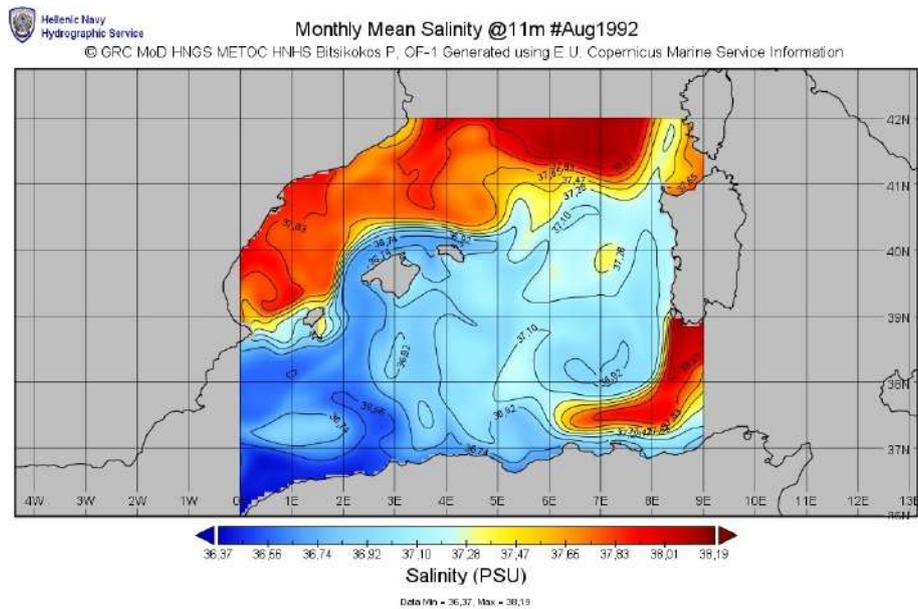
Figure 4. Input Netcdf file plotted. It presents the monthly average Salinity at a depth of 11m for Aug 1992.

VALIDATIONS

The client side application of the Netcdf Sound Velocity Transformer, before constructing the output data sv file, performs several necessary validations of the two input files which are described above:

- Depth variable must have same size in both salinity and temperature file.

- Latitude variable must have same size in both salinity and temperature file.

- Longitude variable must have same size in both salinity and temperature file.

- Time variable must have same size in both salinity and temperature file.

- Both files must have different name and path.

- Acceptable variable name for salinity variable must be either vosaline or so.

- Acceptable variable name for temperature variable must be votemper or thetao.

- Name of the output file and absolute path must be unique (different the input files).

Moreover, important to mention is that the application can handle cases where *scale_factor* and *add_offset* attributes exist in the data set variables of salinity and temperature. In these cases these values are transformed with the formula.

Final value = Value * scale_factor + add_offset

Where the attributes "*scale_factor*" and "*add_offset*" should both be of the type intended for the unpacked data (float values).

RESULTS

The output file is a new .nc file containing a new float value variable which is defined as sv and units are meters/second using the Mackenzie's formula.

```
netcdf file:/C:/Users/bitsikokos/Desktop/NCfiles/ResultFile.nc {
  dimensions:
    time = 36;
    depth = 74;
    lat = 145;
    lon = 169;
  variables:
    float sv(time=36, depth=74, lat=145, lon=169);
      :long_name = "sound velocity";
      :units = "m/s";
      :missing_value = 1.0000000200408773E20; // double
    float lat(lat=145);
      :units = "degrees_north";
      :axis = "Y";
    float lon(lon=169);
      :units = "degrees_east";
      :axis = "X";
    float time(time=36);
      :units = "seconds since 1970-01-01 00:00:00";
      :axis = "T";
    float depth(depth=74);
      :long_name = "depth";
      :units = "m";
      :axis = "Z";
  // global attributes:
  :title = "Example Data";
```

Figure 5. Output Netcdf file description.
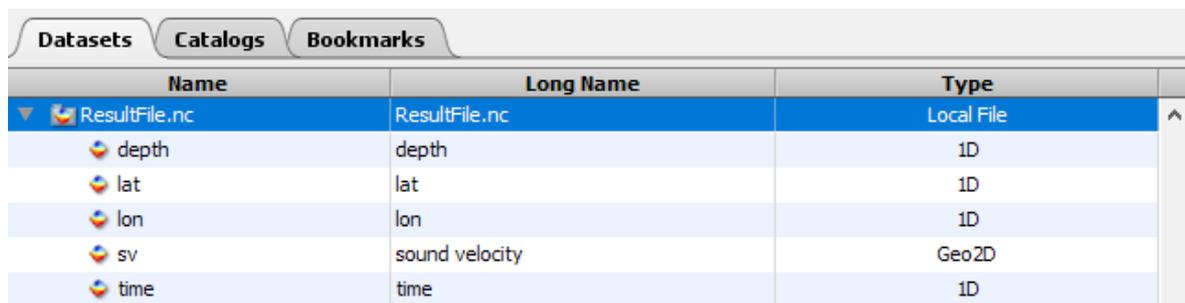
Figure 6. Caption of variable description for Result.nc sound velocity file from Panoply

The thematic maps in jpeg format show the areas where the combination of temperature and salinity values, produce mid and large scale acoustic fronts at various depths. The calculator combines all temperature and salinity values, correlating them both spatially and timely with the Mackenzie sound speed equation and produces the new Netcdf file which contains the sound velocity results per each point. The plotted final product is shown below.
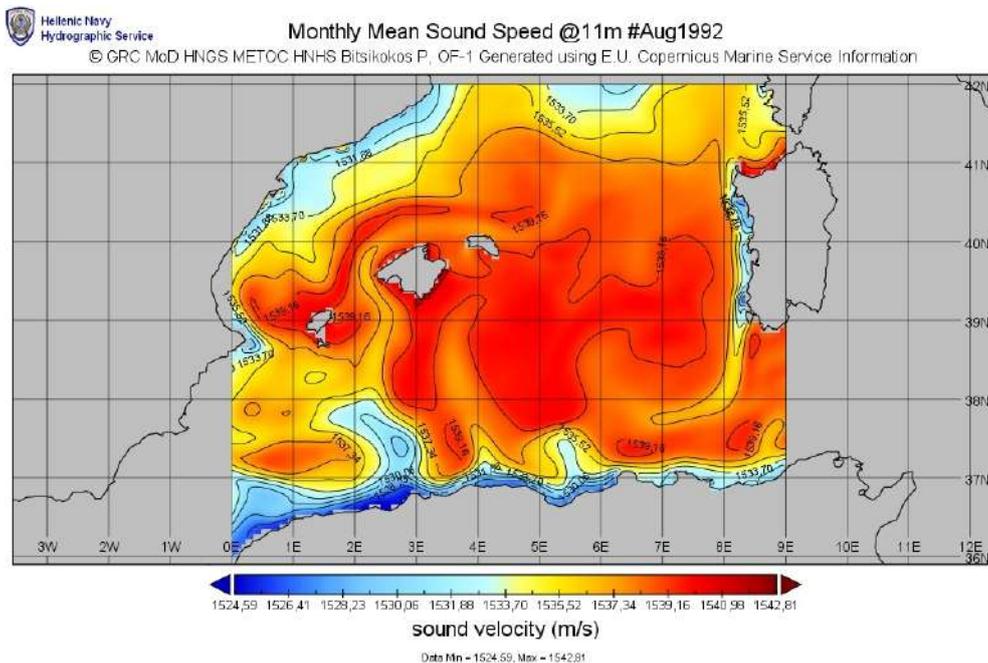
Figure 7. Output Netcdf file plotted. It presents the monthly average Sound Speed at a depth of 11m for Aug 1992.

As depicted at fig.7, the production of a Sound Speed thematic map, validates the effect of temperature gradients to the sound speed. Additionally the effects of salinity spatial distribution are also accounted for, allowing for a more robust soundscape / acoustic environment analysis of the area of interest (AOI).

CONCLUSIONS

The current project is utilized as a calculation tool for marine researchers and specialists, thus for future reference, message queue mechanisms could be used in order to process larger files and support a larger software ecosystem.

The **NetcdfSoundVelocityTransformer.jar** can also be used as a support java reference library for other netcdf file processing java projects. For further information/queries and details please contact the authors.

As graphically presented above, a robust analysis for the soundscape of an AOI can be produced by creating sound velocity thematic maps. The use of these maps by all levels of command, allows the commander to have a detailed description of the subsurface acoustic environment. Therefore, the thorough knowledge of the environment, procures the Environmental Knowledge and Operational Effectiveness (EKOE) principle, thus allowing the optimal use of the assets involved.

AKNOWLEDGEMENTS

REFERENCES

[1] Simoncelli, S., Fratianni, C., Pinardi, N., Grandi, A., Drudi, M., Oddo, P., & Dobricic, S. (2014). Mediterranean Sea physical reanalysis (MEDREA 1987-2015).

[2] Copernicus Monitoring Environment Marine Service (CMEMS) <https://doi.org/10.25423/MEDSEA_REANALYSIS_PHYS_006_004> (15 Feb 2019)

[3] GLOBAL_REANALYSIS_PHY_001_030 <http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_PHY_001_030> (16 Feb 2019)

[4] Product User Manual <http://marine.copernicus.eu/documents/PUM/CMEMS-MED-PUM-006-004.pdf> (15 Feb 2019).

[5] K.V. Mackenzie, Nine-term equation for the sound speed in the oceans (1981) J. Acoust. Soc. Am. 70(3), pp 807-812.

[6] NASA Panoply netCDF, HDF and GRIB Data Viewer < https://www.giss.nasa.gov/tools/panoply/> (10 Feb 2019).

Risk Sensitive Shifted Rayleigh Filter for Underwater Bearings-Only Target Tracking Problems

Nikhil Sharma[a]   Ranjeet Kumar Tiwari[b] and Shovan Bhaumik[b]

[a]Department of Mechanical Engineering, Indian Institute of Technology Patna, India;
[b]Department of Electrical Engineering, Indian Institute of Technology Patna, India.

ABSTRACT

Tracking with bearings-only measurement is a challenging filtering problem since many years. It is observed that the angle measurements do not always adhere to the assumed statistics during the estimation interval. Sometimes, for a particular interval of time, an un-modeled measurement spike or exploding variance is observed, which corrupts the estimation accuracy severely. In this work, we develop a robust shifted Rayleigh filter (SRF) based on exponential quadratic cost criteria for a bearings-only tracking problem with such corrupted measurements. The performance of the proposed filter in terms of the root mean square error and the track-loss is compared with the SRF and the risk-sensitive unscented Kalman filter (RS-UKF). It has been observed that the proposed filter performs better than the RS-UKF and the SRF when measurement spike is present for a specific interval of time.

Keywords: Bearings-only tracking (BOT), shifted Rayleigh filter (SRF), risk-sensitive filtering, state estimation, measurement spike.

1. INTRODUCTION

Bearings-only tracking (BOT) is a non-linear filtering quest which has a wide range of applications, which includes aircraft surveillance, target tracking, navigation, to name a few.[1,2] BOT is highly significant in a war-like scenario where observer position is not to be revealed, thus asserting the use of only passive sensors. Therefore, the only information available to estimate kinematic states is noise corrupted bearings. The problem becomes challenging due to the unobservability of the system (without observer maneuver) and the non-linear measurement equation. The problem is often termed as target motion analysis since the primary objective is to estimate the position and velocity of the target.[1]

While working on real-life BOT scenarios, we found that the measurement does not strictly adhere to the assumed statistics. A 'sudden spike' or explosion of measurement variance is observed, which last for a certain number of iterations. One possible reasoning for such a reaction is that passive sensors mainly deal with acoustic positioning systems and due to the availability of multiple acoustic paths between source and observer, the measurements are affected and under such situations, traditional filters fail miserably. To mitigate this effect, we lean towards a class of robust filters known as the risk-sensitive filters. The risk-sensitive filters originated from the risk-sensitive control law optimize an exponential cost criterion to arrive at the estimator equations and is reported to be more robust in the presence of modeling error[3] and un-modeled bias in the system.[4] Since exponential quadratic cost criterion penalizes higher order moments, which were ignored in conventional quadratic cost, the risk-sensitive filters are more robust towards process and measurement uncertainties.[5] For estimation under such cost criteria, closed-form solution is available for linear Gaussian system, which is known as risk-sensitive Kalman filter.[3,6] For non-linear systems, no closed form solution exists and various forms of approximations based risk-sensitive filters have been developed for such systems. These approximate solutions include the risk-sensitive extended Kalman filter (RS-EKF),[7] the risk-sensitive unscented Kalman filter (RS-UKF),[8] the risk-sensitive particle filter,[9] adaptive grid risk-sensitive filter (AGRSF),[10] etc.

It has been reported that the SRF is one of the best estimators when it comes to bearings-only tracking problems,[11].[12] It exploits the essential nature of non-linearity involved in BOT problems and calculates the exact posterior density using a Gaussian prior.[13] At the end of each iteration, this posterior is approximated as a Gaussian distribution to prepare for the next iteration, which accounts for the only approximation used in the SRF algorithm. We expect if a robust form of the SRF can be formulated, it will perform with more accuracy in presence of the process and measurement uncertainties. From this motivation, this work formulates a robust SRF, which uses the risk-sensitive cost criteria to deal with un-modeled measurement spike or the exploding variance as shown in Figure 1. The results are compared with the risk-sensitive unscented Kalman filter (RS-UKF) and the standard SRF. A huge track divergence is observed while using the risk-neutral filters under un-modeled measurement spike whereas a risk-sensitive filter has performed with a low failure rate. It has also been seen that the proposed risk-sensitive shifted Rayleigh filter (RS-SRF) is more accurate than the existing RS-UKF.

2. PROBLEM FORMULATION

The bearings-only tracking problem for a target moving with constant velocity is given by discrete time equation

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} - \boldsymbol{\mho}_{k-1,k} + \mathbf{w}_{k-1}, \tag{1}$$

where $\mathbf{x}_k$ is the relative state vector between the target and observer and is expressed as $[x_k, y_k, v_{x_k}, v_{y_k}]^T$. x_k, y_k are the position components and v_{y_k}, v_{y_k} are the velocity components of the relative state vector. $\mathbf{F}$ is the state transition matrix defined as

$$\mathbf{F} = \begin{bmatrix} I_{2\times2} & \Delta T I_{2\times2} \\ O_{2\times2} & I_{2\times2} \end{bmatrix},$$

$\mathbf{w}_{k-1}$ is a zero mean Gaussian distributed process noise with covariance matrix $\mathbf{Q}$. The covariance matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} = \begin{bmatrix} \frac{\Delta T^3}{3} I_{2\times2} & \frac{\Delta T^2}{2} I_{2\times2} \\ \frac{\Delta T^2}{2} I_{2\times2} & \Delta T I_{2\times2} \end{bmatrix} \tilde{q},$$

where ΔT is the sampling time and $\tilde{q}$ denotes the process noise intensity.[14,15] $\boldsymbol{\mho}_{k-1,k}$ defines a vector of deterministic inputs from observer with state vector $[x_k^o, y_k^o, v_{x_k}^o, v_{y_k}^o]^T$ and is given by[15]

$$\boldsymbol{\mho}_{k-1,k} = \begin{bmatrix} x_k^o - x_{k-1}^o - \Delta T v_{x_{k-1}}^o \\ y_k^o - y_{k-1}^o - \Delta T v_{y_{k-1}}^o \\ v_{x_k}^o - v_{x_{k-1}}^o \\ v_{y_k}^o - v_{y_{k-1}}^o \end{bmatrix}.$$

The only measurement available in BOT scenario is the angle with respect to true north direction which is modeled by the equation

$$\theta_k = \tan^{-1} \left[\frac{x_k}{y_k} \right] + v_k, \tag{2}$$

where v_k denotes the measurement noise with probability density $v_k \sim \mathcal{N}(0, R)$.

2.1 Received Measurement

In this work, it is considered that the received measurements do not adhere with Eqn. (2) and a spike like measurement that lasts for a specific sampling time intervals is received as shown in Figure 1. As we do not use any real data, the following expressions are used to model the measurement spike.

$$\theta_i = \begin{cases} aA_b \frac{(i-t_s)}{\lfloor \frac{t_s+t_e}{2} \rfloor - t_s} + \theta_{t_s-1}; & \text{if } t_b < i \leq \lfloor \frac{t_s+t_e}{2} \rfloor, \\ -aA_b \frac{i - \lfloor \frac{t_s+t_e}{2} \rfloor}{t_e - \lfloor \frac{t_s+t_e}{2} \rfloor} + \theta_{\lfloor \frac{t_s+t_e}{2} \rfloor - 1}; & \text{if } \lfloor \frac{t_s+t_e}{2} \rfloor < i \leq t_e, \\ \tan^{-1} \left[\frac{x_i}{y_i} \right] + v_k; & \text{else.} \end{cases} \tag{3}$$

Here, a can take values either of $+1$ and -1 with equal probabilities. Due to this we observe equal number of both rising and falling peaks for the spike. A_b is the bias amplitude, t_s and t_e denote the sampling time instants for spike start and end respectively. Note that the spike amplitude A_b and the time duration of the spike $t_e - t_s$ do not change and remain same for all Monte-Carlo runs.

3. RISK SENSITIVE SHIFTED RAYLEIGH FILTER

3.1 Augmented measurements

The shifted Rayleigh filter introduces an augmented measurement, $\boldsymbol{\alpha}_k$, which is the relative position vector between the observer and the target, and is given by[16]

$$\boldsymbol{\alpha}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\gamma}_k. \tag{4}$$

In this equation, $\mathbf{H} = [\boldsymbol{I}_{r \times r} \ \boldsymbol{O}_{r \times (n-r)}]$, where r is the dimension of augmented measurement space, n is the order of the state vector ($n = 4$ in our case), and $\boldsymbol{I}_{r \times r}$ and $\boldsymbol{O}_{r \times r}$ are identity and zero matrix, respectively. We also define $\boldsymbol{\beta}_k = \Pi(\boldsymbol{\alpha}_k)$ as projection of $\boldsymbol{\alpha}_k$ onto a unit circle ($r = 2$) or a unit sphere ($r = 3$), as given in.[16] γ_k is the noise associated with augmented measurement model in (4) whose covariance $\mathbf{R_a}$ is expressed as follows[16]

$$\mathbf{R_a} = \sigma_\theta^2 [\hat{\mathbf{x}}_{k|k-1}^2 + \hat{y}_{k|k-1}^2 + var(\mathbf{x}_{k|k-1}) + var(\mathbf{y}_{k|k-1})] \boldsymbol{I}_{r \times r}, \tag{5}$$

where $var(.)$ denotes the covariance.

3.2 Solution approach

The cost function, which is optimized to formulate RS-SRF, is defined as[3]

$$
\begin{aligned}
J_k(\hat{\mathbf{x}}_k) &= \mathbb{E}\left[\exp\{\mu_1 \sum_{i=0}^{k-1}(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T(\mathbf{x}_i - \hat{\mathbf{x}}_i) + \mu_2(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T(\mathbf{x}_k - \hat{\mathbf{x}}_k)\}\right] \\
&= \int_{\mathbb{R}^n} \exp\left\{\mu_1 \sum_{i=0}^{k-1}(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T(\mathbf{x}_i - \hat{\mathbf{x}}_i) + \mu_2(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T(\mathbf{x}_k - \hat{\mathbf{x}}_k)\right\} p(\mathbf{x}_k|\boldsymbol{\alpha}_k)d\mathbf{x}_k.
\end{aligned}
\tag{6}
$$

3.3 Risk-sensitive filtering with augmented measurements

We define an *information state*, σ_k, in order to work out the estimate recursively that minimizes the cost function in Eqn. (6). Information state is a function of available information set that completely summarizes the past of the system in a probabilistic sense.[2] One such candidate for the information state is the un-normalized conditional density function in risk sensitive sense, which is given by[3]

$$\sigma_k = \exp\left\{\mu_1 \sum_{i=0}^{k-1}(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T(\mathbf{x}_i - \hat{\mathbf{x}}_i)\right\} p(\mathbf{x}_k|\boldsymbol{\alpha}_k). \tag{7}$$

Substituting in (6), the cost function can now be compactly re-written as

$$J_k(\hat{\mathbf{x}}_k) = \int_{\mathbb{R}^n} \exp\left(\mu_2(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T(\mathbf{x}_k - \hat{\mathbf{x}}_k)\right) \sigma_k d\mathbf{x}_k, \tag{8}$$

and the minimum cost estimate is expressed as

$$\hat{\mathbf{x}}_k = \arg\min_{r \epsilon \mathbb{R}^n} \int \exp(\mu_2(\mathbf{x}_k - r)^T(\mathbf{x}_k - r))\sigma_k d\mathbf{x}_k. \tag{9}$$

We will exploit the linear property of the process equation and the augmented measurement in Eqn. (4) to derive the risk sensitive estimate. Since we have used the conditional pdf as the information state, it follows the recursion

$$\sigma_{k+1} = \int p(\boldsymbol{\alpha}_{k+1}|\boldsymbol{\alpha}_k, \mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{x}_k)\sigma_k d\mathbf{x}_k. \tag{10}$$

Here, $p(\boldsymbol{\alpha}_{k+1}|\boldsymbol{\alpha}_k, \mathbf{x}_{k+1})$ is the distribution of predicted measurement at time $k+1$ given measurements upto time k. $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ denotes the prior state distribution, both of these densities can be found out if density of process noise $\mathbf{w}_k$ and measurement noise $\boldsymbol{\gamma}_k$ respectively are known.

The final risk sensitive equations for system defined by Eqn. (1) and Eqn. (4) can be obtained by assuming a Gaussian distributed conditional density σ_k, following the recursion given by Eqn. (10) and utilizing Eqn. (9) to find the minimum cost estimate.[3] The final risk sensitive equations are

$$
\begin{aligned}
\hat{\mathbf{x}}_{k+1|k} &= \mathbf{F}\hat{\mathbf{x}}_{k|k} \\
\boldsymbol{\Sigma}_{k+1|k} &= \mathbf{F}[(\boldsymbol{\Sigma}_{k|k}^{-1} - 2\mu_1 I)^{-1}]\mathbf{F}^T + \mathbf{Q} \\
\hat{\mathbf{x}}_{k+1|k+1} &= \hat{\mathbf{x}}_{k+1|k} + \boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T + R)^{-1}[\boldsymbol{\alpha}_{k+1} - \mathbf{H}\hat{\mathbf{x}}_{k+1|k}] \\
\boldsymbol{\Sigma}_{k+1|k+1} &= (I - \boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T + \mathbf{R_a})^{-1}\mathbf{H})\boldsymbol{\Sigma}_{k+1|k},
\end{aligned}
\tag{11}
$$

where the term $\boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_{k+1|k}\mathbf{H}^T + \mathbf{R_a})^{-1}$ is referred to as the Kalman gain, $\mathbf{K}_{k+1}$.

3.4 Risk-sensitive shifted Rayleigh filter (RS-SRF)

The derivation for the risk-sensitive shifted Rayleigh filter will closely follow the approach given in[16, 17] as the expressions with augmented measurement are available in linear form. In the derivation of SRF equations, the assumption of direct accessibility of the augmented measurements is maintained. However, since the range is not available with us, filtering equations are altered by evaluating moments of $\mathbf{x}_k$ conditioned on $\boldsymbol{\beta}_k$. Writing $\boldsymbol{\alpha}_k = r_k\boldsymbol{\beta}_k$, where $r_k = \|\boldsymbol{\alpha}_k\|$ and $\boldsymbol{\beta}_k$ is defined in section 3.1. If $\hat{\mathbf{x}}_{k|k}$ is approximated as a Gaussian distributed random variable, it is legit to write $\mathbf{x}_k$ as

$$
\mathbf{x}_{k+1} = (I - \mathbf{K}_{k+1}\mathbf{H})\hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}r_{k+1}\boldsymbol{\beta}_{k+1} + \boldsymbol{\xi}_{k+1},
\tag{12}
$$

where $\boldsymbol{\xi}_{k+1}$ follows the probability density $\mathcal{N}(0, (I - \mathbf{K}_{k+1}\mathbf{H})\boldsymbol{\Sigma}_{k+1|k})$. Taking the conditional expectation of $\mathbf{x}_{k+1}$ given $\boldsymbol{\beta}_{k+1}$, we have

$$
\begin{aligned}
\hat{\mathbf{x}}_{k+1|k+1} &= E(\mathbf{x}_{k+1}|\boldsymbol{\beta}_{k+1}) \\
&= (I - \mathbf{K}_{k+1}\mathbf{H})\hat{\mathbf{x}}_{k+1|k} + E[r_{k+1}|\boldsymbol{\beta}_{k+1}]\mathbf{K}_{k+1}\boldsymbol{\beta}_{k+1}.
\end{aligned}
\tag{13}
$$

Similarly, the posterior covariance is calculated as[13]

$$
\begin{aligned}
cov(\mathbf{x}_{k+1}|\boldsymbol{\beta}_{k+1}) &= E[(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1})^T(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1})|\boldsymbol{\beta}_{k+1}] \\
&= (I - \mathbf{K}_{k+1}\mathbf{H})\boldsymbol{\Sigma}_{k+1|k} + cov[r_{k+1}|\boldsymbol{\beta}_{k+1}]\mathbf{K}_{k+1}\boldsymbol{\beta}_{k+1}\boldsymbol{\beta}_{k+1}^T\mathbf{K}_{k+1}^T.
\end{aligned}
\tag{14}
$$

Eqn. (13) and Eqn. (14) can be expanded to get the final equations for risk sensitive SRF[16] as follows:

$$
\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} - \mho_{k-1,k},
\tag{15}
$$

$$
\boldsymbol{\Sigma}_{k|k-1} = \mathbf{F}[(\boldsymbol{\Sigma}_{k-1|k-1}^{-1} - 2\mu_1 I)^{-1}]\mathbf{F}^T + \mathbf{R_a},
\tag{16}
$$

$$
\hat{\mathbf{x}}_{k|k} = (I - \mathbf{K}_k\mathbf{H})\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k E[r_k|\boldsymbol{\beta}_k]\mathbf{V}_k^{1/2}\boldsymbol{\beta}_k,
\tag{17}
$$

$$
\boldsymbol{\Sigma}_{k|k} = (I - \boldsymbol{\Sigma}_{k|k-1}\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_{k|k-1}\mathbf{H}^T + \mathbf{R_a})^{-1}\mathbf{H})\boldsymbol{\Sigma}_{k|k-1} + cov(r_k|\boldsymbol{\beta}_k)\mathbf{K}_k\mathbf{V}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{\beta}_k^T\mathbf{V}_k^{1/2}\mathbf{K}_k^T,
\tag{18}
$$

where $E[r_k|\boldsymbol{\beta}_k]$ and $cov[r_k|\boldsymbol{\beta}_k]$ are given in.[16]

4. SIMULATION RESULTS

A typical measurement spike is presented in Figure 1. The spike is included by modifying the mathematical model of measurement equation within a certain range of time steps as expressed by Eqn. (3). The process noise intensity $\tilde{q}$ is taken as[18] 2.142×10^{-6} km^2/min^3. The sampling time taken between successive measurements is 1 second. The risk-sensitive parameter μ_1 is considered to be -5 for this simulation work. Measurement spike amplitude is $15°$ and the observation period lasts for 30 minutes.

The values of all parameters used and the initialization of the different filters considered are according to the method given in.[18] Further, two metrics, the root mean square error (RMSE) and the track-loss, are defined, as illustrated in,[18] to compare the performance of the different filters considered in this simulation. The track-loss is considered to occur when $RMSE^{pos}_{k_m} > T_l$ km, where

$$RMSE^{pos}_{k_m} = \sqrt{(\mathrm{x}^t_{k_m} - \hat{\mathrm{x}}^t_{k_m})^2 + (\mathrm{y}^t_{k_m} - \hat{\mathrm{y}}^t_{k_m})^2},$$

k_m is the final time step and T_l is the threshold value, which is taken as 1 km for this work. Note that diverged tracks and corresponding Monte-Carlo runs are excluded while calculating the position RMSE.
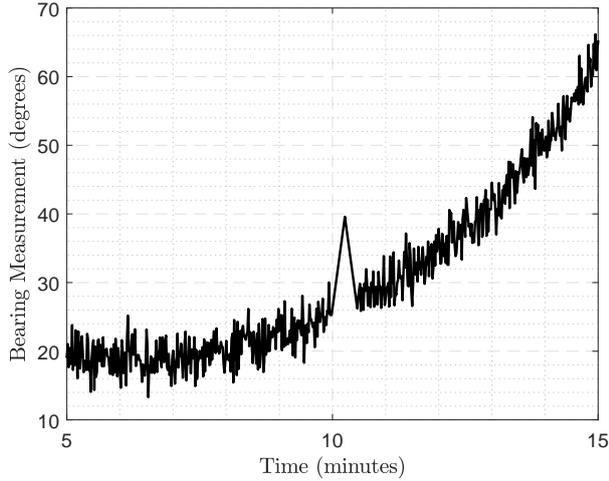

Figure 1: Measurement Spike
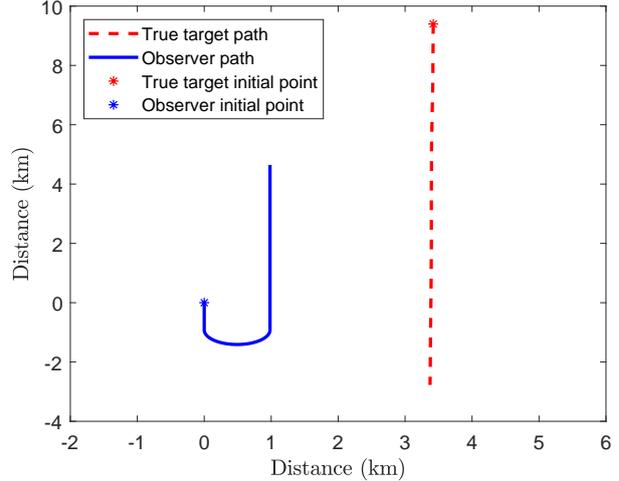
Figure 2: Observer and target dynamics

4.1 Scenario

We have chosen the RS-UKF[8] and the standard SRF for comparison. The BOT scenario considers a linear process model (constant velocity target) and is presented in Figure 2. The observer traverses with an initial course of 180° with respect to true north at a velocity of 8 knots. From 1st to 6th minute the observer maneuvers from $180°$ to $0°$ at constant rate. The target moves with a constant velocity of 15 knots at a course of $180°$ from true north.

The measurement spike is shown in Figure 1 and is included between $t_s = 600$ and $t_e = 660$ sampling time instants for a total of 60 seconds. In this work, we have assumed that there is no ambiguity in its time of occurrence and maximum amplitude, thus for all Monte-Carlo runs, the glitch occurs at the same time. Maximum amplitude is taken as 15°. Since the measurements were highly corrupted during measurement spike phase, we tried using only process update in Kalman filter recursions and using measurement update only when measurements are convincing. This also lead to considerable divergence while tracking.

On comparison with the RS-UKF, we found that it takes much more time to settle down than the RS-SRF even with low initialization errors. The final RMSE in position was better for the RS-SRF as shown in Figure 3. The results with track-loss and terminal RMSE in position and velocity are mentioned in Table 1. An appreciable difference in track-loss is evident from the table.

Result: It was observed that the risk neutral filters ultimately failed leading to 100% track-loss. With the robust approach a 10000 Monte Carlo run simulation cycle have shown 670 cases of track-loss in the RS-UKF whereas for the specified value of risk sensitive parameter, the same is 0 in case of RS-SRF. Table 1 shows the comparison between the RS-SRF, RS-UKF and SRF in terms of the track-loss, RMSE position and RMSE velocity. The RMSE position is plotted in Figure 3. Since the SRF and UKF lead to remarkably high track-loss,

they are not presented in the figures. Superior performance of the RS-SRF in presence of measurement spike is evident over other mentioned filters.

Table 1: Terminal RMSE in position and velocity, and % track-loss for different filters

|	$\mathbf{RMSE^{pos}_{k_m}}$(m)	$\mathbf{RMSE^{vel}_{k_m}}$(m/s)	Track-loss
RS-SRF	292	1.28	0.0%
RS-UKF	308	1.31	6.7%
SRF	—	—	100%
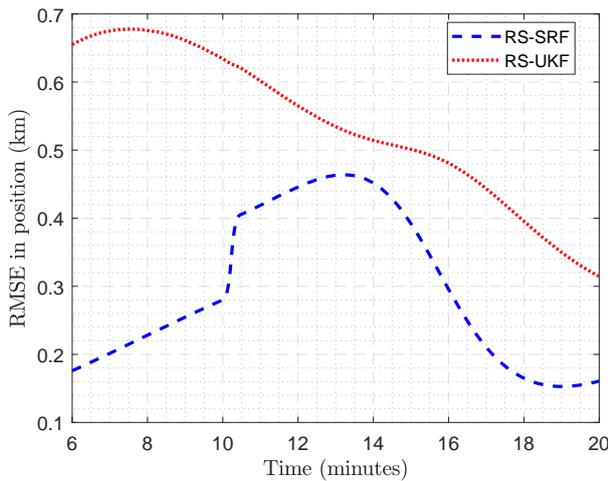
Figure 3: RMSE position
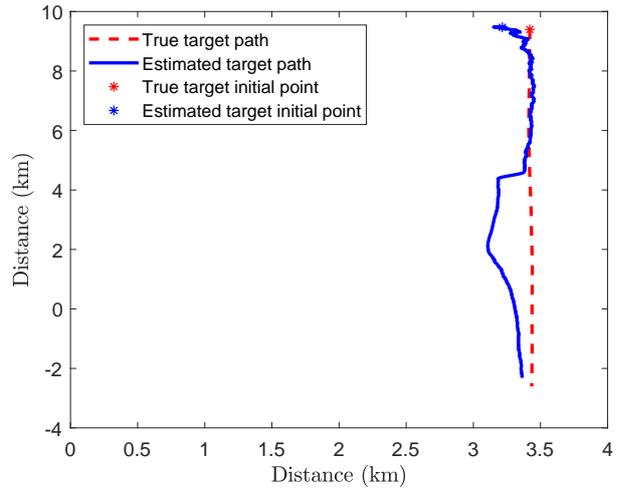
Figure 4: Truth and estimated target

5. CONCLUSIONS

Conventional moment matching filters have shown high track divergence while dealing with measurement spike and thus lose their applicability in underwater passive BOT systems. In this technical work, the shifted Rayleigh filter has been modified by switching from the existing quadratic error criterion to exponential quadratic error criterion. To validate the performance of proposed filter, an underwater real life BOT problem including measurement spike is solved using risk neutral SRF, RS-UKF and the proposed RS-SRF. The simulation results suggest the use of RS-SRF if measurement spike is likely to occur.

ACKNOWLEDGMENTS

REFERENCES

[1] Ristic, B., Arulampalam, S., and Gordon, N., "Beyond the Kalman filter," *IEEE Aerospace and Electronic Systems Magazine* **19**(7), 37–38 (2004).

[2] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., "Estimation with applications to tracking and navigation: theory algorithms and software," (2004).

[3] Boel, R. K., James, M. R., and Petersen, I. R., "Robustness and risk-sensitive filtering," *IEEE Transactions on Automatic Control* **47**(3), 451–461 (2002).

[4] Bhaumik, S., "Improved filtering and estimation methods for aerospace problems," *PhD Thesis, Dept of Engineering, Jadavpur University* (2008).

[5] Ford, J., "Risk-sensitive filtering and parameter estimation," *Aeronautical and Maritime Research Lab, Melbourne (Austrailia)* (1999).

[6] Speyer, J. L., Fan, C.-H., and Banavar, R. N., "Optimal stochastic estimation with exponential cost criteria," *Proceedings of the 31st IEEE Conference on Decision and Control* , 2293–2299, IEEE (1992).

[7] Jayakumar, M. and Banavar, R. N., "Risk-sensitive filters for recursive estimation of motion from images," *IEEE transactions on pattern analysis and machine intelligence* **20**(6), 659–666 (1998).

[8] Bhaumik, S., Sadhu, S., and Ghoshal, T., "Risk-sensitive formulation of unscented Kalman filter," *IET control theory & applications* **3**(4), 375–382 (2009).

[9] Sadhu, S., Bhaumik, S., Doucet, A., and Ghoshal, T. K., "Particle-method-based formulation of risk-sensitive filter," *Signal Processing* **89**(3), 314–319 (2009).

[10] Bhaumik, S., Srinivasan, M., Sadhu, S., and Ghoshal, T. K., "Adaptive grid solution of risk sensitive estimator problems," 356–360, IEEE (2005).

[11] Arulampalam, S., Clark, M., and Vinter, R., "Performance of the shifted Rayleigh filter in single-sensor bearings-only tracking," *10th International Conference on Information Fusion* , 1–6, IEEE (2007).

[12] Arasaratnam, I. and Haykin, S., "Cubature Kalman filters," *IEEE Transactions on automatic control* **54**(6), 1254–1269 (2009).

[13] Clark, J. M., Robbiati, S., and Vinter, R. B., "The shifted Rayleigh mixture filter for bearings-only tracking of maneuvering targets," *IEEE transactions on signal processing* **55**(7), 3218–3226 (2007).

[14] Leong, P. H., Arulampalam, S., Lamahewa, T. A., and Abhayapala, T. D., "A Gaussian-sum based cubature Kalman filter for bearings-only tracking," *IEEE Transactions on Aerospace and Electronic Systems* **49**(2), 1161–1176 (2013).

[15] Radhakrishnan, R., Singh, A. K., Bhaumik, S., and Tomar, N. K., "Quadrature filters for underwater passive bearings-only target tracking," 1–5, IEEE (2015).

[16] Clark, J., Vinter, R., and Yaqoob, M., "Shifted Rayleigh filter: A new algorithm for bearings-only tracking," *IEEE Transactions on Aerospace and Electronic Systems* **43**(4), 1373–1384 (2007).

[17] Özelçi, A. C., "Tracking and estimation algorithms for bearings only measurements," *PhD Thesis, Dept of Electrical Engineering, Imperial College, London* (2012).

[18] Radhakrishnan, R., Bhaumik, S., and Tomar, N. K., "Gaussian sum shifted Rayleigh filter for underwater bearings-only target tracking problems," *IEEE Journal of Oceanic Engineering* (99), 1–10 (2018).

11
Additional material

The MSAW'19 was located at

Villa Marigola Congress Centre
Via S. Biaggini 1
19032, Lerici
La Spezia (SP), Italy

1st Maritime Situational Awareness Workshop - MSAW 2019
Science and Technology meet operational needs
Villa Marigola, Lerici (La Spezia), Italy
8–10 October 2019

La Spezia → **VENERE AZZURA**: Bus L or S (every 15 min.) for 30/45 min.
Lerici or San-Terenzo → **VENERE AZZURA**: 15 minutes walk

VENERE AZZURA → Villa Marigola:
Free shuttle service continuously from 7h50 to 10h00
Walk, approx. 5-10 minutes through stairs, with a 50m elevation gain

No car is allowed at the Villa
Parking is **free** of charge:
▪ Parking areas P
▪ Parking slots along the street
(white, blue or yellow)

The workshop dinner was organized as a no-host gathering at **Ca' Borghese Restaurant**, Via Arena 1, 19031 Ameglia, La Spezia (SP), Italy.

11.1 Programme

Tuesday 8 October

08:00	**Registration & Welcome coffee**	*Main Hall*
09:00	Welcome and Opening Talk	*Main Conference Room*
	Catherine Warner, CMRE Director	
09:30	Keynote presentation	
	Maritime surveillance: radar technologies and scenario characteristics	
	Alfonso Farina[1], Sergio Gallone[1], Massimo C. Comparini[2,3]	
	[1]Leonardo SpA, [2]e-GEOS, [3]Telespazio	
10:30	*Coffee break*	*Garden*

RADAR SYSTEMS — *Main Conference Room*
Chair: R. Streit

11:00	**RANGER: Radars and early warning technologies for long distance maritime surveillance**	
	L. Karagiannidis, D. Dres, E. Protopapadakis, F. Lamole, F. Jacquin, G. Rigal, E. Ouzounoglou, D. Katsaros, A. Karalis, L. Pierno, C. Mastroeni, M. Evangelista, V. Fontana, D. Gaglione, G. Soldi, P. Braca, S. Sarlio-Siintola, E. Sdongos, A. Amditis	
11:30	**Fifteen years of HF-Surface-Wave radar for maritime surveillance**	
	J. Widera, T. Helzel, M. Kniephoff, L. Petersen	
11:50	**Beyond the horizon: High frequency surface wave radar**	
	F. Lamole, F. Jacquin, G. Rigal	
12:10	**A MIMO architecture for SaR application**	
	F. Prodi, L Pierno, A. Farina	
12:30	**Passive radar on moving platforms for maritime and coastal surveillance**	
	P. Wojaczek, F. Colone, D. Cristallini, P. Lombardo, D. O'Hagan	
12:50	*Lunch*	*Garden*
13:50	Keynote presentation (remote)	*Main Conference Room*
	Re-examining fusion-sensemaking-decision-making interdependencies again	
	Jim Llinas, University at Buffalo	
14:50	*Break*	

ARTIFICIAL INTELLIGENCE FOR MSA — *Main Conference Room*
Chair: J. Garcia Herrero

15:00	**A geospatial complex event processing engine for abnormal vessel behavior detection suitable for maritime surveillance**	
	M. Tsogas, P. Parthymos, M. Moutzouris, N. Patlakas, G. Karagiannis, A. Kostaridis, D. Diagourtas	
15:20	**Detection of maritime anomalous behavior in a successful MARISA North Sea trial**	
	A. Mohamoud, J. van de Pol, E. den Breejen, K. Veltman, J. van der Velde, T. Mannucci, H. Hildmann	
15:40	*Coffee break*	*Garden*
16:10	**Achieving maritime situational awareness using knowledge graphs: a study**	
	J. Everwyn, B. Zanuttini, A-I. Mouaddib, S. Gatepaille, S. Brunessaux	
16:30	**Maritime Situation Awareness through data analytics, machine learning and risk assessment based on ship trajectories**	
	F. Opitz, C. Mohrdieck, K. Dästner	
16:50	**IDCP as kernel element for force level recognized maritime picture generation**	
	J. Ziegler	
17:10	Poster Session & Demos	*Gallery*
18:00	Icebreaker	*Garden*

1st Maritime Situational Awareness Workshop - MSAW 2019
Science and Technology meet operational needs
Villa Marigola, Lerici (La Spezia), Italy
8–10 October 2019

POSTERS

1. Acoustic clutter removal, Hong Kong ship channel
W. Jobst, L. Whited, D. Smith

2. Underwater sound speed Netcdf calculator
P. Bitsikokos, I. Bitsikokos

3. Risk sensitive shifted Rayleigh filter for underwater bearings-only target tracking problems
N. Sharma, R.K. Tiwari, S. Bhaumik

4. Maritime anomaly detection
E. D'Afflisio, L.M. Millefiori, P. Braca

5. AIS-based learning of maritime traffic patterns for long-term prediction and anomaly detection
N. Forti, L.M. Millefiori, P. Braca

6. Heterogeneous information fusion for multitarget tracking using the sum-product algorithm
G. Soldi, D. Gaglione, F. Meyer, F. Hlawatsch, P. Braca, A. Farina and M. Z. Win

7. Enabling situational awareness in millimeter wave massive MIMO systems
R. Mendrzik, F. Meyer, G. Bauch, M. Z. Win

8. Arms smuggling investigation: contextual reasoning and uncertain information fusion with evidential networks
P. Kowalski, A-L. Jousselme

9. Automatic object classification for low-frequency active sonar using convolutional neural networks
G. De Magistris, P. Stinco, J. R. Bates, J. M. Topple, G. Canepa, G. Ferri, A. Tesei, K. Le Page

10. Cooperative beam tracking by multi-sensor data fusion in V2X networks
M. Brambilla, L. Combi, A. Matera, D. Tagliaferri, M. Nicoli, U. Spagnolini

11. Combining data from unmanned underwater vehicles and satellites for covert marine rapid environmental assessment
I. Borrione, P. Oddo, A. Russo and E. Coelho

12. AIS support to reduce collision risk for glider operations in high traffic areas
G. Pennucci, M. Oddone, D. Cecchi, B. Garau, L.M. Millefiori, G. Arcieri, P. Braca, A-L. Jousselme

DEMO BOOTHS

FFI
www.ffi.no

MARINETRAFFIC
www.marinetraffic.com

NODAL POINT
www.nodalpoint.com

STUDIOMAPP
www.studiomapp.com

EC MARISA project
www.marisaproject.eu

EC RANGER project
www.ranger-project.eu

1st Maritime Situational Awareness Workshop - MSAW 2019
Science and Technology meet operational needs
Villa Marigola, Lerici (La Spezia), Italy
8–10 October 2019

Wednesday 9 October

08:30	**Registration & Welcome coffee**	*Main Hall*
09:00	Keynote presentation	*Main Conference Room*
	Push and pull in digitalization: Technology drivers for sensor data fusion	
	Felix Govaers, Fraunhofer FKIE	
10:00	*Coffee break*	*Garden*

RADAR SIGNAL PROCESSING — *Main Conference Room*
Chair: A. Farina

10:30	**Synthesis of constant modulus radar signals in spectrally crowded scenarios**	
	A. Aubry, A. De Maio, M.A. Govoni, L. Pallotta	
10:50	**Data fusion engine based on belief propagation in the RANGER experiment**	
	D. Gaglione, G. Soldi, P. Braca	
11:10	**A constrained least squares approach for 2D PBR localization**	
	A. Aubry, V. Carotenuto, A. De Maio, L. Pallotta	
11:30	*Break*	
11:40	**Photonics applied to coherent radar networks for border security**	
	S. Maresca, G. Serafino, F. Scotti, L. Lembo, P. Ghelfi, A. Bogoni	
12:00	**A marine radar dataset for multiple extended target tracking**	
	J.S. Fowdur, M. Baum, F. Heymann	
12:20	**Ship length estimation using common radar field entries**	
	E. Protopapadakis, M. Bimpas, D. Dres	
12:40	*Lunch*	*Garden*
13:40	NITEC Challenge Winner	*Main Conference Room*
	Dual-use of AIS data, combining AIS tracking with	
	social network analysis for increased maritime network awareness	
	Dagfinn Vatne, FFI, introduced by F. Marziani, NATO NCIA	
14:10	*Break*	

ANOMALY DETECTION & BEHAVIOURAL ANALYSIS — *Main Conference Room*
Chair: S. Horn

14:20	**Traffic density measures for mapping maritime patterns of life**	
	K. Landmark, M. Aronsen	
14:40	**A data driven approach to maritime anomaly detection**	
	D. Zissis, K. Chatzikokoladis, M. Vodas, G. Spiliopoulos, K. Bereta	
15:00	**Verification of sensor data in a maritime multi-sensor network**	
	M. Broetje, G. Battistello, M. Ulmke	
15:20	**Aggregated risk assessment from multi-source data fusion**	
	F.C. Daffinà, T. Stahl, D. Quattrociocchi, M. Zavagli, S. Chesworth, R. Migliorini, G. Sear	
15:40	**EC project presentation: MARISA**	*Main Conference Room*
	Maritime Integrated Surveillance Awareness	
	Valeria Fontana	
16:00	*Coffee break*	*Garden*
16:30	**Behavioral analysis for maritime safety**	*Main Conference Room*
	E. Messel, M.G. Bjørndal	
16:50	**Context-enhanced maritime surveillance optimization**	
	S. Horn	
17:10	**Context information analysis from IMM filtered data classification**	
	D. Sanchez Pedroche, D. Amigo Herrero, J. Garcia Herrero, J.M. Molina Lopez	
19:30	Bus pick-up / drop-off	*La Spezia, Piazza Europa & Lerici, Piazza Bacigalupi*
20:00	No-host dinner	*Restaurant Ca' Borghese*
		Via Arena, 1, 19031 Ameglia SP

RANGER

European Commission
Horizon 2020
European Union funding
for Research & Innovation

Wednesday 9 October

DEEP LEARNING AND CLASSIFICATION   *Small Conference Room*
Chair: D. Zissis

10:30	**Vessel identification using convolutional neural networks** A. Reche Martinez, A. Baldacci
10:50	**Operationalizing ship detection using deep learning** T. Brenn, L-P. Gjøvik, G. Rasmussen, T. Bauna, M. Kampffmeyer, R. Jenssen, S. Anfinsen
11:10	**Monitoring marine protected areas using data fusion and AI techniques** K. Bereta, A. Millios, K. Chatzikokolakis, D. Zissis
11:30	*Break*
11:40	**Custom state-of-the-art CNN algorithm for ship detection and segmentation** G. Bottini, M. Corsi, F.C. Daffinà, S. Tilia, T. Stahl, D. Quattrociocchi
12:00	**Industry is the key for R&D transition to operational end-users - An innovative MSA project example powered by AI, led by CS** M. Degache, L. Gardenal, J-C. Charles, V. Marié, B. Soloch, O. Tran
12:20	**Automatic recognition of underwater acoustic signature for naval purposes** E. Artusi, F. Chaillan

BIG DATA ARCHITECTURES & SYSTEMS   *Small Conference Room*
Chair: C. Ray

14:20	**LIDARs usage in maritime operations and ECO – Autonomous shipping, for protection, safety and navigation for NATO allies awareness** A. Pantazis
14:40	**Big data infrastructure for heterogeneous sources and data fusion services applied to maritime surveillance** G. Vella, G. Barone, V. Latino, D. Messina, V. Morreale
15:00	**A cloud architecture for big data analytics and real-time anomaly detection in global maritime applications** M.A. Biscardi, M. Cinque, M. Corsi, F.C. Daffinà, R. Della Corte, A. Di Martino, C. Perrotta, S. Recano, D. Quattrociocchi
15:20	**Development and Integration of coastal surveillance radar system over four seas** A-A. Mert, S. Üzümcü, F. Atay

MARITIME SITUATIONAL AWARENESS AND ETHICS   *Small Conference Room*
Chair: F. de Rosa

16:30	**Redefining situation awareness for the maritime information warfare domain** A-L. S. Lapinski
16:50	**Collaborative Space-based Maritime Situational Awareness (CSMSA)** G. Thomas
17:10	**An ethical framework for maritime surveillance technology projects** S. Sarlio-Siintola, T. Tammilehto, S. Siintola

Thursday 10 October

08:30	**Registration & Welcome coffee**	*Main Hall*
09:00	Keynote presentation **Ocean of Things** John Waterston, DARPA	*Main Conference Room*
10:00	*Coffee break*	*Garden*

MARITIME UNMANNED SYSTEMS — *Main Conference Room*
Chair: R. Been

10:30	**Maritime cyber situational awareness elaboration for unmanned vehicles** O. Jacq, P. Merino Laso, D. Brosset, J. Simonin, Y. Kermarrec, M-A. Giraud
10:50	**Smoothing optimal RRT* trajectories for recovery of an AUV by a moving surface vessel** S. Williams, X. Wang, D. Angeley, C. Gilliam, B. Moran, T. Jackson, R. Ellem, A. Bessell
11:10	**How to measure distance between the elements of an underwater robotic swarm by power LEDs in unknown sea water conditions** R. dell'Erba
11:30	**Copter currents high resolution surface current fields measured by a small UAV** J. Horstmann, R. Carrasco, M. Streßer

11:50	**EC project presentation: COMPASS2020** **Coordination of Maritime Assets for Persistent and Systematic Surveillance** Nexhat Kapidani	*Main Conference Room*
12:10	*Lunch*	*Garden*

PERSISTENT SURVEILLANCE — *Main Conference Room*
Chair: G. Soldi

13:10	**Spire / ICEYE – Arctic maritime awareness platform** F. Guillaume, Q. Gollier, C. Winquist, S. Andersson
13:30	**BalSAR: A stratospheric balloon-borne SAR system and its use for maritime surveillance** M. Martorella, E. Aboutanios
13:50	**Persistent maritime surveillance against underwater contacts using a wave gliders: Fleet composition and effectiveness** R. Kessel, C. Hamm, M. Taillefer

14:10	**EC project presentation (remote): ARCSAR** **Arctic and North Atlantic Security and Preparedness Network** Bent-Ove Jamtli	*Main Conference Room*
14:30	CMRE presentation	*Main Conference Room*
15:20	*Coffee break*	*Garden*
15:50 17:00	Panel & Closure	*Main Conference Room*

Legend

| |
|---|
| Keynote and plenary presentation |
| Regular sessions |
| Poster, demo and panel sessions |
| EC project presentations |
| Social events |

RANGER

European Commission | Horizon 2020 European Union funding for Research & Innovation

11.2 Survey

General satisfaction
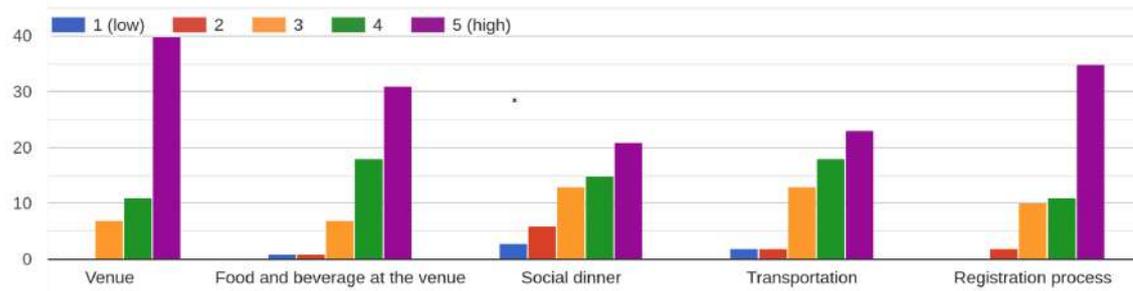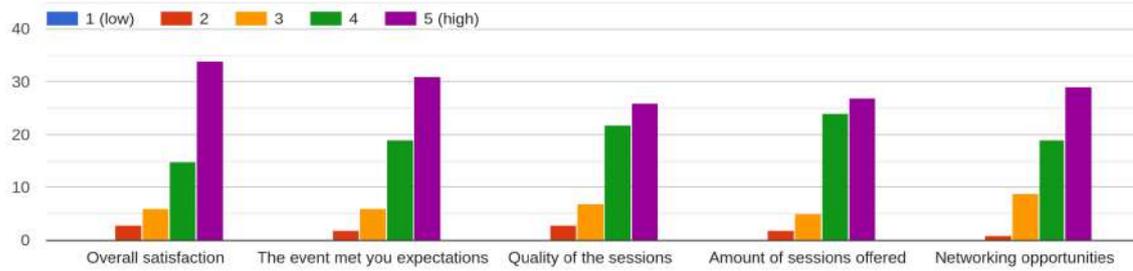
Given your experience, how likely would you participate in future editions?

58 responses

- Very likely
- Likely
- Medium likelihood
- Unlikely
- Very unlikely
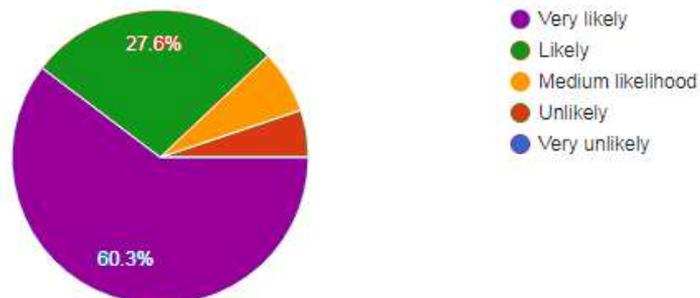
Topics for future events

Which topic would you like to see covered in future events?[1]

- Operational Aspects of MSA; aspects of the Navys / NATO in addition to the coast guard
- Operational aspects and human-machine teaming
- Interoperability
- Something more on AUV
- Space-based MSA
- Collaboration in space for international global maritime awareness, global info sharing and collaboration for maritime security
- More artificial intelligence. Any topic related to naval architecture
- Maritime exploration
- Data Protection and Privacy
- More MSA operational implementations (actual usage within C4IS)
- Unmanned vessels
- Earth observation / Synthetic Aperture Radar
- Maritime security
- Radar architectures for sea surveillance
- Awareness of rubber small craft
- Data mining
- Distributed detection
- The environmental monitoring aspect
- Make the link between research, operation and business more obvious
- Machine Learning and Deep Learning
- Some more theoretical exercise from academia
- Situational Awareness of Military Higher Echelon Units and synergies to Civil MSA
- More interaction between AI and Tracking, Data Fusion principles
- Intersection of Game Theory, AI and ML for maritime
- Anomaly detection and pattern of life recognition
- Operational needs for port surveillance
- Underwater robotics and communications
- Radar
- Anti-piracy with sensors on-board commercial ships
- Satellite Observation
- Search and rescue

---

[1] Most relevant answers among 32 received

11.3 Participation
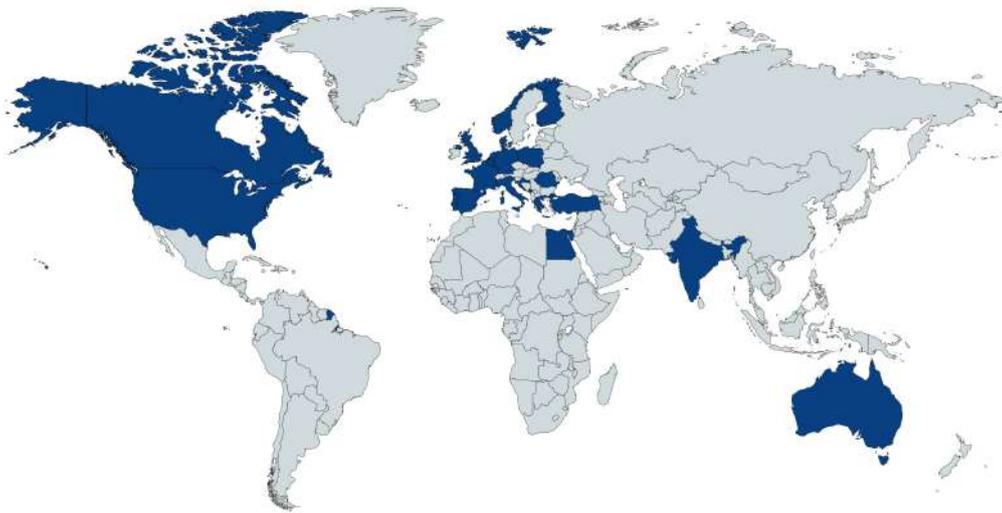
24 countries

Australia	2	Germany	14	Norway	12
Belgium	1	Greece	15	Poland	2
Canada	6	India	1	Portugal	3
Croatia	1	Italy	43	Romania	2
Denmark	2	Luxembourg	1	Spain	9
Egypt	1	Mauritius	1	Turkey	3
Finland	2	Montenegro	1	United Kingdom	7
France	20	Netherlands	3	United States	17

170 attendees

Document Data Sheet

Security Classification		Project No.
RELEASABLE TO THE PUBLIC		DKOE

Document Serial No.	Date of Issue	Total Pages
CMRE-CP-2020-001	January 2021	364 pp.

Author(s)

Anne-Laure Jousselme, Paolo Braca, Leonardo Millefiori, Francesca de Rosa, Maximilian Zocholl, Murat Uney, Clément Iphar, Domenico Gaglione, Giovanni Soldi, Nicola Forti, Enrica D'Afflisio

Title

Proceedings of the Maritime Situational Awareness Workshop 2019

Abstract

The Maritime Situational Awareness Workshop (MSAW) was held 08 – 10 October 2019 in Lerici, La Spezia (Italy). CMRE organized this workshop with the objective of fostering the cross-fertilization of ideas from scientific and military domains, toward the design and implementation of future solutions tailored to Maritime Situational Awareness (MSA) operational needs. Centred on the theme of Science and technology meet operational need, the call for papers was distributed amongst the international MSA community during the early months of 2019. In response, CMRE received 48 submissions and several contributions from the European Commission (EC) funded MSA projects. The workshop brought together about 170 participants from 23 countries including scientists, engineers, researchers from scientific communities with national and international authorities, end users and operators, and industrial representatives. Overall the audience was composed of 14% academia, 39% applied research institutes, 31% industry and 16% operational community. This document contains the proceedings of the MSAW'19 papers accepted for presentation.

The MSAW'19 was co-sponsored by the EC project RANGER (Radars Long Distance Maritime Surveillance and SAR Operations) and NATO Allied Command Transformation as part of the CMRE Data Knowledge Operational Effectiveness programme.

Keywords

Maritime Situational Awareness, Radar signal processing, Artificial Intelligence, Anomaly detection and behaviour analysis, Deep learning and classification, big data architectures and systems, MSA and ethics, Maritime Unmanned Systems, Persistent surveillance

Issuing Organization

NATO Science and Technology Organization
Centre for Maritime Research and Experimentation

Viale San Bartolomeo 400, 19126 La Spezia, Italy

[From N. America:
STO CMRE
Unit 31318, Box 19, APO AE 09613-1318]

Tel: +39 0187 527 361
Fax:+39 0187 527 700

E-mail: library@cmre.nato.int