Proc. of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)
16-18 November 2022, Maldives

Application of an Automotive Assurance Case Approach to Autonomous Marine Vessel Security

Luis-Pedro Cobos
Research Engineer Connected Autonomous Vehicle Safety and Security
Horiba-MIRA
England, United Kingdom
luis-pedro.cobos@horiba-mira.com
ORCID: 0000-0002-1333-7767

Tianlei Miao
Research and Engineering Department of Electrical Engineering, Waves: Core
KU Leuven
Bruges, Belgium
tianlei.miao@kuleuven.be
ORCID: 0000-0002-8669-0465

Kacper Sowka
Centre For Future Transport And Cities
Coventry University
England, United Kingdom
sowkak@coventry.ac.uk
ORCID: 0000-0001-6954-8865

Garikayi Madzudzo
Cyber Security Lead
Mobility Innoavtion Hub
Horiba-MIRA
England, United Kingdom
garikayi.madzudzo@horiba-mira.com
ORCID:0000-0002-1156-3875

Alastair R. Ruddle
Vehicle Resilience Technologies
VRES
Horiba-MIRA
England, United Kingdom
alastair.ruddle@horiba-mira.com
ORCID: 0000-0003-4425-0979

Ehab el Amam
ASV Consultant
RH Marine
Rotterdam, Netherlands
Ehab.elamam@rhmarine.com

Abstract—The increase of autonomy in autonomous surface vehicles development brings along modified and new risks and potential hazards, this in turn, introduces the need for processes and methods for ensuring that systems are acceptable for their intended use with respect to dependability and safety concerns. One approach for evaluating software requirements for claims of safety is to employ an assurance case. Much like a legal case, the assurance case lays out an argument and supporting evidence to provide assurance on the software requirements. This paper analyses safety and security requirements relating to autonomous vessels, and regulations in the automotive industry and the marine industry before proposing a generic cybersecurity and safety assurance case that takes a general graphical approach of Goal Structuring Notation (GSN).

Keywords— Autonomous marine vessels, Autonomous Driving, Assurance Case, Vehicle Safety, Cybersecurity, Automotive Standards

I. Introduction

The degree of automation within navigation systems means that many manned vessels operate autopilot, ensuring that the journey adheres to the planned trajectory. This is enabled by connecting GNSS (Global navigation satellite system) receiver with autopilot and gyro-compass, which passes all relevant information to the autopilot. Then, a new course can be computed and compared with the current course, which establishes the basis for calculating the correction to the rudder/thruster. An autonomous sailing system principally consists of environment perception, situational awareness, decision making, and controlling, which enables more intelligent functions like environment detection, obstacle recognition, collision avoidance, and control.

An effective means of perceiving the environment is a necessary precondition of safe sailing. For Autonomous or Surface Vehicles (ASV, aka. Autonomous Sailing Vessels) typically, there are multiple sensors used onboard, such as the automatic identification system (AIS), through which ships can broadcast static information (e.g., identification number, ship name), dynamic information (e.g., position, speed), and other relevant information to each other. In addition, the vessels can utilise exteroceptive sensors such as radar, lidar, and camera. In conjunction with these, AI-based methods are usually required for the obstacle recognition, classification, and tracking. Relying on a single sensor has limitations since different types of sensors have different cover ranges, applicable scenes, and measurement accuracy, thus sensor fusion can lead to a better perception by utilizing data from multiple sensors.

Another important aspect of the ASVs is a robust collision avoidance system, with the effective collision avoidance algorithms to ensure the safety during sailing. In addition to its primary purpose, collision avoidance also includes other concerns such as the dynamics of the ship, external environmental disturbances, movement and intention prediction of target vessels, and the compliance of International Regulations for Preventing Collisions at Sea (COLREGs)[1]. Considering all of the above, this usually implies a significant computation load, especially in complex scenarios. Ensuring real-time performance is a critical factor in a practically viable system.

An assurance case is a document that outlines the process and methodology that will be used to provide assurance on the software requirements. It is important to include a high-level overview of the system and its components, as well as a general description of the environment in which it will be operating. The assurance case should also include any assumptions that have been made, such as those pertaining to development methodology, user interface, or external dependencies.

With regard to the automotive industry, this example can be seen as analogous to a vehicle with SAE level 4 autonomy (for reference Table 1). The naval vessel industry does not have standards similar to Safety of the intended function, or functional safety, neither do they have an exclusive cybersecurity standard. Using the FMEA and the attack tree it is possible to extrapolate these concepts and help build on the assurance case. Despite the lack of a specific standard, these elements are still taken into consideration throughout the life cycle and there is documentation in place to certify any claims made.

Table 1: Society of Automotive Engineers (SAE) classification for the levels of Autonomy

Feature	Driver Support Features			Automated Driving Features		
Levels	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
Control scope	Warnings and only momentary assistance	Steering OR brake/ throttle support	Steering AND brake/ throttle support	Can drive vehicle for limited conditions	Can drive vehicle for limited conditions	Can drive vehicle under all conditions
Driver input	Always required	Always required	Always required	May be requested to resume	Not required under some conditions	Never required
Sample systems	Blind-spot warning, AEB, LDW	Lane keeping OR ACC	Lane keeping & ACC	Traffic jam pilot	Automated local taxi (fixed route)	Driving anywhere, under all conditions

II. AUTONOMOUS SURFACE VEHICLES

ASVs are gaining increasing attention worldwide due to the potential benefits of improving operational safety and efficiency. An ASV is defined as a vessel that has achieved some level of autonomy in its deployment[2]. The system on board or at a remote site can assist human operators in navigation or completely take over the decision-making and control during the sailing. Lloyd's Register defines six autonomy levels (AL) for ASVs, ranging from AL1 for ships with data collated for onboarding decision making through to AL6, which denotes a fully autonomous ship with autonomous decision-making capability enabling itself to finish tasks without any crew [3]. To achieve higher ALs, many studies and commercial projects have been carried out in the last decade. E.g., the Norwegian University of Technology and Science conducted a project, Autosea, aiming at sensor fusion and collision avoidance for ASVs [4]. Another research project, Remote Operations of Machinery and Automation Systems (ROMAS) was established by DNV GL, Høglund, Fjord1, and the Norwegian Maritime Authority in 2018 to explore the idea of moving the engine control room from the ship to a shore-based center [5].

In contrast to the automotive industry, where the autonomous system is typically composed of three parts: perception, decision making, and control [6], autonomous sailing systems follow the architecture of the guidance navigation control (GNC) system [7]. Although these systems are designed differently, their functionalities are similar. The navigation sub-system collects information from sensors to support the guidance system. The guidance system is engaged to detect and solve the conflict simultaneously, namely decision-making. The control system implements the planned actions. Plus, additional functional modules can be added to achieve full autonomy, such as environmental modeling, localization, obstacles detection, classification, and tracking.

Since adequate perception of the environment is a necessary precondition for safe sailing, an ASV typically boasts various exogenous sensors used onboard, such as radar, LiDAR, and camera [8], to perceive the surrounding environment. Radar uses a rotating antenna to sweep a narrow beam of microwaves around the water surface surrounding the ship to the horizon, detecting targets by microwaves reflected from them. Usually, radars with multiple bands are used to cover different ranges and distances [9]. Further, LiDAR and cameras/infrared cameras have become popular for ASVs in conjunction with deep learning (DL) methods due to their success in the application of autonomous driving [10][11][12]. Additionally, the automatic identification system (AIS) is a ship-tracking system that uses transceivers on ships and is used by vessel traffic services (VTS) [13]. Ships can broadcast static information (e.g., MMSI number, IMO number, ship name), dynamic information (e.g., position, speed), and other relevant information through AIS. Endogenous sensors, such as IMU and GNSS, are used to locate the own ship. Although these sensors have already been in practical use for navigation, a single sensor always has limitations due to its limited cover range, applicable scenes, and measurement accuracy. Therefore, sensor fusion is further required to combine the functionalities of various sensors effectively [14][15]. All the perception steps aim to obtain a better and broader detection of the environment, which enhances the situational awareness of the autonomous sailing system.

Another critical aspect of the ASVs is the collision avoidance (CA) module with effective and robust CA algorithms [7]. With the processed data from the perception module and situational awareness of the current sailing situation, the CA module is designed to provide a safe and feasible solution. Apart from the above primary purpose, CA also includes other concerns such as the dynamics of the ship in question, the external environment disturbances and movement and intention prediction of target ships. Furthermore, the compliance with International Regulations for Preventing Collisions at Sea (COLREGs) makes it more challenging [16]. In addition, as a computational algorithm, it usually means a significant computation load, especially in complex scenarios, after considering all the above concerns. Even though many studies have been conducted to propose effective and efficient CA algorithms [17][18], designing a CA system considering all aspects is still challenging[19].

Conventional module engineering deals with these modules separately including the data processing and transmission between each module. An information flow diagram of an ASV is presented in Figure 1. Raw data from sensors need to be pre-processed, e.g., filtering and enhancing, and then passed to other modules. Methods, such as artificial intelligence (AI) technologies are used for follow-up obstacle recognition, classification, tracking, and collision avoidance. Another way is end-to-end (ETE) learning [20], which trains a neural network using a certain data format (e.g., raw data from radar or LiDAR) as input and outputting the desired results directly (e.g., obstacle detection results or the control commands directly). All the parameters are trained jointly. ETE learning can be used for functionalizing a single module or a wrap of multiple modules. However, low explainability is a common problem that makes it not fully accepted by the industry yet.

There is no doubt that ASVs have become a part of the digital world, with owners and equipment manufacturers able to remotely monitor, manage and actively intervene in maintenance and operations. However, to achieve complete autonomy, there is also a new challenge in the transmission, processing, analysis, and storage of the data. Real-time information gathering from the surroundings is mandatory, and communications between each module, as well as the ship and land, must be broadband and have low latency. Redundancy is usually added over the distributed system to cope with the damages and errors of individual hardware and modules. The whole autonomous system should also be sufficiently robust to the different scenarios and potential hazards.

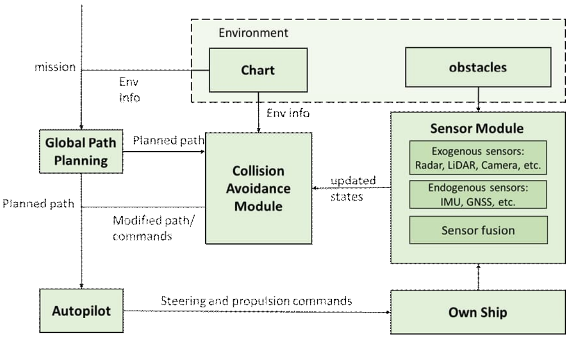

Figure 1: Information flow of an ASV [6,18]

III. Safety and Security

A. Marine Industry

Ships are the most secure and ecologically friendly mode of commercial transportation[21]. Almost all shipping operations have long been committed to safety, due to the innate possibility of danger. The shipping industry was one of the first to establish universally accepted worldwide safety standards. Because shipping is fundamentally international, it is governed by several United Nations bodies, including the International Marine Organization (IMO), which has produced a comprehensive set of worldwide maritime safety laws. On the other hand, all governments demand that ships and other maritime constructions flying their flag meet specific criteria. A non-governmental regulating organization known as classification society, is an organized operation in the marine industry, the vessel and offshore structure development. The society oversees developing rules for the design and classification of ships and offshore buildings. IMO and the Governments are influenced by classification societies, that appeal to the user base, when deciding on regulations. Ship surveyors, naval architects, and a wide range of certified marine engineers are used by classification societies. These professionals oversee monitoring ship building and maintenance, as well as conducting required surveys of ships currently in operation to ensure that standards are fulfilled[22]. To improve stability, safety, and cleaner emissions, classes are designed to control the structure and design of all vessel types. To that aim, classification societies agree on technical standards, supervise designs, and double-check computations to guarantee that the rules are followed. Qualified personnel are assigned to inspect ships and structures during construction and commissioning, as well as to survey vessels (including submarines) on a regular basis to verify that they continue to meet all requirements. They are also in charge of classifying offshore constructions such as oil rigs, platforms, and other structures. Propulsion systems, navigation devices, pumps, valves, and other equipment are all included by this survey. The 3 largest classification societies are DNV, (Det Norske Veritas,) Nippon Kaiji Kyokai, and ABS (the American Bureau of Shipping)[23]. The main Regulations followed by all classification societies are:

- SOLAS (International Convention for the Safety of Life at Sea) Established in 1974 and updated to date defines and clarifies the minimum standards of safety equipment on board
- MARPOL (International Convention for the Prevention of Pollution from Ships) Available since 1978 discloses the requirements necessary to prevent pollution from occurring yet accidental or as a result of routine operations
- COLREG (Convention on the International Regulations for Preventing Collisions at Sea) Accepted since 1972, and mainly achieves what the road safety rules do on cars.
- ISPS (The International Ship and Port Facility Security Code) From 2002 onward is the gold standard of marine vessel safety it has been updated to also include cybersecurity.

B. Comparison with Automotive Industry

After analyzing the safety and security concerns surrounding autonomous vessels, it has become clear that there are many overlapping aspects with the automotive industry. Thus, it may prove beneficial to integrate methods and approaches utilized within the automotive industry in order to enhance the safety and security assurance practices within the autonomous vessel industry. In particular, there is tangible potential in extrapolating the cybersecurity case integrating safety concerns as used in the automotive domain (cite) for use within the assurance process of autonomous sailing vessels.

Regarding safety the automotive industry is bounded by the UN-ECE type approval regulations plus ISO 26262 for Functional Safety and ISO 21448 for Safety of the intended function (SOTIF). These regulations also not just are legally binding but needed to compare and be compatible with all other vehicles. [24] Other than the regulations automotive vehicles are set on consumer tests like NCAP to see if they are more than the standards and regulations and how they rank. So even when there are many regulations, standards, and consumer tests on automotive, marine manages to achieve the same objective through its organizations and has been doing it since before cars. Tools like a FMEA (Failure Modes and Effects Analysis) are used in vehicle safety in addition to a HARA (Hazard Analysis and Risk Assessment), ASIL (Automotive Safety Integrity Level), as for functional safety (FS) these tools are applied to a Goal Structured Notation (GSN) safety case. Considering that for the current scenario there is an existing FMEA, it will be used to create the safety objectives ala FS trying to not forget the vital parts of SOTIF[25]. Therefore, by narrowing the ideas from the FMEA it is possible to convert them into goals or claims that fit the method for the assurance case. The FMEA will also help to counterclaim and argument these goals and eventually deduce evidence regarding its status.

Security and cybersecurity specifically become harder to manage as technology and automatization develop farther, the vehicle industry normally follows ISO21434 and the mandatory UN regulation 155. Both the standard and the regulation require a safety case, and to do so a holistic approach that can consider the links with safety is effective. When it comes to the marine industry, IMO Resolution MSC. 428(98) introduced in 2021 has various cybersecurity requirements for ships. The main difference is that IMO requires the owners to assess the cyber risks while the automotive industry expects it from the developers. With an attack tree, things that are linked to security become more apparent and logically perceivable. The threats of an attack tree can seamlessly transfer into the assurance case.

IV. Methods and Tools

It is considered that an assurance case for automotive cybersecurity should ideally provide the following basic characteristics and requirements. This is based on emerging technological trends in the automotive industry, as well as current functional safety engineering practices. It should address aspects beyond traditional safety, including availability of mission-critical (rather than safety-related) functions, privacy issues, fraudulent financial transactions, and indirect safety implications (such as kidnapping). The aim of this is to produce a more explicitly "adversarial" case than has traditionally been used in functional safety, in a similar style to the way legal cases are presented and examined in a court of law. It is expected that such a cybersecurity case would effectively be the first draft of a dynamic assurance case that would be updated through the operational life of the vehicle. Ongoing assurance activities will also be needed to complement the product launch assurance, in order to ensure that cybersecurity assurance is maintained over the operational lifetimes of vehicles as outlined in UNECE regulations and ISO/SAE 21434.

Systems will need to cope with the inevitability of threats that were unforeseeable at design time, and must be able to address aspects beyond traditional safety concerns. These include availability of mission-critical (rather than safety-related) functions, privacy issues, fraudulent financial transactions, and indirect safety implications (such as kidnapping), among others. The goal is to create a case that is more openly "adversarial" than what has previously been utilized in functional safety, comparable to how legal cases are presented and reviewed in a court of law. This cybersecurity case is planned to serve as the initial draught of a dynamic assurance case that will be updated during the vehicle's operating life. In addition to product launch assurance, ongoing assurance operations will be required to guarantee that cybersecurity assurance is maintained throughout the operating lifespan of vehicles, as defined in UNECE standards and ISO/SAE 21434.

The concept of risk is a combination of the likelihood of an event and the severity of its impact on the stakeholders, such that low severity with a low likelihood represents a low risk and high severity with a high likelihood represents a high (and probably unacceptable) risk. In cybersecurity, threats, and attacks on the vehicle, perpetrated by malicious individuals, could lead to a variety of possible outcomes, including safety impacts as well as non-safety risks. Risk analysis is the process of identifying and analyzing potential issues that could have a negative impact for the stakeholders. It should be noted that the severity of impact can only be assessed at system level, where the impact on the stakeholders can be assessed, whereas the likelihood depends on the individual likelihoods of actions in the chain of events that lead to the specific outcome.

A defining process and what made the cybersecurity case something relevant to OEMs is ISO/SAE 21434 and UNECE Regulation 155, that is heavily backed up by the ISO standard. And even though the standard does not actually identify a specific method to build the case it has a clear objective by conveying that it shall be created to provide the argument for the cybersecurity of the item or component, supported by work products, and that it can be created by combining customer supplier cybersecurity cases but most also support post-development. The standard has a clause focused on operations and maintenance, in this section the relation of prerequisites for post development are relevant, and how it must be used in the instant of delivering updates[26][27]. Section 6 specifically mentions the need of a cybersecurity case while Section 9 details the concept phase, by (1) defining the operational environment, (2) specifying the cybersecurity goals and claims and (3) specifying the cybersecurity requirements. Section 10 aims to verify the cybersecurity requirements, identify and manage vulnerabilities; and provide the evidence that it complies with cybersecurity.

A. Attack Trees

Attack trees are a long-established method for aiding the security analysis of a wide variety of systems, particularly within the IT domain with respect to cybersecurity. Within the automotive domain, attack trees have proven to be a popular method of performing aiding threat enumeration and related activities such as risk assessments and even penetration testing, with the attack tree being explicitly mentioned in the ISO/SAE 21434 standard. Briefly, attack trees consist of a directed acyclic graph which breaks down nebulous attack goals into more immediately understandable individual components following a hierarchal tree structure[28]. This allows the analyst to more intuitively understand how individual threats or vulnerabilities can combine to facilitate larger scale attacks targeting a wide variety of damage scenarios.

As an extension to the base attack tree, Attack-Defense Trees (ADT) have also become a popular technique for analysing threats and identifying counter measures in cybersecurity. This provides a methodical, graphical way of modelling possible cybersecurity threats to systems while also enabling the explicit consideration of countermeasures and the attacks which can be mounted against them, allowing for a more expressive and wider-reaching cybersecurity assessment. Like attack trees, ADTs are represented in a logical way and follow the node flow in one direction. For each node of an ADT that has more than one branch the relationships between the subsidiary branches may be either disjunctive (OR) or conjunctive, the latter using either a simple AND or a sequential AND (SAND). The SAND approach provides a more compact representation a specific sequence of steps that may be needed to mount an attack. ADT techniques have many advantages, for example they are easy to understand and can be easily shared and explained to people with little experience in security, and can often be reused to address similar threats[29].

B. *Goal Structured Notation and Confidence Maps*

The GSN safety case structure works towards a “goal”, which is a claim, through a strategy and context, which are arguments, then leads to a solution, that is evidence. The “vanilla GSN” mentioned also has the possibility of extensions that make the safety case building more effective. These extensions include maintenance of arguments, modular safety cases, assurance case patterns, eliminative argumentation and more. As a technique that is easily understood and able to present advanced concepts, it has been appearing with increasing frequency in the domains of safety and security [30].

In eliminative argumentation there are three potential types of doubt; doubts about the claim, doubts about the evidence, or doubts about the inference used to link the claim and the evidence. The objective is therefore to identify the relevant doubts about claims, evidence and inferences, and to provide counter arguments against the identified doubts where possible to increase confidence in the assurance case[31].

A graphical representation of an eliminative argument is described as a confidence map, as it details the identified doubts concerning an argument and also shows whether these doubts can be countered or if they remain, thus illustrating the confidence that can be attributed to the argument. It should be noted that not all doubts will have the same importance, and appropriate weightings should be. To maintain the clarity of the safety case it has been recommended that the confidence map should be separate from, but linked to, the safety case [32].

V. Proposed Methodology

A generic Cybersecurity Assurance Case is proposed here that takes the general graphical approach of GSN, whilst applying the additional symbolism of confidence maps from eliminative argumentation[31], and also integrating the structure of ADTs to augment and amplify the arguments. The symbols used in the generic illustration presented here are summarized in Figure 2, with the meaning behind each symbol elaborated below.

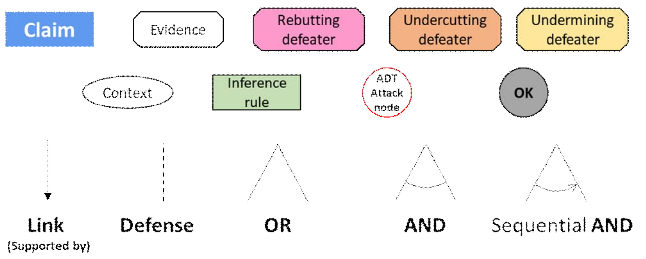

Figure 2 Key to symbolism used in the proposed methods

Arguments taking account of the potential for unforeseeable cybersecurity risks and the requirements for a through-life cybersecurity management system (denoted by CSMS) are present in the method. The system description, which provides the context for the assurance case, is indicated by a white ellipse. The claims are represented by blue rectangles, which are justified by inference rules represented by green rectangles. The claims may be challenged by rebutting defeaters, the inference rules by undercutting defeaters, and the evidence by undermining defeaters[33]. These challenges may be responded to by using further claims, inferences, and evidence. The need for a robust threat analysis approach is also included, along with the treatment of a threat judged to be of inherently acceptably low risk, and considering a cybersecurity management system (CSMS). Lines of argument that are considered to have been acceptably resolved are indicated by the grey circles.

The diagram method illustrates approaches for threats involving attack trees that could contain disjunctive and conjunctive relationships (the latter including both simple and sequential AND possibilities). These could be addressed either by outright elimination of possible attack steps (denoted by "ATK" and represented by circles with red boundaries), or at least mitigating the vulnerability to reduce the anticipated likelihood of success to a sufficiently low level to achieve an acceptable level of residual risk. These requirements for defense are indicated by dashed lines terminating in claims for possible successful elimination (denoted "ELIM" in a blue box) or mitigation countermeasures (denoted by "MIT" in a blue box), which therefore represent sub-claims that must be supported by appropriate evidence.

It should be noted that if an OR node occurs in an attack tree fragment then all of the options must be addressed with suitable countermeasures in order to achieve complete resolution. If there is an AND or SAND node, however, then mitigating any of the contributors could be sufficient to achieve the necessary risk reduction.

To understand the order and usage of the method, here are some simple guidelines and an easy way to remember the shapes and usage of the method:

- A context {white oval} can be placed anywhere it defines or limits the contextualization, by formatting the statements and ideas that limit and contain argument.
- A goal {blue rectangles} are the claims of an argument and the main debatable object of this assurance case. They are the initial structure and continue to be the main point of diversion within the assurance case. When an ADT branch happens, mitigations and defenses become claims, and therefore goals, by themselves. They are followed by inferences or rebutting defeaters.
- Evidence {White polygons} provide factual information that validate. They are bound to follow goals or defeaters before being validated by an “OK”
- Inference Rules {green rectangles} present a possible scenario that support a claim and tend to start using the word “if”. As the name implies an inference rule is a condition that is supporting a claim through reasoning the basis of the evidence. They follow either goals or rebutting defeaters.
- Rebutting defeaters {pink polygons} challenge the claims, by claiming a counter-condition or scenario, they tend to start with the word “Unless”. They follow claims and are followed by inferences when an acceptable condition can be deduced or

inferred; or may be followed by an ADT where necessary to develop the argument through attack.

- Undercutting defeaters {Orange polygons} challenge the inferences and tend to start with "but", they aim to point out how the inferences could be less of effective or weaken the logical conclusion on which the inference was reasoned. They follow inference rules and are to be followed by evidence to clarify. They follow inferences and are to be followed by another inference or evidence.
- Undermining defeaters {Yellow Polygons} challenge the evidence and tend to start with "However" or "Still". The lessen the effectiveness, power, or ability of the evidence by pointing out how that evidence might be void. They follow evidence and should be followed by further evidence.

The idea is that this approach could be implemented into a future cybersecurity case framework, and more specific examples are developed in the line of the research.

A. *Requirements*

The method explained is a holistic assurance case that considers elements of safety and security, to produce a logical and efficient diagram to follow. The process leading to be able to generate all of these requirements relies on organizational safety and security, the culture of safety and security within manufacturing process and proper life cycle management from concept phase till postproduction phase; this industrial and organizational processes are bigger than the scope of this document, that just aims to focus on a new approach on assurance cases. As a regular assurance case it requires some information bestowed upon it before hand, then it can be summarized in the following points:

- Identifying hazards: As implied if there are no hazards identified there will be no assurance case as the goals and claims aim to ensure they are as hazardless as possible. (Activities and methods necessary could be, but not limited to: HAZOP, Fault Tree Analysis, etc.)
- Extent of harm: Deciding who might be harmed, when, and how. (Activities and methods necessary could be, but not limited to: Asset definition, damage scenarios, ASIL analysis, etc.)
- Evaluation of risks: All the risks associated with the identified hazards. (Activities and methods necessary could be, but not limited to: Risk Matrix, TARA, Risk determination analysis etc.)
- Mitigation strategies: Deciding on the necessary control and security measures, necessary to counter, deter, stop or mitigate threats. (Activities and methods necessary could be, but not limited to: Attack path and feasibility rating, adversary-driven state-based system security evaluation, quantitative cyber risk reduction estimation methodology, etc.)
- Recording Evidence: As not all evidence should be implied, many should come from statistics or testing. Documentation must be clear on recording those findings and implementing them. (Activities and methods necessary could be, but not limited to: Vehicle & component Testing, Statistical Data, etc.)
- Evaluating and monitoring: Progress and changes should be maintained under observation as ongoing basis, any change should be considered and assessed properly in accordance with all previously mentioned requirements. (Activities and methods necessary could be, but not limited to: product Life cycle assessment, product update process, etc.)

Even when the above-mentioned requirements are essential for building an assurance case it is also important to consider legal requirements and following certain standards[27].

B. *Step by Step Build up*

In order to make the process tangible and the possibility to create an assurance case in this style, three phases are to be stablished. The second phase or building phase should be repeated iteratively until the process is completed.

1) *Start Phase*

This phase is the beginning part, it is important to set the main goal and have all the sub goals clear, plus all context must be defined. Information for the goals, context and any argumentation or evidence should be clear before proceeding to the next step.

2) *Building phase*

This phase is made of several steps, these steps are to be repeated iteratively until all branches are concluded. In this part of the process all open elements must be reconsider until all branches are deemed OK, and therefore closed.

1. Interconnect any non-connected context that is relevant to the "Goal" {Blue rectangle} on the topmost part that it is not yet completed.
2. Set any strategies in the form of inference rules {Green Polygon} relevant to this "Goal"
3. Put any rebuttal that the claim "Goal" may face as a rebuttal defeater {Pink Polygon}
4. Check if any inference can be undercut by an undercutting defeater {Orange polygon}; when an undercutting defeater follows an inference, the inference gives context and works as a strategy, like in the vanilla GSN, letting the undercutting defeater be more direct.
5. Try to link any relevant evidence {White Polygon} to any open defeater or inference
6. Check if the evidence can be undermined if it can be it shall be placed with an undermining defeater {Yellow Polygon}
7. Review open rebuttal defeaters if they invoke a possible attack path develop the attack tree regarding it
8. Develop the attack tree until attacks are mitigated or eliminated, any mitigation or elimination of a threat will become a sub goal {Blue rectangle} and should be reviewed next as a starting goal.
9. Check if there is any evidence {White Polygon} or inference {Green Polygon} relevant to an open defeater and link it.

10. Check if evidence that could not be undermined or inferences that where not challenged can be deducted as safe as possible with an “OK” {grey circle}
11. Make sure the current branch has reached an “OK” and restart the building phase from a goal in any unclosed branch.

3) Verification and maintenance phase

During this phase it is important first to review the resulting diagram everything has been correctly validated with paths being ensured with an “OK” through validation or deduction. Any change, upgrade or update of the system should be analyzed in this phase and when required return to the building phase.

VI. Example using the methodology

This example was created in a secondment at RH Marine where the inner workings of an ASV were analyzed and the documentation of an attack tree and a FMEA were used to produce the example. It consists of an AI that should maintain a self-diagnose system capable of understanding and identifying threats, risks and hazards.

A. Marine Vessel SafetyIdentification of potential hazards

This section identifies the safety hazards, in reference to any hazard that may cause a physical harm or loss inherently from the AI. The main hazards are the following:

- The loss of the Vessels Ability to maintain its position
- Problems relevant to the ability to self-drive
- Unavailability of the control systems
- Communication problems
- Failure of collision avoidance (Other vessels and shore)
- Non functionality of alarms and problem detection systems
- Incorrect lighting or identification
- Unreadiness to depart from docked
- Safety of the crew and evacuation methods

B. Marine Vessel Security: Identification of Threats

Regarding security threats the most important factors to consider and that are the backbone to an attack tree regarding these issues are:

- Confidentiality Issues: This considers that the information regarding the vessel loses the secrecy of the private information it possesses
- Attacks on the AIS: May be due to spoofing either the closest point of approach (CPA) or Search and Rescue Transponder (SART). Also, an inaccurate understanding of the weather will affect how the AIS responds.
- GPS Attacks: Anything from spoofing to eavesdropping on the location can have dire consequences in helping the vessel orientate
- Radar System Attacks: Any potential attack on the radar system will cause the effect of blinding the vessel
- Access to the Network or Server Issues: The server provides access to information pertinent to decision making, limiting this access disables control and service
- Limiting the Availability: Reducing availability functions
- Problems with the physical bridge or Workstation: This would not permit a manual override to save the vessel.

C. Example Diagram

Description and diagram of the applied assurance case, the example is shown as figure 3 and detailed in the further images

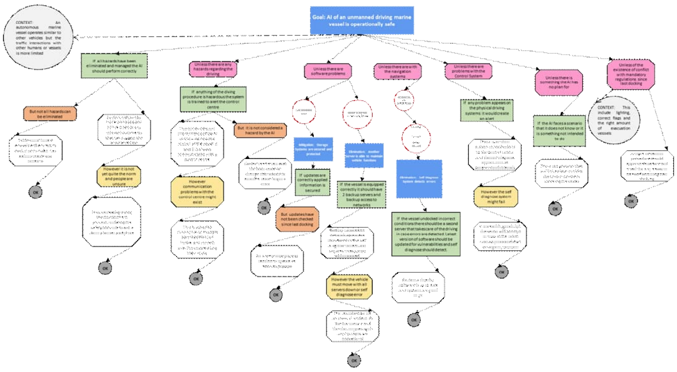
Goal: AI of an unmanned driving marine vessel is operationally safe

Figure 3 Applied method into an assurance case

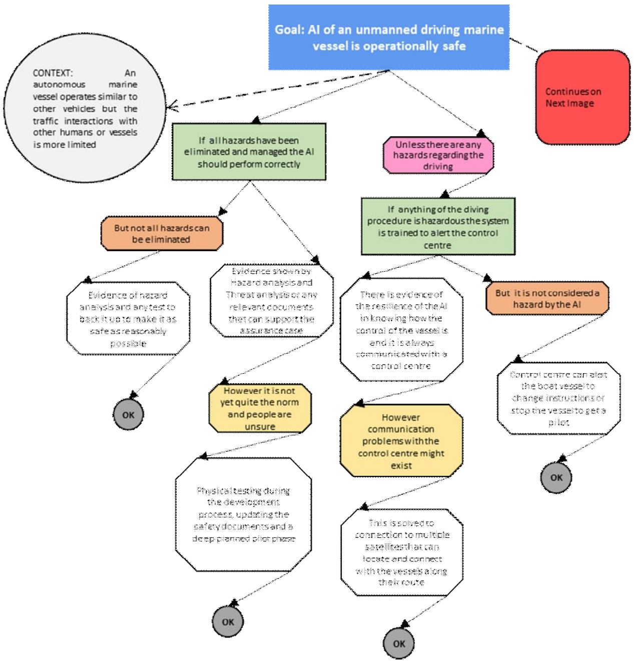
Goal: AI of an unmanned driving marine vessel is operationally safe
CONTEXT: An autonomous marine vessel operates similar to other vehicles but the traffic interactions with other humans or vessels is more limited
Continues on Next Image
If all hazards have been eliminated and managed the AI should perform correctly
Unless there are any hazards regarding the driving
But not all hazards can be eliminated
If anything of the diving procedure is hazardous the system is trained to alert the control centre
Evidence of hazard analysis and any test to back it up to make it as safe as reasonably possible
Evidence shown by Hazard analysis and Threat analysis or any relevant documents that can support the assurance case
There is evidence of the resilience of the AI in knowing how the control of the vessel is and it is always communicated with a control centre
But it is not considered a hazard by the AI
OK
However it is not yet quite the norm and people are unsure
However communication problems with the control centre might exist
Control centre can alert the boat vessel to change instructions or stop the vessel to get a pilot
OK
Physical testing during the development process, updating the safety documents and a deep planned pilot phase
This is solved to connection to multiple satellites that can locate and connect with the vessels along their route
OK
OK

Figure 4 Enhanced detail of figure 3 part 1

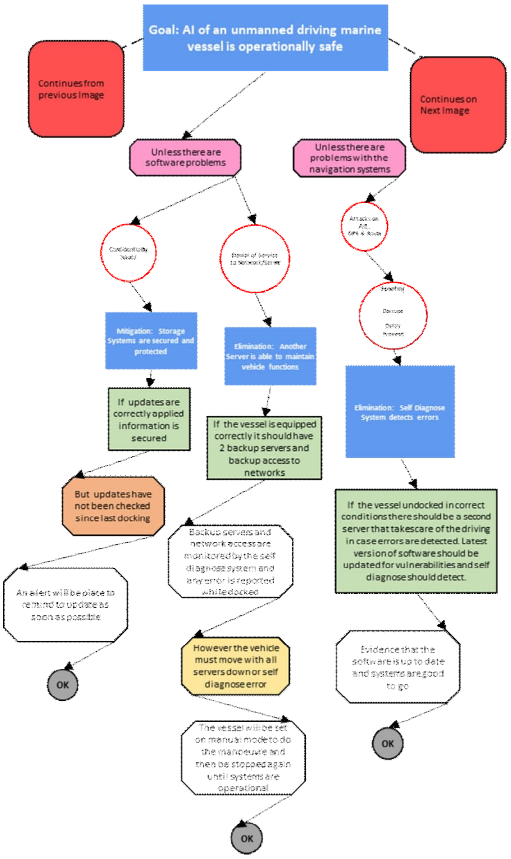
Goal: AI of an unmanned driving marine vessel is operationally safe
Continues from previous Image
Continues on Next Image
Unless there are software problems
Unless there are problems with the navigation systems
Confidentiality Issues
Denial of Service to Network/Server
Attacks on AIS, GPS & Radar
Spoofing
Corrupt
Delay
Prevent
Mitigation: Storage Systems are secured and protected
Elimination: Another Server is able to maintain vehicle functions
Elimination: Self Diagnose System detects errors
If updates are correctly applied information is secured
If the vessel is equipped correctly it should have 2 backup servers and backup access to networks
If the vessel undocked in correct conditions there should be a second server that takescare of the driving in case errors are detected. Latest version of software should be updated for vulnerabilities and self diagnose should detect.
But updates have not been checked since last docking
Backup servers and network access are monitored by the self diagnose system and any error is reported while docked
An alert will be place to remind to update as soon as possible
However the vehicle must move with all servers down or self diagnose error
Evidence that the software is up to date and systems are good to go
OK
The vessel will be set on manual mode to do the manoeuvre and then be stopped again until systems are operational
OK
OK

Figure 5 Enhanced detail of figure 3 part 2

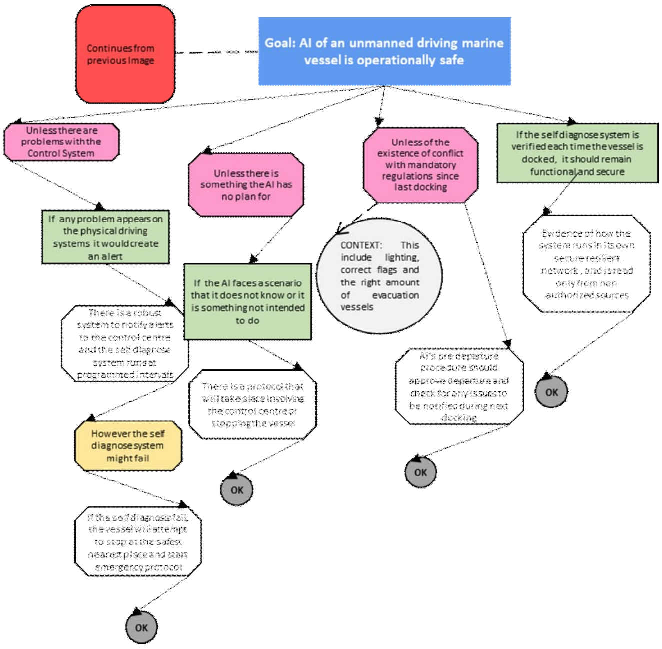
Continues from previous Image
Goal: AI of an unmanned driving marine vessel is operationally safe
Unless there are problems with the Control System
Unless there is something the AI has no plan for
Unless of the existence of conflict with mandatory regulations since last docking
If the self diagnose system is verified each time the vessel is docked, it should remain functional and secure
If any problem appears on the physical driving systems it would create an alert
If the AI faces a scenario that it does not know or it is something not intended to do
CONTEXT: This include lighting, correct flags and the right amount of evacuation vessels
Evidence of how the system runs in its own secure resilient network , and is read only from non authorized sources
There is a robust system to notify alerts to the control centre and the self diagnose system runs at programmed intervals
There is a protocol that will take place involving the control centre or stopping the vessel
AI's pre departure procedure should approve departure and check for any issues to be notified during next docking
OK
However the self diagnose system might fail
OK
OK
If the self diagnosis fail, the vessel will attempt to stop at the safest nearest place and start emergency protocol
OK

Figure 6 Enhanced detail of figure 3 part 3

VII. Evaluation of Example

An FMEA sourced from RH Marine provided most of the information needed to produce the assurance case, with the relevant safety and security context established each of the desired goals can be challenged with a relevant threat scenario. Figures 4 and 6, alongside the sub goals after the attacks on Figure 5, show a consistent method of representing threats resolved via the FMEA within the diagram. For the FMEA analysis, it is important to enumerate as many of the possible failures, and the potential solutions (being backed up by the FMEA) logically deducing the "OK" by adhering to the documentation.

Even when the hazards and threats are identified, the ability to link them to safety and security goals becomes significant, especially in terms of clearly demonstrating the explicit relationship between safety and security in the big picture. In Figure 5 the effects of the attack tree are in full view and directly influence the subgoals that appear as mitigations or eliminations. With regard to the possibility enumerating every possible attack mentioned in an attack defense tree within the assurance case, this is not the intended purpose of an assurance case. Instead, the assurance case succinctly demonstrates all relevant claims through logical deduction and reasoning, and attempting to fully reason through every little detail in an attack defense tree would prove cumbersome. Instead, the most relevant potential attacks can be integrated into a more compact attack defense tree by summarizing various branches within a more contextually sound and abstract manner, with the main concern being that each branch is logically proven safe and secure.

Significantly, by comparing and contrasting the method with additional documentation it can be shown that the method takes in consideration relevant data to demonstrate assurance. Within the automotive industry it is important to rely on a GSN Safety Case, ASILs and HARAs, and thus in the naval industry it is expected that any analysis or assurance would still comply with its relevant regulations. Additionally, the management of the data and the recollection of such can be considered acceptable by taking in account how the information from the FMEA and the ADT were represented, it also demonstrates that with respect to the available information it can be assured that it is as safe and secure as reasonably possible by clearly demonstrating the safety and security threats and how they are linked and mitigated. This becomes apparent with the use of the relevant attacks show in Figure 5. Thus, if the evaluation is performed taking in all the relevant consideration outlined, the result is acceptable as an assurance case, that can be interpreted as a safety case or cybersecurity case.

VIII. Conclusion

The amount of automotive vehicles outweighs the marine vessels, making the quantity of sold each year more than the quantity of marine vessels sold, and therefore by use the traffic is different, leading to different regulations. But regarding safety and security the objectives remain the same, keeping the users and the people as safe as reasonably possible be it from intentional or non-intentional harm. Using a method as the one explained in the paper it becomes more apparent how safety and security link; but it also helps tackling things from a different perspective by using inductive reasoning to reach the same conclusion as the documentation it relies on. When the methods goals cannot be closed using inductive reasoning it might be an indication that while creating the documentation some assumptions were made and must be addressed to close the goals. This method has been used in automotive examples such as a traffic recognition system and updating a safety critical system. The idea that it can be apply to other industries and still preserve the core objective of assuring safety and security might open the door to trying to apply methods from other industries to deepen assurance specially as AI driven vehicles start to take over.

Acknowledgment

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 812788 (MSCA-ETN SAS). This publication reflects only the authors' view, exempting the European Union from any liability. Project website: http://etn-sas.eu/.

References

[1] Y. He, Y. Jin, L. Huang, Y. Xiong, P. Chen, and J. Mou, "Quantitative analysis of COLREG rules and seamanship for autonomous collision avoidance at open sea," Ocean Eng., vol. 140, no. October 2016, pp. 281–291, 2017, doi: 10.1016/j.oceaneng.2017.05.029.

[2] M. Schiaretti, L. Chen, and R. R. Negenborn, "Survey on Autonomous Surface Vessels: Part I - A New Detailed Definition of Autonomy Levels," *Lecture Notes in Computer Science*, pp. 219–233, 2017, DOI: 10.1007/978-3-319-68496-3_15.

[3] "Cyber-enabled ships: ShipRight procedure – autonomous ships," info.lr.org. http://info.lr.org/l/12702/2016-07-07/32rrbk

[4] E. F. Brekke et al., "The Autosea project: Developing closed-loop target tracking and collision avoidance systems," Journal of Physics: Conference Series, vol. 1357, p. 012020, Oct. 2019, doi: 10.1088/1742-6596/1357/1/012020.

[5] "The future of remotely operated machinery - Industry insights - DNV," *DNV GL*, 2019. https://www.dnv.com/expert-story/maritime-impact/The-future-of-remotely-operated-machinery.html (accessed May 19, 2022).

[6] D. Gonzalez, J. Perez, V. Milanes, and F. Nashashibi, "A Review of Motion Planning Techniques for Automated Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 4, pp. 1135–1145, Apr. 2016, DOI: 10.1109/tits.2015.2498841.

[7] Y. Huang, L. Chen, P. Chen, R. R. Negenborn, and P. H. A. J. M. van Gelder, "Ship collision avoidance methods: State-of-the-art," Safety Science, vol. 121, pp. 451–473, Jan. 2020, DOI: 10.1016/j.ssci.2019.09.018.

[8] R. G. Wright, "Intelligent Autonomous Ship Navigation using Multi-Sensor Modalities," TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation, vol. 13, no. 3, pp. 503–510, 2019, doi: 10.12716/1001.13.03.03.

[9] E. R. Ibbetson, "Civil Marine Radar Developments in the Eighties," Journal of Navigation, vol. 44, no. 3, pp. 373–376, Sep. 1991, doi: 10.1017/s0373463300010195.

[10] D. Wang, J.-G. Wang, and K. Xu, "Deep Learning for Object Detection, Classification and Tracking in Industry Applications," Sensors, vol. 21, no. 21, p. 7349, Nov. 2021, doi: 10.3390/s21217349.

[11] S. Van Baelen, G. Peeters, H. Bruyninckx, P. Pilozzi, and P. Slaets, "Dynamic Semantic World models and increased situational awareness for Highly Automated Inland Waterway Transport," Frontiers in Robotics and AI, vol. 8, 2022.

[12] Qiao, D., Liu, G., Lv, T., Li, W. and Zhang, J., 2021. Marine Vision-Based Situational Awareness Using Discriminative Deep Learning: A Survey. Journal of Marine Science and Engineering, 9(4), p.397.

[13] A. Goudossis and S. K. Katsikas, "Towards a secure automatic identification system (AIS)," Journal of Marine Science and Technology, vol. 24, no. 2, pp. 410–423, May 2018, doi: 10.1007/s00773-018-0561-3.

[14] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," Information Fusion, vol. 14, no. 1, pp. 28–44, Jan. 2013, doi: 10.1016/j.inffus.2011.08.001.

[15] T. Miao, E. E. Amam, P. Slaets, and D. Pissoort, "Multi-Target Tracking and Detection, fusing RADAR and AIS Signals using Poisson Multi-Bernoulli Mixture Tracking, in support of Autonomous Sailing," Proceedings of the International Naval Engineering Conference & Exhibition (INEC). Institute of Marine Engineer, Science & Technology (IMarEST); Zenodo, 2020: 1-13.

[16] H.-C. Burmeister and M. Constapel, "Autonomous Collision Avoidance at Sea: A Survey," Frontiers in Robotics and AI, vol. 8, Sep. 2021, doi: 10.3389/frobt.2021.739013.

[17] T. A. Johansen, T. Perez, and A. Cristofaro, "Ship Collision Avoidance and COLREGS Compliance Using Simulation-Based Control Behavior Selection With Predictive Hazard Assessment," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 12, pp. 3407–3422, Dec. 2016, doi: 10.1109/tits.2016.2551780.

[18] Y. Huang, L. Chen, and P. H. A. J. M. van Gelder, "Generalized velocity obstacle algorithm for preventing ship collisions at sea," Ocean Engineering, vol. 173, pp. 142–156, Feb. 2019, doi: 10.1016/j.oceaneng.2018.12.053.

[19] T. Miao, E. E. Amam, P. Slaets, and D. Pissoort, "An improved real-time collision-avoidance algorithm based on Hybrid A* in a multi-object-encountering scenario for autonomous surface vessels," Ocean Engineering, vol. 255, p. 111406, Jul. 2022, DOI: 10.1016/j.oceaneng.2022.111406.

[20] M. Bojarski, D. D. Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J, Zhao, K. Zieba, "End to end learning for self-driving cars". arXiv preprint arXiv:1604.07316, 2016.

[21] T. Notteboom, A. Pallis, and J. Rodriguez, "Port Economics, Management and Policy" (1st ed.). Routledge. 2022

[22] J. Grinstead, "Marine safety regulation in transport Canada: The winds of change". Marine Technology and SNAME News, 33(3), 211–217. 1996

[23] B.-H Song, K.-H Lee, and W.-K Choi, "A Study on the Advancement of the Legal System for Small Fishing Vessels to Ensure Marine Safety". Journal of the Korean Society of Marine Environment and Safety, 24(7), 875–888, 2018

[24] [1] L.-P. Cobos, A. Ruddle, and G. Sabaliauskaite, "Requirements for a Cybersecurity Case Approach for the Assurance of Future Connected and Automated Vehicles," no. Vehits, pp. 626–633, 2021, doi: 10.5220/0010478606260633.

[25] [1] L.-P. Cobos, A. R. Ruddle, and G. Sabaliauskaite, "Cybersecurity Assurance Challenges for Future Connected and Automated Vehicles," in Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021), 2021, doi: 10.3850/978-981-18-2016-8.

[26] M. Mohamad, A. Åström, Ö. Askerdal, J. Borg, and R. Scandariato, "Security assurance cases for road vehicles: An industry perspective," ACM Int. Conf. Proceeding Ser., 2020, doi: 10.1145/3407023.3407033.

[27] Y. G. Dantas, V. Nigam, and H. Ruess, "Security Engineering for ISO 21434," pp. 1–15, 2020, [Online]. Available: http://arxiv.org/abs/2012.15080.

[28] K. Sowka, L. P. Cobos, A. Ruddle, and P. Wooderson, "Requirements for the Automated Generation of Attack Trees to Support Automotive Cybersecurity Assurance," SAE Tech. Pap., no. 2022, pp. 1–15, 2022, doi: 10.4271/2022-01-0124.

[29] B. Kordy, S. Mauw, S. Radomirović, and P. Schweitzer, "Foundations of Attack--Defense Trees," in Formal Aspects of Security and Trust, 2011, no. C, pp. 80–95.

[30] T. Kelly and R. Weaver, "The Goal Structuring Notation – A Safety Argument Notation," Elements, 2004

[31] J. B. Goodenough, C. B. Weinstock, and A. Z. Klein, "Eliminative Argumentation: A basis for arguing system confidence in system properties with induction," Proc. - Int. Conf. Softw. Eng., no. February, pp. 1161–1164, 2015

[32] R. Hawkins, T. Kelly, J. Knight, and P. Graydon, "A new approach to creating clear safety arguments," Adv. Syst. Saf. - Proc. 19th Safety-Critical Syst. Symp. SSS 2011, no. MoD 2007, pp. 3–23, 2011, doi: 10.1007/978-0-85729-133-2_1.

[33] The Assurance Case Working Group, "Assurance Case Guidance ' Challenges , Common Issues Version 1 The Assurance Case Working Group (ACWG)," no. January, p. 54, 2018.