ASSURING AUTONOMY
INTERNATIONAL PROGRAMME

DEMONSTRATOR PROJECT

Final report

SUCCESS: Safety assurance of cooperating construction equipment in semi-automated sites

OCTOBER 2020

Final Technical Report

SUCCESS

Safety assurance of Cooperating Construction Equipment in Semi-automated Sites

Funding period: 2018-12-17 – 2020-10-31

SUCCESS Project Team

MÄLARDALEN UNIVERSITY SWEDEN   VOLVO   Safety Integrity

Project leader: Sasikumar Punnekkat
sasikumar.punnekkat@mdh.se
+46 21 107324

November 15, 2020

A Demonstrator Project funded by the AAIP Program
University of York & Lloyd's Register Foundation, UK

ASSURING AUTONOMY
INTERNATIONAL PROGRAMME
Assuring the safety of autonomous systems
A partnership between Lloyd's Register Foundation and the University of York

Contents

Summary

The SUCCESS project[1] aims to enable the verification and certification of safety-relevant software-intensive production sites which can be considered as a complex system-of-systems (SoS) with many collaborating constituent systems.

The SUCCESS project is primarily focused on the support for SoS hazard analysis, safety arguments, digital twin-based safety analysis, dynamic reconfiguration and risk management, update of contracts and safety arguments, and targeted the processes and requirements from relevant standards. SUCCESS developed a framework for SoS safety-assurance that was evaluated for individual and emergent behaviours of machines. The demonstrator was evaluated in a range of quarry site operations that are carried out with different kinds of machines such as autonomous hauler, wheel loader, excavator, primary/mobile crusher and secondary crusher. We also considered a synthetic, but realistic and generic, use case that concerns transportation and data flow in Industry 4.0.

SUCCESS addressed challenges at multiple stages. At the design and development stage, the safety analysis is carried out through the identification of hazards, the assessment of risks, and the control of hazards' risks. Based on the hazard analysis results, the safety requirements were derived to prevent or mitigate the identified hazards. Safety contracts have also been derived for uncertainty sources. We constructed the safety cases and associated safety contracts with them. The safety requirements and hazard mitigation recommendations were implemented in the digital twins as code scripts. This not only gives the possibility to detect deficiencies, such as additional hazards and risks but also to identify, monitor, evaluate, and resolve deviations from specified behaviours during the operational phase. For systems with enhanced automation and connectivity, there is a need to deal with unknowns and uncertainties during operational phase. The dynamic safety assurance during the operational phase, in particular, the risk management and update of safety cases, were also carried out. By applying the advanced SoS research in industrial demonstrators, we have evaluated the applicability of the research, and also obtained guidance on future research directions.

SUCCESS has made significant research contributions to enhance the state-of-the-art, in particular, related to the following themes:

- SoS hazard analysis: Specific contributions include: (1) enriching System Theoretic Process Analysis (STPA), (2) adapting the Hazard and Operability (HAZOP) technique, (2) leveraging simulator-based digital twins to perform verification and validation to gain confidence in SoS production site. In order to perform the SoS hazard analysis, we also looked into aspects such as enhanced automation, digitization and connectivity of the processes, in addition to the specific aspects of the electric site use case.
- Safety arguments: Related research addressed the following main areas: (1) creation of safety arguments by considering the emergent SoS architecture for the advanced production site, (2) reflection of the architectural levels of things, fog and cloud in the safety arguments, (3) association of performance and safety indicators in terms of contracts with safety cases, (4) support for monitoring the safety arguments modelled in the PolarSys OpenCert platform against specified parameter values obtained during the operational phase and dynamically updating safety contracts and safety cases, and (5) tailoring of safety arguments in production lines.
- Dynamic risk management: Key results include: (1) a proposal to enforce static, dynamic, time-based and conditional geofences by defining different parameters such as geometric areas and conditions, for ensuring safety during normal flow of operations and resolving the failure cases, (2) a novel approach for platooning of autonomous haulers in production sites, (3) using the simulator-based digital twins to perform dynamic risk management, and (4) a novel reconfiguration approach in which tradeoffs between production safety and operational performance, in circumstances of site changes, failures and uncertain conditions are identified and resolved.

For the demonstrations in SUCCESS, we have extended and adapted the Volvo CE training simulators. They serve as digital twins of various machines used at the quarry site. The Volvo CE platforms used for training the operators of articulated haulers, excavators, and wheel loaders are connected to the quarry site scenario to demonstrate the functionality and behavior of manually-driven machines. We interacted closely with a number of research projects such as, FiC and FORA

[1] http://www.es.mdh.se/projects/530-SUCCESS

projects,to obtain synergistic effects. The SUCCESS project has provided excellent means to test and evaluate our ideas and results, as well as to collect the industrially relevant requirements important for the formulation of further research challenges.

1 Background, Objectives and the Project Team

Safety critical systems are traditionally developed as monolithic systems with emphasis on large scale verification and validation efforts. However, in recent years the industry had realized the importance of using component-based design for cost effectiveness, quicker market innovations and faster time to market. These are often referred to as systems-of-systems (SoS) and they combine characteristics of autonomous and in-dependently developed complex components with emergent behaviour and evolutionary nature. SoS can be evolutionary with dynamic system definitions, distributed in nature and often uses some communication infrastructure, involve multiple stakeholders, intended to work even in unpredictable operating environments. When the application context is safety/mission critical, they are also referred as Cooperative Cyber-Physical Systems (CCPS) since they combine cyber- and physical parts to provide intended services. The AAIP program targets robots and automation systems (RAS) which can be considered as equivalent to CCPS or SoS in terms of intended objectives, functionalities delivered, operational context, design and assurance challenges.

Since Humans and machines may be co-existing in RAS, it is important to obtain design-time safety assurance evidence to guarantee that the system manages risks acceptably. Therefore, in many domains, developers must prepare an explicit safety case, combining evidence with a safety argument, while in other domains developers must show that their processes and work products conform to a relevant standard. The design-time evidence enables developers to foresee the possible safety-related problems that can arise from the interaction between the system and the environment, to show that these interactions do not pose an unacceptable risk. Intelligent methods and tools are the need of the hour to help in safety certification which is rather expensive. SUCCESS has been addressing the above issues in the context of Volvo's 'Electric Site', which brings many innovations to the processes in quarries. [2].

1.1 The Volvo Electric Site Project

A quarry site as we focus on in this research is producing gravel in different granularity, typically used for road construction, railway track beds, construction of buildings and many more (see Figure 1). The gravel is produced in the following steps in the production process today using only human-operated machines.

Figure 1: The Volvo Electric Site Project

1. **Blasting**: The rocks are blasted to receive material for production.

[2] www.volvoce.com/global/en/this-is-volvo-ce/what-we-believe-in/innovation/electric-site/

2. **Pre-Crushing**: Often the rocks are too big after blasting and are crushed using a mobile crusher (primary crusher). The primary crusher is typically fed by an excavator.

3. **Pile Building**: The smaller rocks produced by the primary crusher, are piled by a wheel loader.

4. **Transporting**: The pre-crushed rocks are loaded on rigid haulers using the wheel loader. The rigid haulers transport the material to a secondary crusher.

5. **Secondary Crushing**: Pre-crushed rocks are further crushed into smaller granularity based on the customer needs. The secondary crusher is a static factory.

Further production steps might be necessary depending on the products requested by the customers.

At VCE a team has been working on the electric site research project that aims to transform the quarry and aggregates industry by reducing carbon emissions by up to 95% and total cost of ownership by up to 25%. To achieve this, an autonomous battery-driven hauler called 'HX' for transporting material at the site has been developed and planned to be used as a fleet. Other prototype machines that make-up the electric site system include a hybrid wheel loader and a grid-connected excavator. New technology encompasses machine and fleet control systems and logistic solutions for electric machines in quarries. Majority of the equipment and machines on a modern site are already connected and can enable collaborative emergent functionality, which has been the foundation for the Electric site. This can enable efficient production processes, provided one can assure the safety and similar concerns related to such heavy machines operate autonomously together with humans.

1.2 Objectives of the SUCCESS Project

Current safety standards and processes focus on single machines where the driver/operator in the machine is assumed to take appropriate actions in case of critical situations. In hazard analyses, the criticality of a risk can be reduced, when the operator has the possibility to put the machine into a safe state, for example by pushing the emergency button.

In the case of autonomous machines, critical situations may be missed, if safety assurance is applied for each single machine ignoring different application scenarios. Furthermore, human control of single machines is not provided if no appropriate control mechanisms are available in the control room. Accidents may happen, if the autonomous machines enter wrong areas on the site where humans are working. Identifying the state of the autonomous vehicles correctly may be difficult for workers.

Not only the workers at the site need to be considered, but also non-informed external humans (or rescue vehicles) that may enter the site. In case of an emergency situation, where a worker needs to be rescued from the site, the entering emergency team is at risk if the autonomous machines are not switched off.

The main focus of SUCCESSS is the assurance of safety for this quarry site case, by conducting research on hazard/safety analysis methods, adapt them to the Electric Site requirements and incorporate best practices into a new assurance methodology also applicable to similar classes of SoS.

1.3 The SUCCESS Project Team

The SUCCESS Demonstrator Project is lead by Mälardalen University(MDH) with Volvo CE and Safety Integrity AB(SAFI) as partners. The Project Team members are:

Prof. Sasikumar Punnekkat. SUCCESS Project Leader. Expert on dependable real-time systems, fault-tolerant computing, software engineering, functional safety, software reliability, and software testing. PhD from U. of York 1997.

Prof. Hans Hansson. Expert on component-based design of safety-critical real-time embedded systems, modeling and analysis of real-time communication, real-time testing and debugging, execution-time analysis, development of automotive control SW, and formal modeling of timing and probability. PhD Uppsala U. 1992.

Dr. Faiz Ul Muram. Postdoctoral Researcher at MDH, PhD from University of Vienna; under Uwe Zdun 2017 and post-doctoral experience with AMASS project (EU-ECSEL). Expert on safety assurance, process engineering, model checking, formal methods and model-driven engineering.

Dr. Muhammad Atif Javed. Postdoctoral Researcher at MDH; PhD from U. of Vienna under Uwe Zdun 2016. Expert on process engineering, software design and architecture, hazard and mishap risk management, assurance cases, traceability, IoT, reconfigurable systems, and empirical software engineering.

Dr. Anas Fattouh. He received a Diploma of Engineering in Automatic Control and Industrial Electronics from Aleppo University in 1992, a DEA degree and a PhD degree in Automatic Control from Automatic Laboratory of Grenoble, INPG, France in 1997 and 2000 respectively. He is expert in control, simulators.

Site Safety Engineer, Volvo Autonomous Solutions. Responsible for assuring safety for autonomous site projects at Volvo Autonomous Solutions. Research on System of Systems Safety; Leading Volvo Construction Equipment's efforts as part of the SUCCESS project. Stephan is expected to complete his PhD during 2021 as an industrial PhD student at MDH.

Dr. Henrik Thane. CTO of Safety Integrity AB. Software safety assessor/auditor Adj. Professor in Functional Safety. Expert on standards like: ISO26262, EN50128, IEC61508, and EN62061. Member of national committees for IEC61508 and EN50128; Ph.D. from KTH, 2000.

2 The Research of the Project

2.1 Scientific Approach and Results

The SUCCESS project focuses on the safety analysis and certification of production systems and the enhanced automation, digitalization and connectivity of processes. In the following sections, we present our key focuses in relation to the advancement and demonstration of safety critical production sites, and state-of-the-art, together with an account of the project achievements.

.

2.1.1 Identification of Specific Scenarios:

Handover: In any collaborating SoS, handing over control to machine and back to humans or to another machine could potentially result in hazardous behaviours. A specific scenario of interest we identified for further investigation in SUCCESS is the switching between control. In particular, we focused on the remote takeover of an individual autonomous machine by an operator with a remote control. This is interesting for our studies as it requires smooth and safe take over between different controllers.

Human-robot interaction: In the case of construction machinery due to the sheer huge size of the machines, human machine interactions are much riskier in comparison to robots in factories, at the same time the slow speeds of operation and inertia could potentially help humans to get some time to react. Another aspect is the high noise levels in open surface mines, which could keep the human operators unaware of approaching machinery. In the quarry site, it is of interest to restrict the entry of humans in such areas due to safety reasons. On the other hand, the travelling of machines/robots towards areas in which the humans are present could be restricted.

Movement control in site zones: The congested zones occur when multiple haulers simultaneously arrived at the loading, dumping, charging or parking zones from different paths. In such zones, there is a need to control the movement. A machine can enter in the congested zone with high speed and collide with another machine even with the presence of a collision-avoidance system, as the machine is approaching the same location and has not enough distance to stop. Accordingly, a specific scenario of interest we consider is to ensure the presence of only one machine in the specific points, for example, in the loading point, to perform successful and safe loading operation. The automated loading is compromised if two or more haulers intend to arrive in the loading point.

Managing operational risks with geofences: The virtual boundaries around geographic zones (i.e., geofences) can serve as an active countermeasure against operational mishap risks. In the SUCCESS project, we define the static, dynamic, time-based and conditional geofences by defining different geometric areas. For instance, the capsule shape can be used to widen the boundary for collision avoidance of hazardous vehicles not equipped with or have faulty obstacle detection devices, failed hardware, or transporting dangerous materials (e.g., explosive, toxic, etc.).

Flexible production system: The dynamic reconfiguration is a fundamental consideration for advanced safety-critical production systems. In the SUCCESS project, a scenario of interest is to dynamically reconfigure the quarry production site in consequence to market changes (i.e., low, normal and high demands), hazardous conditions and system failures. Since the operational changes, such as adding/removing machines may adversely impact the production safety and operational performance, we also focus on the quantitative assessment through configuration analytics to determine the corresponding impacts of changes on the safety, performance and production demands.

2.1.2 SoS Hazard Analysis

In SoS, besides the behaviour of an individual system, the emergent behaviour of systems that comes from their individual actions and interactions must be considered.

STPA: The System-Theoretic Process Analysis (STPA) provides an approach to specify the system to be analyzed and a method to identify potential accidents and their possible causes. STPA is a safety analysis method based on the systems thinking approach STAMP (Systems-Theoretic Accident Modeling and Processes) presented by Nancy Leveson [Nan12]. In safety analysis methods like FMEA or FTA, the failure of each component and its impact to the system are analyzed and assessed [Cif15]. Unlike FMEA or FTA, STPA is assuming that accidents are the results of unsafe control actions.

Conducting STPA requires four main steps [Lev18]:

1. Define Purpose of the Analysis
2. Model the Control Structure
3. Identify Unsafe Control Actions
4. Identify Loss Scenarios.

During the first step, the system of interest is defined and potential accidents and hazards related to the application scenarios of the system of interest are identified. This assumes that all relevant knowledge is available, which might not be the case in early development stages. In the second step, the control structure is developed which is an input to the analysis of each control action to identify potential unsafe control actions in step 3. Based on this information, potential loss scenarios are derived, i.e. finding the root cause of the hazards in step 4.

System-of-systems heavily rely on communication between the involved systems. STPA is targeting risks related to the communication channels, i.e. when messages are not received, unintendedly sent or a wrong contents is broadcasted.

For applying a hazard and safety analysis for our case, we see STPA of having a potential compared to the other described methods. STPA comes with an approach to define the architecture by developing the control structure diagram. This can help us to describe our SoS in way that is useful for conducting a safety analysis.

In the SUCCESS project we have applied STPA to the quarry case and studied how to apply STPA, which hazards can be identified on system-of-systems level and which hazards cannot be identified. In this phase, we studied possible interesting use cases at the site that can reveal advantages and shortcomings of this method.

HAZOP: The Hazard and Operability (HAZOP) analysis is an inductive technique for identifying and analysing the potential hazards and operational concerns of a system. HAZOP was initially developed to analyse chemical process systems, but later extended for other types of complex systems, for instance, nuclear power plants, rail systems and air traffic management systems [DFVA10, FBGCGGPBP17]. There exist some attempts to perform HAZOP analysis on software related systems [RMS08, MSCR10, SFD$^+$11]. However, previous research has not considered the emergent behaviours in an SoS. In the SUCCESS project, we have performed the SoS HAZOP analysis for the identification and elimination of potential hazards in the advanced quarry production. The SoS architecture is first described for the quarry production. However, to be able to perform the HAZOP analysis, a set of guide words (e.g., early/late, slower/faster, part of, other than, reverse, omission, before/after, etc.), parameters (e.g., position, distance, detect, speed, etc.), system inputs and outputs, a list of messages and their flow are determined. Subsequently, a list of system parameters are compared against a list of guide words. The HAZOP analysis is reported in the worksheets containing matrix or columns, in which the different items and proceedings are recorded.

The SUCCESS project also performed the HAZOP analysis in the Industry 4.0 context. This has not been considered in the published research. In this context, besides the modular, dynamic and reconfigurable nature of Industry 4.0, the architectural levels of the things, fog and cloud computing are considered. The performed analysis not just focuses on the individual behaviour of autonomous machines operating in Industry 4.0, such as Automated Guided Vehicles (AGVs) and robots, but also emergent interactions between them, and with fog/cloud server or other working equipment. The command/control functions and the ways for transmitting, processing and storage of big data are mapped to the Industry 4.0. This, however, significantly increases the safety assurance challenges, for instance, the critical incidents can occur in Industry 4.0, if correctly and timely communication is not established. We also targeted the autonomous driving hazards like collision of AGV with static and dynamic obstacles. Based on the hazard analysis results, we derived the safety requirements and safety contracts for the materials transportation and data flow in Industry 4.0 context.

In the SoS production, the capability to adjust production capacity and functionality is regarded as essential. The published works only focus on the reconfiguration of individual components of a single system [GT06, PHST12, PST13, BL17]. Besides that, they have not considered the dynamic reconfiguration of safety-critical production systems. In the SUCCESS project, we have targeted the dynamic reconfiguration in consequence of market changes, environmental conditions and system failures. We have applied the SoS HAZOP analysis to identify deviations from design intent and operational aspects related to reconfiguration of safety-critical production systems. The performed analysis does not only focus on the advanced production in the quarry site, but considers also the principal characteristics for highly reconfigurable production systems, in particular, modularity, scalability, customizability, convertibility, integrability and diagnoseability [KGG18].

FTA: One of the most commonly used deductive analysis approach for modelling, analysing and evaluating failure paths in a large complex dynamic systems is FTA [XA08]. Similar to the HAZOP and STPA analysis, the fault trees have not previously investigated for an SoS. In the SUCCESS project, to develop the fault trees, we selected the human injury, machine damage and mission failure as the top undesired events or mishaps. Specifically, the fault trees are developed based on the hazards and their potential effects understood from the other analysis techniques (HAZOP and STPA), in which the identified hazards can serve as the top undesired events. The FTA process starts with a top undesired event or mishap and attempts to find out what nodes of a system, combination of events, or component behaviour lead to the occurrence of this top event. The cause–effect relationships between the components of a system and their events are achieved based on the operating principle and fault mechanisms of the system by using logic gates (e.g., AND-gate, OR-gate, etc.). After establishing a top event, sub-undesired events are identified and structured that is referred to the top fault tree layer. The logic between every event is investigated, in particularly the type of gates and their specific inputs are formulated. All possible reasons including human errors, and environmental influences are evaluated level-by-level until all relevant events are found. In the fault trees, the resulting behaviour of particular failures will always reach a top mishap scenario.

Digital twin based safety analysis: The simulation-based digital twins are regarded as fundamental for systems with enhanced automation, digitalization and connectivity, such as those based on the Industry 4.0. It is one of the four pillars of Industry 4.0. In particular, a novel enabler of information transparency could be through the concept of digital twins: leverage the virtual equivalents such as simulation models for assessing and improving the real products.

The initial effort is although required to configure the digital twins, but they can reduce the development time. At the operational stage, the digital twins can be used in conjunction with the real systems, to deal with the unknowns and uncertainties. In the context of our other work, our aim is to support the hazard analysis and derivation of safety requirements/contracts with digital twins. Based on the performed hazard analysis, the mitigation mechanisms are established; they are translated into the safety requirements. However, during the design and development phase, all hazards and causal factors may not be identified. To establish a basis for evaluation and improvement of performed operations, an appropriate simulation environment needs to be configured [SPAR18]. In the SUCCESS project, we have leveraged the digital twin for performing verification and validation of the production site; it features realistic models of the machines and processes of Volvo quarry site. In the Volvo CE simulators-based digital twins, the hazard mitigation mechanisms and safety requirements are implemented as code scripts. Accordingly, the digital twins are used to perform verification and validation to gain confidence in production site. The digital twin based analysis served as a resource to discover additional hazards. A detailed list of parameters were used to support automated operations of the scenario that can be accessed and changed during operational phase, when necessary. This not only gives the possibility to detect deficiencies in site, such as additional hazards and risks but also identify, monitor, evaluate, and resolve deviations from specified behaviours during the operational phase.

2.1.3 Safety Arguments

To demonstrate the acceptable safety of production operations, safety cases are constructed to provide comprehensive, logical and defensible justification of the safety of a production system for a given application in a predefined operating environment. We utilise the PolarSys OpenCert platform that provides support for modelling and visualizing the safety arguments in Goal Structuring Notation (GSN) [The18]. It stores the argument (i.e., model and diagram) in CDO[3] such that different distributed stakeholders can access them and work on the same safety arguments concurrently.

Safety arguments for active/present configuration: In the SUCCESS project, the safety arguments are constructed for SoS. In this context, the modular, dynamic and reconfigurable nature of systems, and involved architectural levels are taken into consideration. However, the safety arguments constructed for variable systems reflect several alternatives to choose from [SFD$^+$11, HK07, HKP09, dOBM$^+$15, NNG19], each of which may exhibit divergent impacts on production safety and operational performance. Previous studies have not supported the variability management of safety case fragments in a family/line. The seamless integration between argumentation and variability management activities is needed for the tailoring of argument fragments in the production line. We used the BVR tool for orthogonal variability management [VHC$^+$15]. The safety arguments modelled in the OpenCert tool are customised with the three editors: VSpec (extended feature model), resolution (configuration) and realization (derivation). Besides the adaptation of changes in production line, the generation of argumentation (safety case) models and diagrams is supported.

Update of safety arguments: The safety arguments constructed at system design and development phases might be invalidated during operation. The automated systems and flexible manufacturing underlines the need for update of safety arguments to respond to the observed reality [DPH15, MJH19]. The published studies have not explicitly supported the update of safety arguments. The approach proposed in the SUCCESS project with focus on end-to-end tracebility and support of a tool framework can provide a significant boost for the designers to avoid the culture of paper safety at the expense of actual system safety [HC09]. Based on the gathered data from simulators based digital twins, the safety contracts constructed for uncertainty sources are monitored, deviations between the intended and actual behaviour are tracked and evaluated, and the safety contracts and safety cases modelled in the OpenCert platform are updated. Their update is carried out based on the optimal actions for which the thresholds regarding the performance degradation and upgradation are specifically taken into consideration. In contrast to the matching of parameter names and their values/ranges for safety contracts, the text-based matching is performed to alter the description of safety case elements. The required changes in safety cases are tracked and then the update command is issued. The assurance (safety) cases are updated on the CDO server, which is accessed by the OpenCert argumentation editor connected to the CDO server.

2.1.4 Risk Management During the Operational Phase

Flow of operations to ensure safety at site zones: In the SUCCESS project, the 'queue', 'pause' and 'exit' restrictions are defined to control risks in various site zones, such as loading and unloading. In particular, the essence of all these server commands sent to vehicles are to adjust the vehicle speeds to acceptable levels in relation to the context. Let us consider the loading zone. The successful and safe loading operation requires the presence of only one hauler H1 in the loading point. For entry in the region, there is a need to communicate with the server that is triggered when the located hauler H2 touches the sensor at entry to the loading zone. Since the hauler H1 is already present in the loading point, the hauler H2 requesting permission to enter is given a command to be in a 'queue'. After the completion of current loading, an 'exit' command is given to the loaded hauler H1, which then start moving. Next, the waiting hauler in queue (H2) is given the permission to enter; its maximum speed limit is set to 20 km/h. The other hauler H3 may also arrive in the meanwhile and instructed to be in the queue at next level; H3 moves to the place of H2.

Managing operational risks with geofences: There has been increasing attention in geofences that are enforced in various domains, such as smart city, healthcare, road transport, smartphones, security, forestry and aerospace [ZKS$^+$17]. However, the published studies have not considered the geofences for automated transportation and production sites. In the SUCCESS project, we have enforced the static, dynamic, periodic and conditional geofences in particular contexts. They can serve as an active countermeasure against operational mishap risks. For example, the loss of communication with a server, messages containing less, or wrong data are safety risk. In such cases, the control action

[3] http://www.eclipse.org/cdo/

is in place, i.e., the movement of autonomous haulers is still restricted in geofenced areas. As another example, consider a subsystem failure, depending on the severity risk factor, dynamic geofences are enforced as a mitigation strategy. The travelling in nearby area is blocked and the autonomous haulers in travelling path, including in standstill mode, such as loading point are commanded to drive away to reduce the risk to an acceptable level. In case of path problem, to create a new path compliant with the conditional geofence, the autonomous hauler wait for the human-driven hauler or another machine, such as wheel loader to formulate an alternative travel path, and then follows it.

Dynamic reconfiguration: The dynamic reconfiguration is a fundamental consideration for SoS safety-critical production systems. It can be performed in consequence of market changes, environmental conditions and system failures. Previous studies have not considered the dynamic reconfiguration of safety-critical production systems. Accordingly, in the SUCCESS project, the dynamic reconfiguration of safety critical SoS is targeted. The transformation to the low, normal and high production demands is explicitly evaluated. In this context, the adding/removing of machines and travelling paths may negatively influence the production safety and operational performance. The configuration analytics acts as a balancer and an active countermeasure against arbitrarily making decisions. Tradeoff factors that can be affected by the conflicting strategy are determined; they include the speed, distances, load capacity and time taken for loading, transportation, unloading and charging. The reconfigurations in consequence of subsystem failures have been considered. To avoid the mishap risks, switching to another configuration is performed. Let us consider the direct loading, in case the direct loading cannot be performed in a safe manner, i.e., the risk is not acceptable, the direct loading is terminated and just truck loading is performed. Other factors that lead to this situation are: primary crusher is jammed or humans are present in a loading area. The mishap risk is although controlled in these situations, but the operational performance is influenced. However, in circumstances when the achieved benefit is higher and the compromise remain within an acceptable region, the intended choice in favour of an alternative is tolerated.

3 Key Outcomes of the Project

The main contributions of the project are

- Identification of shortcoming of existing Hazard analysis techniques: Based on the categorization of system-of-systems hazards described by Redmond [RMS08], we are able to classify the hazards found by STPA. Because STPA is utilizing the control structure diagram as an input for identifying unsafe control actions, the main focus is on the interfaces between the involved systems and sub-systems and therefore interface hazards can be identified. Resource Hazards and Proximity Hazards are not directly derivable from the control actions, because such information is not part of the control structure diagram. Hazards related to reconfiguration, i.e. Reconfiguration Hazards, can only be identified to a limited extent. A reconfiguration can be context related, e.g. new routing of HX, additional decision points and added or removed machines. Another type of reconfiguration could be related to the behavior, e.g. messages and controls are sent with different timing. Use cases at the site might be changed because of new tasks. The environment in the quarry site is dynamic and therefore it is possible that reconfiguration is necessary from one day to another. Such information is not part of the control structure diagram and therefore not possible to identify. Instead, if the site is reconfigured as described, the complete control structure diagram needs to be revisited and the analysis to find unsafe control actions must be conducted again with the new context and behavior information. Because of the number of messages, data and control actions, this analysis would require huge efforts. Interoperability hazards, which are hazards caused by a different understanding of message between sender and receiver can be identified by STPA, if such a guide-word is added to the analysis of unsafe control actions.

- Digital twins based verification and assurance of safety requirements: We leverage simulators based digital twins to perform verification and validation to gain confidence in the production site by incorporating safety requirements in them. During this process, we checked the specifications on accuracy and precision requirements. Furthermore, additional hazards are detected. Functions, behaviour, and communication capabilities are mirrored in the digital twins, which requires initial investment, but can shorten the development time.

- Up-to-date safety arguments for SoS: To resolve the gaps between the ‘world as imagined’ and the reality ‘world as observed’, operational data is utilised. In particular, we retrieved a detailed list of parameters that provided the means to identify, monitor, evaluate, and resolve deviations from specified behaviours during the operational phase. The safety cases modelled in the OpenCert platform and contracts associated with them are updated.

- Geofences for safety: The geofences are categorised into static, dynamic, periodic and conditional. We have defined them over various zones at the site, different machines, other actors at the site such as humans, and even around specified paths of movements. These geofences are of different geometric characteristics, such as circle, rectangle and capsule geometry.

- Dynamic reconfigurations aspects: The dynamic reconfiguration of the safety critical production site is performed in consequence of market changes, environmental conditions and system failures. In this regard, the hazard analysis for safe reconfiguration, safety-performance tradeoffs and customization of safety-critical systems with production lines are specifically considered.

- Platooning strategy for automated transportation in safety-critical production sites: A strategy for automated vehicle platoons in the production sites is proposed and implemented in which the production efficiency and fault tolerance aspects are specifically considered.

3.1 List of publications:

The research in SUCCESS has resulted in the following peer-reviewed publications (pre-prints of them are included in Appendix A of this report):

1. A State-based Extension to STPA for Safety-Critical System-of-Systems (Nov 2019), Stephan Baumgart, Joakim Fröberg , Sasikumar Punnekkat, International Conference on System Reliability and Safety (ICSRS-2019)
https://ieeexplore.ieee.org/document/8987632

2. System of Systems Hazard Analysis Using HAZOP and FTA for Advanced Quarry Production (Nov 2019), Faiz Ul Muram, Muhammad Atif Javed, Sasikumar Punnekkat, International Conference on System Reliability and Safety (ICSRS-2019)
https://ieeexplore.ieee.org/document/8987613

3. Enforcing Geofences for Managing Automated Transportation Risks in Production Sites (Sep 2020) Muhammad Atif Javed, Faiz Ul Muram, Anas Fattouh, Sasikumar Punnekkat Workshop on Dynamic Risk managEment for Autonomous Systems (DREAMS 2020)
https://link.springer.com/chapter/10.1007/978-3-030-58462-7_10

4. A Process to Support Safety Analysis for a System-of-Systems (Oct 2020) Stephan Baumgart, Joakim Fröberg , Sasikumar Punnekkat 31st International Symposium on Software Reliability Engineering (ISSRE 2020) http://2020.issre.net/industry-accepted-papers

5. Towards Dynamic Safety Assurance for Industry 4.0 (Oct 2020), Muhammad Atif Javed, Faiz Ul Muram, Hans Hansson, Sasikumar Punnekkat, Henrik Thane, Journal of Systems Architecture (JSA) https://www.sciencedirect.com/science/article/pii/S1383762120301788

6. Dynamic Reconfiguration of Safety-Critical Production Systems (Sep 2020) Faiz Ul Muram, Muhammad Atif Javed, Hans Hansson, Sasikumar Punnekkat 25th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2020)
http://www.es.mdh.se/publications/5901-Dynamic_Reconfiguration_of_Safety_Critical
_Production_Systems,
Accepted for PRDC 2020- the actual presentation will happen in Dec 2021 since PRDC 2020 was joined with PRDC2021 due to Corona situation

Additional related publications:

1. An Actor-based Design Platform for System of Systems (Jul 2019), Marjan Sirjani and Giorgio Forcina and Ali Jafari and Stephan Baumgart and Ehsan Khamespanah and Ali Sedaghatbaf, COMPSAC 2019: Data Driven Intelligence for a Smarter World

2. Safe Design of Flow Management Systems Using Rebeca (2020), Giorgio Forcina, Ali Sedaghatbaf, Stephan Baumgart, Ali Jafari, Ehsan Khamespanah, Marjan Sirjani, Pavle Mrvaljevic, Journal of Information Processing (JIP), https://doi.org/10.2197/ipsjjip.28.588

3. Defining a Method to Perform Effective Hazard Analysis for a Directed SoS Based on STPA, Stephan Baumgart, Joakim Fröberg , Sasikumar Punnekkat, Third Swedish Workshop on the Engineering of Systems-of-Systems 2018

We are also currently working on few more research articles for future submissions

- Plannned submission- Platooning. We provide a novel platooning strategy for automated transportation in critical production sites. It is implemented in Volvo CE quarry site simulators. In this paper, the performance and fault tolerance aspects are specifically considered.

- Planned submission - Process and guidelines for Safety anlysis of System of Systems(SoS)

3.2 Thesis Projects

3.2.1 PhD Project of Stephan Baumgart

Many existing hazard analysis methods are applied in industry and are incorporated in the tailored development processes which support the required by functional safety standards. The existing industrial development processes in the earth moving machinery domain target the development of single human-operated machines. There is limited support for developing collaborative systems and identifying how failing of the indented behavior may lead to critical accidents.

The main goal of Stephan's research is to guide practitioners on how to analyze the safety for system-of-systems to aid developing such systems. We provide four main contributions in this phd thesis. 1) We study industrial systems-of-systems, how functional safety standard compliance is achieved and which challenges exist. 2) We propose a technique to capture relevant inputs to support a preliminary hazard analysis. 3) We propose a model-based approach to support a System-of-Systems Hazard Analysis and 4) we propose a process to support a SoS-centric safety analysis.

Stephan has presented his PhD proposal in September 2020 and is aiming for the PhD thesis defense during early 2021.

3.2.2 Ongoing Master Thesis: Using UAV as a Redundant System Safety on Autonomous Sites

In the current set up of a site, the site operator located in an office at the site shall monitor the operation of the fleet of autonomous vehicles. This is a challenging task and the cameras located at the track are static. Typical critical scenarios are for example unauthorized people enter the autonomous operating zone or operation of other vehicles in proximity to the autonomous machines. The goal of an going master thesis project is to utilize an UAV to provide additional monitoring features.

One challenge is the data extraction from video feeds. The specified objects need to be identified and tracked as shown in Figure 2. This thesis's primary goal is to provide a structured evaluation of potential algorithms for identifying and tracking objects to identify safety-critical situations.

Due to the Corona pandemic, the Master student started later than expected and the finalized thesis will be delivered in January 2021.

3.2.3 Completed Master Thesis Projects

Two related master thesis projects at MDH using the Volvo simulation setup were supervised by Anas Fatouh a) to test a path planning algorithm for an autonomous wheel loader[4] and to validate the flow management system that was developed and formally verified by an actor-based framework (AdaptiveFlow)[5].

[4] http://mdh.diva-portal.org/smash/record.jsf?pid=diva2:1454119

[5] http://mdh.diva-portal.org/smash/record.jsf?pid=diva2:1442883

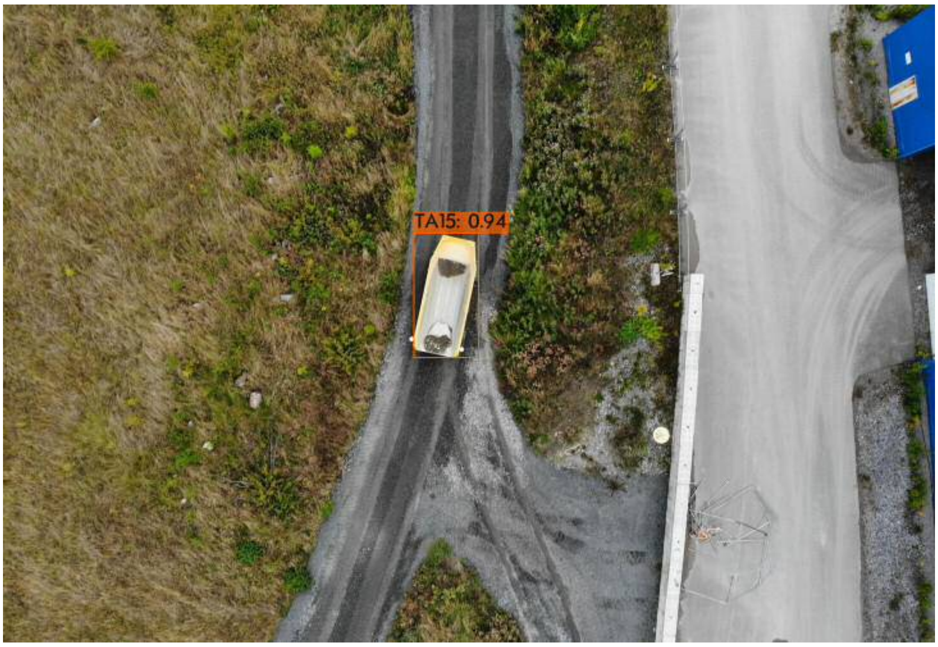


Figure 2: UAV camera picture and localizing Autonomous MAchine

4 Digital-twin Based Demonstrator

Demonstration can be used to show stakeholders the execution of the implemented system and how it meets its requirements. Moreover, it can be used to explore the behaviour of the system at the boundaries in assessing the safety implications. In the context of SUCCESS project, the Volvo Electric Site, partially simulated in the Volvo Simulators, is considered as a case study. The site is simulated in normal operation mode as well as in different faulty scenarios.

4.1 The Volvo Electric Site as a Case Study

In the Volvo Electric Site research project (ref https://www.volvoce.com/global/en/this-is-volvo-ce/what-we-believe-in/innovation/electric-site/) , the production processes on a quarry site are analysed to identify potentials for automation. The transport of material from the primary crusher to the secondary crusher is automated by utilizing a fleet of autonomous HX shown in Figure **??**. This means, that both human-operated machines run in the same production site together with the fleet of autonomous HX machines. The targeted production process at the quarry site is subdivided into the following production zones for better analysis purposes:

1. **Pre-Crushing**: The rocks are pre-crushed using the mobile Primary Crusher (PCR), which is fed by a human operated Excavator (EXC).

2. **Loading HX**: This can either be done by direct loading from the PCR or by using a human operated Wheel Loader (WL).

3. **Dumping at SCR**: Once an HX is loaded, it travels to a dump spot where the HX dump its load onto a pile which feeds the Secondary Crusher (SCR).

4. **Charging**: Since the HXs are battery-drive, they require charging of their batteries at the charging stations (CH). Two alternative chargers are available.

5. **Parking**: For maintenance purposes or for stop of operation, a parking area (PA) is available beside the transporting routes.

The autonomous HXs are traveling on predefined routes (black lines) between the production zones. At the Main Decision Point (MDP), the HX wait until receiving a new mission, which specifies where to get loaded including a targeted GPS position for loading. The loading missions differ

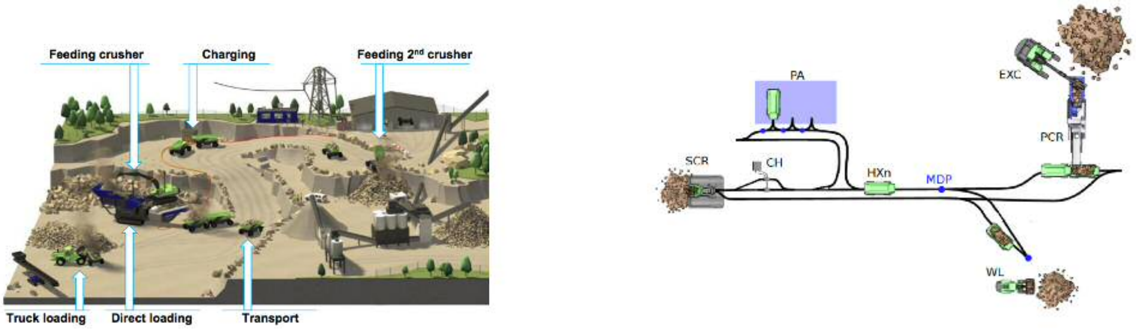


Figure 3: Modelling the Electric Site

depending on getting directly loaded at the PCR or getting loaded by the wheel loader. Loading by the PCR includes queuing up under the PCR conveyor and ones the first HX in this queue is loaded, it receives the mission to transport the pre-crushed rocks to the dump spot at the SCR. Getting loaded by the WL is different, since only one HX is getting loaded by the WL operator. The WL operator can request an HX for loading and send the HX away once loaded using an Operators Interface in the WL. The WL operator can even set new loading spots to reduce the time needed for loading an HX. The positions of the input piles are moving as material is transported away. Therefore, the routing of the HX must allow flexible loading positions. The automated quarry site is a system containing different complex systems collaborating together to realize the targeted production goals. Some of the constituent systems like the wheel loader and the excavator are systems that can be used without being part of the automated quarry site. Furthermore, machines and solutions from other suppliers can be integrated on the site, while a top-level management system needs to keep control over the site and the production targets.

Figure 4 shows a simplified control structure of the electric quarry site. The boxes are systems and subsystems and the lines with arrows are the messages and controls between subsystems. This control structure can be subdivided into four layers, where the top three layers are parts of the controlling components while the lowest layer consists of the physical components, that are controlled through actuators based on sensor feedback.

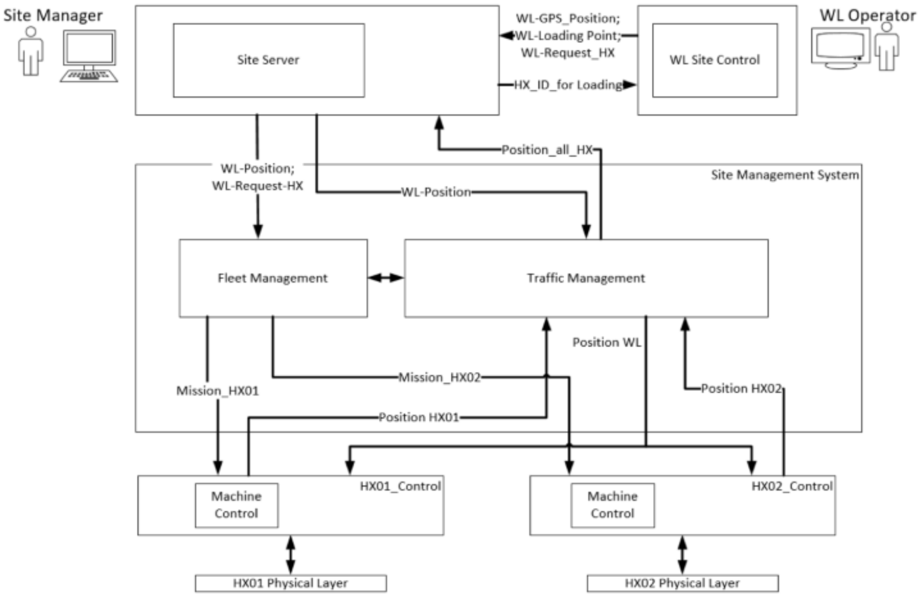


Figure 4: Simplified Control Structure

In the first layer on top the humans, who can control or influence the site, are located. The Site Manager receives all information through the Site Server, where all site-related information like the position of all machines, their status, current mission and status of production are available.

Production steps can be added or stopped during the production to adjust the system. Another worker at the site shown in Figure 3 is the Wheel Loader operator, who can interact with the fleet of HX through a display installed in the WL cabin connected to the WL Site Control. Messages sent from the WL Site Control are the GPS position of the WL WL-GPS Position, a loading point WL-Loading Point, which is also a GPS position indicating where the HX shall stop for getting loaded by the WL. The WL operator can even request a new HX for loading. This request WL-Request HX is sent to the Site Server. The WL operator receives an information, which HX shall be loaded HX ID for loading. The Site Server acts as an interface between the fleet of HX and the other machines used at the quarry site. The second layer of this control structure diagram consists of a computer system which is controlling the fleet of autonomous HX - the Site Management System. This computer system consists of subsystems like Fleet Management and Traffic Management. The Fleet Management sets the missions for each HX, which are tasks an HX shall perform like receive direct loading at the PCR, travel to dump spot or move to the parking area. The Traffic Management system is a system that controls the overall traffic situations on the site to avoid collisions between HX or between HX and human-operated machines. The Fleet Management system receives information from the Site Server like the position of the WL WL-Position and if a new HX is requested by the WL operator WL-Request-HX. This information is translated into missions sent to a specific HX in the fleet. The Traffic Management System receives the position of other machines from the Site Server. At the same time the current position of all active HX in the fleet are sent to the Site Server. Information is shared between the Fleet Management and the Traffic Management, shown by the line with double arrows between these subsystems. The bottom two layers focus on the single HX, in our case HX01 and HX02, while the third layer focuses on the machine control and the fourth layer on the physical parts of each machine. The control layer HX01 Control is a computer system located on the autonomous machine HX01 and is receiving mission data and is interpreting traffic information. HX01 Control is the direct interface to the physical layer of HX01, which can be engine, steering mechanisms, brakes, valves for hydraulic cylinders for unloading the bucket and many more. At the same time sensor data from the physical layer is collected by the Control System of each HX and shared with the Site Management System. The control and reading of sensor data are simplified in Figure 3 by a double arrow between the machine control and the physical layer of each HX. As an example, for sharing sensor data, the position of an HX is sent to the Traffic Management.

4.2 Simulating the Volvo Electric Site in the Volvo Simulator

4.2.1 The Volvo Simulators

Volvo Group Arena[6] at Mälardalen University (MDH) has two Volvo simulators[7] that are used to simulate Volvo construction machines (Wheel Loaders, Haulers, and Excavators) working in different environments (see Figure 5).

The simulators were originally developed by the ORYX Simulations Company[8] for product marketing, operator training, and product development processes. Several scenarios were developed to train the operator driving Volvo construction machines in different situations with increased task complexities (see Figures 6 and 7). At the end of each scenario, a summary of the run scenario is displayed which shows the statistics about the travelled distance, the fuel consumption, the committed errors, and other parameters of interest. Several simulators can be involved in one scenario to train operators to work cooperatively.

At MDH, the simulators are equipped with an Editor (see Figure 8) that permits students and researchers to create new scenarios, modify the machines' behaviours, and add more functionalities to the machines or the scenarios. This allowed the students and researchers to test their developed algorithms in high-fidelity simulators. The Volvo Simulators were used also used in two related master thesis projects.

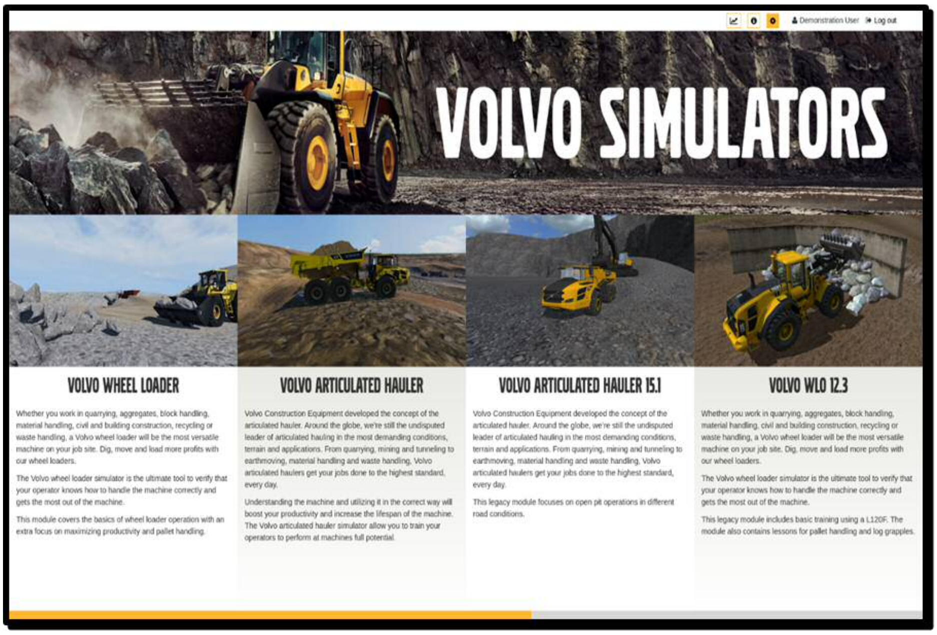
4.2.2 The Volvo Simulators' Event Passing-Reacting Paradigm

The Volvo Simulator platform is designed as a reactive system where each object can have several scripts to react to different events that is passed to it. These events can come from many different sources such that the input signals, the physics system, and other scripts. Objects can have scripts to

[6] http://www.es.mdh.se/projects/543-Volvo_Group_Arena

[7] https://www.volvoce.com/europe/en/services/volvo-services/productivity-services/volvo-simulators/

[8] https://www.oryx.se/



VOLVO WHEEL LOADER

Whether you work in quarrying, aggregates, block handling, material handling, civil and building construction, recycling or waste handling, a Volvo wheel loader will be the most versatile machine on your job site. Dig, move and load more profits with our wheel loaders.

The Volvo wheel loader simulator is the ultimate tool to verify that your operator knows how to handle the machine correctly and gets the most out of the machine.

This module covers the basics of wheel loader operation with an extra focus on maximizing productivity and pallet handling.

VOLVO ARTICULATED HAULER

Volvo Construction Equipment developed the concept of the articulated hauler. Around the globe, we're still the undisputed leader of articulated hauling in the most demanding conditions, terrain and applications. From quarrying, mining and tunneling to earthmoving, material handling and waste handling, Volvo articulated haulers get your jobs done to the highest standard, every day.

Understanding the machine and utilizing it in the correct way will boost your productivity and increase the lifespan of the machine. The Volvo articulated hauler simulator allow you to train your operators to perform at machines full potential.

VOLVO ARTICULATED HAULER 15.1

Volvo Construction Equipment developed the concept of the articulated hauler. Around the globe, we're still the undisputed leader of articulated hauling in the most demanding conditions, terrain and applications. From quarrying, mining and tunneling to earthmoving, material handling and waste handling, Volvo articulated haulers get your jobs done to the highest standard, every day.

This legacy module focuses on open pit operations in different road conditions.

VOLVO WLO 12.3

Whether you work in quarrying, aggregates, block handling, material handling, civil and building construction, recycling or waste handling, a Volvo wheel loader will be the most versatile machine on your job site. Dig, move and load more profits with our wheel loaders.

The Volvo wheel loader simulator is the ultimate tool to verify that your operator knows how to handle the machine correctly and gets the most out of the machine.

This legacy module includes basic training using a L120F. The module also contains lessons for pallet handling and log grapples.

Figure 5: The Volvo Simulators

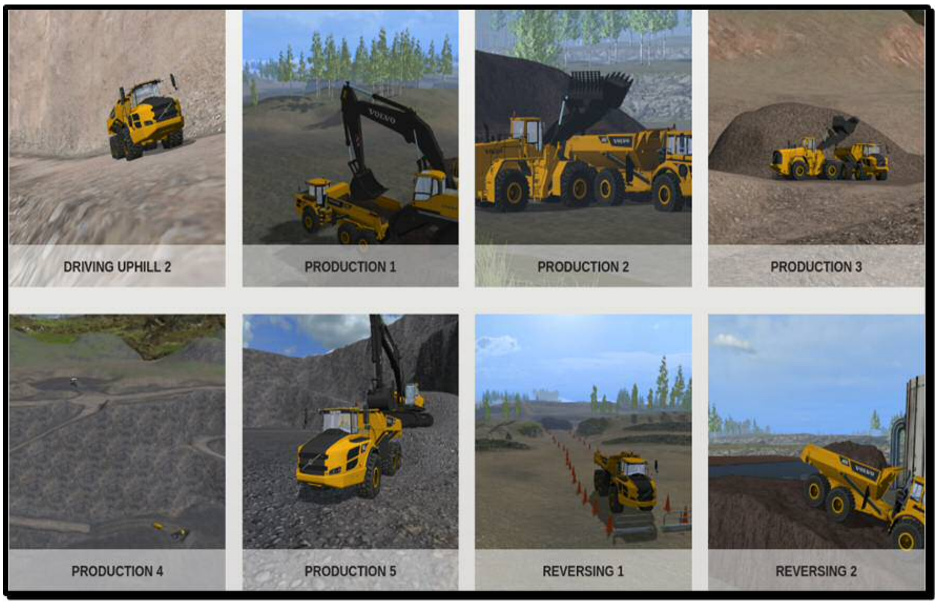
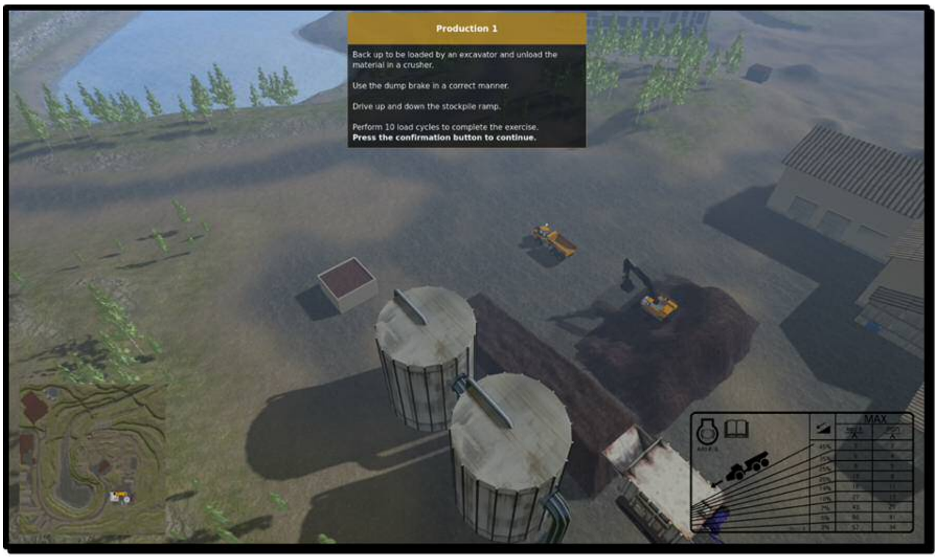


Figure 6: Examples of available scenarios



Production 1

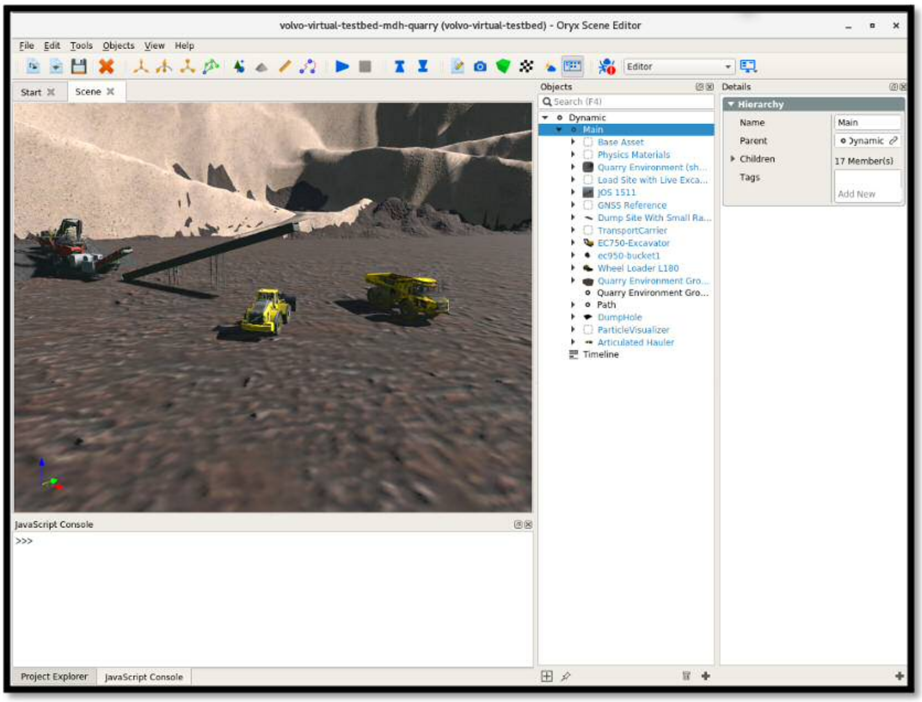
Back up to be loaded by an excavator and unload the material in a crusher.

Use the dump brake in a correct manner.

Drive up and down the stockpile ramp.

Perform 10 load cycles to complete the exercise.
Press the confirmation button to continue.

Figure 7: A production Scenario



volvo-virtual-testbed-mdh-quarry (volvo-virtual-testbed) - Oryx Scene Editor

Dynamic
Main
Base Asset
Physics Materials
Quarry Environment (sh...
Load Site with Live Exca...
JOS 1511
GNSS Reference
Dump Site With Small Ra...
TransportCarrier
EC750-Excavator
ec950-bucket1
Wheel Loader L180
Quarry Environment Gro...
Quarry Environment Gro...
Path
DumpHole
ParticleVisualizer
Articulated Hauler
Timeline

Name	Main
Parent	Dynamic
Children	17 Member(s)
Tags	Add New

Figure 8: The volvo simulator editor

react to built-in events or newly created events and they can also pass events to other objects. One of the built-in events is the onPlay event which is emitted when the simulation is started. An object can react to this event by having a script with handleEvent(onPlay) function that specifies how the object will react to that event by, for example, initialize the object's variables or moving the object on a predefined path. Objects can communicate through passing events to each other through the function sendEvent(targetObject, eventName, optional arguments).

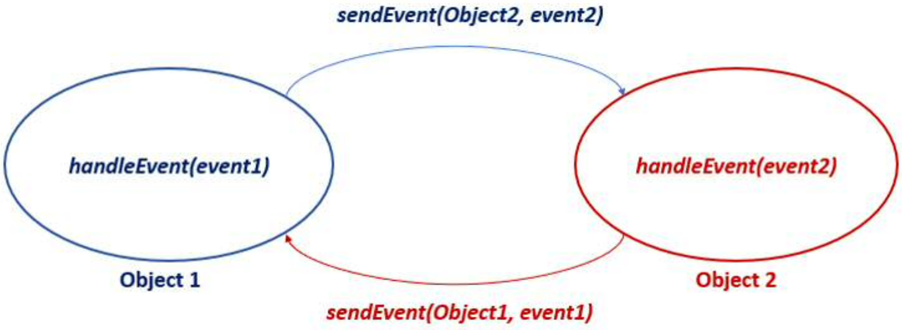


Figure 9: The Event Passing-Reacting Paradigm

The event passing-reacting paradigm (See Figure 9) will be used in the next subsection to partially model the Volvo Electric Site in the Volvo Simulator.

4.2.3 Modelling the Volvo Electric Site in the Volvo Simulator

The Electric Site case study explained in Section 3.1 and Figure 3 is modelled in the Volvo Simulator as shown in Figure 10.

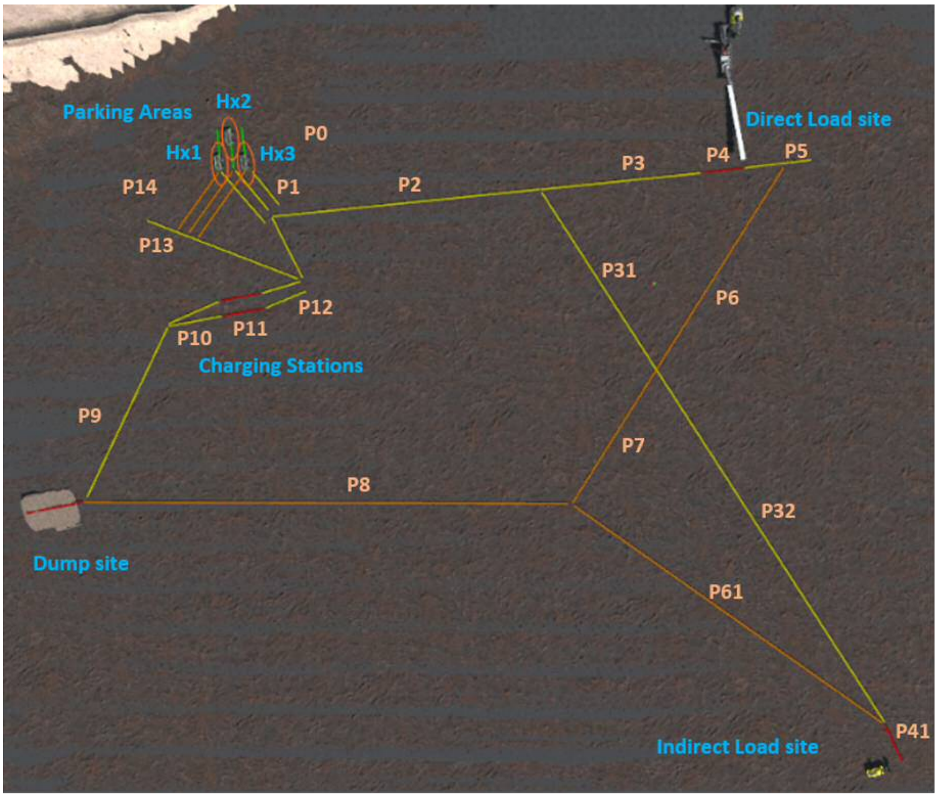
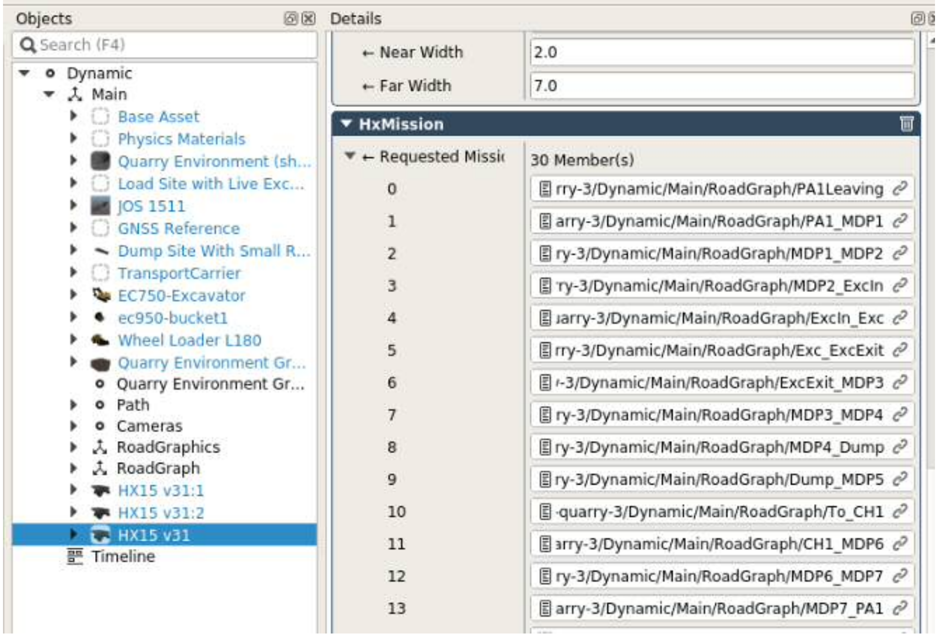


Figure 10: The Volvo Electric Site in the Volvo Simulator

The model consists of:

- Three parking areas.
- One direct loading site with a specified loading time.
- One indirect loading site with a specified loading time.
- One dump site with a specified dumping time.
- Two charging stations with a specified charging time.
- Predefined paths between different sites with specified properties (capacity, speed, type, waiting time).
- Three autonomous loaders (Hxs) with specified properties (battery level, . . .). Each Hx is equipped with the following components:
    - Hx autonomous driver that drives the Hx on a path (defined by a set of points) with a max allowable speed.
    - A path planning algorithm that generates a path between two defined points.
    - Collision avoidance that keeps the Hx far by a specified distance from other Hxs.

An interface for the autonomous loaders (Hxs) was developed during the SUCCESS project that allows the site manager to assign a mission to an Hx, specify its parameters, and to record its status (see Figure 11).



← Near Width	2.0
← Far Width	7.0

← Requested Missi	30 Member(s)
0	rry-3/Dynamic/Main/RoadGraph/PA1Leaving
1	arry-3/Dynamic/Main/RoadGraph/PA1_MDP1
2	ry-3/Dynamic/Main/RoadGraph/MDP1_MDP2
3	ry-3/Dynamic/Main/RoadGraph/MDP2_ExcIn
4	uarry-3/Dynamic/Main/RoadGraph/ExcIn_Exc
5	rry-3/Dynamic/Main/RoadGraph/Exc_ExcExit
6	r-3/Dynamic/Main/RoadGraph/ExcExit_MDP3
7	ry-3/Dynamic/Main/RoadGraph/MDP3_MDP4
8	ry-3/Dynamic/Main/RoadGraph/MDP4_Dump
9	ry-3/Dynamic/Main/RoadGraph/Dump_MDP5
10	quarry-3/Dynamic/Main/RoadGraph/To_CH1
11	arry-3/Dynamic/Main/RoadGraph/CH1_MDP6
12	ry-3/Dynamic/Main/RoadGraph/MDP6_MDP7
13	arry-3/Dynamic/Main/RoadGraph/MDP7_PA1

Figure 11: Mission interface for Hxs

The fleet manager, which was developed during the SUCCESS project, consists of a set of nodes placed at the end of each path. Each node detects the arrival of an Hx (pathPlayed event) and handles this event through the function handleEvent(pathPlayed) (see Figure 12)). The handleEvent(pathPlayed) function reads the Hx mission, waits for the mission completion (based on the path type: loading, unloading, or charging), and gives the Hx the permission to play the next path when it is available.

4.3 Running Different Scenarios on the simulator

In this section, for illustration purpose, we describe execution of three scenarios for the Volvo Electric Site shown in Figure 9 using the fleet management logic depicted in Figure 10. The first scenario runs the site in a normal operation mode while the second and third scenarios runs the site in faulty

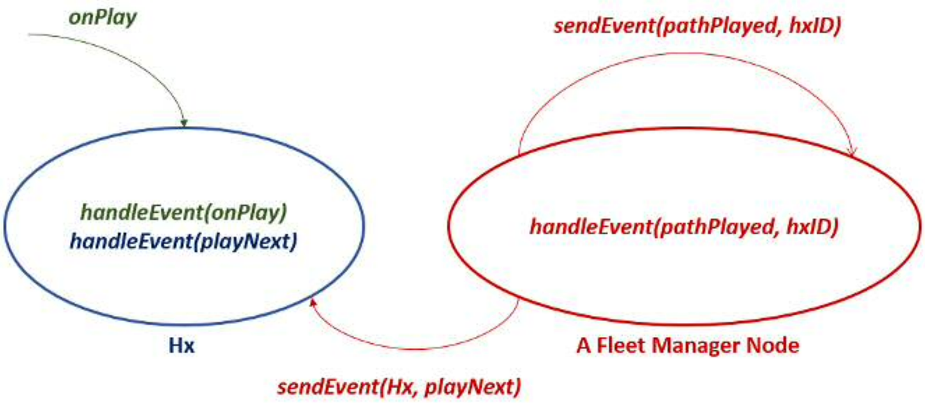


Figure 12: Events of the Fleet Manager

Table 1: Paths and their properties

Path number	Path type	Speed (m/s)	Waiting time (s)	Capacity (No. of Hx)
P0	Parking	20.0	0.0	1
P1	Track	20.0	0.0	1
P2	Track	50.0	0.0	3
P3	Track	50.0	0.0	3
P4	Loading	20.0	20.0	1
P5	Track	20.0	0.0	1
P6	Track	20.0	0.0	3
P7	Track	20.0	10.0	3
P8	Unloading	50.0	40.0	3
P9	Track	50.0	0.0	3
P10	Track	20.0	0.0	1
P11	Charging	20.0	40.0	1
P12	Track	20.0	0.0	1
P13	Track	20.0	0.0	1
P14	Track	20.0	0.0	1
P31	Track	50.0	0.0	3
P32	Track	50.0	0.0	3
P41	Loading	20.0	20.0	1
P61	Track	50.0	0.0	3

operation modes where a wrong mission is assigned to the Hx (which is identify as an unsafe control action that could lead to different kinds of hazards and accidents depending on which parts of the mission are incorrect [1]). The paths in the scenario and their properties are given in Table 1.

4.4 Scenario 1: Normal Operation Mode

In this operation mode, two Hxs were assigned the missions given in Table 2.

Table 2: Scenario-1 of HXs

Hx\Path	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Hx1	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
Hx2	P0	P1	P2	P31	P32	P41	P61	P8	P9	P10	P11	P12	P13	P14	P15	

The Hxs perform their missions successfully as shown in Figure 13 and 15 (the mission index). Figure 14 shows a snapshot of Hxs while performing their missions.

Figure 15 shows the distance between the Hxs and their speeds while performing their missions which shows no accident or deadlock.

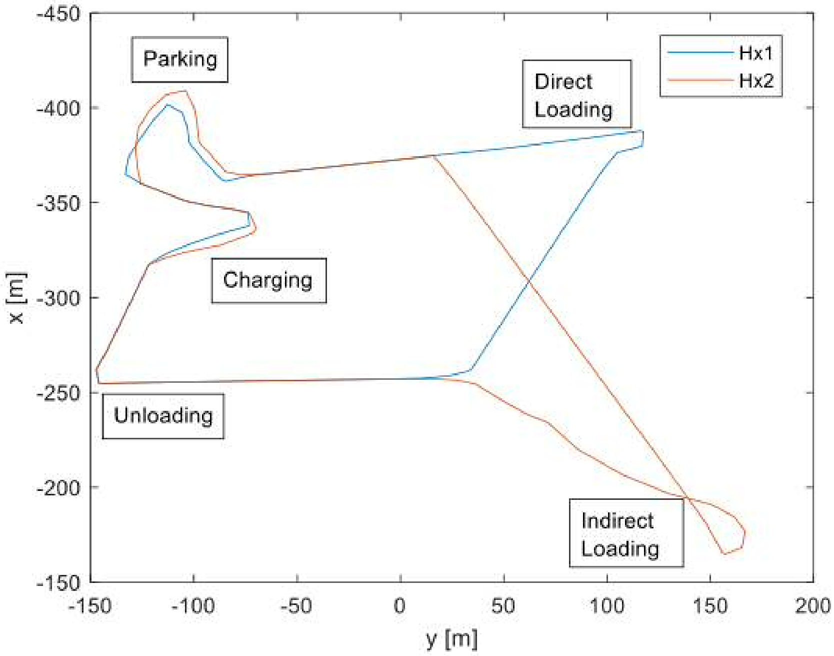


Figure 13: Paths played by the Hxs in Scenario-1

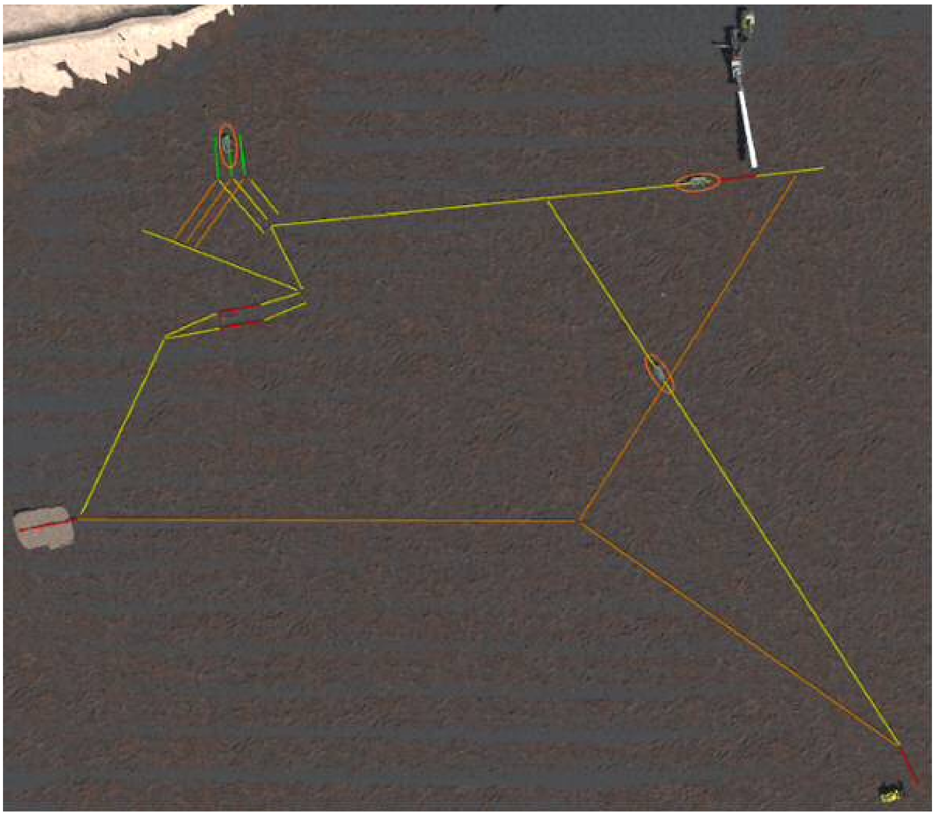


Figure 14: Hxs while performing Scenario-1 (Recording available at https://youtu.be/fj6mu4020VM)

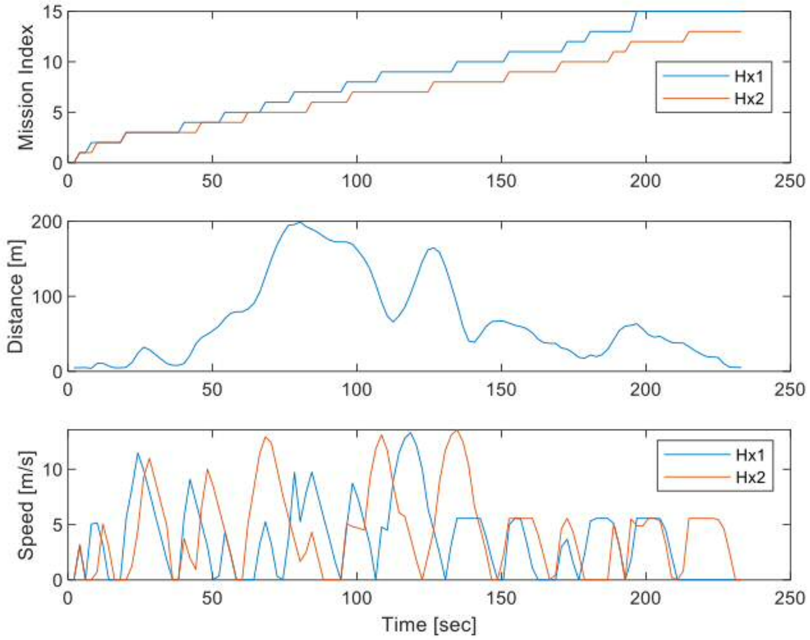


Figure 15: The distance between Hxs and their speeds while performing their missions in Scenario-1

4.5 Scenario 2: Faulty Operation Mode 1

In this operation mode, two Hxs were assigned the missions given in Table 3. Hx2 is assigned the wrong mission P5. The Hxs perform their missions successfully as shown in Figure 16 and Figure 21 (the mission index), however, Hx2 needed to find a new path (unplanned path) to complete its mission which could lead to an accident if there are other operators on the new path.

Table 3: Scenario-2 of HXs

Hx\Path	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Hx1	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
Hx2	P0	P1	P2	P31	P5	P41	P61	P8	P9	P10	P11	P12	P13	P14	P15	

Figure 17 shows a snapshot of Hxs while performing their missions.

Figure 21 shows the distance between the Hxs and their speeds while performing their missions which shows no accident or deadlock, however, Hx2 needed more time to execute its mission.

4.6 Scenario 3: Faulty Operation Mode 2

In this operation mode, two Hxs were assigned the missions given in Table 4. Hx1 is assigned the wrong mission P61.

Table 4: Scenario-3 of HXs

Hx\Path	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Hx1	P0	P1	P2	P3	P4	P5	P6	P7	P61	P9	P10	P11	P12	P13	P14	P15
Hx2	P0	P1	P2	P31	P32	P41	P61	P8	P9	P10	P11	P12	P13	P14	P15	

The simulations reveal occurrence of an accident as shown in Figures 19 and 20. Although the Hx is equipped by a collision avoidance algorithm, it had not enough time to break as it moves on a high-speed path and is facing the other Hx.

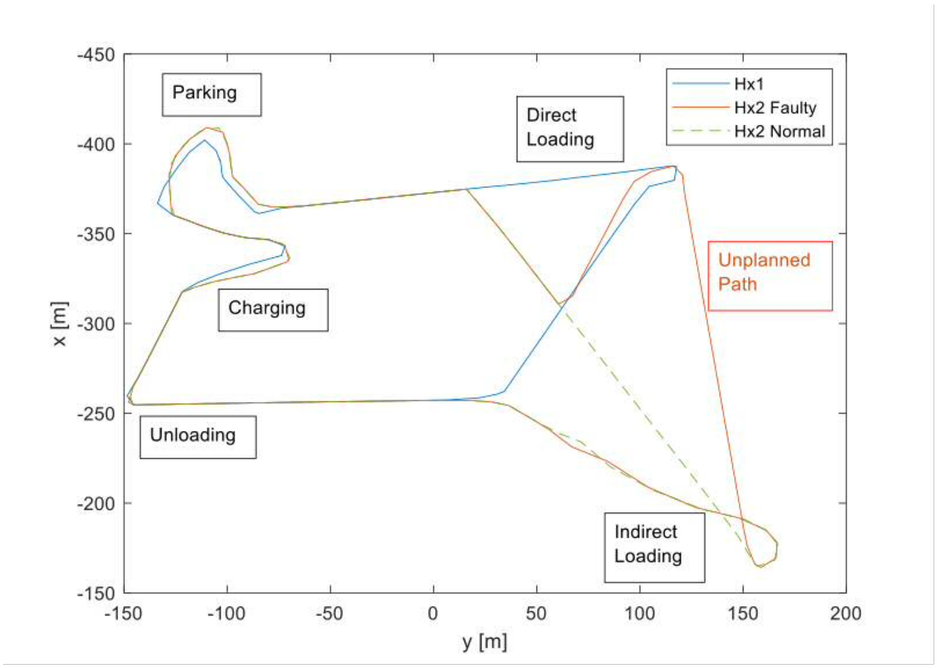


Figure 16: Paths played by the Hxs in Scenario-2

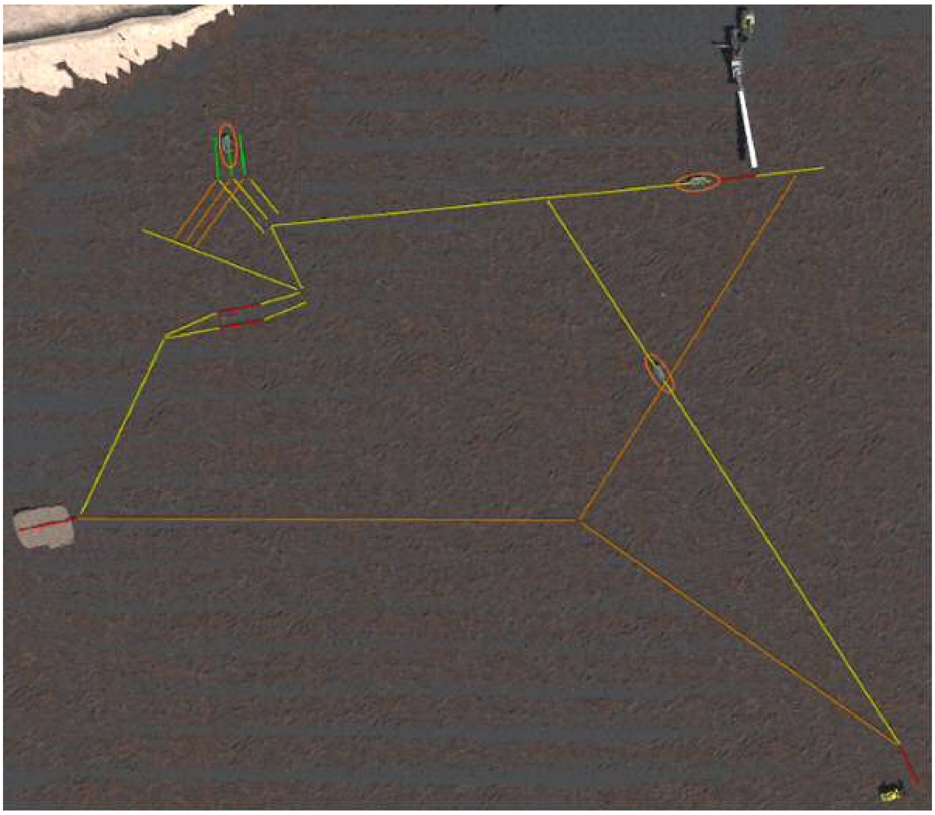


Figure 17: Hxs while performing Scenario-2 (Recording available at https://youtu.be/NKqs4m5zkkQ)

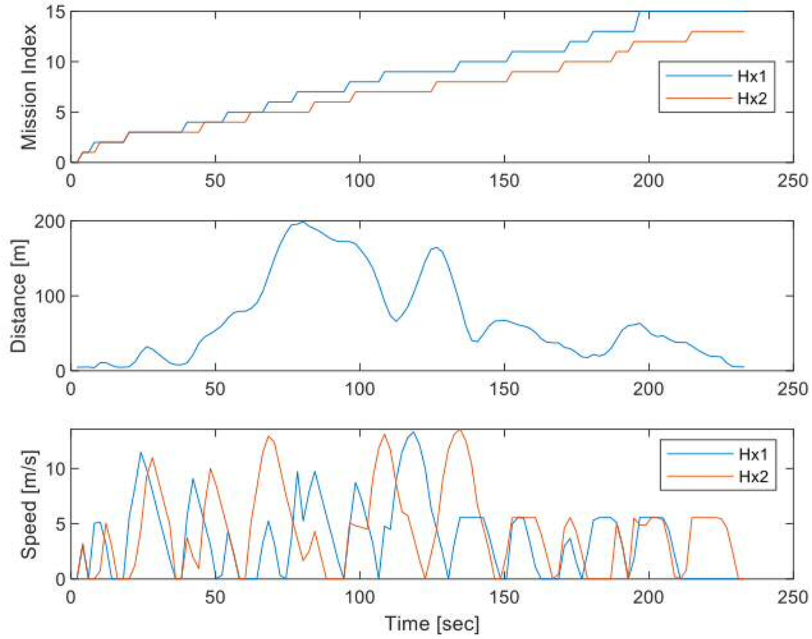


Figure 18: The distance between Hxs and their speeds while performing Scenario-2

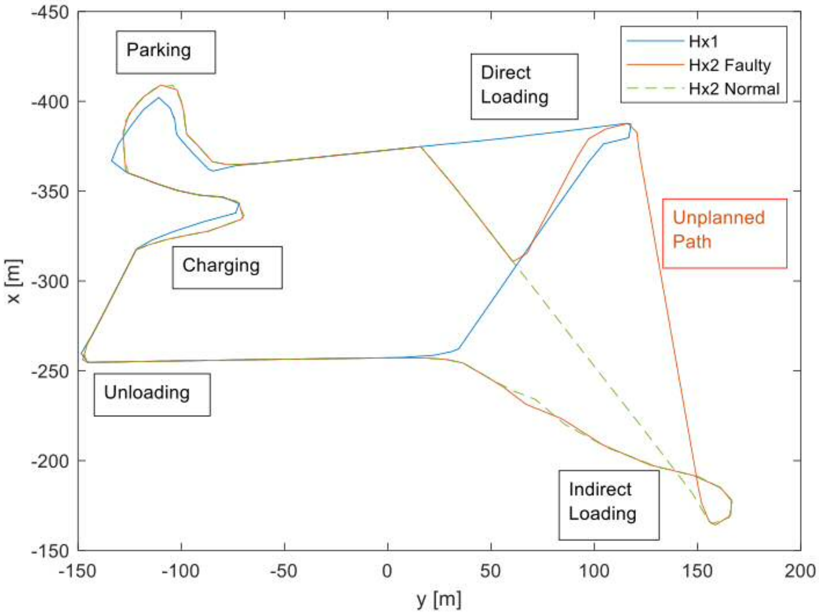


Figure 19: Paths played by the Hxs during Scenario-3

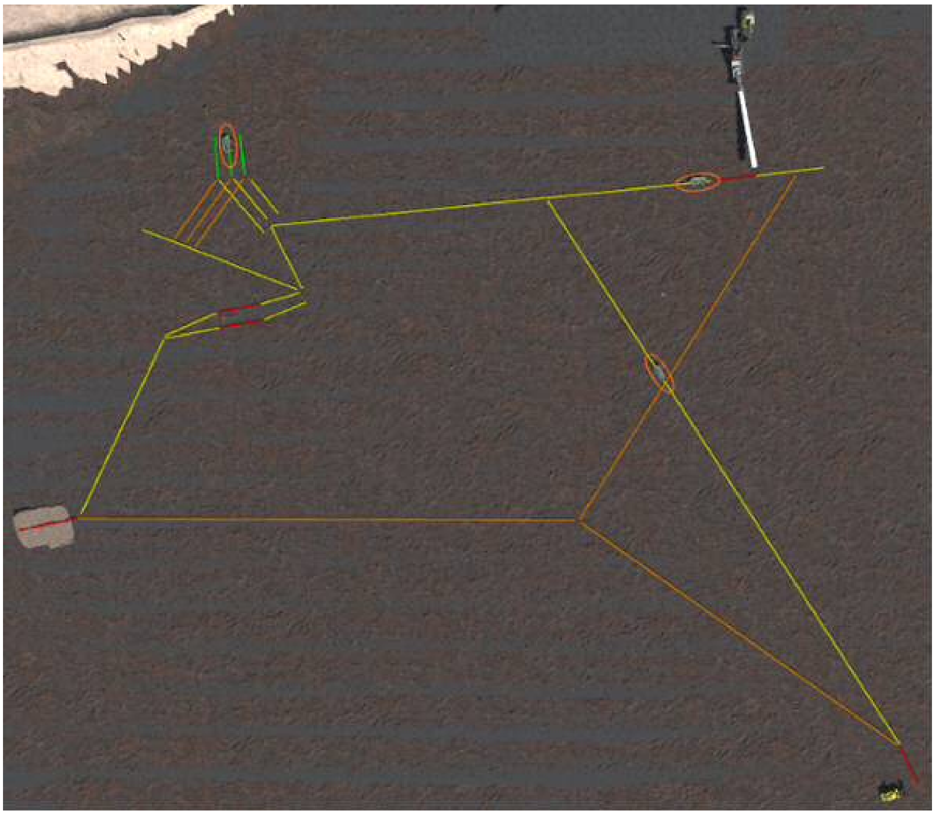


Figure 20: Hxs while performing Scenario-3 (Recording available at https://youtu.be/$_I$ $-$ $hyGwOoYM$))

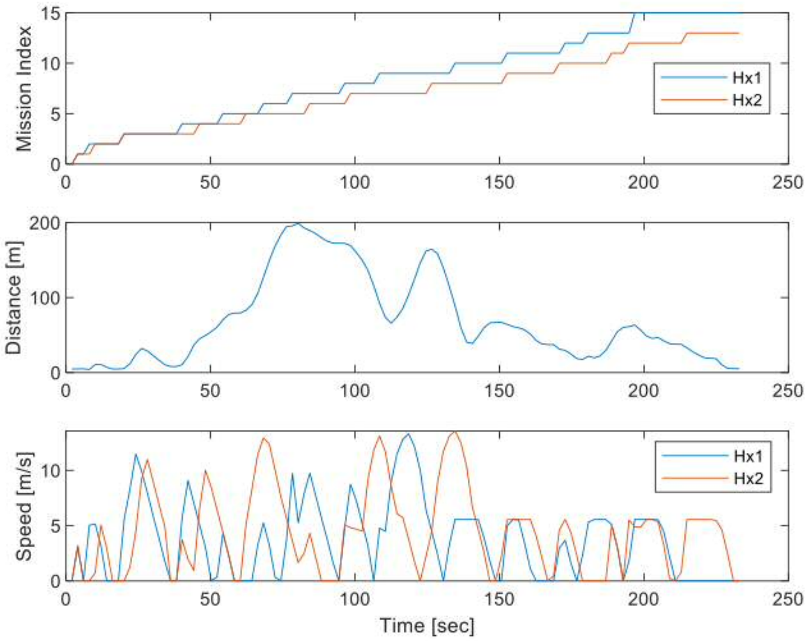


Figure 21: The distance between Hxs and their speeds while performing Scenario-3

5 Contributions to the Body of Knowledge

We has provided 4 contributions to the Body of Knowledge mainly based on the research and experiences from the demonstrator Project. Short summaries of these contributions are provided in the following subsections. The details are attached as part of the Appendix-B

5.1 STPA – Challenges to apply in System of Systems Hazard Analysis

Performing hazard analysis of system-of-systems (SoS) is hard due to emergent behaviours. One potentially useful and popular approach to analyze the safety for complex systems is the System Theoretic ProcessAnalysis (STPA). We tried to use STPA for hazard analysis of SUCCESS Project use case of Electric Site. However, STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations. We have described the STPA process and the problems we faced while applying it in this document.

5.2 End-to-End Tool Framework for Safety Analysis

To provide an end-to-end tool framework for safety analysis, different tools are integrated in the AMASS [AMA16] platform that facilitate (i) modelling of standards and safety process in EPF Composer[9], (ii) modelling of systems, contract-based design and different model-based analyses in Polarys CHESS toolset[10], (iii) assurance case modelling in Polarys OpenCert tools platform[11], and (iv) formal verification of assumption guarantee contracts with OCRA[12]. Figure 22 shows the framework for safety analysis along with integration of involved tools. The end-to-end tool framework for safety analysis consists of following main steps.

- Generation of process-based arguments
- Generation of product-based arguments
- Update of argument fragments during operational phase

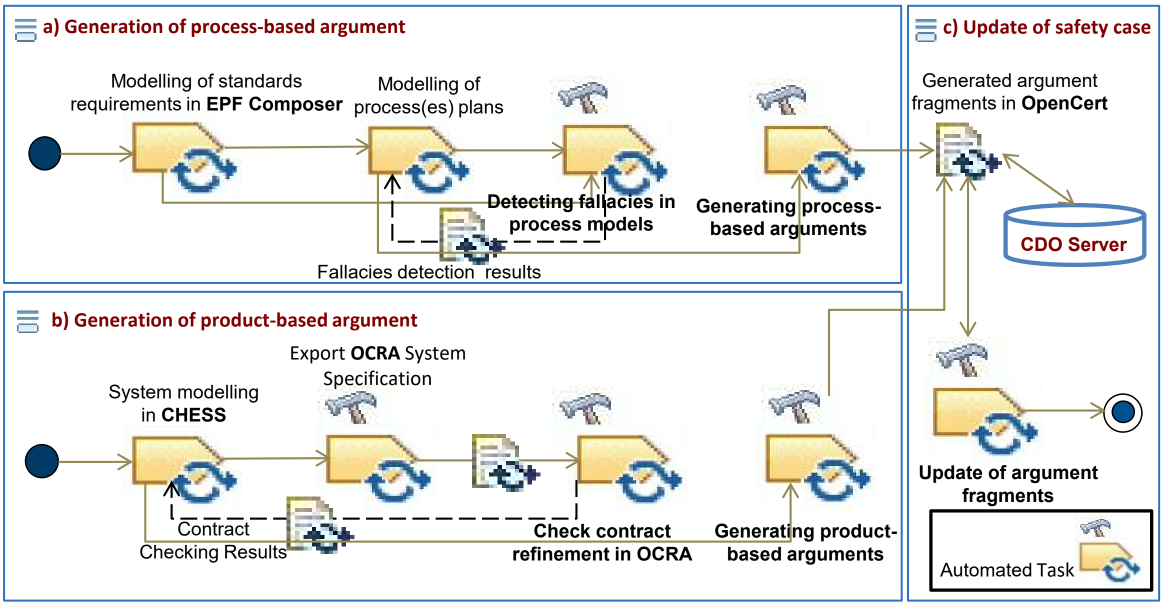

a) Generation of process-based argument
Modelling of standards requirements in EPF Composer
Modelling of process(es) plans
Detecting fallacies in process models
Generating process-based arguments
Fallacies detection results
c) Update of safety case
Generated argument fragments in OpenCert
CDO Server
b) Generation of product-based argument
System modelling in CHESS
Export OCRA System Specification
Contract Checking Results
Check contract refinement in OCRA
Generating product-based arguments
Update of argument fragments
Automated Task

Figure 22: Overview of the tool framework

[9] https://www.eclipse.org/epf/
[10] https://www.polarsys.org/chess/index.html
[11] https://www.polarsys.org/projects/polarsys.opencert
[12] https://ocra.fbk.eu/

Generation of process-based arguments: EPF (Eclipse Process Framework) Composer is used to model the requirements listed in the standards and process plans, as well as to show basic compliance through mapping requirements. Sufficiency fallacies, particularly omission of key evidence in which no or less evidences are provided to support the claim or no valid reasons (rationales) are given for its omission. A plugin is implemented that performed a validation to detect whether the process modelled in EPF Composer contains the sufficient information corresponding to the key evidence for supporting the specific requirement [MGR18]. In case of omitted crucial evidence detail, the feedback is provided regarding detected fallacies and recommendations to resolve them. The process model is modified based on the provided recommendations. Once process model has been modified, the process-based arguments (model and diagram) are automatically generated from process and are visualized via the argumentation editor in OpenCert platform.The generated process-based arguments model and diagram are not only saved locally in a new project into the current workspace, but also stored in the corresponding destination assurance case in the CDO server.

Generation of product-based arguments: In the PolarSys CHESS (Composition with Guarantees for High-integrity Embedded Software Components Assembly) toolset, an editor is implemented to model all phases of system development, for instance, SysML Block Definition Diagram (BDD) and Internal Block Diagram (IBD) can be used to model the system hierarchical architecture i.e., blocks, ports and connections. CHESS toolset also supports the modelling of contracts (i.e. the assumption and the guarantee properties) and their association with components and system requirements. The integration of CHESS with Othello Contracts Refinement Analysis (OCRA) verification engine allows the validation of component contract assumptions against the specification of other components in the system. The safety case (argument-fragments) can be generated from the selected CHESS model (contract-based architectural specification) [SGCH13]. Argument Generator plugin implemented in OpenCert assumes that the analysed model and the refinement check results are stored in the refinement analysis context. The generated set of product-based argument-fragments stored in the corresponding destination assurance case in the CDO server stated in the OpenCert preferences.

Update of argument fragments during operational phase: As mentioned in Section 2.1.3, the operational data is utilized to monitor and evaluate deviations between the intended behaviour reflected in safety cases and actual behaviour of systems during operation; subsequently, we update the modeled safety cases and associated contracts in the OpenCert platform.

5.3 Safety Requirements in the Context of Evolution and Dynamic Risk Management

Based on the SoS hazard analysis that was carried out as a first step towards safe production site, appropriate mitigation mechanisms such as geofences were established to ensure safety during the normal flow of operations and the failure cases. They were then translated into the safety requirements that are implemented in digital twins. Specifically, the guidance is provided for the enforcement of different kind of geofences for safety assurance in automated transportation/production contexts.

5.4 Guidelines on Safety analysis of System of Systems(SoS)

In Success we studied STPA and it shortcomings when it comes to system-of-systems. One lack we identified is the missing documentation about relevant parts of the system-of-systems. We developed a process called SafeSoS as shown in Figure 23. We utilize a hierarchical approach to document the system-of-systems in order to conduct a safety analysis. We foresee a SoS Macro Level, where general site characteristics are captured. Use cases on what should happen at the site is described as well. The information provided in the Macro Level is then refined in the SoS Meso Level. Here, the details of interaction between constituent systems and humans are specified. Finally, details about a single constituent systems is provided in the SoS Micro Level. This information will help conducting a safety analysis for a system-of-systems. This is an ongoing work and initial thoughts on how to perform safety analysis of systems of systems in an effective and efficient manner has been outlined in our recent publication at ISSRE 2020. This will be further developed and will be covered as part of Stephan Baumgart's PhD thesis or in another detailed future publication.

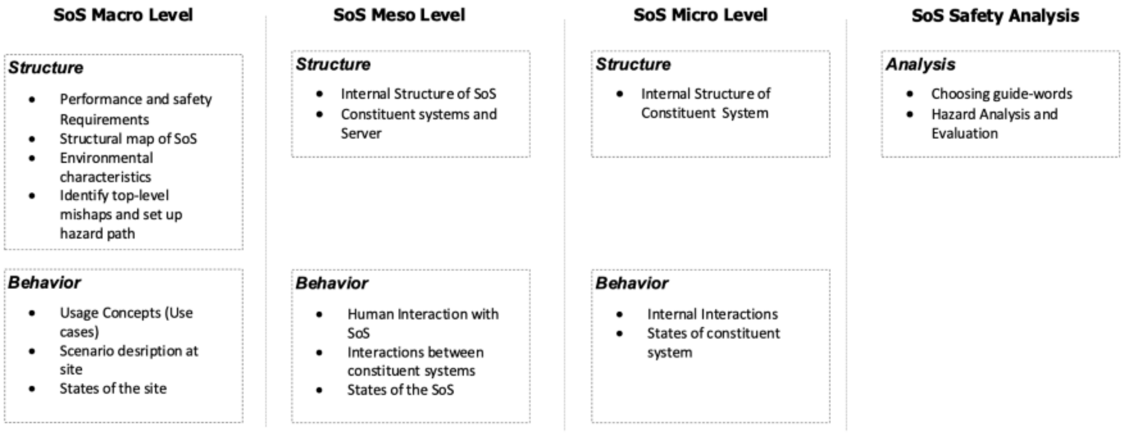

SoS Macro Level
Structure
Performance and safety Requirements
Structural map of SoS
Environmental characteristics
Identify top-level mishaps and set up hazard path
Behavior
Usage Concepts (Use cases)
Scenario desription at site
States of the site
SoS Meso Level
Structure
Internal Structure of SoS
Constituent systems and Server
Behavior
Human Interaction with SoS
Interactions between constituent systems
States of the SoS
SoS Micro Level
Structure
Internal Structure of Constituent System
Behavior
Internal Interactions
States of constituent system
SoS Safety Analysis
Analysis
Choosing guide-words
Hazard Analysis and Evaluation

Figure 23: SafeSoS: Safety Process to support System-of-Systems

6 Assessment Perspective

IEC 61508 is a basic functional safety standard applicable to all kinds of industry. It defines functional safety as: "part of the overall safety relating to the EUC (Equipment Under Control) and the EUC control system which depends on the correct functioning of the E/E/PE safety-related systems, other technology safety-related systems and external risk reduction facilities."

The international standard ISO 26262 for Functional Safety in automotive electronic safety critical systems is being used, at different levels of the automotive industry since its release. It has received wide adoption since it is considered state-of-the-art practice and hence its usage in automotive electronics is informally mandated. However, standards as such, including ISO 26262, need to be interpreted and be well understood for cost-efficient and pragmatic deployment and usage within an organisation. In addition, the standard describes a role for third party assessment and audit. However, ISO 26262 does not cover automated or cooperative vehicles. Results from research is vital for the standard to reflect and be applicable to new technology.

Though there exist certain standards such as ISO13482 on robots and robotic devices, currently there exists no safety standards or safety certification procedures for either mobile logistics robot systems or cooperative mobile robot systems. Furthermore, there have not been any formal specifications published regarding requirements, design, architectures, safety etc.

Machines in the earth moving machinery domain need to fulfil, for example, European regulations. The products require a CE marking to show conformity with health, safety and environmental protection requirements. In other regions in the world, similar regulations exist, and compliance needs to be shown in order to sell products in these markets. A list of standards a new product needs to comply with is maintained at Volvo CE and has evolved over the last decades.

One aspect of these regulations covers the safety of these products including active and passive safety and functional safety. In the earth moving machinery domain, compliance related to relevant machine directives are required. Relevant functional safety standards are ISO 15998, ISO 13849 and ISO 19014 (under development), but again they all focus only on single human operated machines. Current efforts towards automated machines are ongoing like ISO 17757 (Autonomous and semi-autonomous machine system safety). We had constant interactions with the Volvo CE Laws and Regulation Team to obtain feedback related various standardization-related queries, alignments and new ideas. As the list of standards for an earth moving machinery system-of-systems evolve, we expect to have a direct impact on the ongoing standardization work.

Throughout the Project we have been discussing various scenarios with our project partner Safety Integrity to incorporate the safety assessments perspective. This provided us with guidance during our tasks, especially from multiple angles on certification/assurance.

Standardization time schedules are typically far longer than the duration of a short term (less than 2 years) project such as SUCCESS, and the activities involved in influencing them involve many stake holders and focused long term activities.

We plan to further refine the standardization-worthy themes such as SoS safety assurance and dynamic risk management and look for potential standardization venues for presentation of them. Further experiences in joint projects with VCE etc are expected to drive future standardization efforts beyond SUCCESS.

7 Education and Training Perspective

7.1 Industrial Educational Needs

For deployment of the approaches prototyped in the demonstrator in actual production/development involved staff need to understand:

1. Fundamental concepts of assurance
2. The elements of the approach
3. How to use the involved methods and tools
4. How to apply the approach

Additionally one has to provide clear descriptions and training on what the approach provides and what needs to be provided by other means/tools.

The focus of the project was more on clarifying the suitability of approaches and refining them. Hence no specific industrial training sessions were conducted at partner organizations based on SUCCESS results. However through our team members we expect to drive these initiatives in future. Special noteworth aspect is that Safety Integrity AB is specialised in conducting training sessions for various industrial customers and we expect an impact of our experiences and insights carrying forward into their courses.

7.2 Contributions to the Existing Academic Courses at MDH

MDH is conducting two advanced level courses DVA437 (Safety Critical Systems Engineering)[13] and DVA467 (Quality Assurance - Certification of Safety Critical Software Systems)[14], and the post-doc in SUCCESS project Faiz Ul Muram, has been a key instructor in both these courses. The overall objective of DVA437 is to provide the students with knowledge of working with system development of safety-critical applications, as well as development according to a safety standard. DVA437 is primarily offered to the master level students in MDH. Topics covered in the DVA467 includes the relation between the safety-critical systems, development processes according to supplier and manufacturer perspective, overview of standards used as a baseline for certification, compliance management, and modelling methods stemming from state-of-the-art safety standards for developing reusable certification artefacts. DVA467 targets the professionals working in the industry and is given as part of the PROMPT national education alternative for competence development at the industry[15].

7.3 Ongoing/Future Plans for Educational Contributions

The above mentioned courses DVA437 and DVA467 at MDH were primarily developed based on the results of our previous EU funded projects CHESS and AMASS [AMA16]. The work conducted in the context of SUCCESS provides the opportunity to further refine and extend the teaching materials and project assignments of these courses in relation to the system-of-systems and dynamic aspects. Specifically, the emergent behaviours in collaborative systems-of-systems, trade-offs between safety and performance, dynamic reconfiguration, and risk management are regarded as fundamental considerations. Furthermore, the digital twin based simulators that we adapted and extended in SUCCESS is providing important basis to student projects (e.g. thesis projects) and to our follow-up projects.

As an impact of the current COVID-19 epidemic, many industries in Sweden including our main partner Volvo have faced major challenges resulting in partial-total lockdown periods and lay-offs. There is also an increased focus on work-from home as far as possible. In this context, in response to a

[13] https://www.mdh.se/en/malardalen-university/education/course-syllabus?id=27287

[14] http://www.promptedu.se/quality-assurance-certification-of-safety-critical-software-systems/

[15] http://www.promptedu.se

Swedish funding agency (KKS) call, our project proposal ARCS has received funding for development of online courses on Augmented reality and Cyber Security to a) improve the productivity in changed working context as well as to b) support the workforce to improve their skillsets. The subjects of our proposed courses will have also the underpinning of safety and assurance contexts of factories and the interplay between these subjects and safety. Hans Hansson and Sasikumar Punnekkat are part of the ARCS project core team.

The demonstrator is providing important basis for follow-up projects (applications pending) and also for our online professional courses, including several functional safety and cyber security related courses. We expect that the project insights and results will be leveraged on in future course developments. Additionally, we run student projects (e.g. thesis projects) in the Volvo lab at the university that will leverage on results and insights from this project.

8 Collaborations

Here we present a portfolio of projects we have been associated or collaborating with, which had mutually beneficial synergies with SUCCESS in terms of joint research activities, publications and knowledge transfer.

FiC Project: SSF funded Future Factories in the cloud Project aims at providing an efficient and predictable digital infrastructure for computation and communication, and coping with the huge amounts of data needed to provide the envisioned intelligence in Industry 4.0 context. Safety and security are key concerns here and University of York and MDH have been jointly leading the work package on Safety assurance in FiC.

FORA Project: EU ITN-EID Project where 2 PhD students (Zeinab Bakshi Nitin Desai) from MDH conducts research on safety assurance and dependability of FoG computing platform.

ARRAY Project: KKS funded Industrial PhD school, with two related PhD student projects: 1) Björn Leander(ABB) focusing on Security in smart manufacturing and 2) Dino Mustafa (Alten): focusing on security and safety assurance in the context of internet of medical things.

Other PhD Projects Ayhan Mehmed(TTTech, Vienna) has been an industrial PhD student in our research group, with research focus on run-time monitoring for automated driving systems and proposed a safe driving envelope verification concept and related methods for fault tolerance of ADS. Ayhan's PhD defence is scheduled on 23-Nov-2020.

Martin Skoglund (RISE Research Institutes of Sweden) is an industrial PhD student with research focus on testing based assurance of automated vehicles.

9 Beyond SUCCESS

Overall the SUCCESS project provided a strong focus on safety of collaborating system of systems and excellent momentum to our research activities in this key area of industrial relevance. The recently started EU-ECSEL funded InSecTT[16] (Intelligent secure and trustable 'things') project, where MDH is a key partner is a pan-European effort with 54 key European partners aiming to bring the Internet of Things and Artificial Intelligence together in providing secure and trustworthy systems and solutions. In InSecTT, MDH will specifically explore various challenges, including assurance, related to security (and its interplay with safety) in industrial control systems for collaborative modular manufacturing as well as industrial networks. We are also in the process of formulating project proposals together with industrial partners such as ABB, Volvo, Ericsson and the Mining industry to take the SUCCESS results to the next levels of concrete implementations.

Specifically related to MDH, SUCCESS has provided important basis for follow-up activities which include direct collaboration with industrial partners and bringing SUCCESS results and insights closer to innovation and societal value. SUCCESS also played a major role in strengthening the university's profile in safety assurance by bringing focus on to collaborating robots and autonomous systems, which are becoming omnipresent and touching our everyday life closer than before.

[16] https://www.insectt.eu/

References

[AMA16] AMASS. Architecture-driven, Multi-concern and Seamless Assurance and Certification of Cyber-Physical Systems. `http://www.amass-ecsel.eu/`, 2016.

[BL17] Nikita Bhardwaj and Peter Liggesmeyer. A runtime risk assessment concept for safe reconfiguration in open adaptive systems. In *Computer Safety, Reliability, and Security (SAFECOMP) 2017 Workshops, Trento, Italy, September 12*, pages 309–316, 2017.

[Cif15] Cifton A. Ericson. *Hazard analysis techniques for system safety.* Wileys, 2015.

[DFVA10] Jordi Dunjó, Vasilis Fthenakis, Juan A. Vílchez, and Josep Arnaldos. Hazard and operability (HAZOP) analysis. a literature review. *Journal of Hazardous Materials*, 173(1):19 – 32, 2010.

[dOBM+15] André Luíz de Oliveira, Rosana T. V. Braga, Paulo César Masiero, Yiannis Papadopoulos, Ibrahim Habli, and Tim Kelly. Supporting the automated generation of modular product line safety cases. In *Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX, Brunów, Poland, June 29 - July 3*, pages 319–330, 2015.

[DPH15] E. Denney, G. J. Pai, and I. Habli. Dynamic safety cases for through-life safety assurance. In *37th IEEE/ACM International Conference on Software Engineering, ICSE 2015, Florence, Italy, May 16-24*, pages 587–590, 2015.

[FBGCGGPBP17] José Fuentes-Bargues, Maria González-Cruz, C González-Gaya, and Mª Piedad Baixauli-Pérez. Risk analysis of a fuel storage terminal using HAZOP and FTA. *International Journal of Environmental Research and Public Health*, 14:705, 06 2017.

[GT06] Holger Giese and Matthias Tichy. Component-based hazard analysis: Optimal designs, product lines, and online-reconfiguration. In *25th International Conference on Computer Safety, Reliability, and Security (SAFECOMP), Gdansk, Poland, September 27-29*, pages 156–169, 2006.

[HC09] Charles Haddon-Cave. The Nimrod Review: An Independent Review into the Broader Issues surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006. Report, The Stationery Office, London, October 2009.

[HK07] Ibrahim Habli and Tim Kelly. Challenges of establishing a software product line for an aerospace engine monitoring system. In *11th International Software Product Line Conference (SPLC), Kyoto, Japan, September 10-14*, pages 193–202, 2007.

[HKP09] Ibrahim Habli, Tim Kelly, and Richard Freeman Paige. Functional hazard assessment in product-lines – a model-based approach. In *1st International Workshop on Model-Driven Product Line Engineering (MDPLE' 09), Twente, The Netherlands, June 24th*, 2009.

[KGG18] Yoram Koren, Xi Gu, and Weihong Guo. Reconfigurable manufacturing systems: Principles, design, and future trends. *Frontiers of Mechanical Engineering*, 13(2):121–136, Jun 2018.

[Lev18] John P. Leveson, Nancy G.; Thomas. STPA Handbook. page 188, 2018.

[MGR18] Faiz Ul Muram, Barbara Gallina, and Laura Gomez Rodriguez. Preventing omission of key evidence fallacy in process-based argumentations. In *11th International Conference on the Quality of Information and Communications Technology (QUATIC), Coimbra, Portugal, September 4-7*, pages 65–73, 2018.

[MJH19] John McDermid, Yan Jia, and Ibrahim Habli. Towards a framework for safety assurance of autonomous systems. In *Proceedings of the Workshop on Artificial Intelligence Safety 2019 co-located with the 28th International Joint Conference on Artificial Intelligence, AISafety@IJCAI 2019, Macao, China, August 11-12*, pages 1–7, 2019.

[MSCR10] James Bret Michael, Man-tak Shing, Kristian John Cruickshank, and Patrick J. Redmond. Hazard analysis and validation metrics framework for system of systems software safety. *IEEE Systems Journal*, 4(2):186–197, 2010.

[Nan12] Nancy G. Leveson. *Engineering a Safer World - Systems Thinking Applied to Safety*. MIT Press, 2012.

[NNG19] Damir Nesic, Mattias Nyberg, and Barbara Gallina. Constructing product-line safety cases from contract-based specifications. In *34th ACM/SIGAPP Symposium on Applied Computing (SAC), Limassol, Cyprus, April 8-12*, pages 2022–2031, 2019.

[PHST12] Claudia Priesterjahn, Christian Heinzemann, Wilhelm Schäfer, and Matthias Tichy. Runtime safety analysis for safe reconfiguration. In *IEEE 10th International Conference on Industrial Informatics (INDIN), Beijing, China, July 25-27*, pages 1092–1097, 2012.

[PST13] Claudia Priesterjahn, Dominik Steenken, and Matthias Tichy. Timed hazard analysis of self-healing systems. In *Assurances for Self-Adaptive Systems - Principles, Models, and Techniques*, pages 112–151. 2013.

[RMS08] Patrick J. Redmond, J. Bret Michael, and Paul V. Shebalin. Interface hazard analysis for system of systems. In *3rd IEEE International Conference on System of Systems Engineering (SoSE), Singapore, June 2-4*, pages 1–8, 2008.

[SFD+11] Zoë Stephenson, Christian Fairburn, Georgios Despotou, Tim P. Kelly, Nicola Herbert, and Bruce Daughtrey. Distinguishing fact from fiction in a system of systems safety case. In *Advances in Systems Safety–Proceedings of the Nineteenth Safety-Critical Systems Symposium, Southampton, UK, February 8-10*, pages 55–72, 2011.

[SGCH13] Irfan Sljivo, Barbara Gallina, Jan Carlson, and Hans Hansson. Strong and weak contract formalism for third-party component reuse. In *IEEE 24th International Symposium on Software Reliability Engineering, ISSRE 2013, Pasadena, CA, USA, November 4-7, 2013 - Supplemental Proceedings*, pages 359–364. IEEE Computer Society, 2013.

[SPAR18] Michael Schluse, Marc Priggemeyer, Linus Atorf, and Jürgen Rossmann. Experimentable digital twins - streamlining simulation-based systems engineering for industry 4.0. *IEEE Trans. Industrial Informatics*, 14(4):1722–1731, 2018.

[The18] The Assurance Case Working Group. Goal Structuring Notation Community Standard Version 2, January, 2018. [Online] `http://www.goalstructuringnotation.info/`, 2018.

[VHC+15] Anatoly Vasilevskiy, Øystein Haugen, Franck Chauvel, Martin Fagereng Johansen, and Daisuke Shimbara. The BVR tool bundle to support product line engineering. In *19th International Conference on Software Product Line (SPLC), Nashville, TN, USA, July 20-24*, pages 380–384, 2015.

[XA08] Liudong Xing and Suprasad V. Amari. Fault tree analysis. In Krishna B. Misra, editor, *Handbook of Performability Engineering*, chapter 38, pages 595–620. Springer London, London, 2008.

[ZKS+17] Eloise G. Zimbelman, Robert F. Keefe, Eva K. Strand, Crystal A. Kolden, and Ann M. Wempe. Hazards in motion: Development of mobile geofences for use in logging safety. *Sensors*, 17(4):822, 2017.

Appendix A: Publications

2019 4th International Conference on System Reliability and Safety

A State-based Extension to STPA for Safety-Critical System-of-Systems

Stephan Baumgart , Joakim Fröberg†‡, Sasikumar Punnekkat‡
System Architecture Department, Volvo Construction Equipment, Eskilstuna, Sweden
† Research Institutes of Sweden, RISE ICT/SICS Västeråas, Sweden
‡School of Innovation, Design and Engineering, Malardalen University, Väasterås, Sweden
e-mail: stephan.baumgart@volvo.com, joakim.froberg@ri.se, sasikumar.punnekkat@mdh.se

Abstract—**Automation of earth moving machinery enables improving existing production workflows in various applications like surface mines, material handling operations or material transporting. Such connected and collaborating autonomous machines can be seen as a system-of-systems. It is not yet clear how to consider safety during the development of such system-of-systems (SoS). One potentially useful approach to analyze the safety for complex systems is the System Theoretic Process Analysis (STPA). However, STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations. In this paper, we present an approach for enriching STPA to provide the ability to check whether the distributed constituent systems of a SoS have a consistent perspective of the global state which is necessary to ensure safety. In other words, these checks must be capable at least to identify and highlight inconsistencies that can lead to critical situations. We describe the above approach by taking a specific case of state change related issues that could potentially be missed by STPA by looking at an industrial case. By applying Petri nets, we show that possible critical situations related to state changes are not identified by STPA. In this context we also propose a model-based extension to STPA and show how our new process could function in tandem with STPA.**

Keywords-hazard analysis and risk assessment; system-of-systems; autonomous machines; STPA; Safety; Petri net

I. INTRODUCTION

Developing safety critical products requires to understand how the targeted customers use the products. This will help to identify those situations where human errors or failures in the involved systems may lead to critical accidents. Apart from focusing on features to avoid accidents or to reduce the impact of accidents, functional safety focuses on designing the electrical and electronic system (E/E) in such a way that faults in the E/E system will not lead to accidents and the system is put into a safe state. Considering functional safety during the development requires rigor in following development processes defined in the functional safety standards. These standards help developers to avoid critical systematic failures in software and random failures in hardware. Domain specific functional safety standards like ISO 26262 [1] for the automotive domain, ISO 13849 [2] or ISO 19014 [3] for the earth moving machinery domain or the generic functional safety standard IEC 61508 [4] provide guidance for ensuring functional safety during development of safety critical products. As an initial phase, potential hazards related to a product need to be identified and analyzed. Hazard analysis methods applied in development processes in industry are for example Preliminary Hazard Analysis (PHA) [5], Hazard and Risk Assessment (HARA) [1], Fault Tree Analysis (FTA) [6] and Failure Mode and Effect Analysis (FMEA) [7]. PHA and HARA are applied during early phases in the development process to list and evaluate possible hazards related to the product to be developed. FTA and FMEA are applied during later stages in the development process as they require detailed knowledge about the targeted architecture and the used components. The processes described in the functional safety standards as well as the established hazard analysis methods focus on single, human operated machines in the example of earth moving machinery. Currently, there is a paradigm shift in many domains towards adding automation to aid drivers, increase productivity and reduce risks by eliminating human errors. In the earth moving machinery domain, automation of machines enables the improvement of production workflows and the increase the efficiency as it has been shown in the Electric Site Research Project [8]. In this project a fleet of eight autonomous haulers (called HX) are utilized to transport material in an open surface quarry mine. A central server coordinates the fleet of HX and provides missions to each single HX depending on relevant site and individual scenarios. These machines collaborate to achieve common tasks, e.g. transporting material in the quarry site. Additionally, other human-operated machine can be used to interact with the autonomous machines. Such a system can be seen as a system-of-systems. A system-of-systems is defined in [9] as a "system that has operational and managerial independence of its elements." This means, that the involved systems of a

SoS must be able to be operated independent from the SoS to provide a useful purpose. With managerial independence the author emphasize that the involved systems can be "separately acquired and integrated". Periorellis et al. [10] describe that "the purpose of a SoS is to provide a set of enhanced or improved "emergent" services, based on some or all of the services provided by the participating component systems. The provision of these

978-1-7281-4781-9/19/$31.00 ©2019 IEEE

emergent services requires co-operation between the systems." The term system-of-systems (SoS) implies that these individual systems can be grouped and connected to provide services not achievable by one single system alone. System-of-systems rely on communicating between the independent and geographical distributed systems as failing of communication, providing erroneous data or misinterpreting correct data may lead to accidents [11].

One potentially useful approach to analyze the safety for complex systems is the System Theoretic Process Analysis (STPA) [12], which we apply to an industrial case for system-of-systems from the earth moving machinery domain in the scope of this paper. In order to identify all critical situations for a system-of-systems, such a method must be able to deal with emergent and dysfunctional behaviors of a SoS. The objective of this paper is to present an approach for enriching STPA to provide the ability to check whether the distributed constituent systems of a SoS have a consistent perspective of the global state which is necessary to ensure safety.

This paper is structured as follows. We describe the background of our paper in section II. In section III we present a case of automated machines from the earth moving machinery domain. The related work is described in section IV. We describe STPA in section V and apply it to our industrial case. We present an enhancement for STPA to additionally identify inconsistencies in system-of-systems by using Petri nets in section VI. By applying our proposed enhancement to the industrial case, we show how additional critical situations can be found. We analyze and discuss our results in section VII and conclude our paper in section VIII.

II. Background

In this section we provide background information to our work.

A. Hazard Analysis

Hazard analysis methods can be distinguished into two major groups. The first group contains methods that aim to identify and evaluate hazards during early development phases. Typical examples are PHA [13], HARA [1], the Machine Control System Safety Analysis (MCSSA) [3] or the Hazard and Operability Studies (HAZOP) [14]. Each of these methods requires in the first stage to identify the main function of the product that shall be developed. As a second stage the foreseen operation modes shall be identified as for example Idling, Working or Maintenance for the earth moving machinery domain. In brainstorming meetings with experts each operational mode will be analyzed how a failing of a function may lead to accidents. Guide words as proposed by HAZOP can provide further structure to such an analysis. Each identified hazard will be rated by estimating the severity of the accident, the probability this failure could happen and if the humans involved have the possibility to avoid the accident to happen by the controls available. The resulting estimates are used to calculate a rating of a hazard, i.e. SIL [4], ASIL [1] or PL [2], which is necessary for tailoring the development processes required by the functional safety standards. The second group of hazard analysis methods is applied during development to trace the identified top-level hazards and analyze the used architecture and components. Typical examples in this group are the top-down analysis method FTA [6] or the bottom-up analysis method FMEA [15]. FTA is using a tree structure, where the root node is a top-level failure that shall be avoided, and the leave nodes are representing components in the architecture of the system to be developed [16]. FMEA is a safety analysis method, which is using a table to list all safety related components of a system. Typically, FTA provides a list of components to be analyzed. During a FMEA failing of each component is analyzed and if this can lead to system failures. The identified critical component failures are rated in the first stage and potential risk mitigation are identified. The FMEA is repeated to analyze if the applied counter measures will lead to the required risk reduction. When designing complex system-of-systems as in our case, we are interested in hazard analysis methods that are able to deal with emergent and dysfunctional behaviors in a SoS.

B. Petri Nets

Various concept of modeling system specifications and system behavior are available. The goal of our work is to be able to model the states of the involved systems of a SoS and to simulate the interactions to find possible critical scenarios. Petri nets for example provide these required properties. Petri nets (PN) represent a "formal model of information flow" [17]. The graphical representation consists of places (P) depicted as circles and transitions (T) depicted as rectangles. Places and transitions "are connected by directed arcs from places to transitions and from transitions to places." [17] A transition has inputs, when arcs point from places to a transition and outputs where the arcs point from a transition to a place. The behavior of a PN is modeled by using markers which are depicted as dots on the places (P). A transition is "consuming" a predefined number of markers from an input and is generating a predefined number of markers in the output places of this transition. If it is required to simulate timing properties, Timed Petri nets [18] or Stochastic Petri nets [19] can be used. The time a transition needs for transforming the markers from the input places and generating the output markers can be defined. Timing is in our case important since a delayed communication for example may lead to critical situations. In the context of safety critical systems, El Koursi et al. [20] highlight that Petri nets can be used for modeling system specifications to check completeness and consistency and to use simulation to check correctness of safety criteria. We are specifically interested in simulating the behavior of complex system-of-systems to find possible design flaws.

III. Industrial Case - Electric Site

We utilize the electric site research project [8] as a case for our work. In this project a fleet of automated guided vehicles (AGVs) [21] called HX are used to transport material at a quarry site, which is a surface mine for gravel production in our case. The pre-crushed material is transported from a movable primary crusher to a stationary secondary crusher. Along with the fleet of autonomous HX, a human-operated wheel loader and a human-operated excavator are used for loading material onto the HX. In our earlier work we have described and analyzed this complex SoS [22], [23].

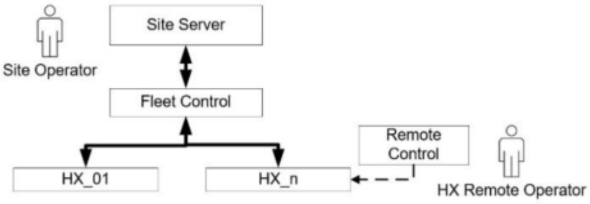


Figure 1. Use Case: Remote Control of HX

The fleet of active HX is controlled by the Fleet Control System, containing features like traffic management or setting missions for each active HX. Each HX is therefore highly dependent on the wireless network and correct commands. In order to be able to activate a HX in the morning, remove a HX for repair purposes or adding a HX to a running production, it is possible at any given instance to control a single HX using a remote control by a HX Remote Operator. The Site Operator is monitoring the quarry site from a control room, where the Site Server is located. In Figure 1 the involved systems and human operators are presented. When designing such a system an in-depth analysis of this scenario is necessary to identify potential hazards leading to critical accidents.

IV. Related Work

We are interested in hazard analysis methods specifically considering system-of-systems and providing support for designing such a system. New approaches have been proposed to analyze hazards for system-of-systems like the System-of-Systems Hazard Analysis (SoSHA) [5], the Interface Hazard Analysis Method [24] or methods utilizing simulations to identify hazards like the Simulation based Hazard Analysis (SimHazan) [25]. These hazard analysis methods assume an integration of existing and already safety certified systems into a system-of-systems. When integrating existing systems into a compound of systems, it is necessary to ensure a safe integration. Furthermore, in many cases human operated machines are integrated into a system-of-systems. The Interface Hazard Analysis Method is focusing for example on the communication channels between the involved systems. In our case, we are designing a system-of-systems including a fleet of autonomous machines. Emergent hazards as described in the taxonomy provided by Redmond [11] may be missed when only considering safety for each single machine.

Instead, we searched for hazard analysis methods supporting the design process of complex safety critical systems. In this process it is important that analysis results are available during early stages in development process to support decision making. We focus in our work on the System Theoretic Process Analysis (STPA), which is a recent approach to analyze safety-critical systems and has grown attention [12], [23], [26]–[30]. We have attempted to use STPA for the Electric Site use case from the earth moving machinery domain [23]. During this exercise, we found several open challenges not clearly solvable by a straight-forward application of STPA and our current research focuses on making the safety analysis efficient by solving those challenges. STPA is aiming to provide inputs during early stages in the development but is not directly considering a quantification of hazards as required by the functional safety standards and as other hazard analysis methods do. Zhang et al. [31] propose an extension to STPA, called STPA-RAM, which adds quantification of identified losses to STPA in order to reduce the number of unsafe control actions and to provide guidance for decision making in industrial projects. The authors utilize Stochastic Petri nets to simulate events and use reliability data from an existing database to calculate the frequency of losses for different cases. Zhu et al. [32] apply Petri nets to formalize the control structure diagram in STPA to support the identification of unsafe control actions and their causal factors. The authors propose a new method called Control Logic Petri Net (CLPN), which is including a Petri net notation and an analysis part to find unsafe control actions. In comparison we are not aiming to replace the control structure diagram of STPA, instead we add a dimension that enables the identification and analysis of inconsistencies in the involved systems of a SoS.

V. System-Theoretic Process Analysis - STPA

To illustrate the application of STPA, we analyze the remote control case and follow the STPA process as described in literature [12].

A. STPA Overview

At first, we provide a short description of STPA. STPA consists of four steps as shown in Figure 2, which we describe in the following section.

STPA - Step 1: During the first step of STPA, the scope of the STPA is set and potential losses and hazards shall be identified. System-level hazards may be derived in brainstorming meetings with experts or by applying hazard identification methods like HAZOP or What-if Analysis. The list of possible system hazards may be extended during later stages when more product knowledge is available.

STPA - Step 2: In Step 2, the control structure of the system is derived. The control structure diagram is a graphical representation of the control actions to aid a structured analysis. The control structure diagram contains the main control elements and control actions between the controllers and the controlled systems.

STPA - Step 3: The control structure diagram is used to apply a structured analysis of each control action and if a failure of the control action would lead to the already listed system-level hazards. STPA uses four guide words for finding such unsafe control actions:

- Not providing causes hazard
- Providing causes hazard
- Too early, too late, out of order
- Stopped too soon, applied too long

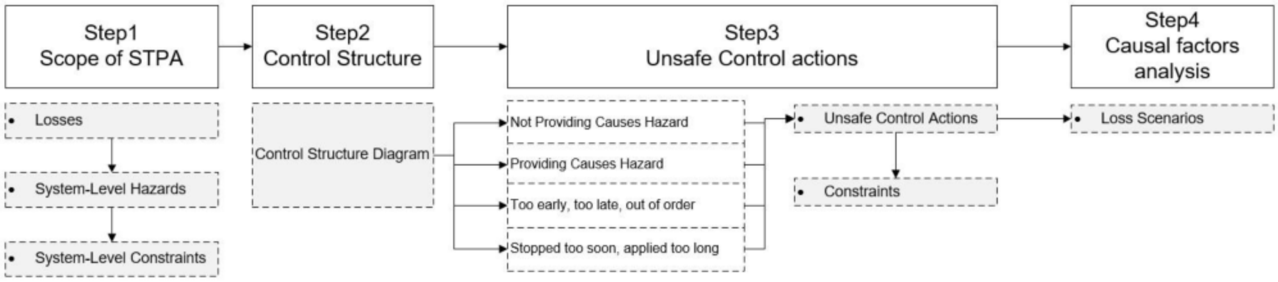


Figure 2. General STPA Process as described in [12]

This means that the following requirements are tested:

- A correct control action is provided.
- A control action is provided at the correct time.
- A control action is provided with correct duration.

STPA - Step 4: In the last step of STPA, possible loss scenarios are identified for each unsafe control action. Reasoning why an unsafe control action would occur and how this could lead to a hazard shall be provided.

STPA – Conclusion: STPA is useful for identifying and analyzing control actions and their causal factors when unsafe control actions are identified. The process of STPA is foreseen to be iterative, i.e. it is possible that further system-level or subsystem-level hazards will be identified during later stages. It is furthermore proposed to add complexity to the control structure diagram during later stages of the development process. This will lead to additional efforts for identifying unsafe control actions in Step 3.

The question is, if STPA is able to deal with emergent and dysfunctional behaviors in the case of system-of-systems. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations.

B. STPA - Application Remote Control Case

In the following we apply STPA to the industrial case described above in section III.

STPA Step 1 - Remote Control Case: For our limited case we have identified two major losses that shall be avoided:

- Loss1: Humans injured or killed
  Situations, where humans are at risk to be injured or killed by the autonomous machines shall be avoided.
- Loss2: Damage of Equipment
  If machines are damaged because of accidents, this may result in a stop of production at the site, which shall be avoided.

Typical SoS hazards in our case can be:

- Hazard 1 (H-1): HX does not maintain safe distance to humans on Site.
- Hazard 2 (H-2): HX enters dangerous area/region
- Hazard 3 (H-3): Squeezing Hazard (e.g. people close to HX)
- Hazard 4 (H-4): Insufficient ability of machinery to be slowed down, stopped and immobilized

STPA Step 2 - Remote Control Case: We simplified the control structure diagram for the purpose of this paper as shown in Figure 3.

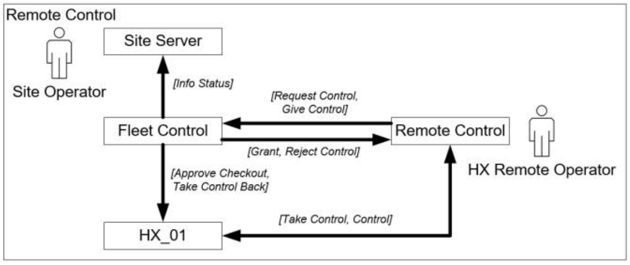


Figure 3. Control Structure Diagram: Remote Control HX 01

The HX Remote Operator sends a request to the Fleet Control server with the purpose to take over the control of a specific HX (HX 01). Fleet Control can decide either to accept (Grant Control) or to reject (Reject Control) the request. At the same time the Fleet Control is sharing information about the active HX with the site server shown by the message Info Status. If the remote control request is accepted, Fleet Control is sending a task (Approve Checkout) to HX 01 to enable the HX to be controlled by the Remote Control. Once this is done, the HX Remote Operator can take control over the HX. The HX Remote Operator can also give back control of HX01 to Fleet Control. Fleet Control will send a request (Take Control Back) to HX 01 that it will listen to controls send from Fleet Control.

STPA Step 3 - Remote Control Case: Each message in the control structure diagram (Fig. 3) is analyzed using the guide words.

We exemplify identifying unsafe control actions by analyzing the messages “Request Control” and “Approve Checkout” in Table I. Applying the first guide word Not providing causes hazard for “Request Control” helps finding the critical situations if the message is either not provided or lost, but this will not directly lead to a hazard. We identify the first unsafe control action (UCA 01) in the situation when the message “Request Control” is provided unintended. This may lead to a situation that a HX is checked out from Fleet Control without awareness of the HX Remote Operator. Humans are at risk, if the machine is moving into dangerous areas, where humans are working (H-2) or if humans are already close by, this may lead to squeezing hazards (H-3). If the signal is delayed (Too early, too late, out of order), this may lead in the worst case to frustration of the operator, but not to hazardous situations. The message "Approve Checkout" is send from the Fleet Control to the HX to indicate, that the HX shall change mode to be controlled by a remote control. We identify, that providing "Approve Checkout" unintended, will lead to a situation where the HX is forced to switch over to be remote controlled. This can lead to critical situations where the HX is moving without a control instance connected to the machine. Altogether, we have identified 15 UCAs for this simplified case during the first brainstorming (Fig. 3).

STPA Step 4 - Remote Control Case: In our case,

"Approve Checkout" might be provided unintended because of a fault in the Fleet Control software or due to a transmission error.

TABLE I. UNSAFE CONTROL ACTIONS

Control Action	Not providing causes hazard	Providing causes hazard	Too early, too late, out of order	Stopped too soon, applied too long
Request Control	Request Control is not provided to Fleet Control [Not Hazardous]	UCA 01: Request Control is sent unintendedly during normal operation. [H-2, H-3]	Request from HX Remote Operator is provided too late. [Not Hazardous]	
Approve Checkout	Approve Checkout is not provided to HX. [Not Hazardous]	UCA 02: Approve Checkout is provided unintended to HX during normal operation. [H-1, H-2, H-3, H-4]		

C. *Conclusion STPA Case Study*

Where is STPA suitable?

STPA is a useful approach to analyze the safety of complex systems. While hazard analysis methods like PHA, FTA and FMEA focus on failures of system functions and their impact, is STPA analyzing possible failures of control actions between the involved systems and sub-systems. This analysis leads to a broader list of possible critical scenarios that require further analysis to list all causal factors. STPA is analyzing the control actions and therefore mostly control and communication related hazards will be identified.

Which critical situations are not captured in STPA? STPA analyzes one single control action a time, which makes it impossible to find critical scenarios which involve for example a combination of control actions, cascading failures or state changes. STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations. It is among others important to check, if the involved systems in a SoS have a consistent perspective of the global state. The states of the involved systems are not considered in the control structure diagram of STPA. Design flaws and casual factors might be missed, if the interaction of state machines is not considered during analysis of the SoS.

VI. STPA ENHANCEMENT

In this section, we present an approach for enriching STPA to provide the ability to check whether the distributed constituent systems of a SoS have a consistent perspective of the global state. In Figure 4 we present our proposed enhancement of STPA. We exemplify three challenges regarding SoS, which require additional analysis efforts:

- Challenge 1: Inconsistent states in SoS We need to be able to even consider the states of the involved systems in a SoS.
- Challenge 2: Communication deadlocks in SoS When analyzing single messages and control actions at a time, it might not be possible to identify if seemingly correct communication will lead to a deadlock.
- Challenge 3: Reachability of Safe States When safe states are already considered, it needs to be checked and analyzed, if specific states can be reached or not. Because states are not considered in the standard STPA analysis, this need to be added.

As shown in Figure 4 we foresee additional formalisms and tools to support an analysis for SoS. Such a method is for example Petri nets, which we apply in the following to identify inconsistencies in a SoS.

A. *Petri Nets - Identifying Critical State Changes*

We apply Petri nets to model the states of the involved systems and simulate state changes and analyze if this may lead to critical situations. The states are depicted as Petri net places, while the transitions between states are shown as directed arcs. Specifically, we use timed transitions to simulate timing aspects and communication delays between the involved systems. The process, we propose for analyzing a SoS using Petri nets is as follows:

1) Model the states of each involved system in a separate Petri net. Prepare interfaces to the other systems using open transitions.
2) Connect all Petri nets to one SoS Petri net and adjust the weight of transitions and arcs and place capacity to enable a workflow as intended.
3) Run simulations of the SoS Petri Net to find possible unintended behavior. Adjust even timing of the transitions for different simulations.

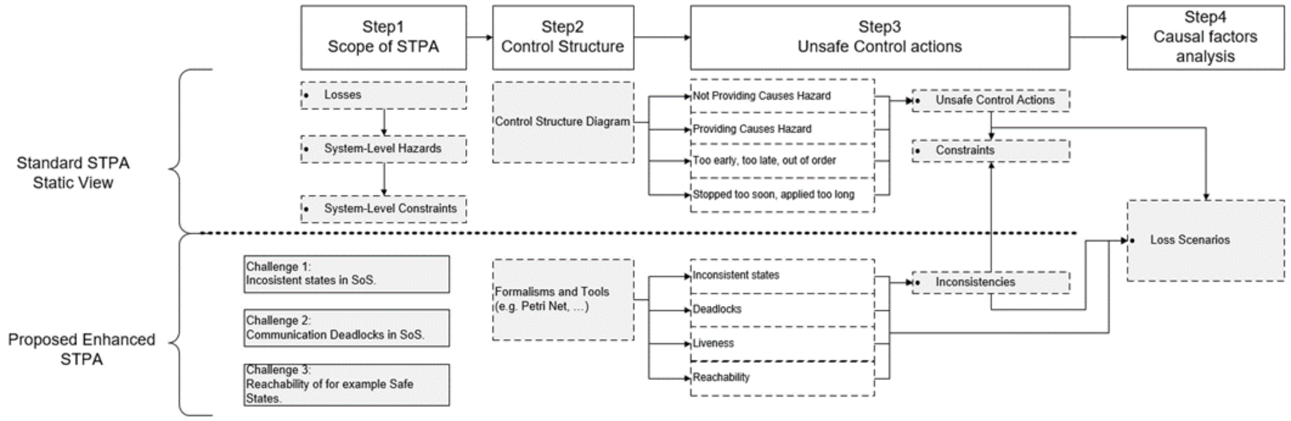


Figure 4. Adding a dimension to STPA to identify inconsistencies in SoS

B. Step 1 - Modeling Single Systems in Petri Nets

The states of each involved system are first modeled separately and prepared for later integration. This helps to manage the complexity of the resulting Petri net. We utilize the remote control case described above but simplify the state machines of the involved systems for the purpose of this paper.

Petri net: HX 01

In Figure 5 the state machine of HX 01 are modeled. While the states of a HX are more complex in reality, we limit our scope and consider only two states, HX 01 Auto for showing that the HX is in autonomous mode and controlled by the Fleet Control. The second state of the HX is HX 01 Remote when the HX is controlled through the remote control. The transitions we consider here are related to switching between the states as shown in Figure 5 and will be triggered externally.

Petri net: Remote control

Now we model the states of the remote control (Fig. 6). Generally, the remote control can either be connected or disconnected to a specific HX, i.e. RC Connected and RC Disconnected. Furthermore, the remote control operator can either give the control back to the Fleet Control Give Control to FC or request the control of a HX Request Control. These states and transitions depend on the overall state of the other systems.

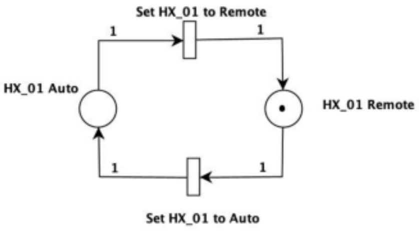


Figure 5. Petri net for HX 01

The Petri net of the remote control is not modeled as a cycle, because changing states is triggered externally by the Fleet Control server. The interfaces to the Fleet Control server are already provided by using the transitions Give Control to FC, triggering that the HX is disconnected from the remote control and Request Control, indicating that the remote control requests be connected to a specific HX.

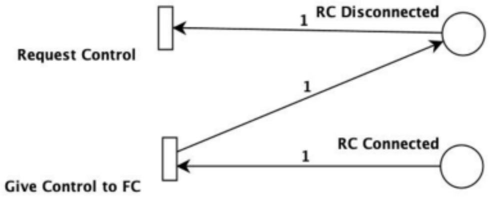


Figure 6. Petri net for Remote Control

Petri net: Fleet Control

The last system in this context is the Fleet Control. A simplified state machine of Fleet Control is presented in Figure 7 consisting of the states FC Control HX 01, indicating that HX 01 is controlled by the Fleet Control server and FC HX 01 Remote, when HX 01 is controlled by the remote control. Once the Fleet Control receives a request from the remote control operator, it can either reject (Reject Control) or grant (Grant Control) control. If the remote control request is granted, HX 01 shall switch mode to be able to be controlled by the remote control.

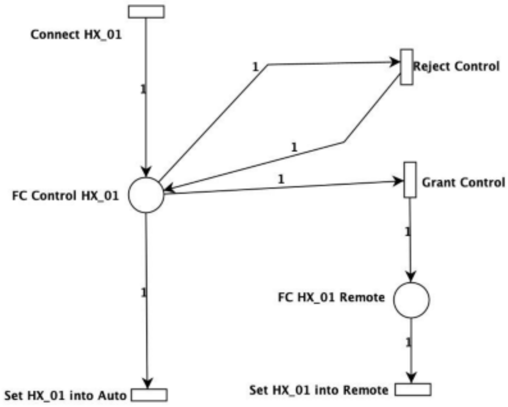


Figure 7. Petri net for Fleet Control

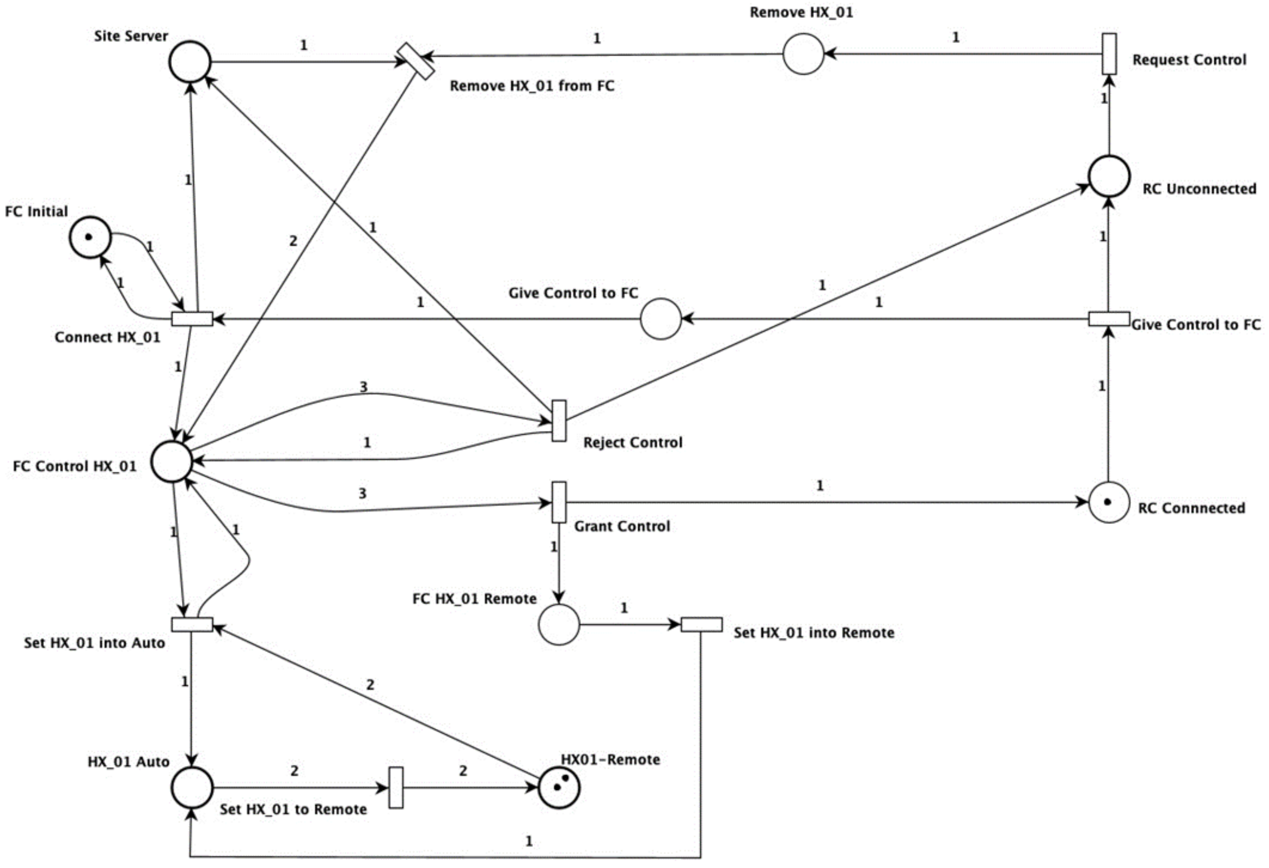


Figure 8. Petri net for the complete SoS

C. *Step 2: Connect all Petri Nets to one SoS Petri Net*

In the separate Petri nets, we have used transitions that transform one marker from the input to one marker at an output to the target place. When connecting the derived Petri nets, we generate a larger Petri net as shown in Figure 8 and the weight of transitions and arcs and the capacity of places need to be adjusted to enable the intended workflow.

As a start situation we model HX 01 to be remote controlled shown by the markers in the related places of each involved system. Fleet Control is in the initial state FC Initial, HX 01 is in state HX 01 Remote and remote control is connected, shown by the marker in state RC Connected. Once the Remote control operator is handing over control to Fleet Control, Fleet Control is changing state to FC Control HX 01. We consider even the site server in this simulation, where the site operator gets information which HX is connected to Fleet Control. Once Fleet Control is in state FC Control HX 01, HX 01 receives a marker, triggering the process in HX 01 moving the marker from HX 01 Remote to HX 01 Auto. This indicates that HX 01 is controlled by Fleet Control. It is now possible to run a simulation with the resulting Petri net, showing that the workflow is functioning as intended.

D. *Step 3 - Run Simulations of the Workflow*

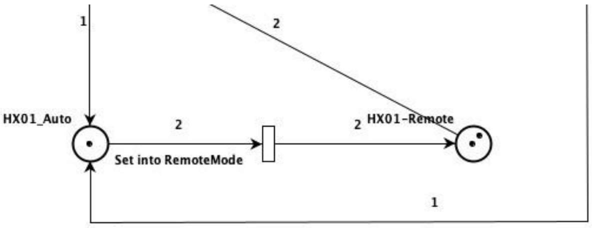


Figure 7. Petri net for a critical scenario

We are interested in situations that are not directly visible and we run the Petri net for a different number of cycles. As one result, we found a deadlock when HX 01 had markers in both states as shown in Figure 9. That HX 01 is in both states at the same time is not realistic, but it might be an indicator for critical controls. This situation needs to be thoroughly analyzed to identify the causal factors and identify possible mitigation strategies. For this case, we found that if the internal states of Fleet Control and the related HX do not match, there is a possibility that a set of different state change requests are send from Fleet Control. Depending on how the HX is managing the incoming state change requests, there is a risk that the HX is set into an incorrect state. A practical example would be, if the remote operator immediately after handing control back to Fleet Control sends a new request to get the control to the same HX. A typical reason could be a human error, when the control is by mistake handed over to Fleet Control. Fleet Control can already be in a state where the command has been sent out to HX 01 to change its state to HX 01 Auto. We even tried time delays for communication with HX 01 and we were able to enforce this scenario. The reason for this behavior is the independence of state machines of the involved systems on the one hand and on the other hand possible communication delays.

It is important to identify such scenarios early during development to add additional states, feedback loops for critical messages and safe states where necessary. The identified inconsistencies are documented as shown in Figure 4. In line with STPA, possible design constraints may be derived. In Step 4 of STPA, loss scenarios are analyzed and documented. The found inconsistencies by simulations will provide additional information during the causal factor analysis in Step 4.

VII. Analysis

When analyzing the case from the electric site project we found, that more information is needed, which is not directly visible in the control structure diagrams of STPA. It is important to understand in which state the HX will be once checked out from the Fleet Control. If the HX will be still active, there is a risk, that it will enter an undefined state once disconnected from the Fleet Control. This additional information about system states will be relevant for a hazard analysis. In our model we have only considered one HX to be controlled by Fleet Control or the remote control. The already complex Petri net will become even more complex when more systems and their states are added. Petri nets are useful to identify critical state-related situations in complex system, but the resulting nets can become very complex reducing the maintainability.

Limitation of Use Case:

The industrial case we have used in this paper is limited to not exceed the scope of this paper. While the states of the HX and Fleet Control are more complex than shown in this paper, we reduced the number of states to highlight how state changes may lead to critical situations. Nonetheless, the complexity of the industrial application makes it even more important to identify inconsistencies of for example the global state.

Correctness of resulting Petri net:

One main challenge with Petri nets we have been facing, is to argue for the correctness of a Petri net. The resulting Petri net, when connecting the Petri nets of the involved systems, needs adjustments in weights of the arcs and transitions as well as adjusting the captivity of places. This might introduce errors in the process flow. We have tried to overcome this challenge by aiming to run a functioning workflow. It is important to choose the right level of abstraction for the Petri net, not to model too many different alternative flows which makes it hard to keep track of targeted workflow.

Complexity of Petri nets:

For large systems with many collaborating subsystems, the associated Petri nets might become very complex and require relevant expertise in building and analyzing them. Model-based engineering methods like SySML are applied by practitioners in industry and SySML state charts can be applied to model the states of the involved systems. Generally, it is possible to transfer SySML state charts to Petri nets as proposed by the authors in [33]. We have not tried this approach to generate the Petri nets presented in this paper. In our work we have shown, that simulation can lead to helpful information for analyzing such critical scenarios, which are not directly derivable from a control structure diagram.

VIII. Conclusion

In this paper we presented a case from the earth moving machinery domain, where autonomous machines are used to transport material in a quarry site. For this system-of-systems, we specifically described a situation where an autonomous machine (HX) is changing from being controlled by a server towards being controlled by remote control operated by a human. We introduced the System Theoretic Process Analysis (STPA), which is a potentially useful approach to analyze the safety for complex systems. STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. In this paper, we presented an approach for enriching STPA to provide the ability to check whether the distributed constituent systems of a SoS have a consistent perspective of the global state which is necessary to ensure safety. We describe the above approach by taking a specific case of state change related issues that could potentially be missed by STPA by looking at an industrial case. By applying Petri nets, we have shown that possible critical situations related to state changes are not identified by STPA. In this context we also proposed a model-based extension to STPA and show how our new process could function in tandem with STPA. This enabled us to simulate the workflow with the goal to find possible flaws in the design. Such information is useful for decision making and development of a SoS.

Acknowledgment

The authors acknowledge the funding support received for this research from the KKS-funded ITS-EASY Post Graduate School for Embedded Software and Systems and the SUCCESS Project supported by the Assuring Autonomy International Programme (AAIP), a partnership between Lloyds Register Foundation and the University of York.

References

[1] International Organization for Standardization, "ISO 26262:2018 - Road vehicles Functional safety," 2018.

[2] ——, "ISO 13849:2015 Safety of machinery - Safety related parts of control systems," 2015.

[3] ——, "ISO 19014:2018 Earth-moving machinery - Functional Safety," 2018.

[4] International Electrotechnical Comission, “IEC 61508:2010 Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems,” 2010.

[5] C. Ericson, *Hazard analysis techniques for system safety*. Wileys, 2015.

[6] International Electronical Commission, “IEC 61025 - Fault Tree Analysis (FTA),” 2006.

[7] United States Department of Defense, *MIL-STD 1629A - Procedures for Performing a Failure Mode, Effect and Criticality Analysis*, 1980. [Online]. Available: http://www.fmea-fmeca.com/milstd1629.pdf

[8] Volvo Construction Equipment, “Electric Site Project.” [Online]. Available: https://www.volvoce.com/global/en/news-and-events/newsand-press-releases/2018/carbon-emissions-reduced-by-98-at-volvoconstruction-equipment-and-skanskas-electric-site/

[9] M. W. Maier, “Architecting principles for systems-of-systems,” *Systems Engineering*, vol. 1, no. 4, pp. 267–284, 1998.

[10] P. Periorellis and J. E. Dobson, “Organisational failures in dependable collaborative enterprise systems,” *Journal of Object Technology*, vol. 1, no. 3, pp. 107–117, 2002.

[11] P. J. Redmond, “A System of Systems Interface Hazard Analysis Technique,” Master’s thesis, 2007. [Online]. Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a467343.pdf

[12] N. G. Leveson and J. P. Thomas, *STPA Handbook*, 2018.

[13] NASA, “NASA Software Safety Guidebook NASA Technical Standard,” p. 389, 2004.

[14] International Electronical Commission, “IEC 61882:2001 Hazard and operability studies (HAZOP studies) Application guide,” 2001.

[15] ——, “IEC60812:2018 Failure modes and effects analysis (FMEA and FMECA),” 2018.

[16] E. Ruijters and M. Stoelinga, “Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools,” *Computer Science Review*, vol. 15, pp. 29–62, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.cosrev.2015.03.001

[17] J. L. Peterson, “Petri Nets*,” *Computing Surveys*, vol. 9, no. 3, 1977.

[18] W. M. Zuberek, “Timed Petri nets definitions, properties, and applications,” *Microelectronics Reliability*, vol. 31, no. 4, pp. 627–644, 1991.

[19] M. A. Marsan, “Stochastic Petri nets: An elementary introduction,” 1990, pp. 1–29.

[20] M. El Koursi and P. Ozello, “Using Petri Nets for Safety Analysis of Unmanned Metro System,” in *SafeComp 1992*. Elsevier, 1992, pp. 135–139. [Online]. Available: http://dx.doi.org/10.1016/S1474-6670(17)49420-1

[21] D. Weyns, T. Holvoet, and K. Schelfthout, “Decentralized control of automatic guided vehicles: applying multi-agent systems in practice,” *Companion to the 23rd*, 2008. [Online]. Available: http: //dl.acm.org/citation.cfm?id=1449819

[22] S. Baumgart, J. Froberg, and S. Punnekkat, “Analyzing hazards in system-of-systems: Described in a quarry site automation context,” in *11th Annual IEEE International Systems Conference, SysCon 2017*.

[23] S. Baumgart, J. Froberg, and S. Punnekkat, “Can STPA be used for a System-of-Systems? Experiences from an Automated Quarry Site,” in *2018 IEEE International Systems Engineering Symposium (ISSE)*, no. 4. IEEE, 10 2018, pp. 1–8.

[24] P. J. Redmond, J. B. Michael, and P. V. Shebalin, “Interface hazard analysis for system of systems,” *2008 IEEE International Conference on System of Systems Engineering*, pp. 1–8, 2008.

[25] R. Alexander and T. Kelly, “Supporting systems of systems hazard analysis using multi-agent simulation,” Safety Science, vol. 51, no. 1, pp. 302–318, 2013.

[26] C. Becker, J. Brewer, L. Yount, D. Arthur, and F. Attioui, “Functional Safety Assessment Of a Generic Electric Power Steering System With Active Steering and Four-Wheel Steering Features,” National Highway Traffic Safety Administration - NHTSA, Tech. Rep. August, 2018.

[27] T. Ishimatsu, N. Leveson, J. Thomas, M. Katahira, Y. Miyamoto, and H. Nakao, “Modeling and hazard analysis using STPA,” in *European Space Agency, (Special Publication) ESA SP*, 2010.

[28] I. Friedberg, K. McLaughlin, P. Smith, D. Laverty, and S. Sezer, “STPASafeSec: Safety and security analysis for cyber-physical systems,” *Journal of Information Security and Applications*, vol. 34, pp. 183–196, 6 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/ S2214212616300850

[29] A. Mallya, V. P. B, M. Adedjouma, M. Lawford, and A. Wassyng, *Computer Safety, Reliability, and Security*, ser. Lecture Notes in Computer Science, A. Skavhaug, J. Guiochet, E. Schoitsch, and F. Bitsch, Eds. Cham: Springer International Publishing, 2016, vol. 9923. [Online]. Available: http://link.springer.com/10.1007/978-3-31945480-1

[30] J. A. Volpe and Van Eikema Hommes, “Assessment of Safety Standards for Automotive Electronic Control Systems (Report No. DOT HS 812 285),” National Highway Traffic Safety Administration, Washington, DC, USA, Tech. Rep. June, 2016. [Online]. Available: https://ntl.bts.gov/lib/59000/59300/59359/812285 ElectronicsReliabilityReport.pdf

[31] J. Zhang, H. Kim, Y. Liu, and M. A. Lundteigen, “Combining systemtheoretic process analysis and availability assessment: A subsea case study,” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 2019.

[32] D. Zhu, S. Yao, and C. Xu, “STAMP-based hazard analysis for computer-controlled systems using petri nets,” *International Journal of Performability Engineering*, vol. 14, no. 9, pp. 1997–2007, 2018.

[33] R. Wang and C. H. Dagli, “An Executable System Architecture Approach to Discrete Events System Modeling Using SysML in Conjunction with Colored Petri Net,” in *2008 2nd Annual IEEE Systems Conference*, 4 2008, pp. 1–8.

2019 4th International Conference on System Reliability and Safety

System of Systems Hazard Analysis Using HAZOP and FTA for Advanced Quarry Production

Faiz Ul Muram, Muhammad Atif Javed and Sasikumar Punnekkat
School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden
e-mail: faiz.ul.muram|muhammad.atif.javed|sasikumar.punnekkat|@mdh.se

Abstract—**The advanced production systems are composed of separate and distinct systems that operate in both isolation and conjunction, and therefore forms the System-of-Systems (SoS). However, a lot of production systems are classified as safety-critical, for example, due to the interactions between machines and involved materials. From the safety perspective, besides the behaviour of an individual system in SoS, the emergent behaviour of systems that comes from their individual actions and interactions must be considered. An unplanned event or sequence of events in safety-critical production systems may results in human injury or death, damage to machines or the environment. This paper focuses on the construction equipment domain, particularly the quarry site, which solely produce dimension stone and/or gravel products. The principal contribution of this paper is SoS hazard identification and mitigation/elimination for the electric quarry site for which the combination of guide words based collaborative method Hazard and Operability (HAZOP) and Fault Tree Analysis (FTA) are used. The published studies on HAZOP and FTA techniques have not considered the emergent behaviours of different machines. The applicability of particular techniques is demonstrated for individual and emergent behaviours of machines used in the quarry operations, such as autonomous hauler, wheel loader, excavator and crusher.**

***Keywords*-hazard analysis and risk assessment, emergent behaviours, system-of-systems, safety and autonomous machines**

I. Introduction

The production systems are typically composed of separate and distinct systems that may not be designed for integration. However, to support the smart production, the characteristics of System-of-Systems (SoS), in particular, operational and managerial independence, evolutionary development, emergent behaviour and geographic distribution are taken into consideration [1], [2]. Compared to an individual system, the system boundary is not clearly defined in SoS and a set of constituent systems might vary over time either as part of normal operation such as another automated vehicle enters in a traffic management system, or otherwise as part of evolutionary development such as traffic management system receives a new version of control system [3]. The SoS hazard identification and mitigation is therefore challenging for which besides the behaviour of an individual system, the emergent behaviour of systems that comes from their individual actions and interactions needs to be considered.

The safety assurance is a regulatory requirement for safety-critical production systems in which an unplanned event or sequence of events may results in human injury or death, damage to machines or the environment. The principal objective of system safety analysis and risks assessment is the identification, elimination or mitigation, and documentation of system hazards, in order to make the system acceptably safe. It has been recognized that the safety analysis is much more cost effective during system design and development than trying to inject safety after the occurrence of an accident or mishap [4]. The functional safety standards, such as ISO 26262 [5], ISO 25119 [6] and IEC 61508 [7] prescribe the adaptation of hazard analysis techniques.

The Hazard and Operability (HAZOP) [4], [8] analysis is widely used to identify possible deviations in systems and subsystems, their possible fault root causes and consequences. It is applicable to all types of systems and equipment [4]. Afterwards, for in-depth analysis, the Fault Tree Analysis (FTA) [9] would be used to develop the fault propagation pathways and to provide a probability for ranking of fault causes before the failures actually occur. The HAZOP and FTA techniques have been combined for risk analysis in fuel storage [10], oil refinery unit [11], and hydrogen refuelling station [12], [13]. Besides the chemical industry, the combination of HAZOP and FTA techniques is used for security vulnerability of web application and infrastructure [14], autonomous service robot [15] and flight conflict at airport [16]. To date, however, the published studies have not considered the HAZOP and FTA techniques for the emergent behaviours of different machines.

This paper focuses on the SoS hazard analysis for the electric quarry site [17], which solely produce dimension stone and/or gravel products. The heavy machines used in the quarry operations such as autonomous hauler, wheel loader, excavator and crusher represent the separate and distinct systems that have not been designed for integration. Due to the autonomous machines, heavy materials and human involvement, the quarry site is regarded as safety-critical. The SoS hazard analysis is performed for which the HAZOP and FTA techniques are used for the identification and elimination of potential hazards in the advanced quarry production. The results obtained from HAZOP and FTA techniques are utilized for elimination or control of identified hazards to demonstrate ultimate, acceptable safety of the quarry site.

The rest of this paper is organized as follows: Section II provides background information on electric quarry site and two hazard analysis techniques, in particular, HAZOP and

978-1-7281-4781-9/19/$31.00 ©2019 IEEE

FTA. Section III describes the quarry production in a smart manner and also performs the SoS hazard analysis. Section IV presents the related work. Section V concludes the paper and discusses future research directions.

II. Baseline and Concepts

A. Electric Quarry Site

This subsection describes an operational quarry site [17]. It falls under the construction equipment domain. The quarry site solely produce dimension stone and/or gravel products of different granularity, which are used for the construction of buildings, roads and railway track beds. The quarry operation is carried out with different kind of machines such as autonomous hauler, wheel loader, excavator, primary/mobile crusher and secondary crusher. In particular, they collaborate together to realize the targeted production goals [18]. The quarry site is subdivided into different production zones.

- Feeding Primary Crusher: The primary crusher breaks the hard and bigger rocks into the smaller rocks. This is done to facilitate the transportation to the secondary crusher. The excavator feeds the raw material to primary crusher, i.e., the rocks that are broken out of the mountain with explosives. The ripper is attached to the excavator or otherwise the dozer to break down the rocks, which may create difficulties for the excavator and/or crusher.
- Direct Loading or Truck Loading: The conveyor belt is attached to the primary crusher. It is therefore possible to directly load the autonomous hauler from the primary crusher or otherwise the rock piles will be formed. The wheel loader is used for making changes in rock piles. The autonomous hauler might also be loaded with the wheel loader.
- Transporting and Dumping: The autonomous haulers travel in the defined path and dumps the loaded rocks in the feeding spot of the secondary crusher. The site management system is responsible for commanding the autonomous haulers. It is composed of three subsystems: (i) user interface visualizes the corresponding information; (ii) fleet management sets missions or tasks for individual autonomous haulers; and (iii) traffic control maintains sufficient distances to avoid collisions.
- Feeding Secondary Crusher: The secondary crusher is a fixed crusher and might be located bit far away. It further crushes the rocks into smaller granularity or fractions to meet the customer demands.
- Charging: The battery-powered autonomous haulers are used in the quarry site. After the completion of mission(s), there is a need to recharge the battery. To be able to recharge the battery, the charging spots have been defined.
- Parking: After the completion of assigned tasks, the machines can be moved to the parking station. If the parking station is not defined, the machines can be parked beside the transportation routes.

B. Hazard Analysis Techniques

The SoS hazard analysis performed in this paper is based on the HAZOP and FTA techniques. This subsection provides an overview of the particular techniques.

1) Hazard and Operability Analysis: The Hazard and Operability (HAZOP) analysis is an inductive technique for identifying and analysing the potential hazards and operational concerns of a system [4]. HAZOP was initially developed to analyse chemical process systems, but later extended for other types of complex systems, for instance, nuclear power plants, rail systems and air traffic management systems [8], [10]. HAZOP analysis is preferably carried out early in the design phase taking different parts into consideration such as software, hardware, procedures and human interactions. The HAZOP analysis sessions are reported in the HAZOP worksheets containing matrix or columns, in which the different items and proceedings are recorded.

The HAZOP analysis process starts with a full description of a system (or a process), which is broken down into system parameters (or steps). Afterwards, all possible deviations are systematically identified by comparing a set of guide words (e.g., more, less and part of etc.) against a list of system parameters or characteristics (e.g., flow of data, pressure and temperature etc.). It might be noted that not all combinations of guide words and parameters are expected to yield sensible or plausible deviations and these combinations can be omitted in the HAZOP worksheets. After the identification of deviations, an assessment is carried out to determine whether particular deviations and their consequences can have negative effects on the system's operation. Finally, the appropriate recommendations are identified that can help to prevent accidents or reduce the associated risk. These steps are repeated for each characteristic and then each node of the system until all hazards are identified.

2) Fault Tree Analysis: The Fault Tree Analysis (FTA) is one of most commonly used deductive analysis approach for modelling, analysing and evaluating failure paths in a large complex dynamic systems such as nuclear power stations, aircraft and chemical processes [9], [19]. Moreover, it can be conducted at different levels of abstraction, such as requirement phase to find out weaknesses in the specification and their impact on the system quality; and the detailed design phase to find weaknesses in design and to identify a direct effect on software safety [4]. The FTA process starts with a top undesired event or mishap and attempts to find out what nodes of a system, combination of events, or component behaviour lead to the occurrence of this top event. It uses a graphical model (i.e. fault tree), which is composed of a top undesired event (outcome), intermediate events, and bottom (basic) events; they are used to describe the internal functional logical (cause–effect) relationship between events. The cause–effect relationships between the components of a system and their events are achieved based on the operating principle and fault mechanisms of the system by using logic gates (e.g., AND-gate, OR-gate, etc.).

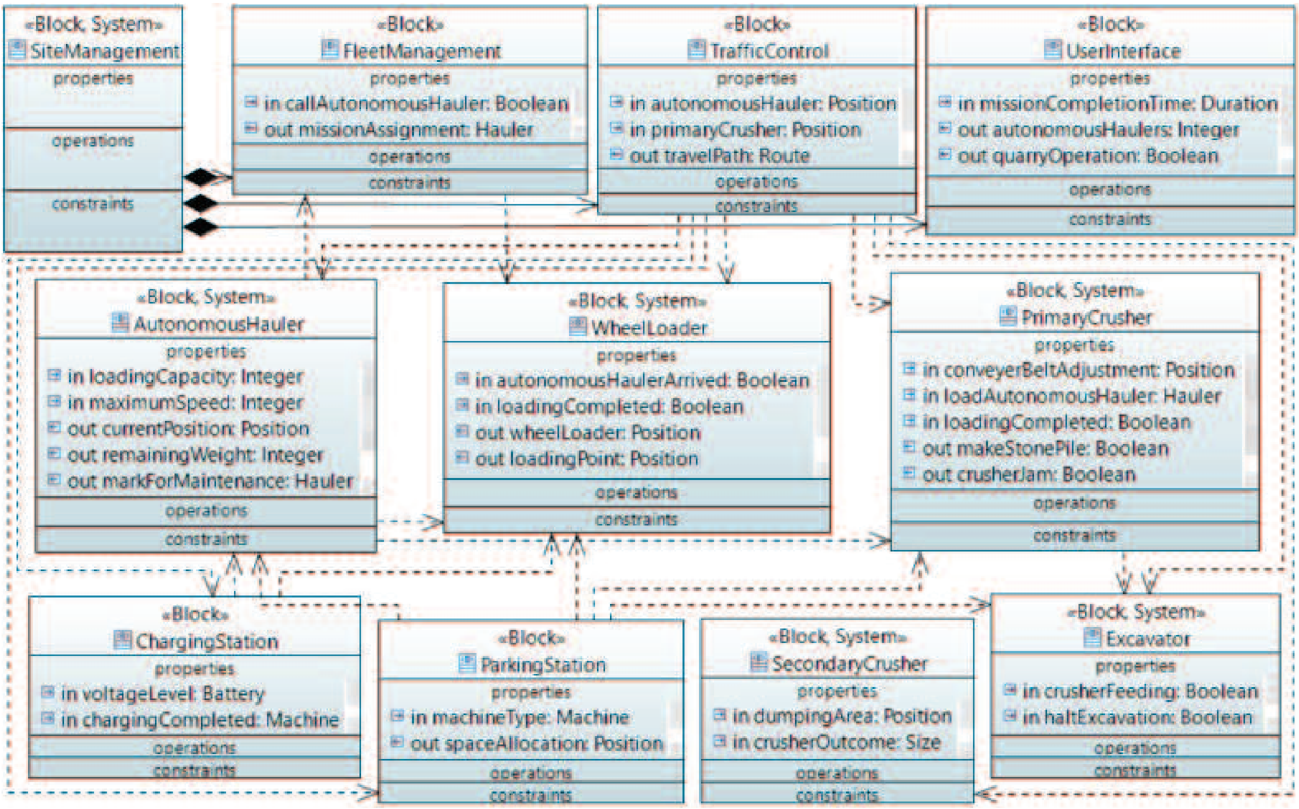


Fig. 1. System-of-Systems Architecture – Advanced Quarry Site

Fault tree development is an iterative analysis process, where the initial structure is continually updated to correspond with design development. During the analysis, those elements which are not contributing in the occurrence of a top undesired event can be eliminated. However, the elements not involved with the occurrence of one undesired event may be involved in the occurrence of another undesired event. A quantitative evaluation can be performed in addition to a qualitative evaluation to measure the probability of the occurrence of a top undesired event and the major faults contributing to this event.

III. Application of Hazard Analysis Techniques to the Advanced Quarry Production

The construction equipment manufacturers aims to provide innovative technological solutions. In the past year, the first emission free quarry site has been made operational [17]. The rocks transportation in quarry site is carried out with the autonomous haulers. The operation of autonomous haulers is similar to the Automated Guided Vehicles (AGVs). After the hauler, the automation of wheel loader could be considered [17], [20]. To support the smart or advanced quarry production, besides the behaviour of individual machines, the emergent behaviour of machines needs to be considered. For example, the automated loading requires emergent interactions of autonomous hauler with crusher or otherwise wheel loader. This section focuses on the SoS hazards analysis. At first, the SoS architecture for the advanced quarry production is described (Section III-A). The PolarSys CHESS[1] Toolset is utilized for the development of system models. After that, two hazard analysis techniques are applied: HAZOP (Section III-B) and FTA (Section III-C). The former establishes the worksheets, while the latter produces the fault trees.

A. System-of-Systems Architecture

The *site management* system serves as a primary controller. From the *traffic control* perspective, the positions of machines are tracked with the Global Positioning Systems (GPS), which are displayed on the site map. The travel paths need to be defined for moving towards the loading, dumping, charging and parking places. The *fleet management* subsystem commands the specific machines to perform their intended operations. For transportation, the missions are assigned to the autonomous haulers. To perform the mission efficiently, the required battery level needs to be determined. This is done before going to the loading place. To adapt the increased transportation demands, besides the direct loading from primary crusher, the parallel loading from wheel loader is considered. The *user interface* subsystem visualizes the status information. The autonomous haulers are moved to the *parking station* after the termination of transportation operation. If the *primary crusher* is building

[1] https://www.polarsys.org/chess/index.html

the rock piles, the direct loading is disabled. It is also possible to turn off the entire quarry operation, in particular, all the machines at the quarry site.

The autonomous vehicles contain the cameras, GPS and LIDAR (Light Detection and Ranging). These sensors are responsible for gathering surrounding information, such as positions, obstacles, and lane or boundaries. This information is processed for controlling the mechanical parts, for example, the drive unit for motion and operation, the steering system for manoeuvring, and the braking system for slowing down the vehicle to avoid collisions and accidents. The interaction platform and other attachments such as batteries for power supply are integrated in the *autonomous hauler*.

Together with the individual behaviour, the emergent behaviour of machines/systems is considered, as shown in Figure 1. The *wheel loader* is able to call the *autonomous hauler*, which informs back the *wheel loader* upon reaching the specified position. To be able to perform the direct loading, the adjustment of conveyor belt or otherwise *autonomous hauler* is desired. The remaining weight is conveyed to the *wheel loader* and *primary crusher*. There is also a need to pause the *primary crusher* for a while so that the next *autonomous hauler* is adjusted under the conveyor belt. If the crusher is jammed or the wait time for a next *autonomous hauler* is increased, the *excavator* is instructed to halt excavation. In the advanced *charging* and *parking* stations, for the assignment of specific places, the kind of machines needs to be determined. Besides that, the remaining battery and machine status might be conveyed.

B. Applying HAZOP Technique

In the context of SoS, a failure may not just lead to a hazard and accident of a system itself. But it can propagate to other systems, which lead to a mishap. This is because different systems have emergent interactions between them. In an SoS quarry production, the critical incidents can occur if correctly and timely communication is not established, for instance, a message is received too late, an incorrect message is transferred, or wrongly interpreted by the receiver. From the loading perspective, the delay of messages can result in severe damages to the machines. The collision of autonomous hauler is possible consequence, especially with the machines not equipped with obstacle detection and collision/avoidance mechanisms. Due to the incorrect mission or travel path assignment, the autonomous hauler may unintendedly enter into the restricted area in which humans are working or hazardous materials are stored. The failure of speed sensors may result in the wrong decisions which, in turn, may leads to the human injury or even life lose or damage to the environment.

TABLE I
A SET OF GUIDE WORDS AND THEIR MEANINGS FOR SOS

Guide Word	Interpretation
Late	A message/data is transferred too late to be used.
Early	A message/data is transferred too early to be used.
No/Not/None /Omission	A message is not transferred. Interaction does not occur at all. None of the design intention is achieved.
More	The message is sent to more objects than intended. Too much or repeated information is transferred.
Less/Part Of	The message is sent to fewer objects (receivers) than intended. Too little information is transferred. Some of the design intention is achieved.
Incorrect/ Other Than	Incorrect message is transferred. Another activity takes place, opposite of what is intended.
Before/After	A message is transferred in a wrong sequence. Something happens before/after the intended order.
Slower/Faster	Activity is (not) done with the right timing.
Reverse	Source and destination objects are reversed.

TABLE II
CONSEQUENCES OF DEVIATIONS

ID	Consequences
C01	Human injuries or life lose
C02	Autonomous Hauler (AH) does not maintain a safe distance from other (autonomous or human operated) machines
C03	AH is unable to complete the mission
C04	AH enters in the restricted areas/region where human are working
C05	Major environmental damage
C06	Machine damage, loss of critical hardware
C07	AH rate of manoeuvre is insufficient to avoid the other obstacles
C08	AH fails to detect the obstacles at sufficient range
C09	AH unable to reduce/manage the speed or apply brake
C10	AH slips and falls during loading and unloading
C11	AH/other machines do not maintain a safe distance from human

TABLE III
EXTRACT OF RECOMMENDATIONS FOR HAZARDS

ID	HAZOP Recommendations
R01	Install roadside Dedicated Short Range Communications (DSRC) devices for better communication
R02	Introduce communication prioritization between machine to server communication, and machine-to-machine
R03	Use efficient networking protocol
R04	Increase number of wireless access point and retransmit the message
R05	Get the information from LIDAR as back up
R06	Take the values from camera
R07	Install additional sensors as back up
R08	Site manager takes the control
R09	Slow down speed motor
R10	Delete the connection with speed evaluator and switching to GPS
R11	Use dynamic filtering and inertial sensors

We have performed a detailed HAZOP analysis for the advanced quarry production. For the hazard analysis, eight systems are taken into consideration; they are further divided into subsystems. The quarry production is carried out in different phases. In the context of an advanced quarry site, there is a need to address both individual and emergent behaviours of particular systems to realize the targeted production goals, as described in Sections II-A and III-A. To perform the HAZOP analysis, a set of guide words, parameters (e.g., speed, position, etc.), system inputs and outputs, a list of messages, and paths are identified. Table I shows a set of guide words and their interpretation. They are used for SoS hazard analysis. Each guide word is applied to all reasonable pairs of parameters, operations or components for determining the deviations. During the HAZOP analysis, the evaluation is carried out to determine whether the combinations makes

TABLE IV
EXTRACT OF THE HAZOP ANALYSIS REPORT FOR COMMUNICATION

Item	Guide Word	Parameter	Deviation	Cause	Consequence [Table II]	Recommendation [Table III]
H01	Late	Communication	Position of wheel loader or primary crusher is sent and/or received late than expected	Communication link between site server and wheel loader/primary crusher, or between site server and autonomous hauler is manipulated, downtime, loss of GPS signal	C02	R01
H02	No	Communication	Position of wheel loader or primary crusher is not transferred to autonomous hauler	Network unavailability, signal transmitter failure, reflection of signals, out of range	C01, C02, C03	R02
H03	More	Communication	Message received twice than expected to site management system	Autonomous hauler repeatedly sends the same message to site management over a determined amount of time	C02, C11	R02
H04	Other Than/ Incorrect	Communication	Incorrect mission or travel path is transferred to autonomous hauler	Wrong command is given by wheel loader and/or site server manager, site management system failed to detect human command	C01, C03, C04, C11	R03
H05	Less/Part of	Communication	Less information about mission is provided to autonomous hauler	Loose communication, intermittent communication	C03, C04	R04
H06	Other Than	Communication	Mission is transferred to the other autonomous hauler	Wrong command given by human, site management system failed to detect human command	C01, C04, C06, C11	R03

TABLE V
EXTRACT OF THE HAZOP ANALYSIS REPORT FOR SYSTEMS AND SUBSYSTEMS

Item	Guide Word	Parameter	Deviation	Cause	Consequence [Table II]	Recommendation [Table III]
H07a	Not	GPS system locate wheel loader position	GPS system fails to locate the wheel loader position, send and receive the location	GPS sensor fails, position estimator fails, communication failure	C02, C03	R05, R11
H07b	Less/Part of	GPS system locate wheel loader position	Send and receive less information of location. Route optimization failure	Wrong reading of GPS sensor, position estimator failure, biased position is calculated and forward	C03, C04	R06, R11
H08a	Not	GPS system locate Autonomous Hauler (AH) position	GPS system fails to locate the AH position, send and receive the location	GPS sensor fails, position estimator fails, GPS receiver fails, network unavailability	C02, C03, C07	R05, R11
H08b	Incorrect	GPS system locate AH position	GPS incorrectly estimates the location and direction. Send incorrect location	GPS sensor failure, wheel speed sensor failure, system controller failure, communication failure	C02, C07	R05, R11
H09a	Other Than/ Incorrect	LIDAR Position Encoder (AH)	Fails to detect obstacles and identify the location. Sends a wrong message	Mirror motor malfunction, position encoder failure, object to far to be detected, light emitter and receiver failure	C05, C06, C02, C08, C11	R07
H09b	More, Less, Other Than	LIDAR locate correct position	Misalignment. Data passed to the state estimator is either corrupted or less	Laser malfunction, light emitter and receiver failure	C01, C05, C06, C08	R07
H10a	No	Camera, Detect object (AH)	Could not detect the obstacles, difficult to localize	Improper lighting, blind spot, object is too far	C01, C02, C05	R07, R08
H10b	Other Than, Part of	Camera, Detect object	Detect object parts, cannot take the whole picture of object	Object is too close or too big. Improper lighting, misalignment, dirty or damaged lens	C05, C06, C01, C08, C11	R07
H10c	Late, Before After	Camera, Detect object	Detect the object late, difficult to localize	Object moves past, high speed, improper lighting, misalignment	C05, C06, C01, C11	R07, R09
H10d	No	Camera, Detect surface	Not able to detect unevenness of the surface	Adverse weather conditions, improper lighting, dirty lens	C10	R07, R08
H10e	No	Camera, Detect lane	Not able to detect lane	Improper lighting, blind spot	C04	R07, R08
H11a	Incorrect	Wheel speed sensor	Speed sensor emits wrong value, encoder feedback unable to be transferred	Speed sensor failure, wheel encoder failure	C09, C02, C01, C08, C11	R07, R10

sense. Afterwards, for the relevant hazards, all possible causes and potential consequences are identified. Table II shows the consequences of deviation. The proposed corrective measures to mitigate the hazards are shown in Table III. This process is repeated deviation by deviation and attribute by attribute until the analysis for SoS quarry production is completed.

The performed analysis not just focuses on the communication failures in SoS, but also external malfunctions and internal systems failures. Table IV shows the reduced hazard analysis results related to transmitting a message, in which the loading point, current position of machines and mission assignment are taken into consideration. Their listed failures concern the site management system, wheel loader/primary crusher, autonomous haulers and communication links. Note that the combinations without plausible deviations are omitted. It can be seen from the HAZOP results that the transformation of incorrect mission or travel path to autonomous hauler that can be caused by the command detection failure leads to the incomplete mission, machine damage or human injuries. This can be prevented by using efficient networking protocol. Table V summarizes the hazards caused because of the environmental influences and internal failures of autonomous hauler and wheel loader systems, or subsystems (e.g., LIDAR, GPS etc.) that are propagated to one or more systems, and in-turn lead to a mishap. The results from the HAZOP analysis have been used for prevention or mitigation of identified SoS hazards.

C. Applying FTA Technique

By focusing on a rigorous and structured methodology, FTA supports system analysts in modelling the unique combinations of fault events, which may cause an undesired event to occur. The comprehensive fault trees are developed based on the hazards and their potential effects understood from the HAZOP analysis, in which the identified hazards can serve as the top undesired events. To develop the fault trees, human injury, machine damage and mission failure are selected as the top undesired events or mishaps. After establishing a top event, sub-undesired events are identified and structured that is referred to the top fault tree layer. The logic between every event is investigated, in particular, the type of gates and their specific inputs are formulated. All possible reasons including human errors, and environmental influences are evaluated level-by-level until all relevant events are found.

Human injuries might occur at different phases of quarrying process, for example, upon the entry of machines or humans in the restricted areas. On the one hand, if an autonomous hauler enters in the restricted area, there is a possibility of collision with the working humans or explosive materials. On the other hand, if a human enters in the dangerous areas such as loading and dumping, there is a possibility of collision with an autonomous hauler or wheel loader. Figure 2 shows how the top mishap human injury is associated with the vulnerability of autonomous hauler at transporting phase. The autonomous hauler may unintendedly enter in the restricted areas due to the communication failures (i.e., emergent interactions), environment influences or obstacles detection failures. If an autonomous hauler or human is entered in the specific areas, besides the detection of obstacles, sufficient distance from human needs to be maintained. But, the failures caused by the battery power, system and mechanical parts may potentially lead to the collisions. The system failure is caused by two intermediate events: software control system failure and navigation safety system failure. In the context of mechanical parts failure, for example, drive unit, steering and brake systems, the autonomous hauler cannot be able to move, turn and apply brakes, respectively. The navigation safety system performance may also be affected for which the reasons include the degradation of GPS, LIDAR, or cameras. The fault tree shows that the resulting behaviour of particular failures will always reach a top mishap scenario.

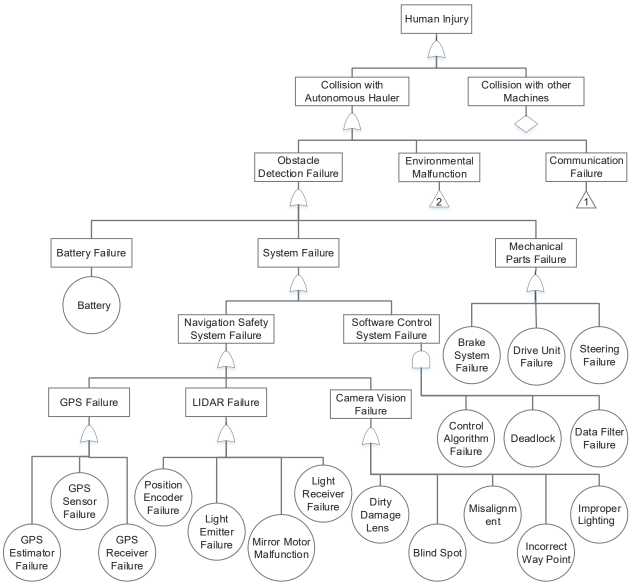


Fig. 2. Top level fault tree of human injury mishap

As we see in Figure 3, the fault tree is further constructed with respect to the communication failure and environmental malfunction. The incorrect mission transfer to autonomous hauler and timing failures are related to the communication failures. If the site manager sent the wrong travel path, or mission and position, system fails to detect or autonomous hauler is out of range to receive commands. Besides the messages sent from site management system to autonomous hauler, the messages from wheel loader/primary crusher to site management system may cause the communication failures. Another reason of communication failure is the link failure (i.e. network unavailability). Therefore, these events are further developed into basic events. The environmental malfunction might be caused by the adverse weather conditions, drivable surface conditions and human behaviours.

In the fault trees, we have removed multiple occurring events and branches for the purpose of avoiding errors and obtaining accurate results. The presence of humans in the

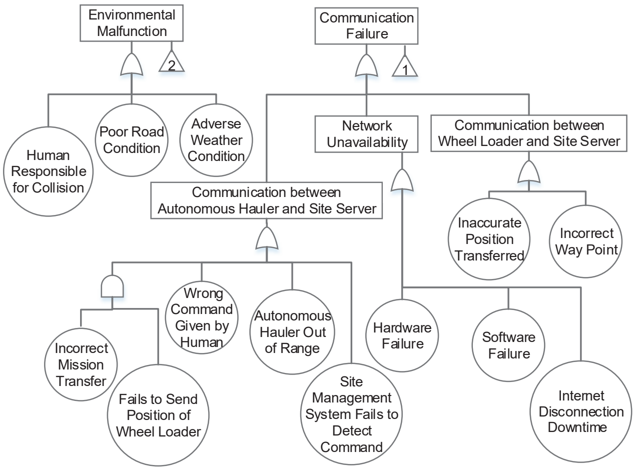

Environmental Malfunction
2
Communication Failure
1
Human Responsible for Collision
Poor Road Condition
Adverse Weather Condition
Network Unavailability
Communication between Wheel Loader and Site Server
Communication between Autonomous Hauler and Site Server
Inaccurate Position Transferred
Incorrect Way Point
Wrong Command Given by Human
Autonomous Hauler Out of Range
Hardware Failure
Software Failure
Incorrect Mission Transfer
Fails to Send Position of Wheel Loader
Site Management System Fails to Detect Command
Internet Disconnection Downtime

Fig. 3. Fault trees of communication failure and environmental malfunction

quarry site is one of the most crucial links. The communication failure is perceived as most vulnerable event among all events. The failure of sensors (e.g., camera, GPS and LIDAR) is the second most common problem for the failure of autonomous vehicles [21]. The results obtained from the FTA demonstrate that human injury, machine damage and mission failure are either caused by a communication failure, or otherwise one or more subsystem components.

IV. Related Work

Two of the studies consider the safety and reliability issues of the AGVs components and their probability of success in completing a prescribed mission. Yan et al. [22] merged Failure Modes and Effects Criticality Analysis (FMECA) and FTA to assess the safety of AGVs. Duran and Zalewski [23] applied the FTA on autonomous ground vehicles to identify hazards related to LIDAR and cameras.

Martin-Guillerez et al. [15] analysed risks for autonomous service robot. At first, the Preliminary Hazard Analysis (PHA) technique is used to identify hazards in the preliminary design. After that, the deviations in UML use cases and sequence diagrams are analysed by applying the HAZOP technique. However, the fault trees based on the PHA and HAZOP-UML hazards are not presented. Wu et al. [16] combined the HAZOP and FTA techniques for analysing the flight conflict at airport. The HAZOP technique is applied for acquiring the deviation and hazards list. For in-depth accident analysis, the fault trees are constructed based on the hazards list, in which the worst result is taken as the top event. Snamchaiskul and Phanrattanachai [14] used the HAZOP and FTA techniques to investigate the security vulnerability of web application and infrastructure. The guide words are proposed to cover the vulnerabilities, such as cross site scripting, SQL injection and script injection. The authors found that the fault tree of vulnerabilities in web applications did not yield much contribution than web infrastructure. Wu et al. [16] and Snamchaiskul and Phanrattanachai [14] have not considered the preventive measures. The aforementioned studies focus on the safety analysis of a single system.

There are few attempts to perform HAZOP analysis on an SoS. Redmond et al. [24] propose an SoS hazard analysis technique, which is a mixture of HAZOP and network analysis. The technique focuses on just one type of hazards, particularly interface hazards, in which one system causes a mishap in another system by transferring a failure over specified interface. Michael et al. [25] introduce a validation framework by combining Goal Question Metric (GQM), HAZOP and network analysis for measuring the sufficiency of software safety requirements with a set of metrics for an SoS missile defense. Stephenson et al. [26] present hazard assessment and safety-case production for Integrated Aircrew Training (IAT). To do that, they adapt product line techniques (feature model) to manage variation between staff training scenarios. The initial data for each system is derived from a differential analysis. The high-level hazard assessment is performed using HAZOP and HAZAN (Hazard Analysis) on training scenario. Then, a low-level exemplars assessment is performed.

Baumgart et al. [18] apply System-Theoretic Process Analysis (STPA) on the quarry site, in which just the control structure diagram is taken into consideration. In comparison to our work, the causes of communication failures, influences of environment and internal failure of the system, as well as the advancements in quarry site have not been investigated. For the SoS hazard and safety analysis, the simulation-based approaches are also proposed. In particular, they are designed to give quantitative assessments of the overall risk present in the system. For example, Blom et al. [27] in airspace system safety, and Mohaghegh et al. [28] in socio-technical systems use Monte Carlo techniques to acquire quantitative statistical measures of the overall safety of a system under specified conditions. Alexander and Kelly [29] present an analysis technique (SimHAZAN) that uses multi-agent modelling and simulation to explore the effects of deviant node behaviour within an SoS. However, the output results of simulation-based approaches contain thousands or millions of run logs, each containing tens of thousands of entries. It is very difficult for a human analyst to read such logs and understand them.

The principal contribution of this paper is SoS hazard analysis, which is performed as a first step towards advanced quarry production. For this reason we applied HAZOP and FTA techniques for the identification of hazards occurred due to the interactions between heavy machines/systems used in the quarry operations. Besides the identification of hazards, their prevention and mitigation had been considered.

V. Conclusion and Future Work

To be able to support the advanced quarry production, besides the individual behaviour of machines used in the quarry operations, such as autonomous hauler, wheel loader, excavator, primary crusher and secondary crusher, their emergent behaviour needs to be considered. Accordingly, this paper focuses on the SoS hazard identification and mitigation/elimination for the quarry production. Two hazard analysis techniques, particularly HAZOP and FTA are applied. The former is applied to identify possible deviations in SoS quarry production, their possible fault root causes and consequences. The latter supports in-depth analysis; the fault trees are constructed based on the hazards and their potential effects understood from the HAZOP analysis. The preventive measures drawn from the hazard analysis are used to eliminate or control the identified hazards for the demonstration of ultimate, acceptable safety of the quarry site.

The simulation environment of machines used in the quarry site is available in the university lab. A site sever is used for the specification of different machines in a site. As future work, we plan to support the dynamic safety assurance. The safety cases will be developed in the PolarSys OpenCert[2] platform. The safety contracts derived from the HAZOP and FTA techniques will be associated with the safety cases. The simulation data is used for the runtime monitoring of safety contracts. The results will be processed for updating the assurance (safety) cases and evidence models developed in the OpenCert platform.

Acknowledgment

This work is supported by SUCCESS (Safety assUrance of Cooperating Construction Equipment in Semi-automated Sites) project via the AAIP (Assuring Autonomy International Programme) and FiC (Future factories in the Cloud) project funded by SSF (Swedish Foundation for Strategic Research).

References

[1] J. T. Boardman and B. J. Sauser, "System of systems - the meaning of of," in *1st IEEE/SMC International Conference on System of Systems Engineering (SoSE), Los Angeles, CA, USA, April 24-26*, 2006, pp. 1–6.

[2] C. B. Nielsen, P. G. Larsen, J. S. Fitzgerald, J. Woodcock, and J. Peleska, "Systems of systems engineering: Basic concepts, model-based techniques, and research directions," *ACM Comput. Surv.*, vol. 48, no. 2, pp. 18:1–18:41, 2015.

[3] R. Alexander, M. Hall-May, G. Despotou, and T. Kelly, "Towards using simulation to evaluate safety policy for systems of systems," in *Safety and Security in Multiagent Systems (SASEMAS) - Research Results from 2004-2006*, vol. 4324, 2009, pp. 49–66.

[4] C. A. Ericson, *Hazard Analysis Techniques for System Safety, 2 edition*. John Wiley & Sons, 2015.

[5] International Organization for Standardization (ISO), "ISO 26262–3:2011-Road vehicles-Functional safety. International Standard, November," 2011.

[6] ——, "ISO 25119–4:2018 tractors and machinery for agriculture and forestry – safety-related parts of control systems–part 4: Production, operation, modification and supporting processes," 2018.

[7] International Electrotechnical Commission (IEC), "IEC 61508-1:2010-Functional safety of electrical/electronic/programmable electronic safety-related systems," 2010.

[8] J. Dunjó, V. Fthenakis, J. A. Vílchez, and J. Arnaldos, "Hazard and operability (HAZOP) analysis. a literature review," *Journal of Hazardous Materials*, vol. 173, no. 1, pp. 19 – 32, 2010.

[9] L. Xing and S. V. Amari, "Fault tree analysis," in *Handbook of Performability Engineering*, K. B. Misra, Ed. London: Springer London, 2008, ch. 38, pp. 595–620.

[10] J. Fuentes-Bargues, M. González-Cruz, C. González-Gaya, and M. Piedad Baixauli-Pérez, "Risk analysis of a fuel storage terminal using HAZOP and FTA," *International Journal of Environmental Research and Public Health*, vol. 14, p. 705, 06 2017.

[11] S. Dacosta, I. Al-Asýari, A. Musyafa, and A. Soeprijanto, "HAZOP study and fault tree analysis for calculation safety integrity level on reactor-c.5-01, oil refinery unit at balikpapan-indonesia," *Asian Journal of Applied Sciences*, vol. 5, 05 2017.

[12] M. Casamirra, F. Castiglia, M. Giardina, and C. Lombardo, "Safety studies of a hydrogen refuelling station: Determination of the occurrence frequency of the accidental scenarios," *International Journal of Hydrogen Energy*, vol. 34, no. 14, pp. 5846–5854, 2009.

[13] E. Kim, K. Lee, J. Kim, Y. Lee, J. Park, and I. Moon, "Development of korean hydrogen fueling station codes through risk analysis," *International Journal of Hydrogen Energy*, vol. 36, no. 20, pp. 13 122 – 13 131, 2011.

[14] S. Pumisake and P. Thitinan, "Using HAZOP and FTA to analyse security vulnerability of web application and infrastructure," in *3rd International Conference on Informatics, Environment, Energy and Applications (IEEA 2014), Shanghai, China*, 2014.

[15] D. Martin-Guillerez, J. Guiochet, and D. Powell, "Experience with a model-based safety analysis process for autonomous service robot," in *7th International Workshop on Technical Challenges for Dependable Robots in Human Environments (DRHE 2010), Toulouse, France*, 2010.

[16] Q. Wu, X. Gan, D. Yao, and Q. Sun, "Fault tree establishment of flight conflict based on the HAZOP method," in *4th International Conference on Machinery, Materials and Computing Technology, ICMMCT 2016, January 23-24, Hangzhou, China*, 01 2016.

[17] Volvo Construction Equipment, "Emission-free quarry," Available at https://www.volvoce.com/global/en/news-and-events/press-releases/2018/testing-begins-at-worlds-first-emission-free-quarry/.

[18] S. Baumgart, J. Fröberg, and S. Punnekkat, "Can STPA be used for a system-of-systems? experiences from an automated quarry site," in *2018 IEEE International Symposium on Systems Engineering (ISSE), Rome, Italy*, no. 4, October 2018, pp. 1–8.

[19] P. Liu, L. Yang, Z. Gao, S. Li, and Y. Gao, "Fault tree analysis combined with quantitative analysis for high-speed railway accidents," *Safety Science*, vol. 79, pp. 344–357, 2015.

[20] R. Lilja, "A localisation and navigation system for an autonomous wheel loader," Master's thesis, Mälardalen University, Västerås, Sweden, 2011.

[21] P. Bhavsar, P. Das, M. Paugh, K. Dey, and M. Chowdhury, "Risk analysis of autonomous vehicles in mixed traffic streams," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2625, pp. 51–61, 01 2017.

[22] R. Yan, S. J. Dunnett, and L. M. Jackson, "Reliability modelling of automated guided vehicles by the use of failure modes effects and criticality analysis, and fault tree analysis," in *5th Student Conference on Operational Research (SCOR), Nottingham, UK, April 8-10*, 2016, pp. 2:1–2:11.

[23] D. Reyes-Duran, E. Robinson, A. J. Kornecki, and J. Zalewski, "Safety analysis of autonomous ground vehicle optical systems: Bayesian belief networks approach," in *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems (FedCSIS), Kraków, Poland, September 8-11*, 2013, pp. 1407–1413.

[24] P. J. Redmond, J. B. Michael, and P. V. Shebalin, "Interface hazard analysis for system of systems," in *3rd IEEE International Conference on System of Systems Engineering (SoSE), Singapore, June 2-4*, 2008, pp. 1–8.

[25] J. B. Michael, M. Shing, K. J. Cruickshank, and P. J. Redmond, "Hazard analysis and validation metrics framework for system of systems software safety," *IEEE Systems Journal*, vol. 4, no. 2, pp. 186–197, 2010.

[26] Z. Stephenson, C. Fairburn, G. Despotou, T. P. Kelly, N. Herbert, and B. Daughtrey, "Distinguishing fact from fiction in a system of systems safety case," in *Advances in Systems Safety–Proceedings of the Nineteenth Safety-Critical Systems Symposium, Southampton, UK, February 8-10*, 2011, pp. 55–72.

[27] H. A. P. Blom, S. H. Stroeve, and H. H. de Jong, "Safety risk assessment by monte carlo simulation of complex safety critical operations," in *Developments in Risk-based Approaches to Safety–Proceedings of the Fourteenth Safety-critical Systems Symposium (SCSC-6), Bristol, UK, February 7-9*, 2006, pp. 47–67.

[28] Z. Mohaghegh, R. Kazemi, and A. Mosleh, "Incorporating organizational factors into probabilistic risk assessment (PRA) of complex sociotechnical systems: A hybrid technique formalization," *Rel. Eng. & Sys. Safety*, vol. 94, no. 5, pp. 1000–1018, 2009.

[29] R. Alexander and T. Kelly, "Supporting systems of systems hazard analysis using multi-agent simulation," *Safety Science*, vol. 51, no. 1, pp. 302–318, 2013.

[2] See https://www.polarsys.org/proposals/opencert

Enforcing Geofences for Managing Automated Transportation Risks in Production Sites

Muhammad Atif Javed( ), Faiz Ul Muram, Anas Fattouh, and Sasikumar Punnekkat

School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden
{muhammad.atif.javed,faiz.ul.muram,anas.fattouh, sasikumar.punnekkat}@mdh.se

Abstract. The key to system safety is the identification and elimination/mitigation of potential hazards and documentation of evidences for safety cases. This is generally done during the system design and development phase. However, for automated systems, there is also a need to deal with unknowns and uncertainties during operational phase. This paper focuses on virtual boundaries around geographic zones (i.e., geofences) that can serve as an active countermeasure for dynamic management of risks in automated transportation/production contexts. At first, hazard analysis is performed using the Hazard and Operability (HAZOP) and Fault Tree Analysis (FTA) techniques. Based on the hazard analysis, appropriate measures, such as geofences for elimination/mitigation of hazards are defined. Subsequently, they are translated into the safety requirements. We leverage on simulation based digital twins to perform verification and validation of production site by incorporating safety requirements in them. Finally, to manage risks in a dynamic manner, the operational data is gathered, deviations from specified behaviours are tracked, possible implications of control actions are evaluated and necessary adaptations are performed. The risk management is assured in situations, such as communication loss, subsystem failures and unsafe paths. This approach provides a basis to fill the gaps between the safety cases and the actual system safety emanating from system/environment evolution as well as obsolescence of evidences. The applicability of the proposed framework is exemplified in the context of a semi-automated quarry production scenario.

Keywords: Geofence enforcement · Risk management · Automated transportation · Safety assurance · Digital twin · Quarry site

1 Introduction

In the safety-critical production sites, an unplanned event or sequence of events can potentially harm humans (injuries or even deaths) or create damages to

S. Bernardi et al. (Eds.): EDCC 2020 Workshops, CCIS 1279, pp. 113–126, 2020.
https://doi.org/10.1007/978-3-030-58462-7_10

machines, property or the environment. The system safety is a basic part of the risk management process that emphasizes the identification of hazardous events and their causes, mechanisms for elimination or mitigation of causes, and documentation of evidences for safety cases. This is primarily done during system design and development phase. The safety analysis and risk management are much more cost effective during system design and development than trying to inject safety after the occurrence of an accident or mishap [5]. However, hazard analysis conducted at system design and development phase may not be sufficient for the evolving systems with enhanced automation, digitalization and connectivity. The capability to manage risks at operational phase is essential for them.

The virtual boundary around a geographic zone, usually called geofence, can serve as an active countermeasure against operational mishap risks. The Global Positioning System (GPS) is used for tracking and navigation purposes and its information is used for triggering alerts in circumstances when the device enters or exits the geographical boundary of a point of interest [12]. During the past decade, there has been increasing attention in geofences. They are enforced in various domains, such as smart city, healthcare, road transport, smartphones, security, forestry and aerospace [16]. This is particularly useful for automated transportation, for example, to control the movement of machines in hazardous situations and areas, though existing studies have not considered them for managing risks in a dynamic manner.

This paper considers geofences as a measure for elimination or mitigation of automated transportation risks. Hazard analysis was carried out as a first step towards safe production site and appropriate mitigation mechanisms such as geofences were established. They were then translated into the safety requirements, which were implemented in digital twins as code scripts. The digital twins were leveraged for performing verification and validation of the production site. The static, dynamic, time-based and conditional geofences were enforced by defining different geometric areas. To carry out the dynamic risk management, the operational data was gathered and the deviations from specified behaviours were tracked. The risk management is assured in situations, such as connection loss, subsystem failures and unsafe movements. Besides that, the gaps between the safety cases and the actual system safety were handled. The applicability of the proposed framework using geofences for dynamic risk management is demonstrated in a simulated quarry production scenario. In the next phase, after studying the effects, accuracy, failure modes and design/standardisation requirements, we plan to implement them in real machines.

The rest of this paper is organized as follows: Sect. 2 presents a risk management framework for automated materials transportation. Section 3 demonstrates the effectiveness of proposed framework in a quarry production scenario. Section 4 discusses related work. Section 5 concludes the paper and presents future research directions.

2 Managing Automated Transportation Risks

Risk management is an overarching process that begins during the earliest phase of a system design and continues throughout its entire life cycle. This section describes our proposed overall framework for managing automated transportation risks, as shown in Fig. 1. The framework consists of essentially three stages. In the first stage, the safety analysis during design and development phase is carried out through the identification of hazards, the assessment of risks, and the control of hazards risk. In the second stage, digital twins are utilized in which measures for elimination or mitigation of hazards are implemented. The geofences are supported and given special consideration for risk management. During the verification and validation with digital twin, additional hazards can be detected. In the third stage, the dynamic safety assurance during operational phase, in particular, the risk management and update of safety cases are carried out.

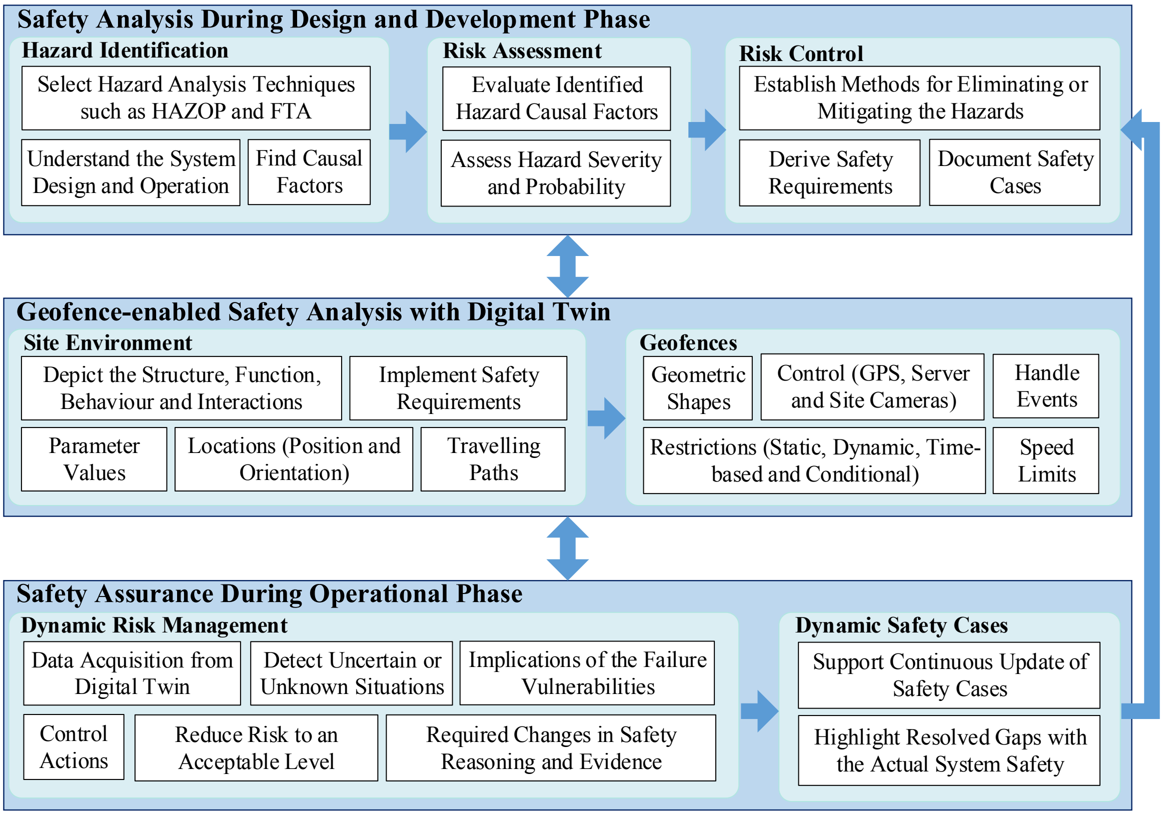

Safety Analysis During Design and Development Phase
Hazard Identification
Select Hazard Analysis Techniques such as HAZOP and FTA
Understand the System Design and Operation
Find Causal Factors
Risk Assessment
Evaluate Identified Hazard Causal Factors
Assess Hazard Severity and Probability
Risk Control
Establish Methods for Eliminating or Mitigating the Hazards
Derive Safety Requirements
Document Safety Cases
Geofence-enabled Safety Analysis with Digital Twin
Site Environment
Depict the Structure, Function, Behaviour and Interactions
Implement Safety Requirements
Parameter Values
Locations (Position and Orientation)
Travelling Paths
Geofences
Geometric Shapes
Control (GPS, Server and Site Cameras)
Handle Events
Restrictions (Static, Dynamic, Time-based and Conditional)
Speed Limits
Safety Assurance During Operational Phase
Dynamic Risk Management
Data Acquisition from Digital Twin
Detect Uncertain or Unknown Situations
Implications of the Failure Vulnerabilities
Control Actions
Reduce Risk to an Acceptable Level
Required Changes in Safety Reasoning and Evidence
Dynamic Safety Cases
Support Continuous Update of Safety Cases
Highlight Resolved Gaps with the Actual System Safety

Fig. 1. A framework for managing automated transportation risks

2.1 Safety Analysis During Design and Development Phase

This subsection presents the system safety risk management process at the design and development phase. The initial step is based on systems context establishment in risk analysis (e.g., system design and operation), identification of hazards and their causes. The next step involves the risk assessment with the aim of evaluating the causes of identified hazards, assessing hazard severity and the probability of occurrence. The last step is the risk control that focuses on the establishment of mechanisms to eliminate or mitigate those causes, which may result in unexpected or collateral damage, and documentation of the safety cases to demonstrate the acceptable safety of the system. The argumentation editor of OpenCert[1] tool platform is used for modelling and visualizing the safety cases, which is based on the Goal Structuring Notation (GSN) [14]. However, Common Assurance and Certification Metamodel (CACM) implemented in OpenCert internally uses the Structured Assurance Case Metamodel (SACM) [11].

The hazard analysis is performed by using HAZOP and FTA techniques. The HAZOP is an inductive technique for identifying and analysing the potential deviations from design intention or operating conditions of a system; whereas, FTA is a deductive analysis approach for modelling, analysing and evaluating failure paths in large complex dynamic systems [5,9]. The hazard analysis not just focuses on the individual behaviour of the used machines in the production site and their emergent interactions, but also the server or other working equipment are considered. The advanced production site can be divided into a set of zones, such as parking, loading, charging, transporting, and unloading (dumping). In production sites, automated vehicles follow defined travel paths to move from one zone to another. However, several potential risks can be found in advanced production sites due to the simultaneous presence of automated vehicles, human-driven machines, and human workers at the same site/zone.

Congested zones occur when multiple machines simultaneously arrived at the loading, dumping, charging or parking zones from different paths. The geofences can provide a mechanism to control the movement of machines in such zones. Usually, the machines' movements are controlled by a fleet management server where the machines inform the server by their locations, and the server permits or restricts them from entering a specific zone. However, if the connection with the server is lost, a machine can enter in the congested zone with high speed and collide with another machine even with the presence of a collision-avoidance system, as the machine is approaching the same location and has not enough distance to stop. In this case, the geofences are particularly useful by enforcing a zero-speed limit at the entry of the congested zone. There is also a possibility that an automated vehicle arrives early at the specific zone and does not maintain the speed limits due to the failure of speed sensors and brakes. Moreover, the failure of obstacle detection and avoidance mechanism caused by Light Detection and Ranging (LIDAR) or camera failure may also lead to a collision of automated vehicles. The hazards like collisions of automated vehicles with static objects (e.g., ladder, stone, fallen pallet), dynamic obstacles (e.g., other automated vehicles, human worker, or other working equipment) appearing in travel paths, and specific zones caused by communication loss, subsystem failures and unsafe paths can also be avoided by enforcing geofences. The safety requirements and mitigation techniques (i.e., geofences) derived from the results of the risk assessment are used for designing and configuring the site environment, which serves as a digital twin of production site.

[1] https://www.polarsys.org/opencert/.

2.2 Digital Twin Based Safety Analysis

The digital twin brings out the virtual depiction of the real-world systems; the functions, behaviour, and communication capabilities are mirrored in the digital twin. The digital twins are perceived as an integral part of systems with enhanced automation, digitalization and connectivity. The automated vehicles are utilised for transportation and distribution of materials in advanced production sites. In addition to the automated vehicles, the human-driven machines can also be considered in the digital twin. Besides that, the emergent interactions with machines, site server and other devices are fundamental aspects in the digital twin. The Volvo CE (Construction Equipment) simulators[2], with unique digital twins of their construction machines, were adapted and extended for use in the case study of this paper. A number of different kinds of machines were selected to build a typical production scenario in the simulator. In the scenario, several zones were defined and the paths between these zones were specified. A detailed list of parameters were used to support automated operations of the scenario that can be accessed and changed during operational phase, if necessary. The safety requirements and hazard mitigation recommendations were also implemented in the scenario as code scripts. This not only gives the possibility to detect deficiencies, such as additional hazards and risks but also identify, monitor, evaluate, and resolve deviations from specified behaviours during the operational phase. There are some risks that can only be identified when a mishap occurs; however, the risks may not be detected if the probability is small, so they never happen.

2.3 Geofence Enforcement for Managing Risks

The geofences can serve as a countermeasure that could either eliminate the encountered hazards or reduce the risk of a mishap to an acceptable level. Geofencing is a virtual boundary (shape and dimension) defined for each zone (e.g., loading, unloading and transportation zones etc.), which in turn are divided into segments or edge paths. The automated vehicles are continuously monitored within the geofencing area. To be able to specify the virtual boundaries around geographic zones (i.e., geofences), the geometric shapes can be drawn in site zones. To control the entry for avoiding hazardous conditions due to the presence of multiple vehicles and uncertainties in collaborative interactions between machines, the communication with server is required; the default parameter value of the specific path point is set to restricted. The speed limits are also specified. For instance, the automated vehicle needs to maintain its position within drawn boundary of a loading point and a zero speed limit. Besides the GPS, the additional devices such as site cameras have also been used for locating machines in geofenced areas. The geofencing can be classified as follows:

[2] https://www.volvoce.com/europe/en/services/volvo-services/productivity-services/volvo-simulators/.

- Static geofences are constant or fixed. In the sites, the loading, dumping and charging areas may not change over time. Besides that, the movement of automated vehicles need to be restricted in various other fixed locations, such as, the storage place of dangerous materials and areas/regions in which humans are working.
- Dynamic geofencing moves over time. The capsule shape is used to widen the boundary for collision avoidance, for example, the vehicles not equipped with or have faulty obstacle detection devices, failed hardware, or transporting dangerous materials (e.g., explosive, toxic, etc.) are hazardous. In addition, the automated vehicles are not allowed to go for loading when maintenance team is present on site. In case of adverse environmental conditions such as slippery surface, the movement of automated vehicles can also be avoided in terms of dynamic geofences.
- Periodic geofences are only active or inactive for specific time periods. Therefore, they can be enforced to restrict the movement in certain areas for a specified time period. An example is termination of operation at the end of the day, so the movement towards areas except parking places is restricted.
- Conditional geofencing: The permissions associated with a geofence depends on certain factors like the number of vehicles can be allowed together, i.e., as a platoon for efficient operation. In addition, to deal with problems, such as path blockage, the movement of automated vehicle with a certain human-driven vehicle can be marked as conditional that follows to formulate an alternative travel path.

2.4 Dynamic Risk Management in Production Site

This subsection discusses the dynamic risk management in a production site. The uncertainty sources that includes the loss of connection with a server, system failures and unsafe paths need to be continuously monitored. The contracts can be derived for the uncertainty sources. In particular, they define the behaviour in a way that the assumptions (conditions) are made on the environment and if they hold then the certain behaviour/properties are guaranteed [2,7]. The site parameter values/ranges are retrieved from the simulation environments for monitoring the uncertainty sources. In circumstances when the deviations from specified behaviours are detected, the implications of failure vulnerabilities are determined and defences against them are performed. Besides the other safety measures, the enforcement of geofences at the operational phase is supported. Dynamic geofences are enforced to reduce mishap risk of emergent and evolving hazards to an acceptable level, for example, the travelling to particular area is blocked and the machines present in area are directed to drive away. The conditional geofence is used as a countermeasure against unsafe paths, the automated vehicle wait for a human-driven vehicle and follows to formulate an alternative travel path. By considering the deviations, implications, and respective control actions, the safety cases modelled in the OpenCert platform are updated. For this reason, the guidance presented in McDermid et al. [8] for safety assurance of automated systems is followed.

3 Case Study

3.1 Electric Quarry Site

This section describes an operational quarry site [15], which solely produces stone and/or gravel in various dimensions. The quarry operation is carried out using different kinds of machines, for instance, an excavator, a mobile primary crusher, a wheel loader, autonomous haulers, and a stationary secondary crusher. In particular, they collaborate together to realize the targeted production goals [9]. The quarry site is subdivided into the following different production zones, as shown in Fig. 2.

- Feeding Primary Crusher and Loading: The excavator feeds the blasted rocks (i.e., the rocks that are broken out of the mountain with explosives) to primary crusher. The primary crusher breaks the blasted rocks into smaller rocks. This is done to facilitate the transportation to the secondary crusher. For the discharging purpose, the conveyor belt is attached to the primary crusher. It is therefore possible to directly load the haulers from primary crusher. If primary crusher starts to build a stone pile, the direct loading is disabled. In such case, the hauler will be loaded with a wheel loader from the stone pile (i.e. indirect loading).
- Transportation: Autonomous haulers and/or articulated haulers are used to transport material in the quarry site. The operation of autonomous haulers is similar to the Automated Guided Vehicles (AGVs). For the perception of surrounding environment, two obstacle detection sensors, in particular, a LIDAR and a camera are mounted whereas the GPS is fitted for tracking and navigation purposes. The data produced by the particular sensors is processed for controlling the mechanical parts, for example, the drive unit for motion and operation, the steering system for manoeuvring, and the braking system for slowing down and stopping the vehicle. The interaction platform and other attachments that include batteries for power supply are integrated into the machines.
- Dumping and Feeding Secondary Crusher: The autonomous haulers move in the defined paths and dump the loaded rocks in the feeding spot of the secondary crusher. The secondary crusher further crushes the rocks into smaller granularity or fractions to meet the customer demands.
- Charging: To perform the mission efficiently, the required battery level needs to be determined. This is done before performing a mission. There are designated charging stations to recharge the battery whenever needed. If the energy consumption of the autonomous hauler is reduced, then battery needs to be less often charged. Energy consumption also depends on distance between different zones.
- Parking: The machines are moved to the parking zone after the termination of transportation operation and for maintenance purposes. For the assignment of specific places, the number and kind of machines needs to be determined.

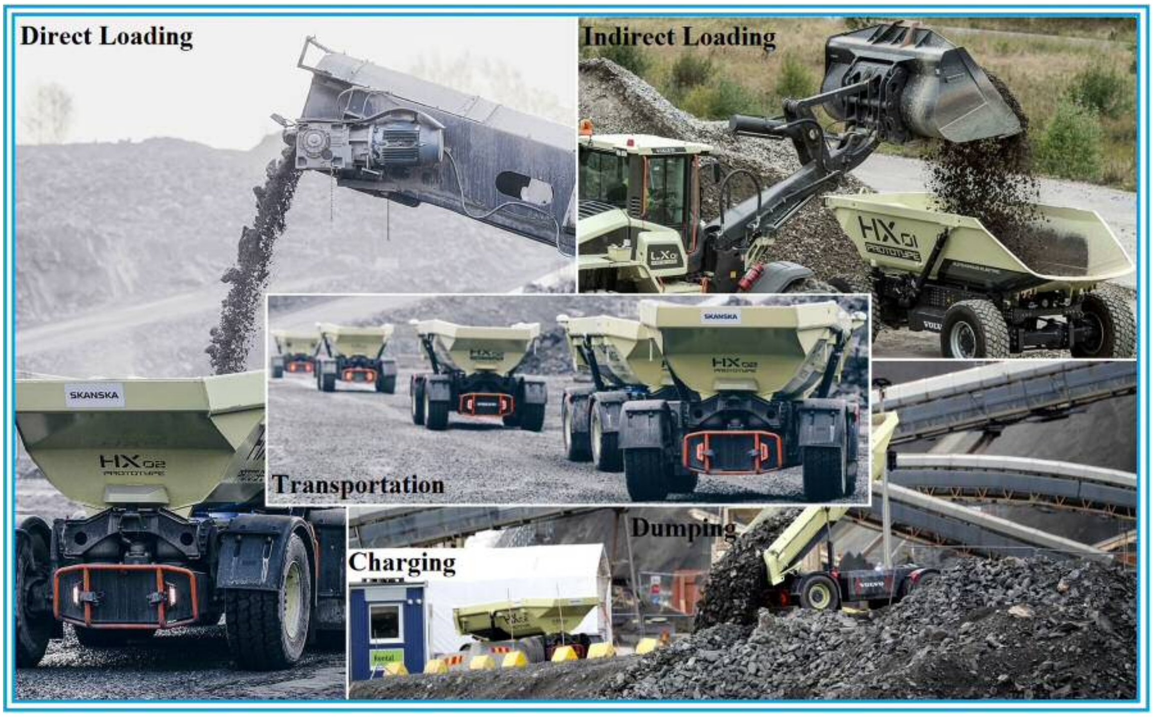


Fig. 2. Automated transportation in quarry site

3.2 Simulation-Based Digital Twin

For designing and configuring the quarry site, we have extended and adapted the Volvo CE simulators, fabricated by Oryx[3]. They serve as digital twins of various machines used at the quarry site such as the mobile primary crusher, the excavator, the wheel loader, autonomous haulers, articulated haulers, and the secondary crusher. The Volvo CE mobile platforms used for training the operators of articulated haulers, excavators, and wheel loaders are connected to the quarry site to demonstrate the functionality and behaviour of manually-driven machines. This means that the connected human-driven machines can operate in conjunction with the other machines in the quarry site. For instance, the rocks transportation in the quarry site can be carried out with the human-driven (articulated) and/or autonomous haulers. The site manager specifies the number and kind of machines to be used in the quarry site and their missions to fulfill the production demands. In the scenarios, the specific spots/zones are marked in the site map, e.g., parking, charging, loading, and dumping, as shown in Fig. 2. The transportation paths are also defined that operating machines use to move between different zones.

We have modelled and implemented static, dynamic, time-based and conditional geofences necessary for safe site operations in the above simulation test-bed. The site management shows the position and movement of all the operating machines; however, the screens placed on the machines show the visual region and machines present in this region. The values regarding timing, location, path

[3] https://www.oryx.se/.

points, load capacity and speed limits are provided to the user interface of the site management. For the emergent interactions and geofences, code scripts have been implemented. In the running mode, the information from machines operating in the quarry site is retrieved, stored, and displayed in the site management system.

3.3 Managing Operational Risks in Automated Transportation

We have defined geofences over a) various zones at the site, b) different machines, c) other actors at the site such as humans, and even d) around specified paths of movements. These geofences are of different geometric characteristics (typical studies use mainly circular geofences represented by a centre point and a radius). We also attach appropriate priority levels to indicate their relative importance/precedence. For example, zones have highest priority, followed by emergency/maintenance team, followed by vehicles and humans. In our simulation framework we can also represent geofences with multiple boundaries marked with different color codes (such as an outer boundary marked in yellow followed by an inner boundary marked in red) to indicate the relative criticality levels when another object reaches these boundaries. Static and common geofences details are available to all actors in the system. Dynamic geofences are enforced in conditions of subsystem failures. Periodic geofences are enforced to stop or control the operation and movement for a certain time-period. The conditional geofences are enforced to adapt new behaviour with an acceptable risk level.

The geofences-enabled safety is achieved through, central server commands, vehicle level actions, multiple checkpoints and a monitoring system; vehicle level actions are typically of two categories, viz., those taken by self for normal actions and those taken in response to failure conditions of self or others. Examples of server commands sent to vehicles include ‘Queue’, ‘Pause’, ‘Exit’, etc. The essence of all these are to adjust the vehicle speeds to acceptable levels in relation to the context. There are many challenges and trade-offs which we explore through our simulation test-bed before arriving at reasonable values for the geofences as well as command/action sequences in case of uncertainties. We now exemplify few typical scenarios of interest.

Normal Flow of Operations – Ensuring Safety at Loading Zone. Let us consider three autonomous haulers. H1 is present inside loading point which is modelled as a geofence, H2 is located at the entry to the loading zone and H3 is approaching towards the loading zone. The successful and safe loading operation requires the presence of only one hauler H1 in the loading point. The automated loading is compromised if two or more haulers arrive in the loading point. There is a need to communicate with the server for entry in geofenced region that is triggered when the located hauler H2 touches the sensor at entry to the loading zone. Since the hauler H1 is already present in the loading point, the hauler H2 requesting permission to enter is given a command to be in a ‘queue’. After the completion of current loading, an ‘exit’ command is given to the loaded hauler H1, which then start moving. Next, the waiting hauler in queue (H2) is given the permission to enter; its maximum speed limit is set to 20 km/h. The other hauler H3 may also arrive in the meanwhile and instructed to be in the queue at next level; H3 moves to the place of H2. The boundary of the loading area is constant/fixed. It should not be violated and the hauler needs to maintain 0 km/h speed while at the loading point. If center of mass of hauler is not maintained, the stones will fall on loading point and the haulers may also get damaged. In this case, the risk is not regarded as acceptable; as a control action, the hauler is given a command to exit and approach again. Besides the GPS position, site cameras are also placed to realize precise point positioning. The geofenced regions in quarry site can involve many uncertainties and therefore continuously monitored.

Resolving the Failure Cases

- **Communication Failures:** The loss of communication with a server is a safety risk. Besides that, the messages containing less, or wrong data can also cause mishaps. When the primary crusher is jammed, the human operators can be called on site, during that period a direct loading command is sent to the autonomous hauler instead of the loading from wheel loader. The transformation of an incorrect mission or travel path to an autonomous hauler can be caused by an incorrect command or timing failure that leads to an incomplete mission, machine damage or human injuries. In such cases, the control action is in place, i.e., the movement of autonomous haulers is still restricted in geofenced areas. The capsule geometry shapes (geofences) around the haulers provide the means for obstacle avoidance. These shapes are drawn in different ranges and different colors based on their criticality level. When an obstacle is detected in yellow range (indicating move with caution at reduced speeds), the slow down or stop measures are taken, the red range is regarded as emergency stopping distance. The haulers can maintain maximum speed limit if no obstacles are detected within the range.
- **Subsystem Failures:** As another example, consider a subsystem failure. There is a possibility that an autonomous hauler arrives early in loading zone and does not maintain new speed limit due to speed sensor and brake failures. In the former case, the focus is shifted to the map to compute the speed, i.e., for detecting distance covered in time frame. In the latter case, besides the steering wheel rotation commands, depending on the severity risk factor, dynamic geofences are enforced.
- **Path Problems:** The travelling in nearby area is blocked and the autonomous haulers in travelling path, including in standstill mode, such as loading point are commanded to drive away to reduce the risk to an acceptable level. In case of path problem, to create a new path compliant with the conditional geofence, the autonomous hauler wait for the human-driven hauler or another machine, such as wheel loader to formulate an alternative travel path, and then follows it.

Update of Safety Cases. Geofences, together with the permissible operations/commands and associated speed limits gets easily translated into a set of safety contracts between the involved actors. Appropriate monitoring mechanisms need to in place to check the associated parameters during operational phase. Failure cases often leads to partial or full non-conformance to contracts or enable new contracts to be considered during such situations. The update of safety cases is carried out based on the risk control actions. The required changes in safety cases and associated safety contracts modelled in the OpenCert platform are tracked and then the update command is launched, so that the gaps are resolved and current system safety is not compromised.

4 Related Work

Baumgart et al. [1] examine the feasibility of System-Theoretic Process Analysis (STPA) for System-of-Systems (SoS). For this reason, a simplified control structure diagram of a quarry site is used. Muram et al. [9] perform the SoS hazard analysis for the quarry site, in particular, the HAZOP and FTA techniques are applied. The research in [4] presents the idea of through-life safety assurance. It is reflected in four activities: identify, monitor, analyse and respond. Jaradat and Punnekkat [7] discuss the monitoring of runtime failures related to the hardware component and failures analysis by comparing with a predefined threshold. The fault trees were used for deriving safety contracts and defining thresholds. The digital twin, however, is not used. In this paper, besides the hazard analysis during design and development phase, the digital twin is used for gaining confidence and managing transportation risks at the operational phase.

Zimbelman et al. [16] use the mobile geofences in forestry to define safe work areas; moving, circular safety zones around people and heavy equipment has the potential to reduce accidents during logging operations. For the tree falling hazard zones around manual fallers, the traditional proximity alerts and the overlap among multiple circular geofences of varying radii are considered. Dasu et al. [3] give a vision of air-traffic control based on geofences. In general, the partition of a sky is needed to allocate the flying space and prevent entry in certain areas. This can be done via geofences. Stevens and Atkins [13] include the operating permissions in geofence specification; access to the airspace is enabled by considering property type or vehicle risk. Nait-Sidi-Moh et al. [10] present the integration of geofencing techniques with the TransportML platform for real-time tracking of mobile devices. The application allows defining adequate and safe itineraries for the registered vehicles. If a vehicle deviates from its geo-corridor, alerts are sent to the in-vehicle computer to warn the driver and to the management center to generate a new alternative itinerary. The use of geofences is analysed in the defence and security sector; for instance, the potential of using geofences as tools to prevent terrorist attacks using hazardous material transportation [12]. Guo et al. [6] developed a model for dynamic geofences. Through a lane-level precise positioning service, the geofences can serve as an efficient and reliable active safety service for a vehicle accident prevention. The published studies have not considered the geofences for automated transportation and production sites. In this paper, the queue, pause and exit restrictions are presented that provides efficient resource for risk control in various areas of production site, such as loading, unloading, charging and parking.

5 Conclusion and Future Work

To support the dynamic risk management, which is perceived as an essential characteristic of production sites with enhanced automation, digitalization and connectivity, the central theme of this paper focuses on three particular aspects: (i) hazard analysis and risk assessment during the design and development phase; (ii) virtual depiction of the real site using simulators-based digital twins; and (iii) risk management and update of safety cases at the operational phase. The hazard analysis during the design and development phase is a fundamental element and absolutely necessary for safety-critical systems. Based on the hazard analysis, which is performed with HAZOP and FTA techniques, the mitigation mechanisms, such as geofences are established; they are translated into the safety requirements. The Volvo CE simulators-based digital twins are leveraged in which the mitigation mechanisms and safety requirements are implemented. During the design and development phase, all hazards and causal factors may not be identified. The intention with the simulators-based digital twins is to perform verification and validation to gain confidence in production site. It served as a resource to discover additional hazards. Finally, the dynamic risk management and safety assurance during the operational phase is carried out. The data from digital twin is gathered to identify and monitor deviations from specified behaviours, evaluate and select the optimal control action and resolve problems. Note that the geofences are used as an active countermeasure against mishap risks in various site areas. The applicability has been demonstrated for the Volvo CE electric quarry site.

The research presented in the paper is primarily based on the exploratory studies we conducted using a simulation test-bed which features realistic models of the machines and processes of Volvo quarry site. We have so far obtained results which satisfy the coarse level specifications on accuracy and precision requirements. However, we need to conduct further detailed evaluations for stabilisation of the geofencing function, as well as implement them as per the mandatory domain-specific safety requirements for potential inclusion in real Volvo machines. It is noteworthy that, the geofencing function is generally applicable to the broad range of scenarios and domains. In the future, we plan to consider additional scenarios and applications based on the Industry 4.0. To the best of our knowledge, geofences-enabled safety mechanisms are not considered in current domain-specific standards. We also aim to highlight its potentials, in due course by having discussions with the relevant standardization bodies at the national level.

Acknowledgment. This work is supported by SUCCESS (Safety assUrance of Cooperating Construction Equipment in Semi-automated Sites) project via the AAIP (Assuring Autonomy International Programme), FiC (Future factories in the Cloud) project funded by SSF (Swedish Foundation for Strategic Research) and Volvo Group Arena.

References

1. Baumgart, S., Fröberg, J., Punnekkat, S.: Can STPA be used for a system-of-systems? Experiences from an automated quarry site. In: 2018 IEEE International Symposium on Systems Engineering (ISSE), Rome, Italy, pp. 1–8 (October 2018). https://doi.org/10.1109/SysEng.2018.8544433
2. Benveniste, A., et al.: Contracts for System Design. Research report RR-8147, INRIA (November 2012). https://hal.inria.fr/hal-00757488
3. Dasu, T., Kanza, Y., Srivastava, D.: Geofences in the sky: herding drones with blockchains and 5G. In: 26th ACM International Conference on Advances in Geographic Information Systems, pp. 73–76 (November 2018). https://doi.org/10.1145/3274895.3274914
4. Denney, E., Pai, G.J., Habli, I.: Dynamic safety cases for through-life safety assurance. In: 37th IEEE/ACM International Conference on Software Engineering (ICSE), Florence, Italy, pp. 587–590 (May 2015). https://doi.org/10.1109/ICSE.2015.199
5. Ericson, C.A.: Hazard Analysis Techniques for System Safety. Wiley, Hoboken (2005)
6. Guo, C., Guo, W., Cao, G., Dong, H.: A lane-level LBS system for vehicle network with high-precision BDS/GPS positioning. Comput. Intell. Neurosci. **2015**, 531321:1–531321:13 (2015)
7. Jaradat, O., Punnekkat, S.: Using safety contracts to verify design assumptions during runtime. In: 23rd International Conference on Reliable Software Technologies, Ada-Europe 2018, Lisbon, Portugal, pp. 3–18 (June 2018). https://doi.org/10.1007/978-3-319-92432-8_1
8. McDermid, J., Jia, Y., Habli, I.: Towards a framework for safety assurance of autonomous systems. In: Workshop on Artificial Intelligence Safety Co-located with the 28th International Joint Conference on Artificial Intelligence, AISafety@IJCAI, Macao, China (August 2019)
9. Muram, F.U., Javed, M.A., Punnekkat, S.: System of systems hazard analysis using HAZOP and FTA for advanced quarry production. In: 4th International Conference on System Reliability and Safety (ICSRS), Rome, Italy, pp. 394–401 (November 2019). https://doi.org/10.1109/ICSRS48664.2019.8987613
10. Nait-Sidi-Moh, A., Ait-Cheik-Bihi, W., Bakhouya, M., Gaber, J., Wack, M.: On the use of location-based services and geofencing concepts for safety and road transport efficiency. In: Matera, M., Rossi, G. (eds.) MobiWIS 2013. CCIS, vol. 183, pp. 135–144. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-03737-0_14
11. Object Management Group: Structured Assurance Case Metamodel (SACM), Version 2.1 (April 2020). https://www.omg.org/spec/SACM/2.1/PDF. Accessed: 9 Aug 2020
12. Reclus, F., Drouard, K.: Geofencing for fleet & freight management. In: 9th International Conference on Intelligent Transport Systems Telecommunications (ITST), pp. 353–356 (2009). https://doi.org/10.1109/ITST.2009.5399328

13. Stevens, M., Atkins, E.: Geofencing in immediate reaches airspace for unmanned aircraft system traffic management. In: 2018 AIAA Information Systems-AIAA Infotech @ Aerospace, Kissimmee, Florida (January 2018). https://doi.org/10.2514/6.2018-2140
14. The Assurance Case Working Group: Goal Structuring Notation Community Standard Version 2, January 2018 (2018). http://www.goalstructuringnotation.info/
15. Volvo Construction Equipment: Emission-free quarry. https://www.volvoce.com/global/en/news-and-events/press-releases/2018/testing-begins-at-worlds-first-emission-free-quarry/
16. Zimbelman, E.G., Keefe, R.F., Strand, E.K., Kolden, C.A., Wempe, A.M.: Hazards in motion: development of mobile geofences for use in logging safety. Sensors **17**(4), 822 (2017)

A Process to Support Safety Analysis for a System-of-Systems

Stephan Baumgart*, Joakim Fröberg†, Sasikumar Punnekkat‡
* Volvo Autonomous Solutions, Eskilstuna, Sweden
Email: stephan.baumgart@volvo.com
† RISE Research Institutes of Sweden, Västerås, Sweden
Email: joakim.froberg@ri.se
‡School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden
Email: sasikumar.punnekkat@mdh.se

Abstract—**Autonomous vehicles grow importance in many domains and depending on the domain and user needs, autonomous vehicles can be designed as stand-alone solutions as in the automotive domain or as part of a fleet with a specific purpose as in the earth moving machinery domain. Contemporary hazard analysis methods primarily focus on analyzing hazards for single systems. Such an analysis requires knowledge about typical usage of a product, and it is evaluated among others if an operator is able to handle a critical situation. Each hazard analysis method requires specific information as input in order to conduct the method. However, for system-of-systems it is not yet clear how to analyze hazards and provide the required information. In this paper we describe a use case from the earth moving machinery domain where autonomous machines collaborate as a system-of-systems to achieve the mission. We propose a hierarchical process to document a system-of-systems and propose the use of model-based development methods. In this work we discuss how to utilize the provided details in a hazard analysis. Our approach helps to design a complex system-of-systems and supports hazard analysis in a more effective and efficient manner.**

Index Terms—**Autonomy, System-of-Systems, Safety Analysis, Hazard Analysis**

I. Introduction

In the past decades the complexity of vehicles, airplanes, trains or heavy machinery has increased significantly, because of a higher utilization of software to provide new customer features. Some of these features support the human drivers or operators like assisting in parking a car (park assist), keeping the lane when driving (lane assist) or ensuring a safe distance to the vehicle in front (adaptive cruise control). We can recognize a trend in industry going from assisting features towards automating tasks of the operation with the intention to reduce the risk for human failure and increasing efficiency. Automation can focus on single vehicles, where the single vehicle is performing a task and collaboration with other systems is not required. In the earth moving machinery domain it is more common that various types of machines collaborate in a repetitive workflow. When automating such machines, the resulting workflows and possible collaborations and interactions must be analyzed thoroughly. Such collaborating systems can be seen as system-of-systems (SoS). When connecting single systems to a system-of-systems, a new level of complexity is added. A failure in constituent system A might not be critical for this system and therefore not identified as safety critical. By sharing this erroneous data through the communication network of connected constituent systems, this may lead to an unforeseen accident with constituent system B as described in [1]. Such systems-of-systems are growing their importance in the truck domain, where platooning of trucks is being explored to improve fuel efficiency [2] or automated vehicles transport material in off-road environments [3]–[5]. Developing system-of-systems and integrating automated vehicles add new dimensions of complexity to the already complex systems and processes. How to achieve safety when developing such a system-of-systems is not yet clear, since existing functional safety standards like ISO 26262 [6] or IEC 61508 [7] also do not explicitly cover system-of-systems.

In this paper, we provide guidance for practitioners on how to document a system-of-systems to aid the safety analysis of such a system. We propose a hierarchical process to document a system-of-systems and provide an example how model-based development practices can be utilized. The practices shown in this paper are based on our experience developing a safety-critical system-of-systems. We primarily focus on the concept stage as described in ISO 21839 [8]. In the concept stage, the system-of-interest shall be defined and analyzed, requirements shall be collected, and risks shall be identified, assessed and appropriate mitigation mechanisms shall be defined.

The paper is structured as follows. In section II we provide the background and related work in the area of system-of-systems and considering safety. As explained above, the provided guidance is related to our experience when designing a system-of-systems and we therefore provide a description of our use case in section III. In section IV we present our approach and describe the different phases. We conclude our paper in section V.

II. Background and Related Work

A. Safety Lifecycle

Developing safety-critical products requires to follow appropriate safety standards, where best practices and approved methods are embedded in a reference process. The term safety in general is related to the “absence of catastrophic consequences on the user(s) and the environment” [9]. The usage scenarios and the included features of the targeted product need to be thoroughly analyzed to identify those situations, where users or bystanders are at risk to get injured or killed or other equipment or the environment can be damaged. This requires thorough analysis to identify what could potentially cause an accident to happen. Such causes can be for example failures in the components in the product or external influences. Analyzing potential failures in one of the embedded systems or the software running on the control units is covered under the term 'functional safety'. Functional safety is defined as the "absence of unreasonable risk due to hazards caused by malfunctioning behavior of E/E systems" [6]. A hazard in this context is a "potential source of harm" [6], meaning that in a specific situation, the hazardous event, this hazard can lead to an accident. A typical example for such a standard is the automotive domain specific functional safety standard ISO 26262 [6]. This standard provides a framework for developing the embedded systems in a car. One part of the framework is a reference process containing process steps of development processes, production, operation, service and decommissioning and other supporting processes. The proposed development process starts with the concept phase, where an *Item* is defined. An item in the context of ISO 26262 is a "system or combination of systems, to which ISO 26262 is applied, that implements a function or part of a function at the vehicle level". An item is limited to a single vehicle and can be a feature like parking brake or steering. Requirements related to the item like the boundary, the targeted behavior seen from the driver's perspective, constraints and dependencies need to be captured. This input is used for conducting the hazard analysis and risk assessment (HARA) and determining the automotive safety integrity level (ASIL). These details are used throughout the development process and closure of the identified hazards need to be shown at the end of the development as part of the safety case, which includes arguments and evidence. This functional safety standard focuses on functions in a single vehicle, i.e., a single system. Clarification on how to develop a system-of-systems is not in the scope of ISO 26262.

B. Systems vs. System-of-Systems

In our work we focus on system-of-systems and therefore we provide a distinction between systems, on which the ISO 26262 focuses on and system-of-systems. ISO 26262 defines the term system as a "set of components or subsystems that relates at least a sensor, a controller and an actuator with one another" [6]. A more general definition of a system is provided in MIL-STD-882E [10]: "The organization of hardware, software, material, facilities, personnel, data, and services needed to perform a designated function within a stated environment with specified results." In the same standard the term system-of-systems is defined as "a set or arrangement of interdependent systems that are related or connected to provide a given capability" [10]. The standard ISO 21841 defines that a system-of-systems consists of a "set of systems or system elements that interact to provide a unique capability that none of the constituent systems can accomplish on its own" [11]. A constituent system in this context is an "independent system that forms part of a system of systems (SoS)" [11].

Various characteristics to highlight the differences between systems and system-of-systems have been listed in literature as for example:

- operational independence of the element [12], boundaries and interfaces [13]
- managerial independence of the elements [12], [13]
- evolution [12]
- emergent behavior [12]–[14]
- geographic distribution [12], operational focus [13]
- autonomy [14]
- belonging [14]
- connectivity [14]
- diversity [14]

A commonly accepted categorization of types of SoS has been proposed by Maier [15]. Maier is using the way a SoS is organized and managed as the parameter to differentiate them. He identifies three types of SoS: 1) Directed SoS, where a master system is coordinating the slave systems in an SoS. 2) Collaborative SoS, where the constituent systems may join a SoS to fulfill the goal of the SoS, and 3) a Virtual SoS, which have no central management or agreed purpose.

Axelsson [16] is providing an extension to the existing definitions by adding the states of the constituent systems, which has an impact if for example such a system is participating in a SoS or if it is not participating and passive with regards to the SoS.

C. Safety and System-of-Systems

In this section, we briefly discuss the literature focusing on safety in a system-of-systems. Hall-May and Kelly [17] utilize a case from the military domain and describe a system-of-systems using model-driven engineering methods and create a safety argumentation using the goal structuring notation (GSN) [18]. Alexander et. al [19] propose a simulation-based hazard analysis as a possibility to handle the complexity of interactions between constituent systems. Focusing on the interfaces and potential cascading failures in a system-of-systems, Redmond described the Interface Hazard Analysis method in [1], [20].

The compliance with existing functional safety standards like ISO 26262 [6] in the context of system-of-systems is described by Saberi et al. [21] through a platooning case from the truck domain and propose a tailored safety lifecycle. The authors highlight, that it is important to understand potential real live scenarios in order to be able to analyze the impact of failures and their potential cascading effect in this context. Axelsson and Kobetski [22] apply the system thinking approach STAMP [23] to analyze risks in a truck platooning case.

Compliance with functional safety standards requires considering critical scenarios during design-time. When self-adaptive collaborating systems are applied in a system-of-systems and no central unit is used to coordinate the activities of the autonomous systems, not all constellations and situations can be considered during design-time. Instead, safety may need to be negotiated at run-time as presented in [24].

III. A Case from the Earth Moving Machinery Domain

We utilize the electric site research project [25] as a case for our work. In this project a fleet of automated guided vehicles (AGVs) called HX are used to transport pre-crushed material from a movable primary crusher to a stationary secondary crusher. Along with a fleet of autonomous HX, a human-operated wheel loader and a human-operated excavator are used for loading material on the HX. In our earlier works we have been analyzing safety in context of certain specified scenarios of this complex SoS [3], [4].

In Figure 1, a typical setup for an automated quarry site is presented. The automated guided vehicles follow predefined tracks on the site. In this configuration there are two alternative possibilities to load a HX with gravel. The first way is to utilize direct loading from the movable primary crusher (PCR), which is filled by an excavator (EXC). Alternatively, the HX can be loaded using a human-operated wheel loader (WL). In order to enable choosing which loading area is relevant, the empty HX wait at the main decision point (MDP) until they get a mission assigned by the fleet control server. A loaded HX transports the material to the stationary secondary crusher (SCR) and unload the contents there. Since the HXs are electrified, they require charging of their batteries at the charging stations (CH).

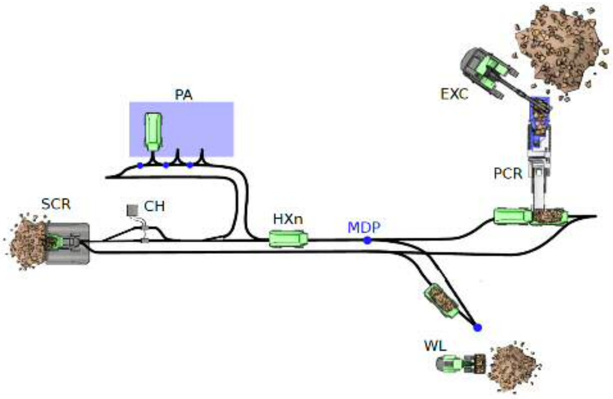


Fig. 1. Automated Quarry Site

Generally, a HX can either be controlled by a remote control or by a fleet control server as shown in Figure 2. While the remote control is maintaining a one-to-one connection and is directly controlling the vehicle movement, the fleet control server is providing vehicle-specific missions to the active HXs, which are interpreted and translated to movements by the on-board system of a targeted HX. The site operator is supervising the activities of the autonomous and human-operated vehicles and possible humans moving inside the restricted area.

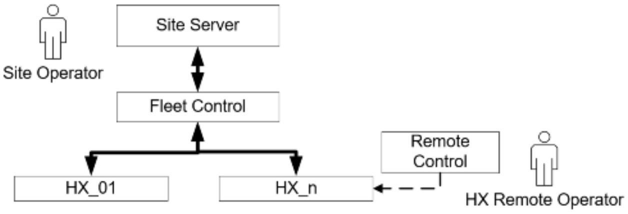


Fig. 2. Automated Quarry Site - Control Structure

The depicted quarry site case is one instance of such a site. When working with this research project we recognize the following dimensions that impact safety activities and argumentation:

1) Structure of SoS and constituent systems: The involved autonomous and human-operated machines, workflows and for example environmental conditions are specific for a site. Since workflows, tasks and environmental conditions may differ from one site to another, the safety and risk avoidance may differ as well. Accordingly, hazards and risks may be missed when reusing safety arguments.
2) Dynamicity at a site: The fleet of HX are operated in an outside and off-road environment with possible impact of changing weather conditions, which require adapting the workflow at a site. Furthermore, changes in the number of machines, changes in the production process, relocating loading and unloading spots and changing routes, may require revisiting the safety arguments.
3) Evolution of the SoS : Machines and systems may evolve over time with changed and adopted features or even new features. Such changes can for example be realized through software updates for improving the capabilities of a vehicle or adjusting the workflow to new conditions. Customer sites can be dynamic so that workflows and tasks evolve over time. This can impact the number and type of machines that are required. Usually, it is challenging to foresee how a site may evolve over time.

In order to be able to support the reuse of development artifacts and related safety analysis and safety concepts, a structured process on how to specify a system-of-systems supporting a safety analysis is necessary.

IV. The SafeSoS Approach

In this section we describe our approach called SafeSoS, which is including concepts for specifying a SoS and using those specifications for performing a safety analysis. We apply the hierarchical levels described by Axelsson [16] to provide a model-centric approach to design the system-of-systems. Axelsson is differentiating between macro analysis, where the scope and the context of the SoS is analyzed. This information is refined in the meso analysis, where information on how the constituent systems form is analyzed. In the micro analysis, the focus is on single constituent systems and how they contribute to the overall SoS goal. We utilize this mindset to structure the information about the SoS.

In Figure 3 the SafeSoS process is shown with descriptions on macro level, meso level and micro level. For each of these levels we distinguish between information w.r.t. structure and behavior and discuss who typically can provide such information. All provided information and requirements on these levels are connected and used in the SoS safety analysis phase.

A. SoS Macro Level

The main goal of the SoS Macro Level of our process is to capture the boundary of the targeted system-of-systems, environmental characteristics and derive use cases and typical scenarios.

1) Macro Level - Structure: The constituent systems planned to be joining the SoS shall be listed. It is also necessary to consider other systems that could possibly enter the SoS operating zone. If for example the constituent systems are not aware of a vehicle entering the operating zone, there is a risk for fatal accidents. Especially, when considering automated vehicles to be part of the system-of-systems, an unknown vehicle entering the automated operating zone, may lead to unpredictable behavior. Another aspect to list is potentially exposed humans, such as informed people, for example those operating a constituent system or controlling the operation, and people that are not informed such as visitors or rescue teams. If possible, environmental conditions also need to be listed and how the capabilities of the SoS are influenced by different conditions. While icy tracks increase the braking distance for vehicles, hot weather conditions may lead to dust and reduced quality of sensor data.

2) Macro Level - Behavior: In the behavior level of the SoS Macro Level, the usage concepts of the SoS shall be described. This can contain use cases on how the SoS is used and how it can be operated. Typical scenarios need to be derived in order to be able to identify those scenarios, where for example humans are at risk. Additionally, the states of the SoS need to be described. In the above-mentioned quarry site, typical states can be morning startup, normal operations or evening shutdown. It is important to identify additional scenarios and use cases that can be relevant like emergency stop of all autonomous machines or their recovery to operation. In this context, the states of the SoS can provide an indication of possible critical situations and need to be captured.

In this initial phase it is useful to interview stakeholders and run brainstorming meetings with developers to understand the processes where the system-of-systems shall be applied. In such a brainstorming meeting, potential losses can be identified and rated to achieve a sorted list based on criticality. Based on the provided information, it can be analyzed which persons are at risk and which scenarios seem to be most critical. It is possible to derive hazard paths based on the identified potential losses.

B. SoS Meso Level

In the SoS Meso Level, the internal perspective of the SoS with focus both on the internal structure and interactions between the constituent systems are captured.

1) Meso Level Structure: The internal structure of the SoS will focus on which constituent systems are participating in a SoS, possible servers and through which channels they communicate. It is for example important to capture, if autonomous constituent systems shall communicate directly to each other or via a coordinating server. The structural dimension of the Meso Level will provide insights about the type SoS, as for example directed SoS or collaborative SoS [15].

2) Meso Level Behavior: In the behavior views of the SoS Meso Level the interaction between the involved humans and the constituent systems shall be described. By the help of this descriptions, possible human errors can be identified. As a second aspect, the interaction between the constituent systems shall be described. This may include additional information about complex messages that are shared between the constituent systems. By the help of these details, the propagation of possible failures can be studied. Finally, details about the states of the constituent systems and their dependencies shall be specified enabling identification of safe states as well as inconsistencies. An example on how the states of constituent systems depend on each other is depicted in Figure 4 using a SySML state chart diagram. In this example a remote control is used to connect to a HX and the server is deciding about the request. System Designers and safety engineers can provide the required information.

C. SoS Micro Level

The SoS Micro Level contains details about a single constituent system. This level also consists of structural and behavioral views.

1) Micro Level Structure: In the structural view of the SoS Micro Level, details about the internal structure of a constituent system are captured. It is important though to focus on those details related to the SoS. In our case it is necessary for example to document how the remote emergency stop feature, that shall stop all active machines at a site when initiated by the site operator, is realized inside the machine.

2) Micro Level Behavior: The behavior level of the SoS Micro level contains details about timing, states or messaging characteristics of a single constituent system with respect to the SoS it is integrated in. The states of the constituent system are directly connected to the states for the complete SoS as described on the SoS Meso Level.

For the Micro Level details, system developers can provide the relevant information and safety engineers may help that all safety related details are provided.

It is necessary to ensure sufficient traceability between the levels, for which we have used state machines. The use cases on SoS Macro Level for example are directly connected to the Human Interactions with SoS on SoS Meso Level.

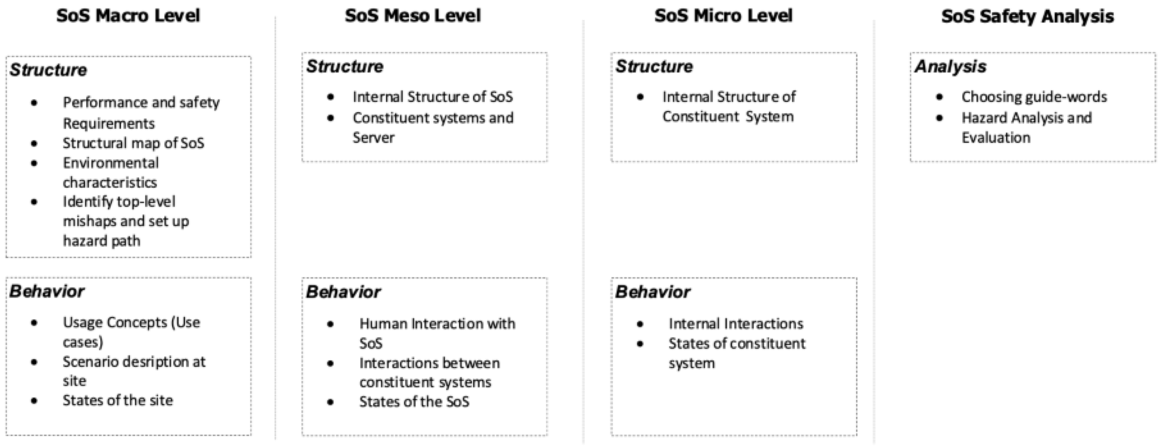

SoS Macro Level
Structure
Performance and safety Requirements
Structural map of SoS
Environmental characteristics
Identify top-level mishaps and set up hazard path
Behavior
Usage Concepts (Use cases)
Scenario desription at site
States of the site
SoS Meso Level
Structure
Internal Structure of SoS
Constituent systems and Server
Behavior
Human Interaction with SoS
Interactions between constituent systems
States of the SoS
SoS Micro Level
Structure
Internal Structure of Constituent System
Behavior
Internal Interactions
States of constituent system
SoS Safety Analysis
Analysis
Choosing guide-words
Hazard Analysis and Evaluation

Fig. 3. SafeSoS: Safety Process to support System-of-Systems

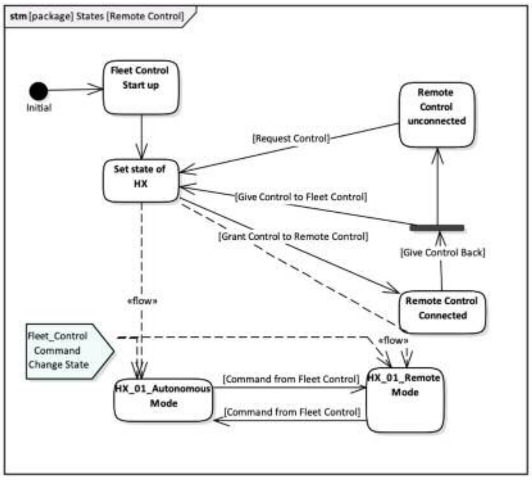

stm [package] States [Remote Control]
Initial
Fleet Control Start up
Remote Control unconnected
[Request Control]
Set state of HX
[Give Control to Fleet Control]
[Grant Control to Remote Control]
[Give Control Back]
Remote Control Connected
«flow»
«flow»
Fleet_Control Command Change State
HX_01_Autonomous Mode
[Command from Fleet Control]
HX_01_Remote Mode
[Command from Fleet Control]

Fig. 4. State Chart - Remote Control Case

D. SoS Safety Analysis

In this section, we describe how a safety analysis for a SoS can be performed utilizing the information provided as part of the SoS specifications. As we described above, the possible humans at risk and possible critical scenarios are important information to understand where humans are at risk. The Meso Level is providing information about control structures, communication channels, timing aspects and potential safe states. The Micro Level in the scope of this paper is solely focusing on the states of a specific constituent system.

Guidewords help to identify critical situations and are commonly used in the literature on hazard analysis. In our work we have used the guidewords from the HAZOP methodology [26] as a starting point. Typical examples of guidewords are

- NO or NOT: This means, that a certain input is not provided or not received (indicating an 'Omission fault'). Depending on the SoS Level, a certain message is not provided by one constituent system or the SoS supervisor is not in an emergency stop when expected.
- MORE: Typically, this guideword would identify too long activation of a valve in the process industry (similar to a 'Commission fault'). In our case, different hazards may be detected using this guideword, depending on the SoS level. On Macro Level, MORE can characterize environmental changes or MORE speed of a specific constituent system. Another aspect that need to be considered is, that MORE constituent systems are used in an SoS, than it was started with for a day or than it was intended for which is indicating failures related to the dynamicity at the SoS.
- LATE: The intention of the LATE guideword is to identify those situations where controls or messages are provided too late (similar to a 'timing fault'). Again, depending on the SoS Levels, different scenarios can be identified. In the SoS Macro Level, the delayed identification of unauthorized personal at a quarry site, may lead to possible accidents when autonomous machines operate. On SoS Meso Level, delayed provision of the GPS position of an autonomous machine, may lead to higher uncertainties about current traffic situations.

The other HAZOP guidewords can also be used, but needs clarification about the meaning on the different SoS levels. It may also be necessary to tailor this list of guidewords to enable a full-fledged SoS safety analysis. From our experience, we added the guideword INCORRECT to our analysis (referring to a 'value fault'), to enable capturing the situation, when missions are sent from the server to the constituent systems. These messages may contain a list of positions and speed profiles for following the track. The message may be received on time and with correct length, but the contents may be wrong, which may lead the autonomous machine not operating as intended. The structural views of the SoS levels described above, help to find hazards related to malfunctioning systems or components of the SoS. The behavior views help to find hazards related to the dynamicity of the SoS.

The results of the analysis performed in our case study was captured using a spreadsheet, where the structuring was pivoted based on use cases and scenarios. The classification helped to derive solutions during development to reduce the risks and provide evidence of mitigation implemented. The spreadsheet was generated in consultation with the design team as well as safety engineers, which resulted in identification of many potential risks (such as, remote control take over), which might have been overlooked otherwise. For individual hazard analysis we have used common methods like FMEA [27] as well as conformance with associated SIL [7]/ASIL [6] requirements. During this process, we also became aware of the many aspects of relevance which demand a more integrated tool-oriented approach for guiding such a safety analysis. Our planned works include development of a tool which can help in capturing essential information at each levels, connect them as well as have an intuitive user interface to the design/safety team. Having inter-operability by providing hooks to other detailed methods of relevance is also planned.

V. Conclusions

The existing hazard analysis methods focus on analyzing hazards for single systems. For the analysis, the usage of the final product needs to be known or assumed and therefore knowledge about application scenarios is essential. Each hazard analysis method requires specific information as input in order to conduct the method. For system-of-systems, how to specify the requirements, capture the essential safety relevant information and analyze hazards, still remain as an open and challenging research domain.

In this paper, we described a case from the earth moving machinery domain where autonomous machines collaborate and form a system-of-systems to meet the mission objective. We described an approach to document system-of-systems and show how the information is used for performing hazard analysis of SoS using guidewords. In the future, we plan to extend our approach and establish the connections and traceability to the individual machine's safety analysis.

Acknowledgments

The authors acknowledge the funding support received for this research from the KKS-funded ITS-EASY Post Graduate School for Embedded Software and Systems and the SUCCESS Project supported by the Assuring Autonomy International Programme (AAIP), a partnership between Lloyd's Register Foundation and the University of York.

References

[1] P. J. Redmond, "A System of Systems Interface Hazard Analysis Technique," Master's thesis, 2007. [Online]. Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a467343.pdf

[2] S. Tsugawa, S. Jeschke, and S. E. Shladovers, "A review of truck platooning projects for energy savings," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 68–77, 2016.

[3] S. Baumgart, J. Froberg, and S. Punnekkat, "Analyzing hazards in system-of-systems: Described in a quarry site automation context," in *2017 Annual IEEE International Systems Conference (SysCon)*. IEEE, 4 2017, pp. 1–8.

[4] ——, "Can STPA be used for a System-of-Systems? Experiences from an Automated Quarry Site," in *2018 IEEE International Systems Engineering Symposium (ISSE)*, no. 4. IEEE, 10 2018, pp. 1–8. [Online]. Available: http://www.es.mdh.se/publications/5246-https://ieeexplore.ieee.org/document/8544433/

[5] S. Baumgart, J. Fröberg, and S. Punnekkat, "A State-based Extension to STPA for Safety-Critical System-of-Systems," in *4th International Conference on System Reliability and Safety*, 11 2019. [Online]. Available: http://www.es.mdh.se/publications/5674-

[6] International Organization for Standardization, "ISO 26262:2018 - Road vehicles – Functional safety," 2018.

[7] International Electrotechnical Comission, "IEC 61508:2010 Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems," 2010.

[8] International Organization for Standardization, "ISO/IEC/ IEEE 21839 -Systems and software engineering - System of systems (SoS) considerations in life cycle stages of a system," 2019.

[9] A. Avizienis, J. C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *Dependable and Secure Computing, IEEE Transactions on*, vol. 1, no. 1, pp. 11–33, 2004.

[10] United States Department of Defense, "MIL-STD-882E," Washington, DC, USA, 2012.

[11] International Organization for Standardization, "ISO/IEC/IEEE 21841 Systems and software engineering — Taxonomy of systems of systems," 2019.

[12] M. W. Maier, "Architecting Principles for Systems-of-Systems," *INCOSE International Symposium*, vol. 6, no. 1, pp. 565–573, 1996.

[13] J. S. Dahmann and K. J. Baldwin, "Understanding the Current State of US Defense Systems of Systems and the Implications for Systems Engineering," in *2008 2nd Annual IEEE Systems Conference*, 2008.

[14] J. Boardman and B. Sauser, "System of Systems - The meaning of of," *Proceedings 2006 IEEE/SMC International Conference on System of Systems Engineering*, vol. 2006, no. April, pp. 118–123, 2006.

[15] M. W. Maier, "Architecting principles for systems-of-systems," *Systems Engineering*, vol. 1, no. 4, pp. 267–284, 1998.

[16] J. Axelsson, "A Refined Terminology on System-of-Systems Substructure and Constituent System States," *2019 14th Annual Conference System of Systems Engineering (SoSE)*, pp. 31–36, 2019.

[17] M. Hall-May and T. Kelly, "Using Agent-based Modelling Approaches to Support the Development of Safety Policy for Systems of Systems," *Proceedings of the 25th International Conference on Computer Safety, Reliability and Security (SAFECOMP '06)*, pp. 330–343, 2006.

[18] T. P. Kelly, "Arguing Safety – A Systematic Approach to Managing Safety Cases," Ph.D. dissertation, University of York, 1998.

[19] R. Alexander, D. Kazakov, and T. Kelly, "System of Systems Hazard Analysis Using Simulation and Machine Learning," pp. 1–14, 2006.

[20] P. J. Redmond, J. B. Michael, and P. V. Shebalin, "Interface hazard analysis for system of systems," in *2008 IEEE International Conference on System of Systems Engineering*. IEEE, 6 2008, pp. 1–8. [Online]. Available: http://ieeexplore.ieee.org/document/4724202/

[21] A. K. Saberi, E. Barbier, F. Benders, and M. Van Den Brand, "On functional safety methods: A system of systems approach," in *12th Annual IEEE International Systems Conference, SysCon 2018 - Proceedings*, 2018, pp. 1–6.

[22] J. Axelsson and A. Kobetski, "Towards a risk analysis method for systems-of-systems based on systems thinking," in *2018 Annual IEEE International Systems Conference (SysCon)*. IEEE, 4 2018, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/8369501/

[23] N. G. Leveson and J. P. Thomas, *STPA Handbook*, 2018.

[24] D. Schneider and M. Trapp, "Runtime Safety Models in Open Systems of Systems," *2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, pp. 455–460, 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5380438/

[25] Volvo Construction Equipment, "Electric Site Project." [Online]. Available: https://www.volvoce.com/global/en/news-and-events/news-and-press-releases/2018/carbon-emissions-reduced-by-98-at-volvo-construction-equipment-and-skanskas-electric-site/

[26] International Electronical Commission, "IEC 61882:2001 Hazard and operability studies (HAZOP studies) — Application guide," 2001.

[27] C. Ericson, *Hazard analysis techniques for system safety*. Wileys, 2015.

ARTICLE IN PRESS

Journal of Systems Architecture xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Systems Architecture

journal homepage: www.elsevier.com/locate/sysarc

Towards dynamic safety assurance for Industry 4.0

Muhammad Atif Javed [a,*], Faiz Ul Muram [a], Hans Hansson [a], Sasikumar Punnekkat [a], Henrik Thane [a,b]

[a] School of Innovation, Design and Engineering, Mälardalen University, Sweden
[b] Safety Integrity AB, Sweden

ARTICLE INFO

Keywords:
Dynamic assurance
Safety cases
Industry 4.0
Supply chain
AGVs
Fog and cloud computing

ABSTRACT

The goal of Industry 4.0 is to be faster, more efficient and more customer-centric, by enhancing the automation and digitalisation of production systems. Frequently, the production in Industry 4.0 is categorised as safety-critical, for example, due to the interactions between autonomous machines and hazardous substances that can result in human injury or death, damage to machines, property or the environment. In order to demonstrate the acceptable safety of production operations, safety cases are constructed to provide comprehensive, logical and defensible justification of the safety of a production system for a given application in a predefined operating environment. However, the construction and maintenance of safety cases in alignment with Industry 4.0 are challenging tasks. For their construction, besides the modular, dynamic and reconfigurable nature of Industry 4.0, the architectural levels of the things, fog and cloud computing have to be considered. The safety cases constructed at system design and development phases might be invalidated during production operations, thus necessitating some means for dynamic safety assurance. Moreover, flexible manufacturing in Industry 4.0 also underlines the need for safety assurance in a dynamic manner during the operational phase. Currently published studies are not explicitly supporting the safety assurance of Industry 4.0, which is the focus of this paper with special emphasis on dynamic safety assurance. At first, the Hazard and Operability (HAZOP) and Fault Tree Analysis (FTA) techniques are used for the identification and mitigation/elimination of potential hazards. Next, based on the hazard analysis results, we derived the safety requirements and safety contracts. Subsequently, safety cases are constructed using the OpenCert platform and safety contracts are associated with them to enable necessary changes during runtime. Finally, we use a simulations based approach to identify and resolve the deviations between the system understanding reflected in the safety cases and the current system operation. The dynamic safety assurance is demonstrated using a use case scenario of materials transportation and data flow in the Industry 4.0 context.

1. Introduction

The main enablers for the fourth industrial revolution (aka Industry 4.0) are enhanced automation and digitalisation, which tend to make manufacturing processes faster, more efficient and more customer-centric. However, the enhanced automation and digitalisation, if not carefully orchestrated, could potentially lead to production failures or mishaps, making them safety-critical. Any industry is categorised as safety-critical if an unplanned event or sequence of events can potentially harm humans (injuries or even deaths) or create damages to machines, property or the environment. The principal objective of system safety is the identification, prevention, mitigation, and documentation of potential hazards. It is a fundamental element and a regulatory requirement for safety-critical systems. The hazard elimination or mitigation is much more cost effective during system design and development phase than trying to inject safety after the occurrence of an accident or mishap [1].

To perform the safety analysis for Industry 4.0, besides the modular, dynamic and reconfigurable nature, the things, fog/edge and cloud levels need to be considered. Transformation to the Industry 4.0 significantly increases the safety assurance challenges [2]. To demonstrate the acceptable safety of Industry 4.0, the assurance (safety) cases are constructed that provide comprehensive, logical and defensible justification of the safety of a production system for a given application in a predefined operating environment [3]. In the context of Industry 4.0, however, the parts of safety cases constructed during system design and

* Corresponding author.
E-mail addresses: muhammad.atif.javed@mdh.se (M.A. Javed), faiz.ul.muram@mdh.se (F.U. Muram), hans.hansson@mdh.se (H. Hansson), sasikumar.punnekkat@mdh.se (S. Punnekkat), henrik.thane@safetyintegrity.se (H. Thane).

https://doi.org/10.1016/j.sysarc.2020.101914
Received 14 January 2020; Received in revised form 7 July 2020; Accepted 13 October 2020
Available online 15 October 2020

development phase may turn out to be incorrect, inapplicable or insufficient during the operation. This can be caused by emergent behaviours and changes performed in consequence of market demands, hazardous conditions or system failures. To exploit the full potential of Industry 4.0, there is a need to support the dynamic safety assurance [4,5].

There exist few studies on dynamic safety assurance. The research in [4] presents the idea of through-life safety assurance. It is reflected in four activities: identify, monitor, analyse and respond. Jaradat and Punnekkat [6] exploit safety contracts for monitoring the failure rates during operational phase. Calinescu et al. [7] develop the partial assurance argument at design phase to which the assurance evidence can be synthesised during self-adaptation. McDermid et al. [5] suggest the continued and proactive assessment of autonomous systems to address the mismatches between the 'world as imagined' system and its operational environment we considered and the 'world as observed' analysis based on the operational data. Jaradat et al. [2] highlight the safety assurance challenges for Industry 4.0. The published studies are not explicitly supporting the dynamic safety assurance by using the operational data.

This paper focuses on the dynamic safety assurance of Industry 4.0, where our specific contributions are: (a) proposing an overall approach incorporating Industry 4.0 specific aspects (b) elaborating on the details of individual steps including suggestions for an overall tool chain and (c) demonstrating the approach using a synthetic, but realistic use case scenario. We now provide an overview of the steps in our approach. First, we performed the hazard analysis, for which the Hazard and Operability (HAZOP) and Fault Tree Analysis (FTA) techniques are utilised for the identification and mitigation of potential hazards. Second, based on the hazard analysis results, we derived the safety requirements to prevent or mitigate the identified hazards. Safety contracts have also been derived for uncertainty sources that provide the means to detect deviations from intended behaviour and perform necessary adaptations at the operational phase. Third, we constructed the safety cases and associated the safety contracts with them. For this reason, the nature of Industry 4.0 and its generic architecture are explicitly taken into consideration. Finally, the operational data is utilised to resolve the deviations between the intended behaviour reflected in safety cases and the actual behaviour. In particular, based on the gathered data, the safety contracts constructed for uncertainty sources are monitored, deviations between the intended and actual behaviour are tracked and evaluated, and the safety contracts and safety cases are updated. The work presented in this paper utilises the OpenCert[1] platform that provides support for modelling and visualising the safety cases in Goal Structuring Notation (GSN) [3]. The usefulness of dynamic safety assurance is demonstrated for the flow of materials and data in an Industry 4.0 oriented use case scenario.

The rest of this paper is organised as follows: Section 2 describes essential background information on Industry 4.0, safety assurance and safety contracts. Section 3 presents the proposed methodology for dynamic safety assurance of Industry 4.0. Section 4 demonstrates the applicability of proposed methodology for materials transportation and data flow in Industry 4.0. Section 5 discusses the related work. Section 6 concludes the paper and presents future research directions.

2. Background

2.1. Industry 4.0

The association of production with the information and communication technologies lead to the Industry 4.0. It is a successor of three industrial revolutions. The first revolution was triggered by the steam and hydroelectric power generation; the machines were developed for the manufacturing of products. The second revolution was driven by the electrification, which paved the way towards the mass production and assembly lines. The third revolution focused on the electronics and information technology; the programmed machines such as robots and Automated Guided Vehicles (AGVs) improved the automation. The fourth revolution is based on the four integral pillars: interoperability, information transparency, technical assistance and decentralised decisions [8]. Interoperability is the capability to exchange data between machines that enables joint collaboration; Internet-of-Things (IoT) is exploited for interconnecting machines, the ultimate intention is whole automation. A novel enabler of information transparency could be through the concept of digital twins [9]: leverage the virtual equivalents such as simulation models for assessing and improving the real products. Technical assistance provides the means to understand and control events that needs to be supported by systems. In this context, the comprehensive visualisations facilitate human operators to solve the problems in short time frames [10]. The intention of decentralised decisions is to incorporate automatic decision making capabilities in machines so that the human intervention in manufacturing process is not required, which makes the operator rather a supervisor.

The manufacturing process is re-defined for Industry 4.0. Its generic architecture consists of three levels: things, fog/edge and cloud. The things and fog/edge represent the local part of the system, whereas the cloud is a remote infrastructure that is typically owned by a third-party service provider. The things interact with the physical environment via different sensing/actuating devices. However, due to their limited storage and processing power, devices from things level rely on the fog or cloud for storage and processing services. The fog acts as a bridge between the things and the cloud. It receives data from the things and perform partial processing, especially for critical functions, or otherwise forward the data to the cloud infrastructure. Likewise, the fog directly instruct to the things, but the cloud forwards commands to the things via fog servers.

2.2. Safety assurance

An assurance (safety) case provides the evidence(s) to support the safety claims such as system behaviour and hazards addressed, and shows why the included evidence is adequate. In particular, the safety cases facilitate the knowledge exchange between different stakeholders, for example, suppliers and acquirers, and between operator and regulator [11]. To document safety cases, several approaches exist, such as free text, tabular structures and graphical notations [12]. Structured Assurance Case Metamodel (SACM) [11] is the Object Management Group (OMG) standard that integrates and standardises the broadly used notations for documenting safety cases including GSN. The work presented in this paper utilises the OpenCert platform that provides support for modelling and visualising the safety cases in GSN. Common Assurance and Certification Metamodel (CACM) implemented in OpenCert internally uses the SACM metamodel. The OpenCert platform also includes a Connected Data Objects (CDO)[2] server, which is a development-time model repository as well as a run-time persistence framework. It stores the safety case (i.e., model and diagram) in a database such that different distributed stakeholders can access them and work on the same safety case concurrently.

The main purpose of GSN is to show how goals (claims about the system behaviour, safety requirement or process plan), are broken down into sub-goals (sub-claims) and supported by available solutions (evidences). The rationale for decomposing the goals into sub-goals is documented by strategy (argument reasoning), whereas the scope and domain in which claim or strategy should be interpreted is done in the context element. Assumption element is an intentionally unsubstantiated statement, which is assumed to be true for a certain goal or strategy. Justification element provides an explanation or rationale

[1] https://www.polarsys.org/opencert/.

[2] http://www.eclipse.org/cdo/.

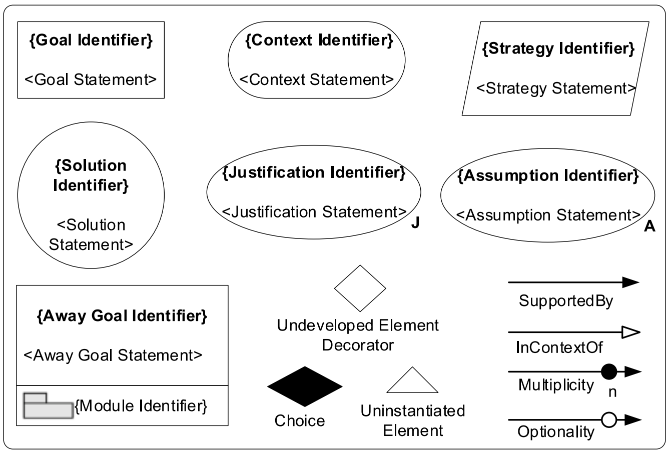


Fig. 1. GSN notations.

why a certain goal or strategy is considered appropriate or acceptable to be used [3]. The argument elements can be linked with one of the two relationships: SupportedBy and InContextOf. The SupportedBy relationship is used to show the inferential or evidential relationships between goals, strategies and solutions elements, whereas InContextOf relationship is used to declare contextual relationships of goals or strategy with context, assumption and justification elements. An away goal referenced to a claim presented in another argument module.

The principal elements of the GSN are shown in Fig. 1, for further details we refer the readers to GSN standard [3]. In order to represent patterns of argument, GSN has been extended to support structural abstraction aspects. Multiplicity relationship shows the n-ary (zero or more) relationships between elements, whereas optionality relationship denotes the alternative (zero or one) relationships between elements. Undeveloped and uninstantiated decorators are applied to the elements that need to be further developed and to be replaced (instantiated) with a more concrete instance, respectively. A choice element is used to denote possible alternatives (1-of-n and m-of-n selection) in satisfying a relationship.

2.3. Safety contracts

The contracts can be used almost everywhere and all phases of system life cycle, for instance, from early requirements to detailed design, for managing the changes in the system or its corresponding assurance case. Contracts were proposed by Bertrand Meyer as the concept of 'design-by-contract' for verification of software program [13]. In particular, the correctness of requirements is described as a contract between a method and its callers. The contracts are made up of two parts: (1) the preconditions that must be true while the initiation of program; and (2) the postconditions that must be true on program completion. A contract can be seen as a pair of assertions C = $\langle A, G\rangle$ i.e. {Assumptions, Guarantees} that describes in a formal way, under which context the design is assumed to operate, and what are its obligations [14]. By using the contracts, the behaviour of the component can be described in a way that the component makes assumptions (conditions) on its environment and if those assumptions A hold then the component will behave as guaranteed G (offers properties). The safety contracts can be associated with the safety case for supporting the dynamic assurance. They provide a mechanism for recording and justifying the agreed relationship between safety case modules that might be illustrated as a pattern in GSN.

3. Dynamic safety assurance of industry 4.0

The intention of the safety assurance of Industry 4.0 is to provide confidence that the risk of potential mishaps and accidents is eliminated or otherwise mitigated to an acceptable level. For this reason, the assurance (safety) case is constructed, in particular, a logical and appealing argument supported by a body of evidence, to justify that an Industry 4.0 oriented application will operate as intended in a defined environment. In the assurance case of Industry 4.0, besides the modular, dynamic and reconfigurable nature, the things, fog and cloud computing levels need to be reflected. Due to the enhanced automation and digitalisation of manufacturing processes, the safety analysis conducted during design and development phase may not be sufficient; consequently, we resolve the deviations from specified behaviours during the operational phase.

Let us consider the autonomous machines operating in Industry 4.0, such as AGVs and robots that are perceived as inter-connected physical objects or things equipped with sensors, actuators and controllers; they are connected via the internet for reasons of data collection and exchange [15]. From the safety perspective, there is a need to determine the potential environments in which they operate, or exposed to, and the movement control measures in certain conditions and areas. As the things have limited power, the cloud and fog services are utilised. The former provides on-demand, broad network access to a shared collection of configurable resources, including storage and processing [16]. The latter extends and brings the cloud platform closer to the things for the resolution of latency, bandwidth and location or context awareness problems, especially for real-time safety-critical functions in Industry 4.0 [17].

An overview of the proposed approach is presented in Fig. 2. The hazard analysis is performed as a first step towards safe operations in Industry 4.0 (see Section 3.1). After that, the safety requirements and safety contracts are derived from the analyses results (see Section 3.2). Next, the construction of safety cases and contracts association is carried out; the safety claims are defined based on the available artefacts, for example, the safety requirements (see Section 3.3). Subsequently, the simulation environments are configured and the operational data is acquired for supporting the data driven production and operations management (see Section 3.4). Finally, to resolve the gaps between the intended and actual behaviour, the safety contracts and safety cases modelled in the OpenCert platform are updated (see Section 3.5). The proposed approach with focus on end-to-end traceability and support of a tool framework can provide a significant boost for the designers to avoid the culture of paper safety at the expense of actual system safety [18].

3.1. Hazard analysis

The factories are often classified as safety-critical, for example, due to simultaneous presence of AGVs, other machines and human workers at the site. The enhanced automation and digitalisation in Industry 4.0 significantly increases the safety issues. To identify hazards, their effects, and causal factors, the hazard analysis is performed. It is the key to prevent or mitigate mishaps. In the context of Industry 4.0, a failure may not just lead to a hazard and accident of a thing/AGV itself. But it can propagate to other things, which lead to a mishap. This is because the things have interactions with one another, and to the fog controller that, in-turn, interact with the cloud infrastructure. In this paper, the hazard analysis is performed by using HAZOP and FTA techniques. The HAZOP analysis is performed to identify possible deviations, operational concerns and their possible fault root causes and consequences. To perform the HAZOP analysis, a set of guide words (e.g., early/late, slower/faster, part of, other than, reverse, omission, before/after, etc.), parameters (e.g., position, distance, detect, speed, etc.), system inputs and outputs, a list of messages and their flow are determined. For in-depth analysis, the fault trees are constructed

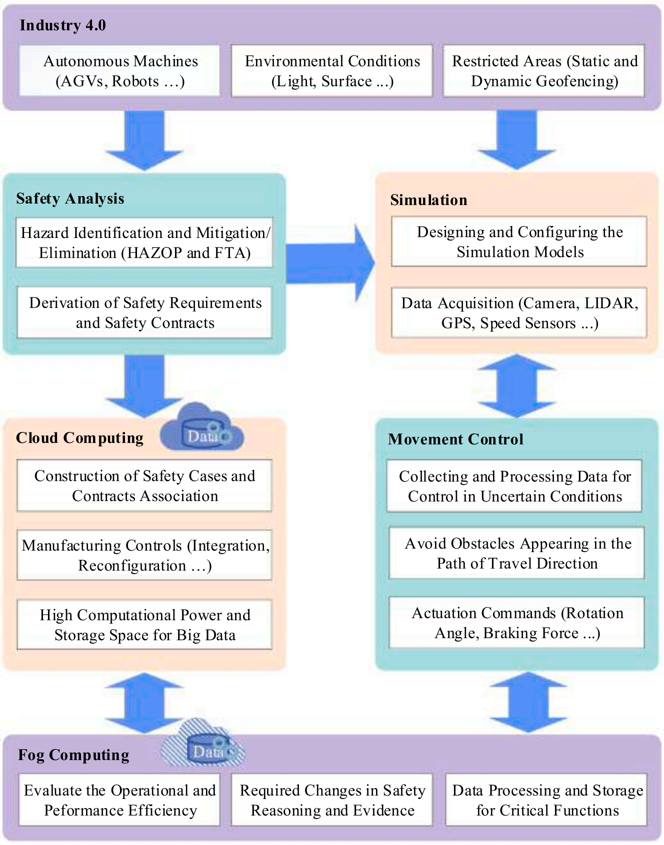

Industry 4.0
Autonomous Machines (AGVs, Robots ...)
Environmental Conditions (Light, Surface ...)
Restricted Areas (Static and Dynamic Geofencing)
Safety Analysis
Hazard Identification and Mitigation/ Elimination (HAZOP and FTA)
Derivation of Safety Requirements and Safety Contracts
Simulation
Designing and Configuring the Simulation Models
Data Acquisition (Camera, LIDAR, GPS, Speed Sensors ...)
Cloud Computing
Data
Construction of Safety Cases and Contracts Association
Manufacturing Controls (Integration, Reconfiguration ...)
High Computational Power and Storage Space for Big Data
Movement Control
Collecting and Processing Data for Control in Uncertain Conditions
Avoid Obstacles Appearing in the Path of Travel Direction
Actuation Commands (Rotation Angle, Braking Force ...)
Fog Computing
Data
Evaluate the Operational and Peformance Efficiency
Required Changes in Safety Reasoning and Evidence
Data Processing and Storage for Critical Functions

Fig. 2. An overview of the proposed approach.

based on the hazards and their potential effects understood from the HAZOP analysis, in which the identified hazards served as the top-level undesired events [19]. After the identification of hazards, their risks are assessed, and mitigation methods are established that are eventually translated into safety requirements to eliminate or mitigate them.

3.2. Derivation of safety requirements and safety contracts

Safety requirements intend to eliminate or mitigate the identified hazards. Stringent safety requirements are either derived from the hazard analysis or extracted from the safety standards. For example, 'the AGV shall detect and avoid both static and dynamic obstacles appearing in the path of travel direction' is derived from the driverless, automated guided industrial vehicles ANSI/ITSDF B56.5 [20] standard. Industry 4.0 comprises of constituent things and infrastructure elements that are manufactured and provided by different organisations. The manufacturers of the constituent systems, such as cameras and LIDARs (Light Detection and Ranging) are supposed to provide the certification that their products are in compliance with relevant standards (e.g., IEC 61508 [21], ISO 13849-1 [22], and IEC 61496-1 [23]) for the application at hand. If the product is integrated according to the provided instructions, then we can state with confidence that the system meets the safety requirements. The derived safety requirements are implemented in digital twin, and are expected to mitigate the hazards and reduce the risk to an acceptable level [24].

The results gained from the hazard analysis are also utilised to derive the safety contracts that can be further supported by evidence to construct the assurance (safety) cases. The safety contracts demonstrate that the utilised properties of things or infrastructure are sufficient to satisfy the safety requirements and identified failure behaviours can be reduced by a variety of mechanisms, such as testing, monitoring and operational constraints etc. Since different stakeholders can be involved in Industry 4.0, the information about the assurance of individual things or infrastructure is required in order to derive the safety contracts. However, to be able to provide the assurance, each safety requirement needs to be satisfied by at least one safety contract, and each contract can be supported by one or more evidence(s). The objective of derivation of safety contracts is to associate these contracts with the safety cases as reference points to support the dynamic safety assurance of Industry 4.0. In particular, safety contracts help in determining the validity of safety case modules and overall safety case by comparing safety requirement with related operational data when things or infrastructure are altered or substituted in the system, for instance, in response to changes in market, environmental conditions and device failures. If no deviations are found during monitoring and comparing of data then the assumptions are satisfied in a specific context, and thus, its guarantees also hold. The safety cases and safety contracts also need to be updated with regard to uncertainty aspects.

3.3. Construction of safety cases and association of contracts

The underlying idea of the assurance of Industry 4.0 is to construct the safety cases; firstly, the safety claims are defined, for example, all identified hazards have been eliminated or sufficiently mitigated. Subsequently, the assurance evidence to support the claims is systematically identified, selected, and linked, such that risk is reduced to an acceptable level. In a similar way to the safety-critical applications, to create the safety cases for Industry 4.0, the root claim is broken down into sub-claims by using argument reasoning. However, the modular, dynamic and reconfigurable nature of Industry 4.0, and its architectural levels of things, fog and cloud are specifically considered. The procedure of breaking down claims to sub-claims is continued until we reached claims that can be connected to solutions for which either a direct evidence is specified or an away goal is created by pointing to an already developed module like LIDAR system. The evidence can be obtained during design and development phase from non-formal sources, such as hazard analysis and testing; or from formal verification, such as the results of model checking [25]. In case of Industry 4.0, there is a need to collect the assurance evidence during the operational phase. For instance, the evidence concerning the undeveloped claims requiring the fog processing time could be obtained. The safety case provides the assurance of potential failure behaviour of things or infrastructure, the associated safety contracts help to maintain and update the safety case based on the incorrect or implicit assumptions and deviations from specified behaviours. The safety case (argumentation model and diagram) are constructed in the assurance editor of OpenCert platform, the safety contracts in the form of assumptions and guarantees are specified in an argument element property field.

3.4. Data driven production and operations management

IoT data will be generated and renewed in the Industry 4.0 at a very high speed. Multiple types of sensors such as LIDAR, Camera, Global Positioning System (GPS) are used for supporting the data driven production and operations management. They are either placed on top of things (such as robots and AGVs) or mounted on appropriate locations in the environment (such as on the walls, pillars), for covering specific areas. Let us consider the scenario of transportation and distribution of materials for which AGVs are used. To establish a basis for evaluation and improvement of performed operations, an appropriate simulation environments needs to be configured. In our demonstrations we used Carla [26], AirSim [27], Webots [28] and Oryx.[3] The mounted cameras provide support for detecting and recognising obstacles appearing in the path of travel direction; message signs and lane markings can

[3] https://www.oryx.se/.

also be considered. However, the LIDAR provides better support for distance measurements and dealing with environmental conditions. Besides the cameras and LIDAR, the GPS would be fitted for tracking and navigation purposes. They transmit data at a rate of 20–60 MB/s, 10–70 MB/s and 50 KB/s, respectively.

In the Industry 4.0, the production operations, for instance, automated loading is carried out in a collaborative manner. This means that the AGV operates in conjunction with other AGVs and machines. The site management is perceived as a fog controller. It gathers the data from different machines and is responsible for managing operations including the mission termination or assignment to the AGVs, and traffic control to avoid collisions. The communication between the AGVs and fog controller also supports entry control, i.e., to permit or restrict entry to geofenced areas. The data produced by sensors placed on specific things and areas is processed by the fog controller. A detailed list of parameters is established to support the data driven production and operations management. This step is fundamental to evaluate the operational and performance efficiency, and to deal with the deviations between the intended and actual behaviours at the operational phase.

The modelling of safety cases and association of contracts is carried out in the OpenCert platform. The safety cases can be stored in the workspace directory, or in the CDO, which is both a development-time model repository and a run-time persistence framework. The repository sessions provide support for obtaining and modifying them. The features provided by CDO satisfy the design- and run-time requirements of the self-adaptive cloud applications [29]. In this sense, the fog controller is not just connected to things and remote resources, but also serve as a bridge between them. The work presented in Seybold et al. [29] discuss the experiences with EMF and CDO in scalable cloud scenarios. The in-browser Eclipse IDE for cloud (Eclipse Che[4]) may also be used. In addition to the safety cases, the cloud level is considered for some manufacturing controls and high computational power and storage space for big data except for real-time safety-critical functions in Industry 4.0.

3.5. Identification and resolution of gaps in safety cases

There is a need for the identification and resolution of gaps between the intended behaviour reflected in safety cases and the actual safety of production operations in Industry 4.0. For this reason, the operational data is utilised. In particular, the safety contracts constructed for uncertainty sources are monitored based on the gathered data. As mentioned in Section 3.3, the safety contracts describe the functional and behavioural properties in the form of assumptions and guarantees. The safety case modelled in the OpenCert platform is traversed to retrieve the elements IDs and associated safety contracts; this information is stored in the data structures. In the safety cases, the contracts are associated in the argument element property field, named as content. To check the validity, the safety contracts are compared against the list of parameter values obtained during the operational phase that includes the distance to the obstacles, location, path points, weight, speed limits, and time frame for transmission, data processing and actuation of given commands. This provide the means to cope with the completeness, consistency and correctness problems.

Based on the comparative analysis of operational data with the assumptions made and predicted guarantees, the evidences are obtained, which are attached to the undeveloped claims, located with their element IDs. Otherwise, the deviations between the intended and actual behaviour are tracked. Let us consider the processing and actuation times, if they are increased, the safe distances to avoid the obstacles is increased, or the speed is reduced in alignment with the determined threshold. The positive impacts of advanced technologies are also considered; they could not only increase the detection capabilities, but also decrease the transmission and processing time frame. In this case, we explicitly evaluate the corresponding impacts on the control actions. The confidence, in certain situations, is gradually increased to resolve the discrepancies. The update of safety contracts and safety cases is carried out based on the optimal actions. The required changes in safety cases are tracked on the fog level and then the update command is issued. The safety contracts and safety cases are updated on the CDO server, which is accessed by the OpenCert argumentation editor connected to the CDO server.

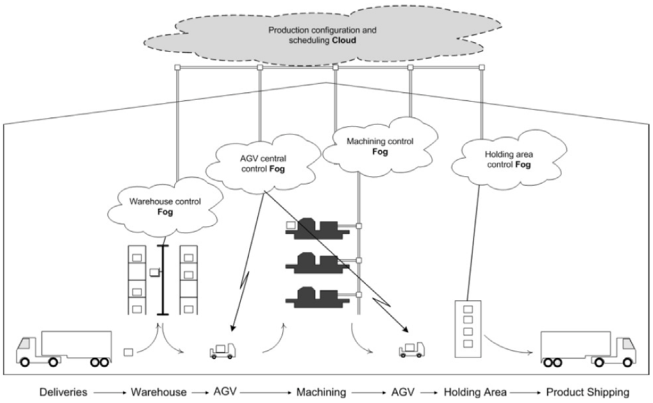


Fig. 3. Materials transportation and data flow in industry 4.0.

4. Case study

To transform the raw materials into the final products, the evolution of traditional supply chain is considered critical for industry 4.0 [30]. For the enhancement of automation and digitisation of supply chain, the alignment with interoperability, information transparency, technical assistance and decentralised decisions is suggested [31]. Let us consider a synthetic but realistic and generic use case that concerns transportation and data flow in the Industry 4.0. The use case chosen in the FiC project[5] and AAIP [32] consists of a number of computer-based machine tools that interacts with each other for product manufacturing. There are few manufacturing controls that are located remotely on the cloud, whereas the others are managed locally at the fog or things levels. The computer-based machine tools make a range of gearbox shafts from metal blanks. The weight of a blank is about 4 kg. They are delivered in pallets of 50 and stored in an automated warehouse until they are required. The finished products are also packed into pallets and taken to a holding area before being shipped to the main assembly plant. The transportation and distribution of pallets are performed with battery-powered AGVs. Fig. 3 shows the conceptual flow of materials and data in the Industry 4.0.

The AGVs are widely used in industries. Their operations are similar to the autonomous haulers that are used for rocks transportation in the advanced quarry site [19]. The AGVs shall be equipped with device(s) to detect and avoid obstacles appearing in the path of travel direction [20]. To carry out the transportation and distribution of materials efficiently, the next generation of AGVs, especially for Industry 4.0 could be equipped with autonomous driving technologies. The cameras and LIDARs provide means for emergent interactions with machines, operating speed increase, and safety of AGVs; they can be mounted on top of the AGVs or otherwise infrastructure level for the detection and recognition of obstacles appearing in the path of travel direction. Since the LIDARs are relatively expensive, if they are mounted on the infrastructure level, then the limited number of LIDARs can fulfil the

[4] https://www.eclipse.org/che/.

[5] http://www.es.mdh.se/fic/.

needs. This is also relevant in situations of failure of obstacle detection devices mounted on top of the things, for instance, the transition is made to other devices that are placed in hazardous zones like loading and unloading. Accordingly, besides the command/control functions, the ways for transmitting, processing and storage of big data are mapped to the generic architecture of Industry 4.0.

4.1. Hazard analysis

The performed analysis not just focuses on the individual behaviour of AGV, but also the emergent interactions of AGVs, with fog/cloud server or other working equipment. The identified hazards can be categorised into: *sub-system failure hazards* occurred due to unavailability, unreliability or incorrectly structured software/hardware component, such as AGV onboard system failure and brake system failure; *deviations from procedures*, for example, incorrect application of LIDAR or inadequate installation of software; *communication failure*, for instance, a message is received too late, an incorrect message is transferred, wrongly interpreted by the receiver, or network unavailability; *operational hazards* such as failure of obstacle detection and collision avoidance mechanism; and *environmental hazards* which involve extreme heat or cold, light, and driving on rigid surfaces. Once hazards are identified and classified, the corrective measures are drawn to eliminate or reduce the risk of the identified hazards to an acceptable level.

The results obtained from the HAZOP and FTA demonstrate that critical incidents can occur in Industry 4.0, if correctly and timely communication is not established. For instance, AGV received a message too late than expected from fog controller concerning position and distance to obstacle. The reception of late message can be caused by AGV fog failure, downtime, or network unavailability, which may lead to the incomplete mission, a collision of AGV with dynamic objects (e.g., other AGVs, human worker, or other working equipment), or with static objects (e.g., ladder, fallen pallet). These collisions can be catastrophic, especially if the AGV carries dangerous materials (e.g., explosive, toxic, etc.). This can be prevented by increasing number of wireless access point or getting the information from another fog. Moreover, the failure of obstacle detection and collision avoidance mechanism (caused by LIDAR or camera failure) may also lead to a collision of AGV. This can be prevented by getting the information from the sensors mounted on the infrastructure level. In this paper, we focus on the hazard like collision of AGV with static objects and dynamic obstacles. It is assumed that if collisions happen while AGV is stopped, then it must be the moving obstacle's fault.

4.2. Derivation of safety requirements

The risk reduction and mitigation measures identified through the hazard analysis are translated into the safety requirements. Additionally, safety requirements are extracted from the automated guided industrial vehicles ANSI/ITSDF B56.5 [20] standard. To address industry-level hazards different safety requirements specified, such as "The AGV shall not collide with obstacles appearing in the path of travel direction" is considered as a main requirement. To achieve this requirement, several other requirements have been specified. Some of them are applicable at the system level, while others are relevant for the software and hardware functions.

SR 1: The AGV shall detect static and dynamic obstacles appearing in the path of travel direction.

> **SR 1.1:** The obstacle and position detection devices shall be sufficiently robust and works in practice.
> **SR 1.2:** The obstacle detection device mounted on AGV shall cover a minimum range of 10 m.
> **SR 1.3:** The obstacle detection device shall detect indoor obstacles with the worst accuracy of ±6cm.
> **SR 1.4:** The GPS fitted on the AGV shall estimate position with the worst accuracy of ±30cm.

SR 2: The retrieved data shall be processed to determine risks and preventive measures within maximum 800 ms.

> **SR 2.1:** The transmission of data retrieved from object detection and positioning devices to the Fog controller shall take maximum 50 ms.
> **SR 2.2:** The Fog controller parse the data to detect moving direction, distance and current speed of the obstacles that shall take maximum 500 ms.
> **SR 2.3:** The Fog controller shall send the parsed data to the AGV within maximum 50 ms.
> **SR 2.4:** The AGV's onboard system shall determine protection measures within maximum 200 ms.

SR 3: The AGV shall avoid obstacles for which deceleration and steering rotation are commanded.

> **SR 3.1:** The actuation time for deceleration and steering rotation shall not be greater than 100 ms.
> **SR 3.2:** If static obstacle is detected in travel path, the maximum speed limit 1.2m/s shall be maintained in hazard zone until the remaining distance of 2 m.
> **SR 3.3:** If dynamic obstacle is detected in travel path or junction, the maximum speed limit 0.3m/s shall be maintained in critical zone until the remaining distance of 2 m.
> **SR 3.4:** The safe distance to the obstacles after stopping shall not be shorter than 1 m.
> **SR 3.5:** If obstacle is detected, the AGV outbreak from the original path to avoid the obstacle for which the smallest deviation from the original path should be chosen as side preference.

4.3. Derivation of safety contracts

After identifying the safety requirements, safety contracts are documented for uncertainty sources. Let us assume that the failure of LIDAR or speed sensor, or transmission delay can be hazardous, then the corresponding safety requirement is addressed by a contract assumption, whereas guarantee shows that the failure is appropriately handled by making the transmission to another detection device, to map detecting distance covered within timeframe and speed is always limited to ≤ 0.3m/s. Table 1 shows a template of derived safety contracts, where *A* and *G* stand for Assumption and Guarantee, respectively.

4.4. Development of safety cases using contracts

Once safety requirements and safety contracts are derived for Industry 4.0, the safety case is created, for which the safety requirement "The AGV shall not collide with obstacles appearing in the path of travel direction" satisfied to SIL 2, is selected as a root claim, as shown in Fig. 4. This claim is decomposed into two sub-claims using a strategy "Arguments over element contracts". In particular, the claim about completeness and correctness of hazards identification to maintain safe distance of AGV from other AGVs, working equipment and human workers, and the claim about the detection and avoidance of static and dynamic obstacles are created in the context of things and infrastructure elements. This procedure is continued until there are elementary claims that can be connected to the available evidence (i.e., InformationElementCitation Property type = "solution"). An away goal (i.e., ArgumentElementCitation Property type = "claim") is created, for the claim "The transmission of data between fog controller and AGV shall take maximum 50 ms", and "Argumentation Reference" property indicates the reference to the module in which it is developed (i.e., Infrastructure module). The supplier of fog services has provided

Table 1
Template of derived safety contracts.

A1: G1:	For the operating AGV, no obstacle detection device failure AND no mechanical parts failure AND no transmission delay AND no obstacles detected within 10 m range of travel path. The higher speed of AGV is maintained.
A2: G2:	For the operating AGV, obstacle detected in 10 m range of travel path AND the obstacle is not moving (i.e., static obstacle is detected). Until the 2 m distance, the speed is always ≤ 1.2 m/s and the smallest deviation with side clearance ≥ 0.5 m is chosen to circumvent the obstacle.
A3: G3:	For the operating AGV, obstacle detected in 10 m range of travel path AND the obstacle is moving (i.e., dynamic obstacle is detected). Until the 3 m distance to the dynamic obstacle, the speed is always ≤ 0.3 m/s and the AGV stops before the remaining distance 2 m.
A4: G4:	The LIDAR is mounted on top of AGV or infrastructure as per the manufacturer recommendations. The LIDAR shall detect indoor obstacles within the range of 10 m with the accuracy of ±6 cm.
A5: G5:	For the operating AGV, LIDAR sensor failure OR wheel speed sensor failure OR transmission delay. The transition is made to another detection device, to map for detecting distance covered in timeframe and the speed is always ≤ 0.3 m/s.
A6: G6:	AGV Fog controller is independent AND accurate AND distinct input data is received within 50 ms timeframe. AGV Fog controller accurately computes the received inputs and provides data in 500 ms time frame.

the assurance evidences for it. The claims, such as "The fog controller parse the data to detect moving direction, distance and current speed of the obstacles that shall take maximum 500ms" are created as undeveloped claim by selecting the property toBeSupported = "true" means that it requires further development. The Assumption and the Guarantee properties for the contract are specified in the Content field of related claims in assurance case editor. For the undeveloped claim, the contract allows that its guarantees, can be supported by artefacts (e.g. the latter referring some verification results, simulation runs).

The claim about "The AGV provides the required properties for handling collisions with static and dynamic obstacles" is decomposed into sub-claims about providing sufficient properties or failures behaviour of mechanical parts, position and obstacles detection devices, and AGV on board systems using three strategies. The mechanical parts of AGV are assumed to be well maintained, and have not any failure, which may not always be the case. Therefore, the claims about mechanical hazards identification and mitigation can be supported by hazard analysis results (e.g., mitigation measures taken to eliminate or prevent the failures), inspection and/or testing reports. The suppliers of camera, LIDAR and GPS are able to provide assurance for their products that support various claims about camera, LIDAR and GPS as detailed in the safety contract. Therefore, we used away goals for them such that these claims are developed in different modules. As we can see in Fig. 4, it is assumed that the LIDAR mounted on AGV will provide point clouds in 10 m range with the worst accuracy of ±6 cm. The confidence in this claim is provided by the LIDAR safety case.

4.5. Transportation and data flow in industry 4.0

In the traditional AGVs, besides the infrared sensors that are mounted to detect obstacles, the camera is placed to give a clue for the lines drawn on the industry floor. This behaviour is analysed with the e-puck robot, which is configured in the Webots simulator. The Radio Frequency Identification (RFID) tags are also used to locate the AGVs. The data produced by eight infrared sensors, camera with limited resolution (52 × 39 pixels), or RFID tags can be easily handled by the AGVs that are perceived as things. In contrast to this, a huge volume of data could be transmitted to the fog level; it acts as a bridge between things and cloud computing levels. The different formats of data from diverse devices for instance high-resolution cameras and LIDAR is gathered from configured simulators. The acquisition and integration of data can be carried out at things (AGVs) level. The fog level distinguishes the real-time safety-critical data for processing and storage itself. Otherwise, it further transmits the data to the cloud. In this sense, the fog improves the computation and communication efficiency by reducing the quantity of data to be transmitted to the cloud. On the other hand, the fog also forwards necessary commands to thing to fulfil the specific production demands. Since the considerable data, including safety cases is stored in the cloud, of course, the ways for retrieval and update are important.

4.6. Dynamic update of safety contracts and safety cases

In addition to the risk control actions, to avoid the culture of paper safety, the update of safety contract and safety case is performed. The safety contracts are associated with the safety case elements, which are retrieved and monitored. The gaps with the contracts, i.e., the assumptions made and guarantees provided indicate that the safety contract is broken. As discussed in Section 3.5, the deviations or mismatches are tracked by comparing the safety contracts with their respective parameters that are gathered from the simulations. This provides a basis for the adaptation of affected parts of safety cases and associated contracts. If the failures, delays and obstacles are not detected, i.e., the assumption "for the operating AGV, no obstacle detection device failure AND no mechanical parts failure AND no transmission delay AND no obstacles detected within 10 m range of travel path" is satisfied. Then the "higher speed of AGV is maintained". The speed limit 14 km/h is reached. In case the assumption "obstacle detected in 10 m range of travel path AND the obstacle is not moving" is satisfied, the predicted guarantee hold "until the 2 m distance, the speed is always ≤ 1.2m/s and the smallest deviation with side clearance ≥ 0.5 m is chosen to circumvent the obstacle". In this way, simulation result related to the satisfied safety contracts is assembled as an evidence to the undeveloped claim "propsAvoidSat".

The assumption A5 "for the operating AGV, LIDAR sensor failure OR wheel speed sensor failure OR transmission delay" is checked. If the

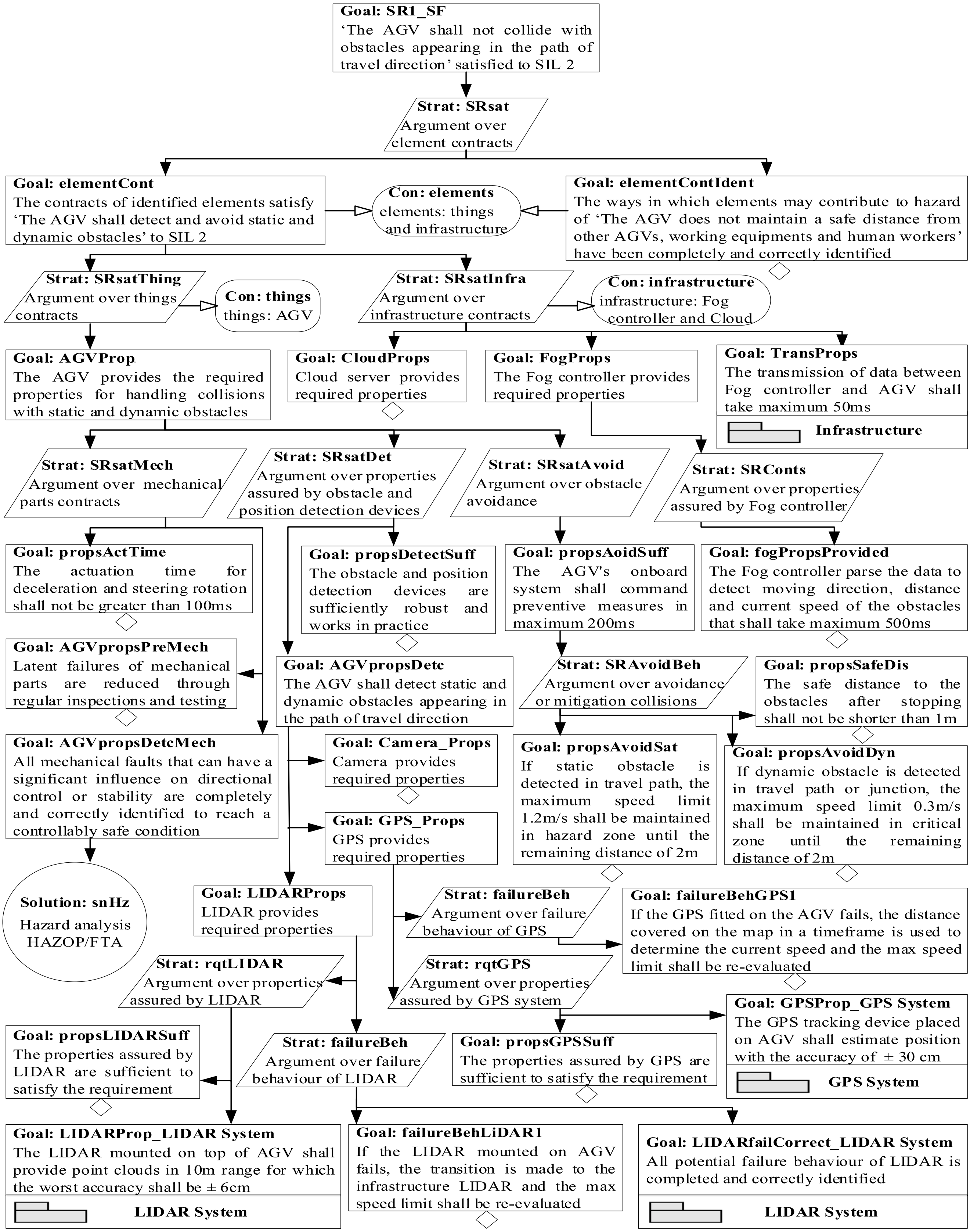

Goal: SR1_SF
'The AGV shall not collide with obstacles appearing in the path of travel direction' satisfied to SIL 2
Strat: SRsat
Argument over element contracts
Goal: elementCont
The contracts of identified elements satisfy 'The AGV shall detect and avoid static and dynamic obstacles' to SIL 2
Con: elements
elements: things and infrastructure
Goal: elementContIdent
The ways in which elements may contribute to hazard of 'The AGV does not maintain a safe distance from other AGVs, working equipments and human workers' have been completely and correctly identified
Strat: SRsatThing
Argument over things contracts
Con: things
things: AGV
Strat: SRsatInfra
Argument over infrastructure contracts
Con: infrastructure
infrastructure: Fog controller and Cloud
Goal: AGVProp
The AGV provides the required properties for handling collisions with static and dynamic obstacles
Goal: CloudProps
Cloud server provides required properties
Goal: FogProps
The Fog controller provides required properties
Goal: TransProps
The transmission of data between Fog controller and AGV shall take maximum 50ms
Infrastructure
Strat: SRsatMech
Argument over mechanical parts contracts
Strat: SRsatDet
Argument over properties assured by obstacle and position detection devices
Strat: SRsatAvoid
Argument over obstacle avoidance
Strat: SRConts
Argument over properties assured by Fog controller
Goal: propsActTime
The actuation time for deceleration and steering rotation shall not be greater than 100ms
Goal: propsDetectSuff
The obstacle and position detection devices are sufficiently robust and works in practice
Goal: propsAoidSuff
The AGV's onboard system shall command preventive measures in maximum 200ms
Goal: fogPropsProvided
The Fog controller parse the data to detect moving direction, distance and current speed of the obstacles that shall take maximum 500ms
Goal: AGVpropsPreMech
Latent failures of mechanical parts are reduced through regular inspections and testing
Goal: AGVpropsDetc
The AGV shall detect static and dynamic obstacles appearing in the path of travel direction
Strat: SRAvoidBeh
Argument over avoidance or mitigation collisions
Goal: propsSafeDis
The safe distance to the obstacles after stopping shall not be shorter than 1m
Goal: AGVpropsDetcMech
All mechanical faults that can have a significant influence on directional control or stability are completely and correctly identified to reach a controllably safe condition
Goal: Camera_Props
Camera provides required properties
Goal: propsAvoidSat
If static obstacle is detected in travel path, the maximum speed limit 1.2m/s shall be maintained in hazard zone until the remaining distance of 2m
Goal: propsAvoidDyn
If dynamic obstacle is detected in travel path or junction, the maximum speed limit 0.3m/s shall be maintained in critical zone until the remaining distance of 2m
Goal: GPS_Props
GPS provides required properties
Solution: snHz
Hazard analysis HAZOP/FTA
Goal: LIDARProps
LIDAR provides required properties
Strat: failureBeh
Argument over failure behaviour of GPS
Goal: failureBehGPS1
If the GPS fitted on the AGV fails, the distance covered on the map in a timeframe is used to determine the current speed and the max speed limit shall be re-evaluated
Strat: rqtLIDAR
Argument over properties assured by LIDAR
Strat: rqtGPS
Argument over properties assured by GPS system
Goal: GPSProp_GPS System
The GPS tracking device placed on AGV shall estimate position with the accuracy of ± 30 cm
GPS System
Goal: propsLIDARSuff
The properties assured by LIDAR are sufficient to satisfy the requirement
Strat: failureBeh
Argument over failure behaviour of LIDAR
Goal: propsGPSSuff
The properties assured by GPS are sufficient to satisfy the requirement
Goal: LIDARProp_LIDAR System
The LIDAR mounted on top of AGV shall provide point clouds in 10m range for which the worst accuracy shall be ± 6cm
LIDAR System
Goal: failureBehLiDAR1
If the LIDAR mounted on AGV fails, the transition is made to the infrastructure LIDAR and the max speed limit shall be re-evaluated
Goal: LIDARfailCorrect_LIDAR System
All potential failure behaviour of LIDAR is completed and correctly identified
LIDAR System

Fig. 4. A safety case for the considered Smart Factory.

LIDAR sensor mounted on AGV is failed then "the transition is made to another detection device". Such devices are placed on the specific zones. To reduce risk to an acceptable level, the speed limit at first stage is set to ≤ 0.3m/s. The "transition is made" to the infrastructure devices, or the other AGV is followed in terms of platoon to accomplish the intended operation. The confidence in terms of the operating speed is gradually increased. In case of speed sensor failure, the maps are used for "detecting distance covered in timeframe and the speed". The additional safety countermeasure have also been considered for the AGVs with faulty (sub)system. Depending on the severity risk factor, the other AGVs in close ranges can be commanded, e.g., to drive away. The safety contract linked to the claim "fogPropsProvided" guarantee that the "AGV fog controller accurately computes the received inputs and provides data in 500 ms time frame". This, however, is dependent on the conditions "AGV fog controller is independent AND accurate AND distinct input data is received within 50 ms timeframe". Due to the transmission delay, the contract assumption A6 does not satisfy and the AGV fog controller may not hold the predicted guarantee G6. As a consequence, the AGV received the command late than expected and therefore corresponding control actions need to be evaluated. In particular, the threshold is determined, slow down and stopping time would be accordingly increased and updated. The alternative to cover the threshold of 200 ms is to reduce the maximum speed limit of AGV to 12 km/h. Besides the performance degradation, the upgradation is taken into consideration. The adaptations in safety contracts and corresponding safety argument elements are triggered in response to the selected control actions. In contrast to the matching of parameter names and their values/ranges for safety contracts, the text-based matching is performed to alter the description of safety case elements.

5. Related work

Denney et al. [4] discusses the dynamic safety cases as an operationalisation of the concept of through-life safety assurance; the data produced from safety management system can be utilised to create continuously evolving assurance argument. The suggested life cycle of dynamic safety cases comprises four main activities: identify the sources of uncertainty that weaken the confidence; monitor the deficits in safety argument(s) by collecting operational data; analyse operational data for the determination of threshold defined for assurance deficits; and respond to the operational events that affect safety assurance. This concept highlights a number of challenges, for example, how to decide the most important subset of uncertainty sources and how to automate them. The research in McDermid et al. [5] proposed the framework for safety assurance of automated systems that also has four major elements: the 'real world' consists of automated system and operational environment; the 'world as imagined' contains models (design and simulation) and results of safety analysis achieved based on them; the 'world as observed' involves the collection and processing of operational data; and the 'safety case' that previously reflects the 'world as imagined' but then updated to respond to the reality 'world as observed'. Jaradat and Punnekkat [6] describes the monitoring of runtime failures related to the hardware component and failures analysis by comparing with a predefined threshold. The fault trees were used for deriving safety contracts and defining thresholds.

Sljivo et al. [33] extends the notion of contract with the strong and weak contracts to handle the multi-context setting of reusable components by distinguishing assumptions on configuration parameters and operational variables. The strong contracts captures properties in all contexts and must be satisfied; whereas the weak contracts not need to be satisfied. Similar to the product line, the configuration aware contracts can be used to assign the minimum requirements. The boolean parameters are defined to capture behaviours in terms of contracts; therefore, a lot of configuration parameters for different contracts are needed. In another study, Sljivo et al. [34] discussed the tool-support for specifying the contracts and argument-fragments generation. Besides the informal description of each of the contracts, the relations to the requirements are added and the status update is performed to proceed with the argument-fragments generation for individual components in the system.

Calinescu et al. [7] present ENTRUST methodology for the engineering of trustworthy self-adaptive software systems: unmanned underwater vehicle and foreign exchange. ENTRUST supports the partial instantiation of argument pattern at design phase, and adaptation of evidence produced by the monitor-analyse-plan-execute (MAPE) control loop to fill in the placeholders from partial assurance argument. For modelling and verification of MAPE models, UPPAAL and PRISM model checkers are utilised. Denney and Pai [35] present a methodology that combines the manually created, aircraft-level argument fragments derived from safety analysis, with an automatically generated lower-level argument fragments derived from formal verification. The formal verification is supported with the AutoCert tool, while the AdvoCATE toolset is utilised to assemble both of the fragments created for Swift Unmanned Aircraft System (UAS). However, the evidence of safe flight and other evidence from operational data, that is operational safety cases are not considered. In another paper, Denney and Pai [36] combines safety assurance arguments with bow-tie diagrams as a common framework to support modelling of UAS safety cases for airworthiness and operational safety. The abstract safety architecture specifying the collection of hazard controls is represented in bow-tie diagrams and an argument architecture is presented as argumentation patterns.

For the systems with enhanced automation and digitalisation, the safety analysis conducted during design and development phase may not be sufficient. The dynamic safety assurance is perceived as fundamental for them [4,5]. The research in [2] highlight the safety assurance challenges for Industry 4.0. This paper, however, is a step towards explicitly supporting the dynamic safety assurance of Industry 4.0. After the safety analysis conducted during design and development phase, the digital twins are configured, the operational data is gathered from them to monitor, evaluate, and resolve deviations from the specified behaviour reflected in safety cases and the actual behaviour.

6. Conclusions and future work

This paper targets the construction and maintenance of safety cases for Industry 4.0. The hazard analysis techniques, particularly HAZOP and FTA, are applied to identify unplanned events or hazardous conditions, their effects on the behaviours and the reliability of transportation and data flow in Industry 4.0. The results gained from the hazard analysis are utilised to derive the safety requirements related to the prevention or mitigation of potential hazards. Safety contracts that support the maintenance and adaptation of safety cases based on operational deviations, incorrect or implicit assumptions and guarantees have also been derived. The OpenCert platform is utilised for safety case construction and association of derived safety contracts with the safety case. To establish a basis for evaluation and improvement of performed operations, simulation environments have been configured. The operational data is gathered from them to monitor the uncertainty sources, track and evaluate deviations with the intended behaviour, and resolve the gaps in safety cases to reflect the actual behaviour of production system. The obstacle detection, data processing and avoidance aspects are specifically taken into consideration. The usefulness is demonstrated for the materials transportation and data flow in Industry 4.0.

In further evaluation, we plan to explore other type of machines that have different characteristics and transformation of application domains to the Industry 4.0. The dynamic safety assurance methodology is generally applicable to various automated machines. The initial effort is although required to derive the safety contracts for uncertainty sources, but it is worthwhile to detect deviations from intended behaviour and perform necessary adaptations at the operational phase. In this process, there is a need to handle the short- and long-term deviations in different ways. The dynamic risk management is crucial for short terms deviations, but the safety contracts and safety cases may just be updated for the frequent or long term deviations.

In another work, the virtual boundaries around geographic zones (i.e., geofences) are used as a measure for elimination or mitigation of automated transportation risks at the operational phase [24]. Specifically, the static, dynamic, time-based and conditional geofences are taken into consideration. In the current circumstances when the different configurations are reflected in the safety cases, problems arise, due to unavailability of operational data for all the parts. Accordingly, we plan to support the safety case configurations as part of the dynamic assurance in Industry 4.0. The configuration analytics will be performed to determine the production safety and operational performance, as well as the tradeoffs between them. Another limitation is that the identification and resolution of gaps in safety cases is dependent on the GSN and the OpenCert platform. In the future, the safety cases modelled in other formats, such as free text and tabular structures could be supported.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by FiC (Future factories in the Cloud) project funded by SSF (Swedish Foundation for Strategic Research) and SUCCESS (Safety assUrance of Cooperating Construction Equipment in Semi-automated Sites) project via the AAIP (Assuring Autonomy International Programme) funded by the UK Lloyd's Register foundation.

References

[1] C.A. Ericson, Hazard Analysis Techniques for System Safety, second ed., John Wiley & Sons, 2015.

[2] O. Jaradat, I. Sljivo, I. Habli, R. Hawkins, Challenges of safety assurance for industry 4.0, in: 13th European Dependable Computing Conference, EDCC' 17, Geneva, Switzerland, September 4–8, 2017, pp. 103–106, http://dx.doi.org/10.1109/EDCC.2017.21.

[3] The Assurance Case Working Group, Goal structuring notation community standard version 2, january, 2018, 2018, [Online] http://www.goalstructuringnotation.info/.

[4] E. Denney, G.J. Pai, I. Habli, Dynamic safety cases for through-life safety assurance, in: 37th IEEE/ACM International Conference on Software Engineering, ICSE 2015, Florence, Italy, May 16–24, 2015, pp. 587–590.

[5] J. McDermid, Y. Jia, I. Habli, Towards a framework for safety assurance of autonomous systems, in: Proceedings of the Workshop on Artificial Intelligence Safety 2019 Co-Located with the 28th International Joint Conference on Artificial Intelligence, AISafety@IJCAI 2019, Macao, China, August 11–12, 2019, pp. 1–7, URL http://ceur-ws.org/Vol-2419/paper_2.pdf.

[6] O. Jaradat, S. Punnekkat, Using safety contracts to verify design assumptions during runtime, in: 23rd International Conference on Reliable Software Technologies, Ada-Europe '18, Lisbon, Portugal, June 18-22, 2018, pp. 3–18.

[7] R. Calinescu, D. Weyns, S. Gerasimou, M.U. Iftikhar, I. Habli, T. Kelly, Engineering trustworthy self-adaptive software with dynamic assurance cases, IEEE Trans. Softw. Eng. 44 (11) (2018) 1039–1069, http://dx.doi.org/10.1109/TSE.2017.2738640.

[8] M. Hermann, T. Pentek, B. Otto, Design principles for industrie 4.0 scenarios, in: 49th Hawaii International Conference on System Sciences, HICSS 2016, Koloa, HI, USA, January 5–8, 2016, pp. 3928–3937, http://dx.doi.org/10.1109/HICSS.2016.488.

[9] M. Schluse, M. Priggemeyer, L. Atorf, J. Rossmann, Experimentable digital twins - streamlining simulation-based systems engineering for industry 4.0, IEEE Trans. Ind. Inf. 14 (4) (2018) 1722–1731, http://dx.doi.org/10.1109/TII.2018.2804917.

[10] D. Gorecky, M. Schmitt, M. Loskyll, D. Zühlke, Human-machine-interaction in the industry 4.0 era, in: 12th IEEE International Conference on Industrial Informatics, INDIN 2014, Porto Alegre, RS, Brazil, July 27-30, 2014, pp. 289–294, http://dx.doi.org/10.1109/INDIN.2014.6945523.

[11] Object Management Group, Structured assurance case metamodel (SACM), version 2.0, 2018, (Last Accessed: 7 July 2020), [Online] https://www.omg.org/spec/SACM/2.0.

[12] S. Nair, J.L. de la Vara, M. Sabetzadeh, L.C. Briand, Classification, structuring, and assessment of evidence for safety - a systematic literature review, in: Sixth IEEE International Conference on Software Testing, Verification and Validation, ICST, Luxembourg, Luxembourg, March 18–22, 2013, pp. 94–103, http://dx.doi.org/10.1109/ICST.2013.30.

[13] B. Meyer, Applying 'design by contract', Computer 25 (10) (1992) 40–51, http://dx.doi.org/10.1109/2.161279.

[14] A. Benveniste, B. Caillaud, D. Nickovic, R. Passerone, J.B. Raclet, P. Reinkemeier, A. Sangiovanni-Vincentelli, W. Damm, T. Henzinger, K.G. Larsen, Contracts for System Design, Research Report, (RR-8147) INRIA, 2012, p. 65, URL https://hal.inria.fr/hal-00757488.

[15] A. McEwen, H. Cassimally, Designing the Internet of Things, first ed., Wiley Publishing, 2013.

[16] Peter. Mell, Timothy. Grance, The NIST definition of cloud computing, recommendations of the national institute of standards and technology, 2011, [Online] https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpub%20lication800-145.pdf.

[17] M.D. Donno, K. Tange, N. Dragoni, Foundations and evolution of modern computing paradigms: Cloud, IoT, edge, and fog, IEEE Access 7 (2019) 150936–150948, http://dx.doi.org/10.1109/ACCESS.2019.2947652.

[18] C. Haddon-Cave, The Nimrod Review: An Independent Review into the Broader Issues surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006, Report, The Stationery Office, London, 2009, URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/229037/1025.pdf.

[19] F.U. Muram, M.A. Javed, S. Punnekkat, System of systems hazard analysis using HAZOP and FTA for advanced quarry production, in: 4th International Conference on System Reliability and Safety, ICSRS, Rome, Italy, November 20-22, 2019, pp. 394–401, http://dx.doi.org/10.1109/ICSRS48664.2019.8987613.

[20] American National Standards Institute/Industrial Truck Safety Development Foundation, Safety standard for driverless, automatic guided industrial vehicles and automated functions of manned industrial vehicles, december, 2019, 2019, [Online] http://www.itsdf.org.

[21] International Electrotechnical Commission, IEC 61508-1:2010-functional safety of electrical/electronic/programmable electronic safety-related systems, 2010.

[22] International Organization for Standardization, ISO 13849-1, safety of machinery — Safety-related parts of control systems — Part 1: General principles for design, 2015.

[23] International Electrotechnical Commission, IEC 61496-1, safety of machinery-electro-sensitive protective equipment - part 1: General requirements and tests, 2012.

[24] M.A. Javed, F.U. Muram, A. Fattouh, S. Punnekkat, Enforcing geofences for managing automated transportation risks in production sites, in: 16th European Dependable Computing Conference, EDCC 2020 Companion Proceedings, Munich, Germany, September 7–10, 2020.

[25] F.U. Muram, H. Tran, U. Zdun, Supporting automated containment checking of software behavioural models using model transformations and model checking, Sci. Comput. Programm. 174 (2019) 38–71, http://dx.doi.org/10.1016/j.scico.2019.01.005.

[26] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, V. Koltun, CARLA: An open urban driving simulator, in: Proceedings of the 1st Annual Conference on Robot Learning, CoRL 2017, Mountain View, California, USA, November 13–15, 2017, pp. 1–16.

[27] S. Shah, D. Dey, C. Lovett, A. Kapoor, Airsim: High-fidelity visual and physical simulation for autonomous vehicles, in: Field and Service Robotics, Vol. 5, 2017, pp. 621–635, http://dx.doi.org/10.1007/978-3-319-67361-5_40.

[28] O. Michel, Webots: Professional mobile robot simulation, Int. J. Adv. Robot. Syst. 1 (1) (2004) 39–42, http://dx.doi.org/10.5772/5618.

[29] D. Seybold, J. Domaschka, A. Rossini, C.B. Hauser, F. Griesinger, A. Tsitsipas, Experiences of models@run-time with EMF and CDO, in: Proceedings of the 2016 ACM SIGPLAN International Conference on Software Language Engineering, SLE '16, October 31–November 1, 2016, Amsterdam, Netherlands, 2016, pp. 46–56, http://dx.doi.org/10.1145/2997364.2997380.

[30] S. Schrauf, P. Berttram, Industry 4.0: How digitization makes the supply chain more efficient, agile, and customer-focused, 2016, [Online] https://https://www.strategyand.pwc.com/gx/en/reports/industry40.pdf.

[31] P. Ghadimi, C. Wang, M.K. Lim, C. Heavey, Intelligent sustainable supplier selection using multi-agent technology: Theory and application for industry 4.0 supply chains, Comput. Ind. Eng. 127 (2019) 588–600, http://dx.doi.org/10.1016/j.cie.2018.10.050.

[32] R. Hawkins, J. McDermid, Safety assurance of autonomy to support the Fourth Industrial Revolution, Research Report,, University of York, 2019, URL https://www.york.ac.uk/media/assuring-autonomy/publications/AAIP%20report%20for%20IfM%20July%202019.pdf.

[33] I. Sljivo, B. Gallina, J. Carlson, H.A. Hansson, Configuration-aware contracts, in: 4th International Workshop on Assurance Cases for Software-Intensive Systems, ASSURE '16, Trondheim, Norway, September 20, 2016, pp. 43–54, http://dx.doi.org/10.1007/978-3-319-45480-1_4.

[34] I. Sljivo, B. Gallina, J. Carlson, H. Hansson, S. Puri, Tool-supported safety-relevant component reuse: From specification to argumentation, in: 23rd International Conference on Reliable Software Technologies, Ada-Europe '18, Lisbon, Portugal, June 18–22, 2018, pp. 19–33, http://dx.doi.org/10.1007/978-3-319-92432-8_2.

[35] E. Denney, G.J. Pai, Automating the assembly of aviation safety cases, IEEE Trans. Reliab. 63 (4) (2014) 830–849, http://dx.doi.org/10.1109/TR.2014.2335995.

[36] E. Denney, G. Pai, Architecting a safety case for UAS flight operations, in: 34th International System Safety Conference (ISSC 2016), Orlando, FL, USA, August 8–12, 2016.

Dynamic Reconfiguration of Safety-Critical Production Systems

Faiz Ul Muram, Muhammad Atif Javed, Hans Hansson and Sasikumar Punnekkat
School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden
Email: faiz.ul.muram|muhammad.atif.javed|hans.hansson|sasikumar.punnekkat@mdh.se

***Abstract*—The current trends of digitalization and Industry 4.0 are bringing ample opportunities for manufacturing industry to fine tune their products and processes at will, to meet changing market needs within short notice. However, the characteristics of advanced production systems, such as dynamic interactions between machines and reconfigurations, if not carefully orchestrated, could potentially lead to production failures or mishaps, making them safety-critical. Previous studies on hazard analysis, safety-performance tradeoffs and assurance cases have not specifically considered the dynamic reconfiguration scenarios in production systems. In this paper, for the hazard identification and mitigation/elimination, the principal characteristics of highly reconfigurable production systems have been given special consideration. Even if the hazard analysis results are incorporated in the initial designs of production systems, operational changes, such as adding/removing machines in response to market demands, system failures, or unanticipated hazardous conditions may still adversely impact the production safety and operational performance. For the operational changes, we perform the quantitative assessment through configuration analytics to determine the corresponding impacts on safety, performance and production demands. After that, the assurance case models are obtained with production line to cope with the potential problems during the dynamic safety assurance. The applicability of the proposed methodology is demonstrated in the context of a quarry site production scenario.**

***Index Terms*—Hazard Analysis, Safety Cases, Reconfiguration, Production Line, Manufacturing Systems, Quarry Site**

I. Introduction

Industry 4.0 targets higher productivity and efficiency in manufacturing by making the machines smarter, production lines more flexible and processes less wasteful, by focusing on enhanced levels of the automation, digitalization and connectivity. The advanced production systems based on the Industry 4.0 are composed of several distinct systems equipped with sensors, actuators and controllers that operate in both isolation and conjunction. For such system of systems (SoS), the capability to adjust production capacity and functionality, in particularly the dynamic reconfiguration in consequence of market changes, environmental conditions and system failures is perceived as fundamental. Koren et al. [1] present six principal characteristics of highly reconfigurable manufacturing systems: modularity, scalability, diagnoseability, customizability, convertibility and integrability. The inability to effectively and safely reconfigure manufacturing or production systems means the loss of business opportunities [2].

The production systems are typically categorized as safety-critical, for example, due to the interactions between autonomous machines and hazardous substances that can potentially harm humans (injuries or even deaths) or create damages to machines, property or the environment. For safety-critical production systems, safety assurance is a regulatory requirement. The hazard analysis provides the basis for safety, in particular, hazard identification and mitigation/elimination is the key to mishap prevention [3]. In order to demonstrate the acceptable safety of production operations, safety cases have been constructed to provide comprehensive, logical and defensible justification of the safety of a production system for a given application in a predefined operating environment [4]. However, some parts of the safety analysis carried out at system design and development phase may turn out to be incorrect/inapplicable during the operational phase of production. Hence, there is a need to handle the mismatches between the constructed safety cases that reflect different configurations and the current system operation, in order to provide continuous assurance of safety.

In the safety assurance literature, despite the following limited efforts from a small number of research centres and individuals, there continues to be a dearth of published studies on hazard analysis for safe reconfiguration [5]–[8], safety-performance tradeoffs [9], [10] and customization of safety-critical systems [11]–[16]. For instance, the hazard analysis for dynamic reconfiguration of production systems and the involved safety-performance tradeoffs have not been taken into consideration. Besides that, there is also a need for the customization of assurance cases in production lines. This paper aims to address the aforementioned key shortcomings, which can be summarised by the following aspects:

1) It presents the hazard analysis for reconfigurable production systems, using the Hazard and Operability (HAZOP) and Fault Tree Analysis (FTA) techniques. For the hazard identification and mitigation/elimination, the principal characteristics of reconfigurable production systems are explicitly considered.
2) Tradeoff analysis is carried out in which the safety is considered upfront and explicitly evaluated together with the increases and decreases in production capacity. For this reason, we consider various factors, including alternative features, travel paths, production failures and number of operating machines in the site.
3) The argument fragments are customized in the production line for which the seamless integration between

argumentation and variability management activities is required. Besides the adaptation of changes in production line, the generation of argumentation (safety case) models and diagrams is supported.

Although during the design and configuration of the production system, derived safety requirements are considered for preventing or mitigating the identified hazards, during its operational phase, changes such as adding/removing machines, system failures, or unanticipated hazardous conditions may still negatively impact the production safety and operational performance. As increased digitalization of manufacturing is becoming a reality, we propose leveraging on simulation-based digital twins. Consequently, the available simulation data for assessing and improving the production reconfigurations, comes handy for safety assurance as well. In circumstances when the achieved benefit is higher and the compromise remain within an acceptable region, the intended choice in favour of an alternative is tolerated. The production line is also extracted specifically considering the information from the production site and documentation that reflects different alternative choices. It is used to determine the usage of specific configuration in the past and make adaptations based on the determined thresholds. The work presented in this paper utilises the OpenCert[1] tool platform to provide support for modelling and visualizing the safety cases in the Goal Structuring Notation (GSN) [4]. As the safety cases constructed for the reconfigurable production system reflects the different alternatives, to cope with the problems, such as unavailability of runtime data due to system failures or reconfiguration changes, for a selected configuration, the assurance (safety case) models are obtained with the production line to further support the dynamic safety assurance. The applicability of the proposed methodology is demonstrated for the construction equipment domain, using the scenario of a quarry site, which solely produce stone and/or gravel products in various dimensions.

The rest of this paper is organized as follows: Section II describes essential background information on hazard analysis, safety assurance and reconfigurable production. Section III provides an overview of the production operations and involved machines in the quarry site. Section IV presents the proposed methodology, including relevant scenarios and aspects of evaluations we performed. Section V discusses related work. Section VI concludes the paper and presents future research directions.

II. Background

This section recalls the background information on which the presented work is based. In particular, a short overview of HAZOP and FTA techniques is presented in Section II-A; the safety cases and their possible representation is summarised in Section II-B; finally, the customized flexibility with production line is discussed in Section II-C.

A. Hazard Analysis

Hazard analysis is an activity, which deals with the identification of hazards, their causal factors and specification of safety goals with the intention to eliminate/mitigate the hazards, to avoid the unacceptable risk [3]. In this paper, we use HAZOP and FTA techniques for performing hazard analysis. The HAZOP analysis is an inductive technique for identifying and analysing the potential deviations from design intention or operating conditions of a system [3]. HAZOP analysis is preferably carried out in the design phase taking different parts into consideration, such as hardware, software, procedures, human error and environment. The HAZOP analysis process starts with a full description of a system (product and/or process), which is broken down into system parameters, such as composition, software data flow, voltage and startup [3]. Next, all possible deviations are systematically identified by comparing a list of parameters or characteristics of a system (e.g., voltage, position and software, etc.) against a set of guide words (e.g., less, early and incorrect, etc.). After the identification of deviations, an assessment is carried out to determine whether particular deviations and their consequences can have negative effects on the system's operation. Finally, the appropriate safety recommendations that can help to prevent accidents and reduce exposure are identified [17].

FTA is a deductive analysis approach for modelling, analysing and evaluating failure paths in large complex dynamic systems [18]. FTA is highly recommended for detailed analysis of an undesired event and evaluation of hazards that are highly safety-critical. The FTA process starts with a top undesired event or mishap, and attempts to find out what intermediate events and bottom (basic) events or combination of basic events (e.g., nodes of a system or component behaviour) lead to the occurrence of this top event. The cause-and-effect relationships between events are defined using logical operators (e.g., AND-gate, OR-gate, etc.).

B. Assurance Case Representation

An assurance (safety) case is a collection of auditable claims, arguments, and evidence created to support the claim that a defined system/service will satisfy its assurance requirements [19]. There are several ways to document safety cases, e.g., free text, tabular structures and graphical notations. The Structured Assurance Case Metamodel (SACM) [19] is the Object Management Group (OMG) standard that integrates and standardizes the broadly used notations for documenting safety cases, including GSN and Claims-Arguments-Evidence (CAE). The work presented in this paper utilises the PolarSys OpenCert tool platform for modelling and visualizing safety cases [20]. OpenCert is an open source assurance and certification tool; its argumentation editor is based on the GSN graphical notations. However, Common Assurance and Certification Metamodel (CACM) implemented in OpenCert internally uses the SACM metamodel.

The main objective of a tree oriented goal structure is to show how goals (claims in CACM/SACM) are broken down

[1] https://www.polarsys.org/opencert/

into sub-goals (sub-claims) until supported by solutions (evidences). Strategy or argument reasoning describes a rationale for decomposing the goals into sub-goals, whereas context describes the operational environment, scope and domain in which the claims or strategy are stated. Undeveloped and uninstantiated decorators are applied to the elements that need to be further developed and to be replaced (instantiated) with a more concrete instance, respectively. Assumption element presents an intentionally unsubstantiated statement, which is assumed to be true for a certain goal or strategy. Justification element provides an explanation or rationale why a certain goal or strategy is considered acceptable. An away goal referenced to a claim presented in another argument module [4]. The argument elements can be linked with one of the two relationships: SupportedBy and InContextOf. The SupportedBy relationship is used to show the inferential or evidential relationships between elements. In particular, inferential relationship (AssertedInference in CACM) declares that there is an inference between goals in the argument, whereas evidential relationship (AssertedEvidence in CACM) shows the connection between a goal and the solution. InContextOf relationship is used to declare contextual relationships of goals or strategy with context, assumption and justification elements. For further details we refer the readers to GSN standard [4].

C. *Reconfigurable Production*

During the past 25 years, many researchers have focused on dynamic reconfiguration [21]. The principal intention of reconfigurable production systems is to enhance the responsiveness in consequence to changes in market, environmental conditions and unexpected machine failures. For the design and operation of highly reconfigurable production systems, the modularity, scalability, diagnoseability, customizability, convertibility and integrability characteristics are taken into consideration [1]. The reconfigurable production systems possess the advantages of both dedicated production lines and of flexible systems. In particular, they focus on customized flexibility that can be supported through production lines. The idea with the production line is the identification and systematization of commonalities and variabilities to concurrently engineer a set of production scenarios; the achievement of a single production scenario is based on the selection and composition of commonalities and variabilities [22], [23].

For reasons of customized flexibility, the Base Variability Resolution (BVR) tool, which is well-known tool built on the OMG's revised submission of Common Variability Language (CVL) is utilised [24]. The BVR tool supports orthogonal variability management for any Meta-Object Facility (MOF)-compliant model. The generation of target configuration models is performed with three editors: (1) VSpec, which is an evolution of the Feature-Oriented Domain Analysis (FODA), usually called Feature Models; (2) Resolution in which the desired inclusion/exclusion choices for the specific configuration are made, multiple resolutions need to be defined for reconfigurable production systems; and (3) Realization that specify the placement and replacement fragments over model elements, the execution of fragments is subject to the selection of assigned VSpec in the resolution.

III. Running Case Study: Electric Quarry Site

This section describes an operational quarry site [25], which falls under the construction equipment domain. The quarry site solely produce dimension stone, gravel and asphalt, which are used for the construction of buildings, roads, railway track beds and shoring up river banks. The quarry operation is carried out using different kinds of machines, such as excavator, mobile primary crusher, wheel loader, stationary secondary crusher and haulers, as shown in Figure 1. *Autonomous haulers* and/or *articulated haulers* are used to transport material in the quarry site. The *excavator* feeds the blasted rocks to primary crusher, i.e., the rocks that are broken out of the mountain with explosives. The *primary crusher* breaks the blasted rocks into the smaller rocks to facilitate the transportation to the secondary crusher. From the discharging perspective, the conveyor belt is attached to the primary crusher. It is therefore possible to directly load the haulers from primary crusher. If primary crusher is building the stone piles, the direct loading is disabled. In such case, the hauler will be loaded with the *wheel loader* from stone pile (i.e. indirect/truck loading).

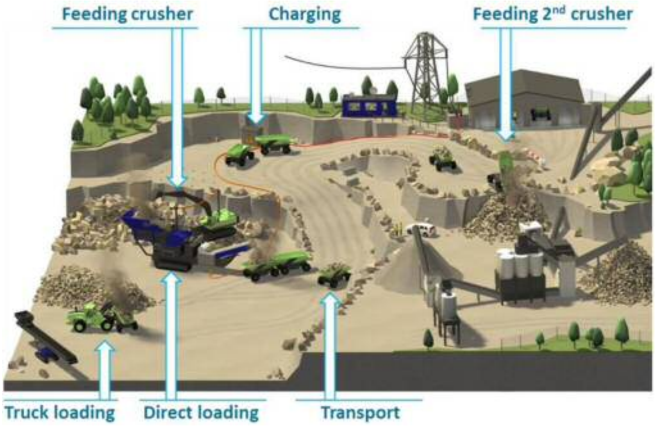


Fig. 1. Quarry Site

The *site management* system serves as a primary controller. It is composed of three subsystems: (i) The *user interface* subsystem visualizes the status information of machines. (ii) From the *traffic control* perspective, the positions of machines are tracked with the Global Positioning System (GPS), which are displayed on the site map. The travel paths are defined for moving towards the loading, dumping, charging and parking places. (iii) The *fleet management* subsystem commands the specific machines to perform their intended operations. For transportation, the missions are assigned to the autonomous or articulated haulers. The operation of *autonomous haulers* is similar to the Automated Guided Vehicles (AGVs). For the perception of surrounding environment, two obstacle detection sensors, in particular, Light Detection and Ranging (LIDAR) and camera are mounted whereas the GPS is fitted for tracking and navigation purposes. The data produced by the particular sensors is processed for controlling the mechanical parts, for example, the drive unit for motion and operation, the steering system for manoeuvring, and the braking system for slowing down and stopping the vehicle. The interaction platform and other attachments that include batteries for power supply are integrated in the machines.

The haulers travel in the defined path and dumps the loaded rocks in the dumping spot for feeding the secondary crusher. The *secondary crusher* further crush the rocks into smaller granularity or fractions to meet the customer demands. To perform the mission efficiently, the required battery level needs to be determined. This is done before going to the loading place. To be able to recharge the battery, the *charging station* is defined. The machines are moved to the *parking station* after the termination of transportation operation.

IV. Dynamic Production Reconfiguration

The enhanced automation, digitalization and connectivity tend to make manufacturing processes faster, more efficient and more customer-centric. However, they significantly increases the safety issues. To identify hazards, their effects, and causal factors, the hazard analysis is performed. For this reason, the HAZOP and FTA techniques are applied. To demonstrate the acceptable safety of production operations, safety cases are constructed; the safety claims are defined based on the available artefacts, for example, the safety requirements. The documentation is primarily done during system design and development phase. The simulators-based digital twins were leveraged to perform verification and validation to gain confidence in production reconfigurations by incorporating safety requirements in them. The dynamic reconfiguration is a fundamental consideration for safety-critical production systems and is performed in consequence of market changes, environmental conditions and system failures. The operational data and production line are used as a base and checkpoint for configuration analytics. For a selected configuration, the assurance case models are obtained with the production line to further support the dynamic safety assurance. An overview of the proposed approach is presented in Figure 2.

A. Hazard Analysis

This subsection presents the hazard analysis process that helps in identifying the hazardous events and their causes, and developing the mechanisms to avoid, control, or otherwise mitigate those causes, which may result in unexpected or collateral damage. The outcomes of hazard analysis are used for the derivation of safety requirements and safety contracts. The derived safety requirements and mitigation techniques are used for designing and configuring the site environment, which serves as a digital twin. Several potential risks can be found in the quarry site due to simultaneous presence of autonomous haulers, human-driven machines and human workers at the same site. In the context of reconfiguration of production systems, hazardous situations due to a single system failure, addition of redundant machines, loading and dumping spots as well as charging station, inadequate update/replacement of software, system or module, deviation in services, communication failures, and environmental influences, which may lead to mishaps or accidents are considered. The HAZOP analysis is applied to identify deviations from design intent and operational aspects related to reconfiguration of safety-critical production systems. To perform the HAZOP analysis, a set of guide words (e.g., early/late, more/less, incorrect, etc.), parameters (e.g., location, speed, distance, travel path, etc.), system inputs and outputs, and type of messages (e.g., command, data and response) are identified. For in-depth analysis, the fault trees are constructed based on the hazards and their potential effects understood from the HAZOP analysis. The performed analysis does not only focus on the advanced production in the quarry site, but considers also the principal characteristics for reconfiguration, in particular, modularity, scalability, customizability, convertibility, integrability and diagnoseability.

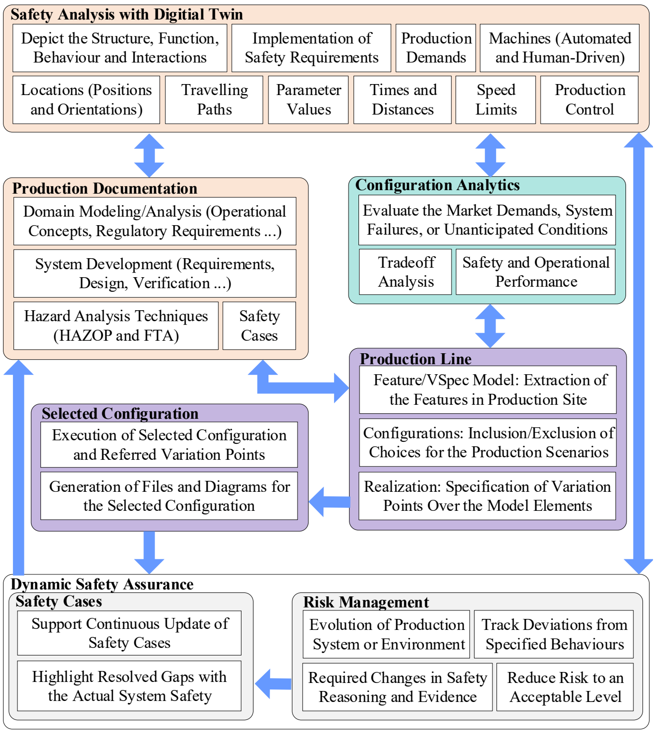

Safety Analysis with Digitial Twin
Depict the Structure, Function, Behaviour and Interactions
Implementation of Safety Requirements
Production Demands
Machines (Automated and Human-Driven)
Locations (Positions and Orientations)
Travelling Paths
Parameter Values
Times and Distances
Speed Limits
Production Control
Production Documentation
Domain Modeling/Analysis (Operational Concepts, Regulatory Requirements ...)
System Development (Requirements, Design, Verification ...)
Hazard Analysis Techniques (HAZOP and FTA)
Safety Cases
Configuration Analytics
Evaluate the Market Demands, System Failures, or Unanticipated Conditions
Tradeoff Analysis
Safety and Operational Performance
Production Line
Feature/VSpec Model: Extraction of the Features in Production Site
Configurations: Inclusion/Exclusion of Choices for the Production Scenarios
Realization: Specification of Variation Points Over the Model Elements
Selected Configuration
Execution of Selected Configuration and Referred Variation Points
Generation of Files and Diagrams for the Selected Configuration
Dynamic Safety Assurance
Safety Cases
Support Continuous Update of Safety Cases
Highlight Resolved Gaps with the Actual System Safety
Risk Management
Evolution of Production System or Environment
Track Deviations from Specified Behaviours
Required Changes in Safety Reasoning and Evidence
Reduce Risk to an Acceptable Level

Fig. 2. An Overview of the Proposed Approach

Table I shows the reduced hazard analysis results related to quarry site reconfiguration. From the scalability perspective, adding a number of autonomous haulers AHs more than intended, can cause congested travel paths, loading and dumping zones as well as charging station. In this particular case, autonomous haulers may not be able to maintain the safe distance of 20 meters from each other, which increases the risk of collision. Therefore, the number of machines, loading and dumping spots/zones should be added by calculating workload, time required to complete a mission and the total number of current machines. In addition, new speed limits and safe distances for autonomous haulers need to be considered. In the context of convertibility, it is necessary to carefully analyse

TABLE I
EXTRACT OF THE HAZOP ANALYSIS REPORT FOR RECONFIGURATION

Item	Guide Word	Parameter	Deviation	Cause	Consequence	Recommendation
H_SC1	More, Other Than	Machines	Adding redundant machines in the site	The site management or human operator does not detect and maintain the time and speed over a workload, distance and number of machines in the site	The congested routes or travel paths, loading and dumping zones as well as charging station on sites; autonomous hauler (AH) does not maintain the safe distance with other machines and human; lead to incomplete mission	Calculate the effectiveness and scalability values of the system; consider the new speed limit and safe distance; use dynamic filtering and inertial sensors
H_CO1	Incorrect	Loading command	The direct loading command sent to AH but the loading from primary crusher is stopped and the stone pile mode is active	The site management system failed to detect the human command, changeover time between configurations and sends the command rather early	The mission may not be completed; AH does not maintain the safe distance from other materials or working equipment; lead to machine damage, loss of critical hardware	Introduce a communication prioritization mechanism between machines and with the site management; AH should focus on the data obtained from obstacle detection devices, reduces the speed limit and maintains the safe distance until it gets the new command
H_CU1	Early	Dumping zone, sequential operations	AH arrives earlier on a dump site pause for unloading the material, it does not make a queue while articulated hauler (ART) is not exited from zone	The high speed due to the failure of speed sensor, wheel encoder or brake system; the site queue command has not been generated when intended	AH does not maintain a safe distance from other machines; machine damage, loss of critical hardware; human injuries or life loss may happen	AH should focus on the data obtained from placed camera and LIDAR, slow down the speed; wait until ART exit from the dump pause
H_DO1	Less/Part of	Wheel loader location (position and orientation)	Less information is sent or received regarding the loading zone; route optimization failure	Due to the GPS sensor failure, imprecise location is calculated or last known location is forwarded	Lead to mission failure; AH may enter in the restricted area where humans are working or dangerous materials are stored	Use the cameras mounted on machines and specific points/zones for detecting the locations; install additional sensors as back up
H_DO2	Inoperative, Failure	Brake system, collision avoidance	AH detected obstacle in a close range and the brake is failed	Loss of power; broken motor; fluid leakage	The collision of AH with static obstacles (e.g., stone) or with dynamic obstacles (moving machines and human workers) may happen	Use auxiliary brake system; an emergency shut-off by site management, with the remote control, or button press by nearby human
H_MO1	Less, More	Obstacle detection	Only part of some objects is detected; AH moves behind an object and not able to detect other objects in the vicinity	Replaced or upgraded LIDAR has a different range or flight of view from previous one; misalignment	AH does not maintain a safe distance with the static and dynamic obstacles; machine damage, loss of critical hardware; human injuries or life loss	Use values from various sensors, e.g., LIDAR and cameras mounted on machines; install additional sensors at zones/spots as back up
H_IN1	Incorrect/ Other Than	Add or update software	Different software (or version) is updated on AHs; software is not in accordance with latest update	Unawareness of software versions, human error, appropriate documentation is not provided for every software in AH	The activation of an unintended movement of AH; collision; entrance in the restricted areas/regions	Draw additional protocols (regulations) that specify that integrator must test all software updates and versions of all machines to determine whether they have the same output

the times required for shutdown, restart and changeover while switching from one configuration to another. Moreover, the number of machines and connected travel paths in each configuration needs to be analysed. For instance, primary crusher discharging through conveyor belt (i.e. direct loading) mode is converted to make stone pile mode, during that period a direct loading command sent (instead of loading from wheel loader) to the autonomous hauler. The transformation of an incorrect mission or travel path to an AH can be caused by an incorrect command or timing failure that leads to an incomplete mission, machine damage or human injuries. This can be prevented by using predictable networking protocols, and switching between modes within a specified time limit. Furthermore, the autonomous hauler should use data from obstacle detection devices and reduce its speed limit until a new command is received.

Customization focuses on the selection of machines or system components based on their adjustment capabilities to facilitate rapid response to unexpected (sub)system failures. Both articulated (human-driven) and autonomous haulers make queues at loading and dumping zones and wait for their turn when the previous hauler is on loading/dumping point for loading/unloading material. The critical incidents can occur if an AH arrives early and does not make a queue due to failure of speed sensors, wheel encoder or brake system. There is a possibility that the queue command is not generated or articulated hauler ART took longer time than expected. This can be prevented by relying on obstacle detection and collision avoidance mechanisms that makes the AH wait until the ART exit after completion of loading or warn the ART by flashing red light and horn sounds. Diagnosability perception for reconfigurability considers the detection and prediction of hazardous events, and time required for identification of alternative solutions. The messages containing less, or wrong data sent from wheel loader/primary crusher to site management system can cause the mishaps. If less information regarding wheel loader location (i.e., position and orientation) is transferred to the AH caused by sensors failure, which in-turn lead to incomplete mission or entrance of haulers into a restricted area. As a control measure, the information can be gathered from cameras mounted on machines and specific points/zones.

In the quarry site, the modular components/systems can be upgraded or replaced to better fit the modern needs and new applications. A LIDAR mounted on the autonomous hauler AH can be replaced by another type of sensors or LIDAR. However, during the replacement, performance specifications, obtainable accuracy and data acquisition on-road environment need to be taken into consideration. It can be seen from the hazard analysis results that the navigation safety system performance might be affected due to incorrect replacement or upgrade of the LIDAR. Furthermore, new systems and components might be integrated within the existing production system and new technologies may also be introduced in the current systems. However, inadequate addition or updates of software due to unawareness of a service engineer might lead to unpredictable behaviour of the AH, such as unintended movement or entrance in a restricted areas/region where a human is working. The control measure that could solve this problem is to draw up additional protocols (regulations) in order to test all software updates and versions of all machines.

As mentioned above, the hazardous events might occur at different phases of quarrying process, such as, loading, transporting, dumping, charging, and parking spots/zones. We have performed an in-depth analysis of hazardous events that possibly occur at loading and dumping zones. To develop the fault tree, machine damage and human injury are selected as the top undesired events. After establishing a top event, sub-undesired events are identified and structured that is referred to the top fault tree layer. All possible reasons, including communication failures (i.e. emergent interactions between primary crusher, autonomous hauler and wheel loader), environmental influences (e.g., adverse weather condition, and surface condition) and failures of system/machines are evaluated level-by-level until all relevant events are found.

B. Simulation-based Digital Twins

The digital twins bring out the virtual depiction of the real-world systems; therefore the functions, behaviour, and communication capabilities are mirrored in the digital twins [26]. They are perceived as an integral part of systems with enhanced automation, digitalization and connectivity. For designing and configuring the production site, the construction equipment simulators are adapted and extended. The safety requirements derived from the results of the hazard analysis were implemented in the digital twins as code scripts. The mobile platforms used for training the operators of articulated haulers, excavators, and wheel loaders are connected to the production site; they operate in conjunction with the other machines. For instance, the rocks transportation can be carried out with the articulated (human-driven) and/or autonomous haulers. A detailed list of parameters were used to support dynamic reconfiguration of the scenario that are accessed and changed during operational phase. The site scenario is expressed in the Extensible Markup Language (XML) file. However, a database is used to store the data related to production operations.

C. Configuration Analytics

Although the safety requirements are implemented in the simulations, the quantitative reasoning or evaluation of configuration at operational phase is deemed essential. The focus of configuration analytics lies on the interplay of selections and the tradeoffs between production safety and operational performance. The safe and optimal configurations are maintained in the production line. It serves as a base and checkpoint for configuration evaluation and improvement. We measure/quantify the impact of alternative choices to select an optimal configuration for which a certain degree of safety of the production operations is assured. The typical scenarios we considered for configuration analytics are market demands, hazardous conditions and system failures.

1) Market Demands: Transformation to the low, normal and high production demands is explicitly evaluated. The site manager specifies the market demands. However, there is a need for configuration analytics that acts as a balancer and an active countermeasure against arbitrarily making decisions. It takes the current demands upfront to select a specific configuration used in the past and make adaptations based on the determined thresholds. Let us consider the loading of haulers. It is carried out with just one, two and even more loading points LPs. According to the used configuration in the past, when there is low demand, direct loading DL is disabled that concerns the loading through conveyor belt, but the stone pile SP mode of primary crusher PC is activated and indirect loading with a wheel loader TL is performed. In a typical scenario, to fulfil the normal demands, parallel loading of autonomous haulers AHs with PC and TL is performed. In circumstances of high demands, the additional loading spots PC and TL are made operational. In case the articulated haulers ARTs are present, the transportation of materials in conjunction, or otherwise with just AHs is performed.

The tradeoff factors that can be affected by the conflicting strategy include the speed S, distances D, load capacity LC and time taken for loading LT, transportation TT, unloading UT and charging CT. The adding/removing of machines and travelling paths may negatively influence the production safety and operational performance. Let us consider the congested

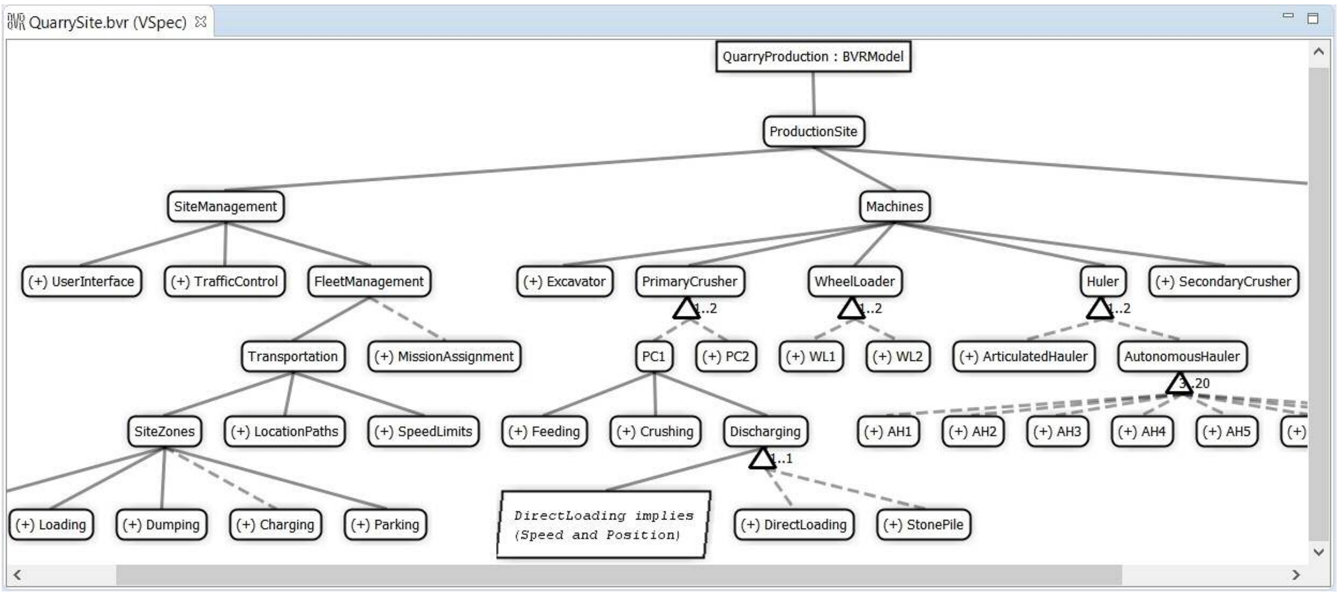


Fig. 3. Extraction of the Features in Production Site

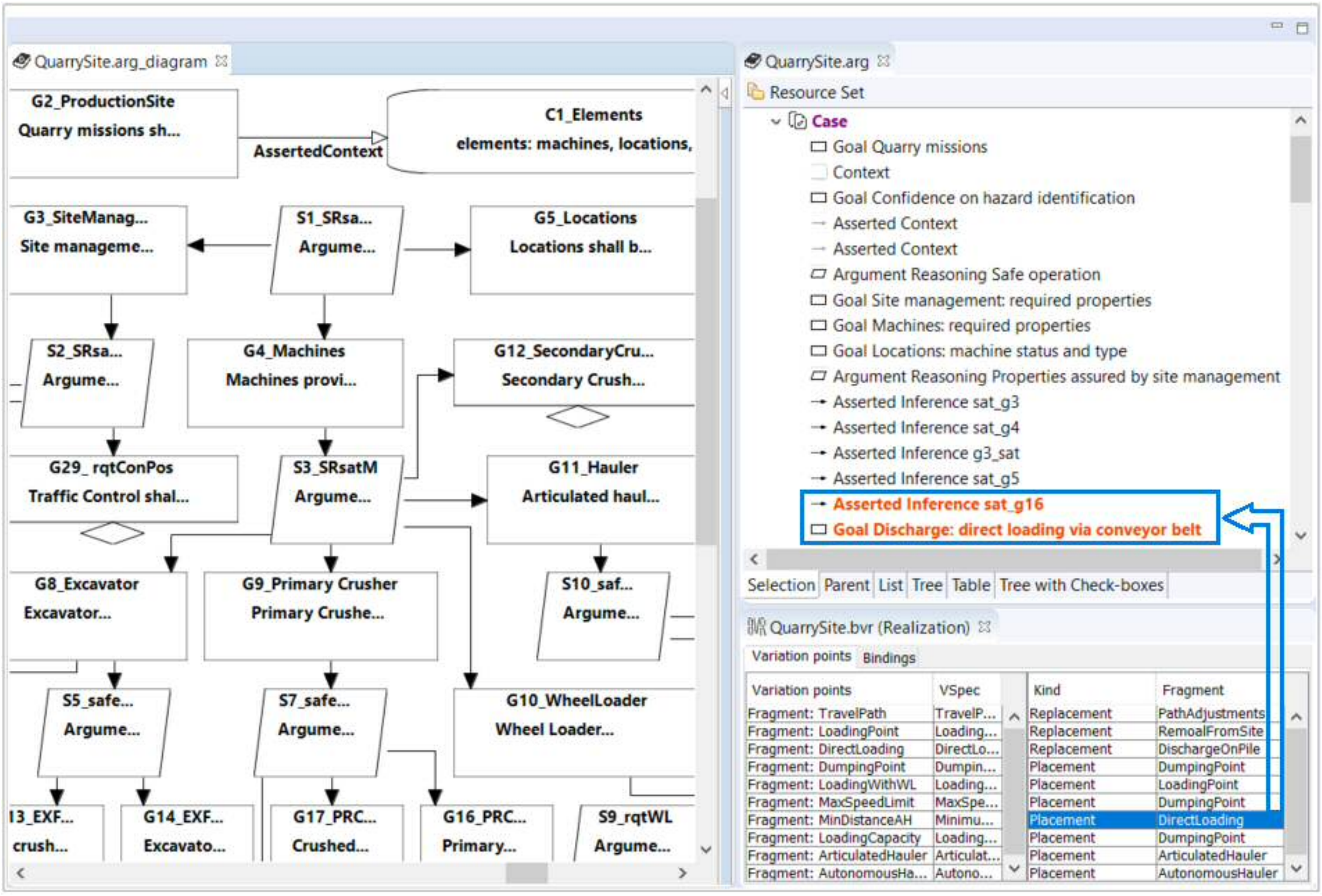
zones that occur when multiple haulers simultaneously arrived in specific locations. The configuration analytics control the movement of machines in production site. Specifically, the distance of haulers with the site locations is maintained in a way so that the waiting times are reduced. The locations of all the haulers are continuously monitored. In case of long waiting queues, the haulers are directed to the alternative locations. In general, the alternatives increase the travel distances, but their selection must not lower the operational performance. The activation or deactivation of machines present in site is commanded when required. Moreover, we assigned appropriate priority levels to indicate their relative importance/precedence. For example, the default loading point LP has highest priority, followed by alternative LP and then AHs. The additional measures include the generation of new speed limits. An example is the threshold delay at LP, when the direct loading queue LQ is empty and the arriving time AT is higher than the LT of hauler. In the simulations-based digital twins, the autonomous hauler AH travels at an average speed of 13 km/h. The risk of the speed increase in certain edge paths is regarded as acceptable when there is no hauler in the moving direction.

2) Hazardous Conditions: Travel paths TP connects the site zones SZ. For gaining confidence and managing tradeoffs, the defined TP is blocked with a static obstacle. In this case, the AHs make a queue and wait for the ART, or another machine, such as wheel loader to formulate an alternative TP, and then follows it. Since the achieved benefit is higher than the compromise, the fixed path in favour of alternative has been tolerated. In case of adverse environmental conditions such as slippery surface, depending on the severity risk factor, slow down of AHs, increase in stopping distance, movement restrictions in certain areas, or termination of transportation operation is performed. To avoid the mishap risks, switching to another configuration is also taken into consideration. Let us consider a scenario in which the SP mode is selected, but later shifted to the DL mode. As a consequence, the stones are fallen from the crusher conveyor belt on the DL point. The site cameras placed for monitoring the position of haulers in LP found the mishap risk potential. Since the DL cannot be performed in a safe manner, i.e., the risk is not acceptable, the DL is terminated and just TL is used. Other factors that lead to this situation are: PC is jammed or humans are present in a loading area. The mishap risk is although controlled, but the operational performance is influenced. After the clearance, a dialogue box pops up, it asks for the permission, the site manager needs to accept, so that the revert to a previous configuration is performed.

3) System Failures: In case of obstacle detection device failures, the transition is made to the other devices; they can be mounted on the AH and infrastructure level. Otherwise, the faulty AH is connected with the faultless AH, i.e., as a platoon. The faulty AH that is sergeant follows the platoon leader faultless AH to accomplish the mission. The additional subsystem failures have also been considered. For instance, there is a possibility that an AH does not maintain new speed limit due to speed sensor and brake failures. In the former case, the focus is shifted to the map to compute the speed, i.e., for detecting distance covered in time frame. In the latter case, besides the steering wheel rotation commands, depending on the severity risk factor, dynamic geofences are enforced. The capsule geometry shapes around the AH are used to widen the boundary for collision avoidance, so that the other AHs in close ranges can be commanded, e.g., to drive away.

D. Production Line

This subsection focuses on the engineering of production line that is used as a base and checkpoint for evaluating



Case
- □ Goal Quarry missions
- Context
- □ Goal Confidence on hazard identification
- → Asserted Context
- → Asserted Context
- ▱ Argument Reasoning Safe operation
- □ Goal Site management: required properties
- □ Goal Machines: required properties
- □ Goal Locations: machine status and type
- ▱ Argument Reasoning Properties assured by site management
- → Asserted Inference sat_g3
- → Asserted Inference sat_g4
- → Asserted Inference g3_sat
- → Asserted Inference sat_g5
- → Asserted Inference sat_g16
- □ Goal Discharge: direct loading via conveyor belt

Variation points	VSpec	Kind	Fragment
Fragment: TravelPath	TravelP...	Replacement	PathAdjustments
Fragment: LoadingPoint	Loading...	Replacement	RemoalFromSite
Fragment: DirectLoading	DirectLo...	Replacement	DischargeOnPile
Fragment: DumpingPoint	Dumpin...	Placement	DumpingPoint
Fragment: LoadingWithWL	Loading...	Placement	LoadingPoint
Fragment: MaxSpeedLimit	MaxSpe...	Placement	DumpingPoint
Fragment: MinDistanceAH	Minimu...	Placement	DirectLoading
Fragment: LoadingCapacity	Loading...	Placement	DumpingPoint
Fragment: ArticulatedHauler	Articulat...	Placement	ArticulatedHauler
Fragment: AutonomousHa...	Autono...	Placement	AutonomousHauler

Fig. 4. Realization of Assurance Cases in Production Line

and improving the production decisions driven by market situations, system failures, or unanticipated conditions. The solution alternatives and configurations used in the past were checked against the decisions to determine whether the realization of each choice would lead to an improvement, or possibly a deterioration, which would then require a tradeoff analysis. The generation of assurance (safety) cases for a selected configuration is carried out with the production line. As mentioned in Section II-C, the production line is engineered with the BVR tool. It comprises of three editors: VSpec (extended feature model), resolution (configuration) and realization (derivation). Similar to Martinez et al. [27], for the extraction of production line, the commands used for VSpec, resolution and realization editors were used from the BVR source code.

For the identification and systematization of commonalities and variabilities, the production documentation and site scenarios information in XML files are traversed. In the VSpec that is shown in Figure 3, the mandatory aspects are connected to the parent via solid lines, whereas the dashed lines represent optionality. Five kind of machines are mandatory in the quarry site: excavator, PC, wheel loader, hauler and secondary crusher. The mission assignment by site management, charging (not required for ART), selection of PC, its discharging mode (DL or SP), wheel loader and hauler are variability aspects. The multiplicity (1..2) is assigned to the PC, wheel loader and hauler kinds. This means that the selection of at least one of them is mandatory. The machines were distinguished based on the IDs. Thus, any PC, wheel loader and AHs can be chosen. The constraints have been used. For instance, the constraint *DirectLoading implies (Speed and Position)* enforces a direct loading restriction; pause state that is zero speed and precise position under conveyor belt must be maintained. The resolutions are generated from the VSpec. After that, the desired inclusion/exclusion choices are specified. The optimal configurations are maintained in the production line that contain a set of feature that are used for fulfilling the market demands and avoiding the certain hazardous conditions and failures. The automatic validation of a specific resolution is performed to confirm whether the resolution/configuration corresponds to the VSpec.

The realization is conducted based on the placements and replacements within the fragment substitutions. The elements of a selected placement are highlighted in orange colour. Note that the realization shown in Figure 4 is based on the argument (safety case) model constructed in the PolarSys Opencert. In the fragment substitution step, besides the assurance cases, the artefacts produced in the different activities of the development process, such as requirements, architectural design, implementation and testing can be mapped. In this way, impact analysis can also be supported. The execution of a fragment substitution is driven by the inclusion of associated feature in the selected configuration; therefore the VSpec features and fragment substitutions are linked. Let us consider a scenario in which the direct loading DL is unsafe, so that it is replaced with the SP and TL. After the execution of a specific configuration, the resolved .arg models are generated. Subsequently, for the achievement of diagrams, the Initialize arg_diagram diagram file command from the OpenCert source code is executed.

E. Dynamic Safety Assurance

In circumstances when the deviations from specified behaviours are detected, the implications of failure vulnerabilities are determined and defences against them are performed [26]. However, the problems arose during the safety assurance of reconfigurable production systems. The safety cases constructed for the reconfigurable production systems reflect several alternatives to choose from, each of which may exhibit divergent impacts on production safety and operational performance. Since the parts of safety cases might become inapplicable or otherwise invalidated during the production operations, to be able to cope with the problems like system failures or reconfiguration changes, the current configuration elements have to be determined. In this paper, the generation of safety case models for an active/present configuration is conducted with the production line. It is necessary to check whether the changes made by the site manager and dynamic risks management have implications (positive or negative) on the site safety and operational performance. For this reason, the tradeoff analysis is performed. The configuration analytics also acts as a balancer and an active countermeasure against arbitrarily making decisions. The measures to deal with unknowns and uncertainties were checked and recorded in the production line, to handle the specific situations in future.

V. Related Work

The discussion of related work concerns three topics: hazard analysis for safe reconfiguration, safety-performance tradeoffs, and configuration management of safety cases.

A. Hazard Analysis of Reconfigurable Systems

Giese and Tichy [5] model the failure propagation by explicitly considering alternative variants. However, to select the optimal variants in system architectures and product lines, minimal hazard probability and sensitivity to estimation errors that concern internal events are computed. Priesterjahn et al. [6] determine reachable component structures for a fixed number of subsequent structural changes. To obtain the risks, the computed hazard probabilities have been associated with the current severity that is encoded in numerical values. If the risk of a reachable component exceeds the system's acceptable risk, the structural change during runtime that result in the particular component structure is blocked. In another research paper, the timing properties related to the failure propagation and structural reconfiguration are considered [7]. Bhardwaj and Liggesmeyer [8] combine design time safety analysis and runtime monitoring to determine risk of potential configurations. The aforementioned works focus on the reconfiguration of individual components of a single system. Previous studies have not considered the dynamic reconfiguration of safety-critical production systems. In this paper, the principal characteristics of highly reconfigurable production systems are supported in the context of safety-critical systems.

B. Tradeoffs between Safety and Performance

Cowing et al. [9] evaluate the alternative risk management strategies for which the predicted operating performance of a critical system is considered. The tradeoffs between immediate productivity and safety are perceived as short terms goals. Despotou and Kelly [10] propose an argument-based approach that justify the design alternatives to facilitate tradeoffs in critical systems. In circumstances when the achieved benefit is equal or higher than the compromise, a design objective in favour of another can be tolerated. However, the compromised objective shall remain within an acceptable region. The published studies have not considered the safety-performance tradeoffs for the modern production systems.

C. Variability Management of Safety Cases

Stephenson et al. [13] focus on hazard assessment and safety-case production for Integrated Aircrew Training (IAT). The feature model is used to reflect variations between staff training scenarios, while the configuration process suggests the corresponding elements that can be traced and validated. Habli and Kelly [11] discuss the configuration management and certification of Aerospace Engine Monitoring Unit (EMU) in a product line. In another research paper, a functional hazard model for product-line development is proposed [12]. It captures the failure conditions, effects, severity classification and safety requirements for specific functional and environmental configurations, which are specified in the product-line context and domain models. Oliveira et al. [14] propose the automatic construction of modular product line safety case based on the model-based safety analysis and feature models. Nešić et al. [16] create a safety case for an arbitrary product line. In particular, the contract-based specification model is used to capture the technical architecture and the corresponding safety requirements of each product configuration. To date, however, the published studies have not supported the variability management of safety case fragments in a family/line.

VI. Conclusion and Future Work

This paper targets the dynamic reconfiguration, which is an essential characteristic of advanced production systems. Due to the dynamic reconfiguration, the safety assurance challenges tend to be significantly harder. To date, however, the published studies on hazard analysis, tradeoffs and assurance cases have not considered the dynamic reconfiguration of production systems. To support the dynamic reconfiguration of safety-critical production systems, this paper focuses on three novel contributions: First, the hazard identification and mitigation/elimination is performed by explicitly considering the principal characteristics of highly reconfigurable production systems. Based on the hazard analysis, which is performed with HAZOP and FTA techniques, the safety requirements are derived. They are implemented in the simulators-based digital twins, with the intention to perform verification and validation, to gain confidence in production site. Second, the tradeoffs between production safety and operational performance, in circumstances of site changes, failures and uncertain conditions are identified and resolved. The required adaptations in site scenarios are performed during operational phase. Third, the assurance cases are included in the production line to facilitate the analytics; the models for configuration are generated to further support the dynamic safety assurance. The applicability of the approach has been demonstrated for a quarry site.

This research is primarily based on the investigations that are carried out using a simulation test-bed. The obtained results satisfy the specifications on safety and performance, as well as their tradeoffs. However, there is a need to take further steps for stabilisation and potential inclusion of reconfiguration feature in real quarry site. The reconfiguration function is generally applicable to the broad range of scenarios and domains. In the future, we plan to consider additional scenarios and applications based on the Industry 4.0.

Acknowledgment

This work is supported by SUCCESS (Safety assUrance of Cooperating Construction Equipment in Semi-automated Sites) project via the AAIP (Assuring Autonomy International Programme) and FiC (Future factories in the Cloud) project funded by SSF (Swedish Foundation for Strategic Research).

References

[1] Y. Koren, X. Gu, and W. Guo, "Reconfigurable manufacturing systems: Principles, design, and future trends," *Frontiers of Mechanical Engineering*, vol. 13, no. 2, pp. 121–136, Jun 2018.

[2] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee, and B. Yin, "Smart factory of industry 4.0: Key technologies, application case, and challenges," *IEEE Access*, vol. 6, pp. 6505–6519, 2018.

[3] C. A. Ericson, *Hazard Analysis Techniques for System Safety, 2 edition*. John Wiley & Sons, 2005.

[4] The Assurance Case Working Group, "Goal Structuring Notation Community Standard Version 2," 2018.

[5] H. Giese and M. Tichy, "Component-based hazard analysis: Optimal designs, product lines, and online-reconfiguration," in *25th International Conference on Computer Safety, Reliability, and Security (SAFECOMP), Gdansk, Poland, September 27-29*, 2006, pp. 156–169.

[6] C. Priesterjahn, C. Heinzemann, W. Schäfer, and M. Tichy, "Runtime safety analysis for safe reconfiguration," in *IEEE 10th International Conference on Industrial Informatics (INDIN), Beijing, China, July 25-27*, 2012, pp. 1092–1097.

[7] C. Priesterjahn, D. Steenken, and M. Tichy, "Timed hazard analysis of self-healing systems," in *Assurances for Self-Adaptive Systems - Principles, Models, and Techniques*, 2013, pp. 112–151.

[8] N. Bhardwaj and P. Liggesmeyer, "A runtime risk assessment concept for safe reconfiguration in open adaptive systems," in *Computer Safety, Reliability, and Security (SAFECOMP) 2017 Workshops, Trento, Italy, September 12*, 2017, pp. 309–316.

[9] M. M. Cowing, M. Paté-Cornell, and P. W. Glynn, "Dynamic modeling of the tradeoff between productivity and safety in critical engineering systems," *Reliab. Eng. Syst. Saf.*, vol. 86, no. 3, pp. 269–284, 2004.

[10] G. Despotou and T. Kelly, "An argument-based approach for assessing design alternatives and facilitating trade-offs in critical systems," *Journal of System Safety*, vol. 43, no. 2, p. 22, 2007.

[11] I. Habli and T. Kelly, "Challenges of establishing a software product line for an aerospace engine monitoring system," in *11th International Software Product Line Conference (SPLC), Kyoto, Japan, September 10-14*, 2007, pp. 193–202.

[12] I. Habli, T. Kelly, and R. Paige, "Functional hazard assessment in product-lines – a model-based approach," in *1st International Workshop on Model-Driven Product Line Engineering (MDPLE' 09), Twente, The Netherlands, June 24th*, 2009.

[13] Z. Stephenson, C. Fairburn, G. Despotou, T. P. Kelly, N. Herbert, and B. Daughtrey, "Distinguishing fact from fiction in a system of systems safety case," in *Advances in Systems Safety–Proceedings of the Nineteenth Safety-Critical Systems Symposium (SSS), Southampton, UK, February 8-10*, 2011, pp. 55–72.

[14] A. L. de Oliveira, R. T. V. Braga, P. C. Masiero, Y. Papadopoulos, I. Habli, and T. Kelly, "Supporting the automated generation of modular product line safety cases," in *Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX, Brunów, Poland, June 29 - July 3*, 2015, pp. 319–330.

[15] I. Sljivo, B. Gallina, J. Carlson, and H. A. Hansson, "Configuration-aware contracts," in *4th International Workshop on Assurance Cases for Software-intensive Systems (ASSURE '16), Trondheim, Norway, September 20*, 2016, pp. 43–54.

[16] D. Nesic, M. Nyberg, and B. Gallina, "Constructing product-line safety cases from contract-based specifications," in *34th ACM/SIGAPP Symposium on Applied Computing (SAC), Limassol, Cyprus, April 8-12*, 2019, pp. 2022–2031.

[17] F. U. Muram, M. A. Javed, and S. Punnekkat, "System of systems hazard analysis using HAZOP and FTA for advanced quarry production," in *2019 4th International Conference on System Reliability and Safety (ICSRS), Rome, Italy, November 20-22*, 2019, pp. 394–401.

[18] L. Xing and S. V. Amari, "Fault tree analysis," in *Handbook of Performability Engineering*, K. B. Misra, Ed. London: Springer London, 2008, ch. 38, pp. 595–620.

[19] Object Management Group (OMG), "Structured Assurance Case Metamodel (SACM), Version 2.1," 2019.

[20] F. U. Muram, B. Gallina, and L. G. Rodriguez, "Preventing omission of key evidence fallacy in process-based argumentations," in *11th International Conference on the Quality of Information and Communications Technology (QUATIC), Coimbra, Portugal, September 4-7*, 2018, pp. 65–73.

[21] G. Fornari and V. A. de Santiago Júnior, "Dynamically reconfigurable systems: A systematic literature review," *Journal of Intelligent and Robotic Systems*, vol. 95, no. 3-4, pp. 829–849, 2019.

[22] M. A. Javed and B. Gallina, "Safety-oriented process line engineering via seamless integration between EPF composer and BVR tool," in *22nd International Systems and Software Product Line Conference (SPLC) - Volume 2, Gothenburg, Sweden, September 10-14*, 2018, pp. 23–28.

[23] M. A. Javed, B. Gallina, and A. Carlsson, "Towards variant management and change impact analysis in safety-oriented process-product lines," in *34th ACM/SIGAPP Symposium on Applied Computing (SAC), Limassol, Cyprus, April 8-12*, 2019, pp. 2372–2375.

[24] A. Vasilevskiy, Ø. Haugen, F. Chauvel, M. F. Johansen, and D. Shimbara, "The BVR tool bundle to support product line engineering," in *19th International Conference on Software Product Line (SPLC), Nashville, TN, USA, July 20-24*, 2015, pp. 380–384.

[25] Volvo Construction Equipment, "Emission-free quarry," Available at https://www.volvoce.com/global/en/news-and-events/press-releases/2018/testing-begins-at-worlds-first-emission-free-quarry/.

[26] M. A. Javed, F. U. Muram, A. Fattouh, and S. Punnekkat, "Enforcing geofences for managing automated transportation risks in production sites," in *Dependable Computing - EDCC 2020 Workshops*, 2020, pp. 113–126.

[27] J. Martinez, T. Ziadi, T. F. Bissyandé, J. Klein, and Y. L. Traon, "Automating the extraction of model-based software product lines from model variants (T)," in *30th IEEE/ACM International Conference on Automated Software Engineering (ASE), Lincoln, NE, USA, November 9-13*, 2015, pp. 396–406.

Appendix B: Contributions to the Body of Knowledge

STPA – Challenges to apply in System of Systems(SoS) Hazard Analysis

Stephan Baumgart (VCE) & Sasikumar Punnekkat (Mälardalen University)

Automation of earth moving machinery enables improving existing production workflows in various applications like surface mines, material handling operations or material transporting. They are used in fleets and integrated with other machines. Such connected and collaborating autonomous machines can be seen as a system-of-systems. It is not yet clear how to consider safety during the development of such system-of-systems (SoS). One potentially useful approach to analyze the safety for complex systems is the System Theoretic ProcessAnalysis (STPA). However, STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations.

Industrial Case - Electric Site

We utilize the electric site research project [1] as a use case for our work. In this project a fleet of automated guided vehicles (AGVs) [2] called HX are used to transport material at a quarry site, which is a surface mine for gravel production in our case. The pre-crushed material is transported from a movable primary crusher to a stationary secondary crusher. Along with the fleet of autonomous HX, a human-operated wheel loader and a human-operated excavator are used for loading material onto the HX. In our earlier work we have described and analyzed this complex SoS [3][4].
The fleet of active HX is controlled by the Fleet Control System, containing features like traffic management or setting missions for each active HX. Each HX is therefore highly dependent on the wireless network and correct commands. In order to be able to activate a HX in the morning, remove a HX for repair purposes or adding a HX to a running production, it is possible at any given instance to control a single HX using a remote control by a HX Remote Operator. The Site Operator is monitoring the quarry site from a control room, where the Site Server is located. In Figure 1 the involved systems and human operators are presented. When designing such a system an in-depth analysis of this scenario is necessary to identify potential hazards leading to critical accidents.

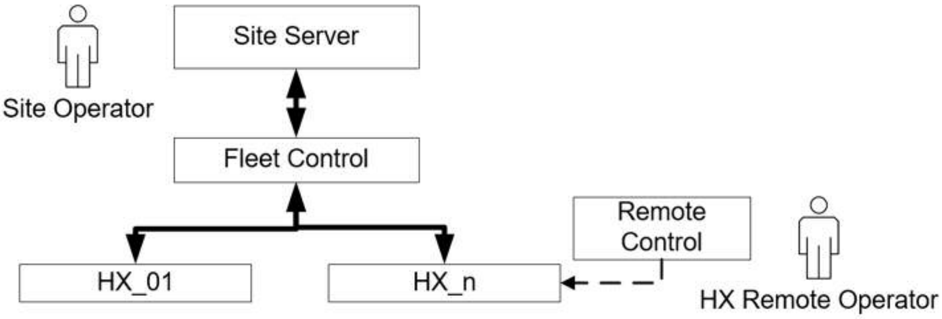


Figure 1 Use Case: Remote Control of HX

System-Theoretic Process Analysis - STPA

To illustrate the application of STPA, we analyze the remote control case and follow the STPA process as described in literature [5].

STPA Overview

At first we provide a short description of STPA.

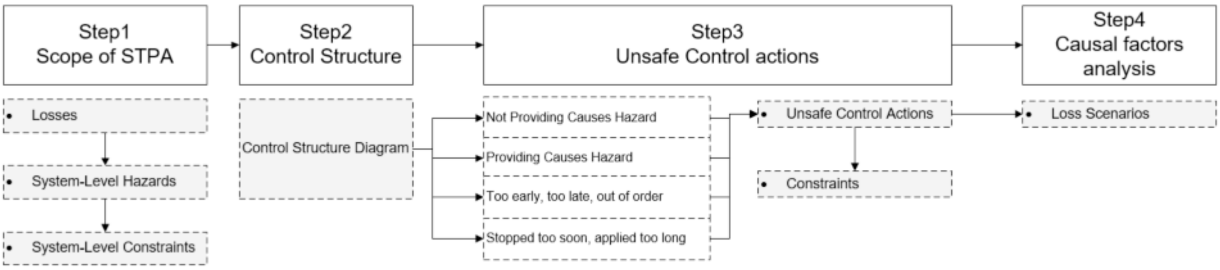


Figure 2 General STPA Process as described in [5]

STPA consists of four steps as shown in Figure 2 which we describe in the following section.

STPA - Step 1: During the first step of STPA, the scope of the STPA is set and potential losses and hazards shall be identified. System-level hazards may be derived in brainstorming meetings with experts or by applying hazard identification methods like HAZOP or What-if Analysis. The list of possible system hazards may be extended during later stages when more product knowledge is available.

STPA - Step 2: In Step 2, the control structure of the system is derived. The control structure diagram is a graphical representation of the control actions to aid a structured analysis.
The control structure diagram contains the main control elements and control actions between the controllers and the controlled systems.

STPA - Step 3: The control structure diagram is used to apply a structured analysis of each control action and if a failure of the control action would lead to the already listed system-level hazards. STPA uses four guide words for finding such unsafe control actions:

- Not providing causes hazard
- Providing causes hazard
- Too early, too late, out of order
- Stopped too soon, applied too long

This means that the following requirements are tested:

- A correct control action is provided.
- A control action is provided at the correct time.
- A control action is provided with correct duration.

STPA - Step 4:
In the last step of STPA, possible loss scenarios are identified for each unsafe control action. Reasoning why an unsafe control action would occur and how this could lead to a hazard shall be provided.

STPA - Conclusion
STPA is useful for identifying and analyzing control actions and their causal factors when unsafe control actions are identified. The process of STPA is foreseen to be iterative, i.e. it is possible that further system-level or subsystem-level hazards will be identified during later stages. It is furthermore proposed to add complexity to the control structure diagram during later stages of the development process. This will lead to additional efforts for identifying unsafe control actions in Step 3.

The question is, if STPA is able to deal with emergent and dysfunctional behaviors in the case of system-of-systems. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations.

STPA - Application Remote Control Case

In the following we apply STPA to the industrial case described above.
STPA Step 1 - Remote Control Case:
For our limited case we have identified two major losses that shall be avoided:

- Loss1: Humans injured or killed
  Situations, where humans are at risk to be injured or killed by the autonomous machines shall be avoided.
- Loss2: Damage of Equipment
  If machines are damaged because of accidents, this may result in a stop of production at the site, which shall be avoided.

Typical SoS hazards in our case can be:

- Hazard 1 (H-1): HX does not maintain safe distance to humans on Site.
- Hazard 2 (H-2): HX enters dangerous area/region
- Hazard 3 (H-3): Squeezing Hazard (e.g. people close to HX)
- Hazard 4 (H-4): Insufficient ability of machinery to be slowed down, stopped and immobilized

STPA Step 2 - Remote Control Case:
We simplified the control structure diagram for the purpose of this paper as shown in Figure 3

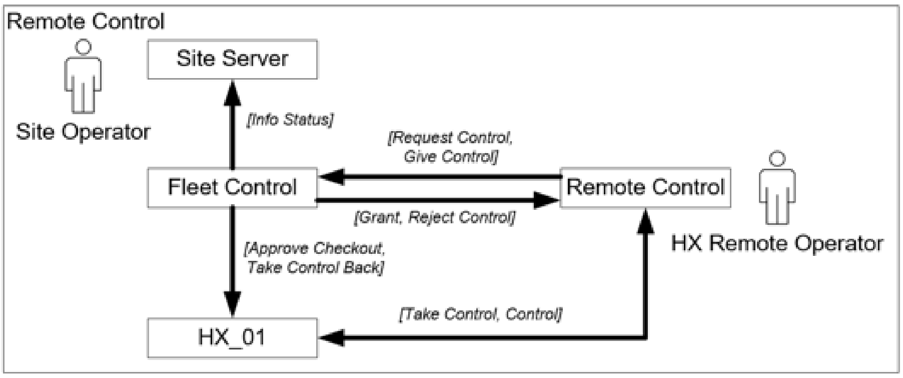


Figure 3 Control Structure Diagram: Remote Control HX 01

The HX Remote Operator sends a request to the Fleet Control server with the purpose to take over the control of a specific HX (HX 01). Fleet Control can decide either to accept (Grant Control) or to reject (Reject Control) the request. At the same time the Fleet Control is sharing information about the active HX with the site server shown by the message Info Status.
If the remote control request is accepted, Fleet Control is sending a task (Approve Checkout) to HX 01 to enable the HX to be controlled by the Remote Control. Once this is done, the HX Remote Operator can take control over the HX. The HX Remote Operator can also give back control of HX 01 to Fleet Control. Fleet Control will send a request (Take Control Back) to HX 01 that it will listen to controls send from Fleet Control.

STPA Step 3 - Remote Control Case:
Each message in the control structure diagram Figure 3 is analyzed using the guide words.

Table 1 Unsafe Control Actions: Remote control case

Control Action	Not providing Causes Hazard	Providing Causes Hazard	Too Early, too late, out of order	Stopped too soon, applied too long
Request Control	Request Control is not provided to Fleet Control [Not Hazardous]	UCA 01: Request Control is sent unintendedly during normal operation. [H-2, H-3]	Request from HX Remote Operator is provided too late. [Not Hazardous]	
Approve Checkout	Approve Checkout is not provided to HX. [Not Hazardous]	UCA 02: Approve Checkout is provided unintended to HX during normal operation. [H-1, H-2, H-3, H-4]		

We exemplify identifying unsafe control actions by analyzing the messages ”Request Control” and ”Approve Checkout” in Table 1. Applying the first guide word *Not providing causes hazard* for ”Request Control” helps finding the critical situations if the message is either not provided or lost, but this will not directly lead to a hazard. We identify the first unsafe control action (UCA 01) in the situation when the message ”Request Control” is provided unintended. This may lead to a situation that a HX is checked out from Fleet Control without awareness of the HX Remote Operator. Humans are at risk, if the machine is moving into dangerous areas, where humans are working (H-2) or if humans are already close by, this may lead to squeezing hazards (H-3). If the signal is delayed (Too early, too late, out of order), this may lead in the worst case to frustration of the operator, but not to hazardous situations.

The message ”Approve Checkout” is send from the Fleet Control to the HX to indicate, that the HX shall change mode to be controlled by a remote control. We identify, that providing ”Approve Checkout” unintended, will lead to a situation where the HX is forced to switch over to be remote controlled. This can lead to critical situations where the HX is moving without a control instance connected to the machine.

Altogether, we have identified 15 UCAs for this simplified case during the first brainstorming.

STPA Step 4 - Remote Control Case:

In our case, ``Approve Checkout'' might be provided unintended because of a fault in the Fleet Control software or due to a transmission error.

Conclusion STPA Case Study

Where is STPA suitable?

STPA is a useful approach to analyze the safety of complex systems. While hazard analysis methods like PHA, FTA and FMEA focus on failures of system functions and their impact, is STPA analyzing possible failures of control actions between the involved systems and sub-systems. This analysis leads to a broader list of possible critical scenarios that require further analysis to list all causal factors.

STPA is analyzing the control actions and therefore mostly communication related hazards will be identified.

Which critical situations are not captured in STPA?

STPA analyzes one single control action a time, which makes it impossible to find critical scenarios which involve for example a combination of control actions, cascading failures or state changes. STPA is essentially suitable to static monolithic systems and lacks the ability to deal with emergent and dysfunctional behaviors in the case of SoS. These behaviors if not identified could potentially lead to hazards and it is important to provide mechanisms for SoS developers/integrators to capture such critical situations. It is among others important to check, if the involved systems in a SoS have a consistent perspective of the global state. The states of the involved systems are not considered in the control structure diagram of STPA. Design flaws and casual factors might be missed, if the interaction of state machines is not considered during analysis of the SoS.

Further details are available in [6].

References

[1] Volvo Construction Equipment, “Electric Site Project.” [Online]. Available: https://www.volvoce.com/global/en/news- and- events/news-and- press-releases/2018/carbon- emissions- reduced- by- 98- at- volvo-construction- equipment- and-skanskas- electric- site/

[2] D. Weyns, T. Holvoet, and K. Schelfthout, “Decentralized control of automatic guided vehicles: applying multi-agent systems in practice,” Companion to the 23rd, 2008. [Online]. Available: http: //dl.acm.org/citation.cfm?id=1449819

[3] S. Baumgart, J. Froberg, and S. Punnekkat, “Analyzing hazards in system-of-systems: Described in a quarry site automation context,” in 2017 Annual IEEE International Systems Conference (SysCon). IEEE, 4 2017, pp. 1–8. [Online]. Available: http://ieeexplore.ieee.org/ document/7934783/

[4] S. Baumgart, J. Froberg, and S. Punnekkat, “Can STPA be used for a System-of-Systems? Experiences from an Automated Quarry Site,” in 2018 IEEE International Systems Engineering Symposium (ISSE), no. 4. IEEE, 10 2018, pp. 1–8. [Online]. Available: http://www.es.mdh.se/publications/5246-https: //ieeexplore.ieee.org/document/8544433/

[5] N. G. Leveson and J. P. Thomas, STPA Handbook, 2018.

[6] S. Baumgart, J. Fröberg and S. Punnekkat, "A State-based Extension to STPA for Safety-Critical System-of-Systems," 4th International Conference on System Reliability and Safety (ICSRS), Rome, Italy, 2019, pp. 246-254, doi: 10.1109/ICSRS48664.2019.8987632.

BoK 3.1 - Identifying potential deviation from required behaviour

Practical guidance – End-to-End Tool Framework for Safety Analysis

Authors: Faiz Ul Muram, Muhammad Atif Javed and Sasikumar Punnekkat (Mälardalen University, SUCCESS Project)

The principal objective of system safety is the identification of hazardous events and their causes, mechanisms for elimination or mitigation of causes, and documentation of evidences for safety cases. This is primarily done during system design and development phase [1]. To demonstrate the acceptable safety of production operations, safety cases are constructed that provide comprehensive, logical and defensible justification of the safety of a production system for a given application in a predefined operating environment. A safety case consists of process-based arguments that can show processes generate trustworthy evidence and product-based arguments that may directly show from the evidence that residual risks for the product are acceptably low. In this regard, contracts can be used for which the behaviour of the component can be described in a way that the component makes assumptions (conditions) on its environment and if those assumptions hold then the component will behave as guaranteed (offers properties). A contract that describes only safety-related properties is referred to as a safety contract. However, the parts of safety cases constructed during system design and development phase may turn out to be incorrect, inapplicable or insufficient during the operation. This can be caused by emergent behaviours and changes performed in consequence of market demands, hazardous conditions or system failures, thus necessitating some means for dynamic safety assurance in these contexts. The safety arguments constructed at system design and development phases might be invalidated during operation. The automated systems and flexible manufacturing underlines the need for update of safety arguments to respond to the observed reality [13]. The safety contracts are used to support the maintenance and adaptation of safety cases based on operational deviations. Jaradat and Punnekkat [8] also exploit safety contracts for monitoring the failure rates during operational phase.

In the SUCCESS project, we have proposed guidelines with focus on end-to-end traceability and support of a tool framework that can provide a significant boost for the designers to avoid the culture of paper safety at the expense of actual system safety [14]. We build upon the end-to-end tool framework for safety analysis different tools, integrated in AMASS [2] platform that facilitate (i) modelling of standards and safety process in **EPF Composer**[1], (ii) modelling of systems, contract-based design and different model-based analyses in **Polarys CHESS toolset**[2], (iii) assurance case modelling in **Polarys OpenCert tools platform**[3], and (iv) formal verification of assumption guarantee contracts with **OCRA**[4].

The end-to-end tool framework for safety analysis consists of following steps. Figure 1 shows the framework for safety analysis along with integration of involved tools.

1. Standards requirements and process modelling in EPF Composer
2. Detecting fallacies in process models
3. Generation of process-based arguments
4. System modelling in CHESS
5. Contract refinement
6. Generation of product-based arguments
7. Update of argument fragments during operational phase

[1] https://www.eclipse.org/epf/
[2] https://www.polarsys.org/chess
[3] https://www.polarsys.org/projects/polarsys.opencert
[4] https://ocra.fbk.eu/

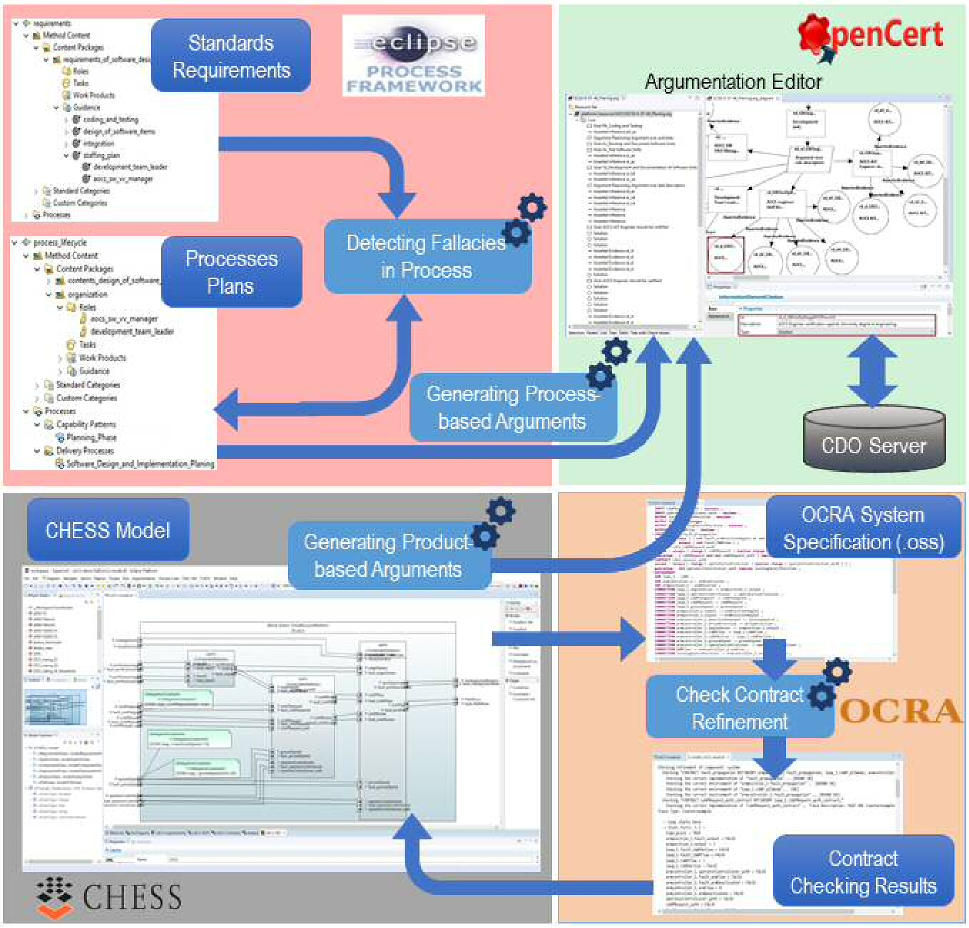
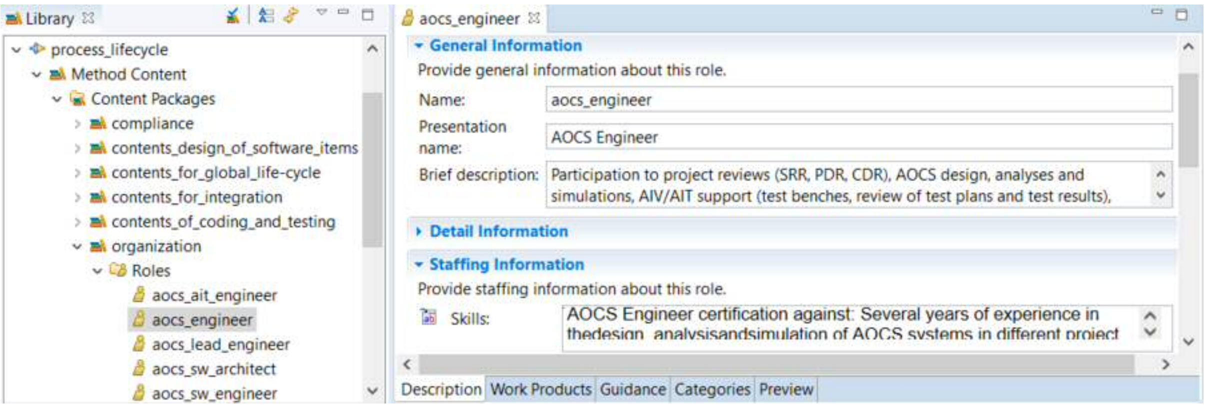

Standards Requirements
eclipse PROCESS FRAMEWORK
OpenCert
Argumentation Editor
Detecting Fallacies in Process
Processes Plans
Generating Process-based Arguments
CDO Server
CHESS Model
Generating Product-based Arguments
OCRA System Specification (.oss)
Check Contract Refinement
OCRA
Contract Checking Results
CHESS

Figure 1. Overview of the framework

1 Standards Requirements and Process Modelling in EPF Composer

EPF (Eclipse Process Framework) Composer is an extensible process framework, based on the Unified Method Architecture (UMA) metamodel, which covers most of the Process Engineering Metamodel (SPEM) 2.0 [6] concepts. EPF Composer has been ported from Eclipse Galileo 3.5.2 to Eclipse Neon 4.6.3 in the context of the AMASS project [3]. EPF Composer is used to model the requirements listed in the standards and safety plans, as well as to show the basic compliance, as presented in [4, 7]. In EPF Composer, *method plugins* are containers of process related information (i.e., Method Content and Processes), while a *configuration* is a selection of sub-sets of library content to be shown in the browsing perspective. To model the standards requirements, the guidance type *Practice* can be customized with an icon in a separate plugin.

The *requirements* plugin represents the standard's requirements and has the variability relationship *Extends* with the previously mentioned plugin. The *process lifecycle* plugin describes the development process (i.e., content elements, processes and categories). To model the qualifications of a role who has the responsibilities to perform relevant tasks correctly and efficiently the *Staffing Information -- Skills* field is used (see Figure 2), whereas to specify certifications or rationales against required tool qualifications *Detail Information -- Key considerations* field of Tools is used. If there are more than one evidence or rationale that correspond to requirements, a semicolon (;) should be used to separate them [5]. To manage basic compliance, standard requirements are copied in the separate plugin. These copied requirements have a variability relationship *Contributes* with original requirements modelled in *requirements* plugin. In addition, the links between process elements (such as tasks, activity) to each *standard requirement* have been established through *References* tab. The outcomes of processes, which are represented in the EPF Composer as work products, provide evidence supporting process and product argumentation.



Provide general information about this role.

Name: aocs_engineer

Presentation name: AOCS Engineer

Brief description: Participation to project reviews (SRR, PDR, CDR), AOCS design, analyses and simulations, AIV/AIT support (test benches, review of test plans and test results),

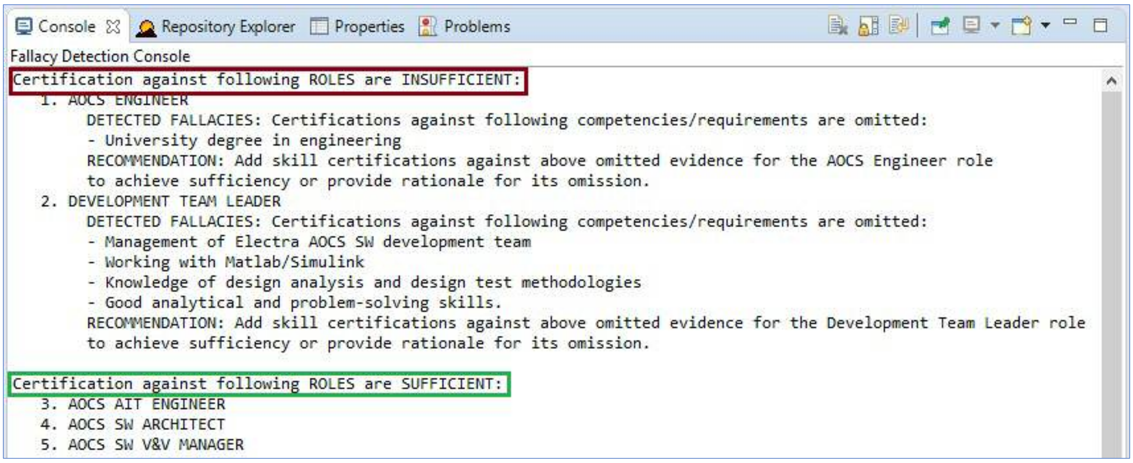
Provide staffing information about this role.

Skills: AOCS Engineer certification against: Several years of experience in thedesign analysisandsimulation of AOCS systems in different project

Figure 2. Modelling of role in EPF Composer

2 Detecting Fallacies in Process Models

A fallacy is a mistake or flaw in the reasoning of an argument. In safety arguments, fallacies exist in different forms. A taxonomy of common fallacies in safety arguments is presented and organized them into three categories namely, relevance, acceptability and sufficiency fallacies [11]. *Relevance* fallacies add no value to an argument and provide irrelevant evidence. *Acceptability* fallacies are those in which an argument provides the unacceptable, contradict or inconsistent evidence to support the claims. Sufficiency fallacies, particularly, *omission of key evidence* in which no or less evidences are provided to support the claim or no valid reasons (rationales) are given for its omission [7]. These fallacies could lead to overconfidence in a system and tolerate certain faults, which in turn contribute to safety-related failures of the system. This risk also affects the process-based arguments. Therefore, a plugin is implemented that performed a validation to detect whether the safety process modelled in EPF Composer contains the sufficient information corresponding to the key evidence for supporting the specific requirement. In case of omitted crucial evidence detail, the feedback is provided regarding detected fallacies and recommendations to resolve them (see Figure 3). Fallacy Detection plugin is invoked from a right-click menu on the ProcessComponent (*Capability Pattern* or *Delivery Process*). The results are printed on the console as well as the TXT file(s) is generated in the selected the target directory. The process can be modified based on the provided recommendations.



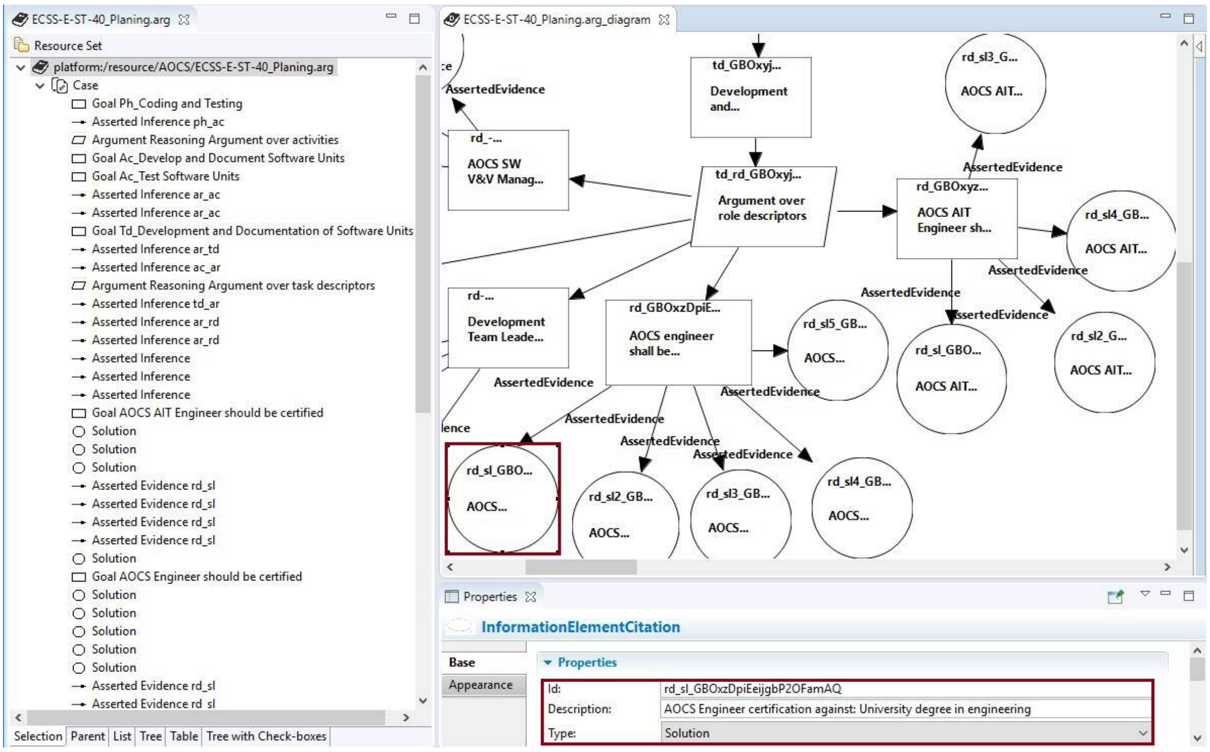
```
Certification against following ROLES are INSUFFICIENT:
   1. AOCS ENGINEER
        DETECTED FALLACIES: Certifications against following competencies/requirements are omitted:
        - University degree in engineering
        RECOMMENDATION: Add skill certifications against above omitted evidence for the AOCS Engineer role
        to achieve sufficiency or provide rationale for its omission.
   2. DEVELOPMENT TEAM LEADER
        DETECTED FALLACIES: Certifications against following competencies/requirements are omitted:
        - Management of Electra AOCS SW development team
        - Working with Matlab/Simulink
        - Knowledge of design analysis and design test methodologies
        - Good analytical and problem-solving skills.
        RECOMMENDATION: Add skill certifications against above omitted evidence for the Development Team Leader role
        to achieve sufficiency or provide rationale for its omission.

Certification against following ROLES are SUFFICIENT:
   3. AOCS AIT ENGINEER
   4. AOCS SW ARCHITECT
   5. AOCS SW V&V MANAGER
```

Figure 3. Printed results on the console

3 Generating Process-based Arguments

The Structured Assurance Case Metamodel (SACM) [12] is the Object Management Group (OMG) standard that integrates and standardizes the broadly used notations for documenting safety (assurance) cases, including Goal Structuring Notation (GSN) and Claims-Arguments-Evidence (CAE). OpenCert is an open source assurance and certification tool; its *argumentation editor* is based on the GSN graphical notations. However, Common Assurance and Certification Metamodel (CACM) implemented in OpenCert internally uses the SACM metamodel. The safety cases can be stored in the workspace directory, or in the Connected Data Objects (CDO)[5], which is both a development-time model repository and a run-time persistence framework. The repository sessions provide support for obtaining and modifying them.

The Process-based Argument Generator plugin takes the ProcessComponent (Capability Pattern or Delivery Process) modelled in EPF Composer as an input and transforms it into arguments (model and diagram). The generated process-based argument model and diagram (see Figure 4) are saved locally in a new project into the current workspace under the name *Argumentation*. They are also stored in the corresponding destination assurance case in the CDO server under the *ARGUMENTATION* folder.


ECSS-E-ST-40_Planing.arg

Resource Set

platform:/resource/AOCS/ECSS-E-ST-40_Planing.arg

Case

- Goal Ph_Coding and Testing
- Asserted Inference ph_ac
- Argument Reasoning Argument over activities
- Goal Ac_Develop and Document Software Units
- Goal Ac_Test Software Units
- Asserted Inference ar_ac
- Asserted Inference ar_ac
- Goal Td_Development and Documentation of Software Units
- Asserted Inference ar_td
- Asserted Inference ac_ar
- Argument Reasoning Argument over task descriptors
- Asserted Inference td_ar
- Asserted Inference ar_rd
- Asserted Inference ar_rd
- Asserted Inference
- Asserted Inference
- Asserted Inference
- Goal AOCS AIT Engineer should be certified
- Solution
- Solution
- Solution
- Asserted Evidence rd_sl
- Asserted Evidence rd_sl
- Asserted Evidence rd_sl
- Asserted Evidence rd_sl
- Solution
- Goal AOCS Engineer should be certified
- Solution
- Solution
- Solution
- Solution
- Solution
- Asserted Evidence rd_sl
- Asserted Evidence rd_sl

ECSS-E-ST-40_Planing.arg_diagram

Properties

InformationElementCitation

Properties

Id:	rd_sl_GBOxzDpiEeijgbP2OFamAQ
Description:	AOCS Engineer certification against: University degree in engineering
Type:	Solution

Figure 4. Generated argument model and diagram

4 System Modelling in CHESS

In the PolarSys CHESS (Composition with Guarantees for High-integrity Embedded Software Components Assembly) toolset, an editor is implemented to model all phases of system development, for instance, requirements definition, software architecture modelling and its deployment to hardware. In particular, SysML Block Definition Diagram (BDD) and Internal Block Diagram (IBD) can be used to model the system hierarchical architecture, for instance, blocks, ports and connections (see Figure 5). A CHESS profile and tool can also be used to exploit different analysis at system level.

[5] https://www.eclipse.org/cdo/

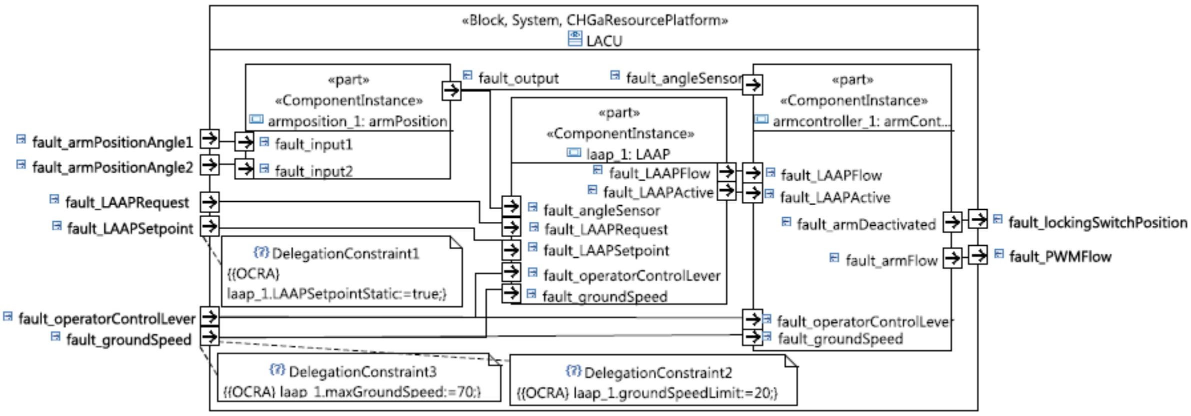
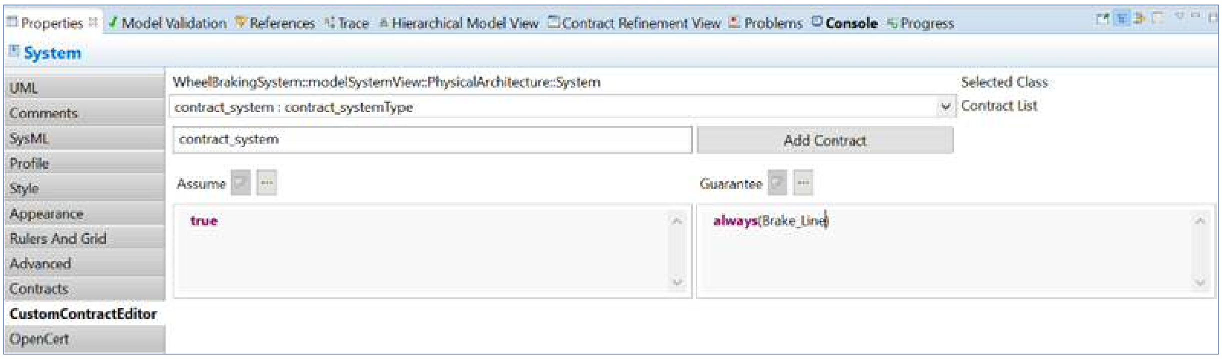


Figure 5. The CHESS Internal Block Diagram (taken from [10])

CHESS toolset supports the modelling of contracts (i.e. the assumption and the guarantee properties) and their association with components and system requirements, as shown in Figure 6. Contracts can either created in a BDD and in a Class/Component Diagram from *Contracts* palette (see Figure 7) or otherwise without its graphical representation by using *ContractEditor+*. In particular, the assumption and the guarantee properties are specified in formal language such as temporal logic. The contract specification can further be enriched by categorising contracts into strong and weak to allow for better support for specification of reusable components behaviour [10]. After that, the contract refinement is performed.



WheelBrakingSystem::modelSystemView::PhysicalArchitecture::System

Selected Class

contract_system : contract_systemType

Contract List

contract_system

Assume

true

Guarantee

always(Brake_Line)

Figure 6. Assume and Guarantee Contract (taken from [5])

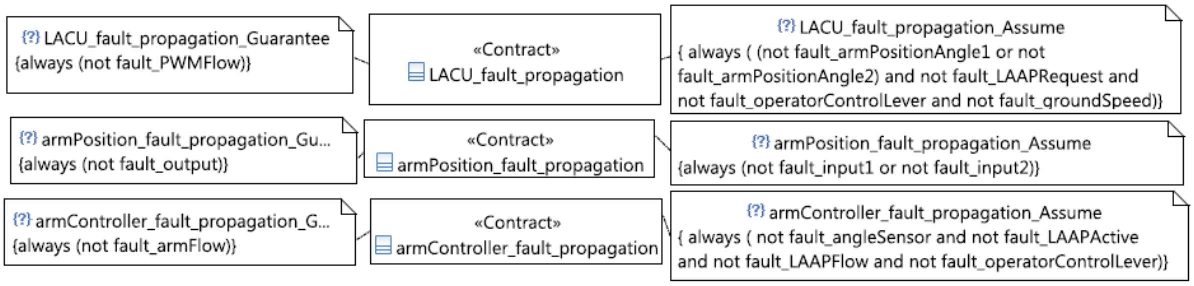
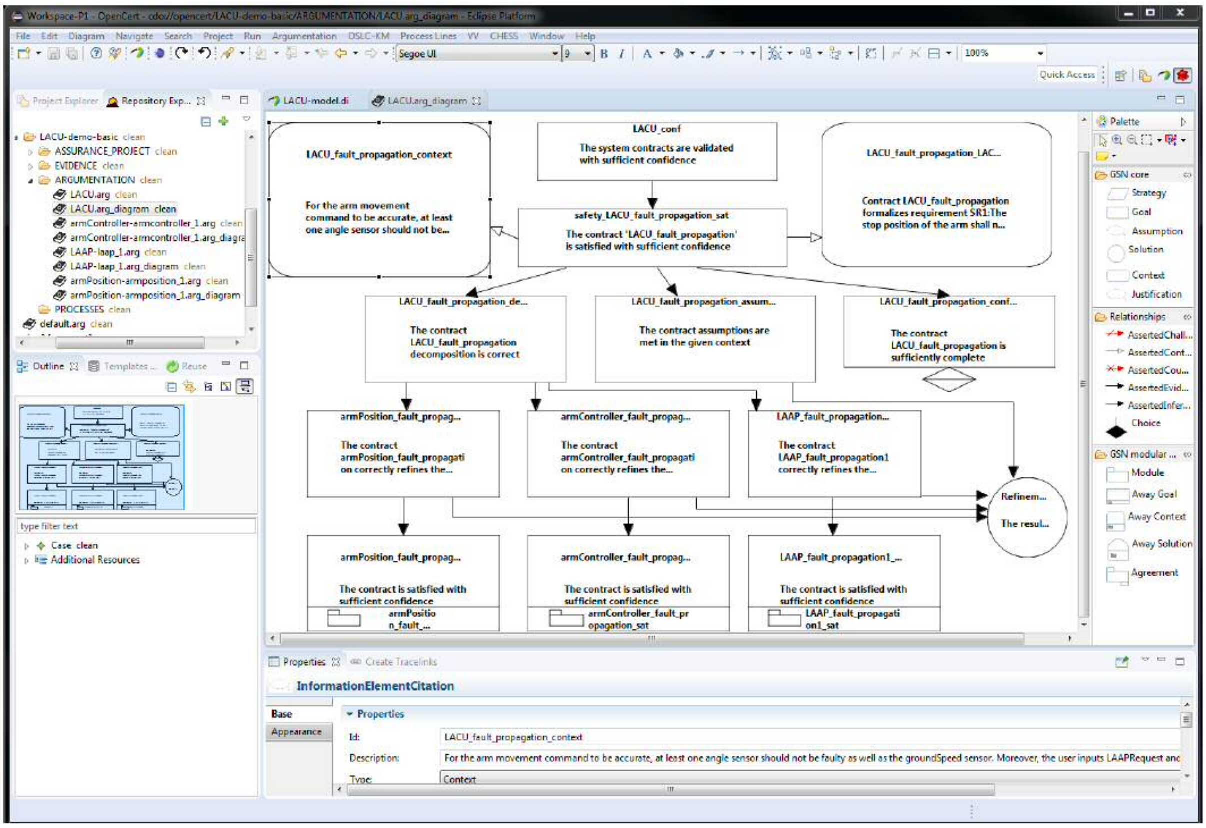

LACU_fault_propagation_Guarantee
{always (not fault_PWMFlow)}
«Contract»
LACU_fault_propagation
LACU_fault_propagation_Assume
{ always ((not fault_armPositionAngle1 or not fault_armPositionAngle2) and not fault_LAAPRequest and not fault_operatorControlLever and not fault_groundSpeed)}
armPosition_fault_propagation_Gu...
{always (not fault_output)}
«Contract»
armPosition_fault_propagation
armPosition_fault_propagation_Assume
{always (not fault_input1 or not fault_input2)}
armController_fault_propagation_G...
{always (not fault_armFlow)}
«Contract»
armController_fault_propagation
armController_fault_propagation_Assume
{ always (not fault_angleSensor and not fault_LAAPActive and not fault_LAAPFlow and not fault_operatorControlLever)}

Figure 7. The strong contracts specified in CHESS (taken from [10])

5 Contract Refinement Checking

Othello Contracts Refinement Analysis (OCRA) is a command-line tool that provides means for checking the refinement of contracts specified in a linear-time temporal logic. The integration of CHESS with OCRA verification engine allows the validation of component contract assumptions against the specification of other components in the system. The CHESS model together with contracts are transformed into an OCRA System Specification (.oss) file readable by OCRA. The contract refinement checking is done by OCRA and the result is back-propagated to the CHESS model. OCRA runs in background or remotely via OSLC (Open Services for Lifecycle Collaboration), and therefore, the user does not interact with them directly. However, OCRA does not distinguish between strong and weak contracts, therefore *weak contract filtering* as a part of reusable component instantiation or weak contract transformation to an appropriate format is supported [7].



LACU_fault_propagation_context

For the arm movement command to be accurate, at least one angle sensor should not be...

LACU_conf

The system contracts are validated with sufficient confidence

LACU_fault_propagation_LAC...

Contract LACU_fault_propagation formalizes requirement SR1:The stop position of the arm shall n...

safety_LACU_fault_propagation_sat

The contract 'LACU_fault_propagation' is satisfied with sufficient confidence

LACU_fault_propagation_de...

The contract LACU_fault_propagation decomposition is correct

LACU_fault_propagation_assum...

The contract assumptions are met in the given context

LACU_fault_propagation_conf...

The contract LACU_fault_propagation is sufficiently complete

armPosition_fault_propag...

The contract armPosition_fault_propagation correctly refines the...

armController_fault_propag...

The contract armController_fault_propagation correctly refines the...

LAAP_fault_propagation...

The contract LAAP_fault_propagation1 correctly refines the...

Refinem...

The resul...

armPosition_fault_propag...

The contract is satisfied with sufficient confidence

armPosition_fault_...

armController_fault_propag...

The contract is satisfied with sufficient confidence

armController_fault_propagation_sat

LAAP_fault_propagation1_...

The contract is satisfied with sufficient confidence

LAAP_fault_propagation1_sat

InformationElementCitation

Properties

Id: LACU_fault_propagation_context

Description: For the arm movement command to be accurate, at least one angle sensor should not be faulty as well as the groundSpeed sensor. Moreover, the user inputs LAAPRequest and

Type: Context

Figure 8. Generated argument-fragment in OpenCert (taken from [10])

6 Generating Product-based Arguments

The safety case (argument-fragments) can be generated from the selected CHESS model (contract-based architectural specification), as shown in Figure 1Figure 8. Argument Generator plugin implemented in OpenCert assumes that the analysed model and the refinement check results are stored in the refinement analysis context [10]. The generated set of argument-fragments stored in the corresponding destination assurance case in the CDO server stated in the OpenCert preferences. The argument-generation is performed for each component and for each validated contract. The set of argument-fragments for each component and applied architectural pattern can be viewed in the selected assurance case.

7 Update of Argument Fragments during Operational Phase

The results gained from the hazard analysis are utilized to derive the safety requirements and safety contracts. The safety contracts derived for uncertainty sources are associated with the safety cases modelled in the OpenCert platform. This provides the means to detect deviations from intended behaviour and perform necessary adaptations at the operational phase. In particular, for the identification and resolution of gaps between the intended behaviour reflected in safety arguments and the actual safety of production operations, the operational data is utilised [9, 15]. Based on the gathered data, the safety contracts constructed for uncertainty sources are monitored, deviations between the intended and actual behaviour are tracked and evaluated, and the safety contracts and safety cases are updated. Their update is carried out based on the optimal actions for which the thresholds regarding the performance degradation and upgradation are specifically taken into consideration [16]. In contrast to the matching of parameter names and their values/ranges for safety contracts, the text-based matching is performed to alter the description of safety case elements. The required changes in safety cases are tracked and then the update command is issued. The assurance (safety) cases are updated on the CDO server, which is accessed by the OpenCert argumentation editor connected to the CDO server.

Environment Setup

The tools used in this guidance are integrated in the OpenCert bundle that can be downloaded from the following link https://www.eclipse.org/opencert/downloads/. For more detailed information, a user manual of the particular tools and a developers' guide to set up the workspaces is provided in [5].

References

1. C. A. Ericson, Hazard Analysis Techniques for System Safety, 2 edition, John Wiley & Sons, 2015.
2. AMASS - Architecture-driven, Multi-concern and Seamless Assurance and Certification of Cyber-Physical Systems, https://amass-ecsel.eu/
3. M.A. Javed, B. Gallina, "Get EPF Composer back to the future: A trip from Galileo to Photon after 11 years". EclipseCon, Toulouse, France, June 13-14, 2018.
4. F.U. Muram, B. Gallina, L. G. Rodriguez, "Preventing Omission of Key Evidence Fallacy in Process-based Argumentations". In: 11th International Conference on the Quality of Information and Communications Technology (QUATIC), Coimbra, Portugal, September 4-7, 2018. pp. 65–73
5. AMASS User guidance and Methodological framework D2.5. https://www.amass-ecsel.eu/sites/amass.drupal.pulsartecnalia.com/files/D2.5_User-guidance-and-methodological-framework_AMASS_Final.pdf (2018).
6. Object Management Group: Software & Systems Process Engineering Metamodel Specification (SPEM), Version 2.0. http://www.omg.org/spec/SPEM/2.0/ (2008).
7. McIsaac, B.: IBM Rational Method Composer: Standards Mapping. Tech. rep., IBM Developer Works (2015).
8. O. Jaradat, S. Punnekkat, "Using safety contracts to verify design assumptions during runtime, in: 23rd International Conference on Reliable Software Technologies (Ada-Europe '18), Lisbon, June 2018.
9. M. A. Javed, F. U. Muram, H. Hansson, S. Punnekkat and H. Thane, "Towards Dynamic Safety Assurance for Industry 4.0", Journal of Systems Architecture (JSA), 2020, ISSN 1383-7621.
10. I. Šljivo, B. Gallina, J. Carlson, and H. Hansson: Strong and Weak Contract Formalism for Third-Party Component Reuse. 3rd International Workshop on Software Certification, pages 359—364. November 2013.
11. W. S. Greenwell, J. C. Knight, C. M. Holloway, and J. J. Pease, "A Taxonomy of Fallacies in System Safety Arguments," in 24th International System Safety Conference (ISSC), New Mexico, July 31-Aug 4, 2006.
12. Object Management Group, 2018. Structured Assurance Case Metamodel (SACM), Version 2.0. https://www.omg.org/spec/SACM/2.0.
13. E. Denney, G.J. Pai, I. Habli, 2015. Dynamic safety cases for through-life safety assurance, in: 37th IEEE/ACM International Conference on Software Engineering, Florence, May 2015, IEEE. pp. 587–590.
14. C. Haddon-Cave, The Nimrod Review: An Independent Review into the Broader Issues surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/229037/1025.pdf
15. M.A. Javed, F.U. Muram, A. Fattouh and S. Punnekkat, "Enforcing geofences for managing automated transportation risks in production sites", in: 16th European Dependable Computing Conference, EDCC 2020 Companion Proceedings, Munich, Germany, September 7–10, 2020.
16. F. U. Muram, M. A. Javed, S. Punnekkat and H. Hansson, "Dynamic Reconfiguration of Safety-Critical Production Systems", in: 25th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC '20), Perth, Australia.

Practical guidance – Safety assurance upon evolution and dynamic risk management (with focus on geofences)

Authors: Muhammad Atif Javed, Faiz Ul Muram, and Sasikumar Punnekkat
(Mälardalen University, SUCCESS Project)

Related to BoK 1.3. – Defining & verifying safety requirements (also 2.4 & 2.5)

Dynamic risk management is an essential characteristic of production site/factory with enhanced automation, digitalization and connectivity [1]. The proposed framework using virtual boundary around a geographic zone, usually called geofence for dynamic risk management is reflected in three stages [2]. In the first stage, the safety analysis during design and development phase is carried out through the identification of hazards, the assessment of risks, and the control of hazards risk [3]. In the second stage, digital twins are utilized. Based on the hazard analysis, the mitigation mechanisms, such as geofences are established; they are translated into the safety requirements, which are implemented in digital twins as code scripts. During the verification and validation with digital twin, additional hazards can be detected. In the third stage, the dynamic safety assurance during operational phase, in particular, the risk management and update of safety cases are carried out.

The geofence can be defined by using different shapes, such as circle, rectangle, capsule (see Figure 1) and freeform etc.); they can serve as an active countermeasure against operational mishap risks. The Global Positioning System (GPS) is used for tracking and navigation purposes and its information is used for triggering alerts in circumstances when the device enters or exits the geographical boundary of a point of interest, as shown in Figure 2.

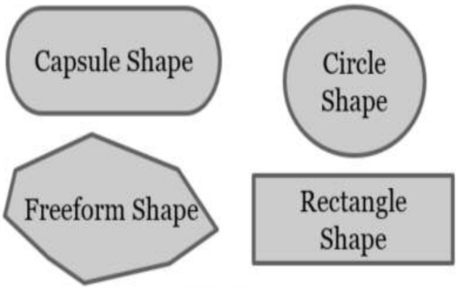


Figure 1. Different shapes for geofences

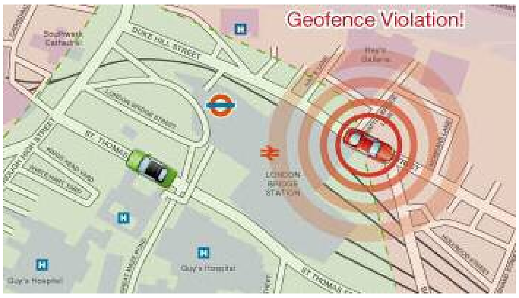


Figure 2. A vehicle is detected in an existing geofenced area [6]

The geofencing are categorised into static, dynamic, periodic and conditional geofences. They are defined over a) various zones at the site/factory, b) different machines, c) other actors at the site such as humans, and even d) around specified paths of movements. The geofences-enabled safety is achieved through, central server commands, vehicle level actions, multiple checkpoints and a monitoring system; vehicle level actions are typically of two categories, viz., those taken by self for normal actions and those taken in response to failure conditions of self or others. There are many challenges and trade-offs which we explore through our simulation test-bed before arriving at reasonable values for the geofences as well as command/action sequences in case of uncertainties [2, 4].

1 Static geofences

The static (constant or fixed) geofences are defined for areas that may not change over time. For instance, the movement of certain objects (humans, robots, vehicles, etc.) need to be restricted in various fixed locations due to safety reasons.
The geofenced regions can involve many uncertainties and therefore continuously monitored. The queue, pause and exit restrictions provide efficient resource for risk control in various areas of

site/factory, such as loading and unloading [2]. The obstacle detection sensors, such as Light Detection and Ranging (LIDAR) and cameras can also be used for monitoring and locating objects (such as humans, robots, vehicles, etc.) that are prohibited in the designated areas due to safety reasons. If an object appeared in the designated restricted area is classified as prohibited, then the respective measures are taken, e.g., the warning can be given via a warning light, alarm or cellphone notification [5].

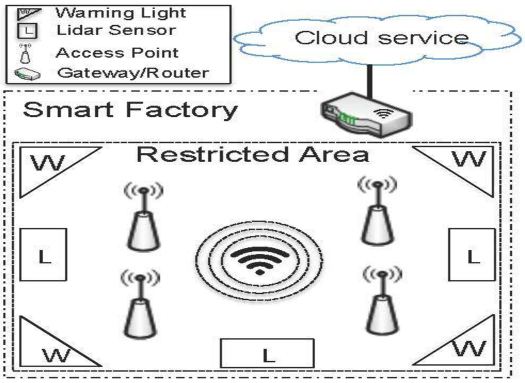


Figure 3 Smart Factory Use Case (taken from [5])

2 Dynamic geofences

The dynamic geofences move over time. They are enforced to reduce mishap risk of emergent and evolving hazards to an acceptable level, for example, the travelling to specific area is blocked and the machines present in area are directed to drive away.

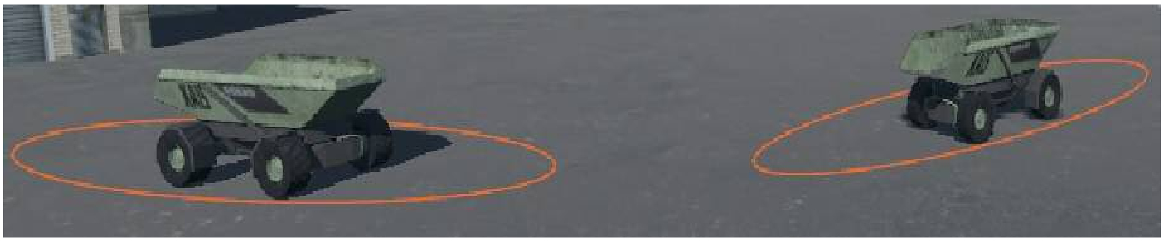


Figure 4. Capsule geometry shapes (geofences) around the haulers

The capsule shape is used to widen the boundary for collision avoidance, for instance, a vehicle or robot could potentially become hazardous due to connection failures, hardware failures, faulty obstacle detection devices, not having obstacle detection capability, and transporting dangerous materials, etc. The capsule geometry shapes (geofences) around the vehicles or robots provide the means for obstacle avoidance. These shapes can be drawn in different ranges and different colors based on their criticality level. When an obstacle is detected in yellow range (indicating move with caution at reduced speeds), the slow down or stop measures are taken, the red range is regarded as emergency stopping distance. The vehicles can maintain assigned speed limit if no obstacles are detected within the range.

3 Periodic geofences

They are only active or inactive for specific time periods. Therefore, they are enforced to stop or control the operation and movement for a certain time-period due to the situations such as site visits. Another example is termination of operation at the end of the day, so the movement towards areas is restricted.

4 Conditional geofencing

The permissions associated with a geofence depends on certain factors like the number of vehicles can be allowed together, i.e., as a platoon for efficient operation. In case of path problem, to create a new path compliant with the conditional geofence, the autonomous vehicle waits for the human-driven vehicle and then follows it to formulate an alternative travel path.

5 Remarks

More details on implementation of such as scheme in the context of the Electric Site of Volvo is described in [2]. However, this is generally applicable to the broad range of scenarios and domains for elimination or mitigation of risks, such as smart city and transportation.

References

1. M. A. Javed, F. U. Muram, H. Hansson, S. Punnekkat and H. Thane, "Towards Dynamic Safety Assurance for Industry 4.0", Journal of Systems Architecture (JSA), 2020.
2. M.A. Javed, F.U. Muram, A. Fattouh and S. Punnekkat, "Enforcing geofences for managing automated transportation risks in production sites", in: 16th European Dependable Computing Conference, EDCC 2020 Companion Proceedings, Munich, Germany, September 7–10, 2020.
3. F.U. Muram, M.A. Javed and S. Punnekkat. "System of systems hazard analysis using HAZOP and FTA for advanced quarry production", in: 4th International Conference on System Reliability and Safety (ICSRS), Rome, Italy, November 20-22, 2019.
4. F. U. Muram, M. A. Javed, S. Punnekkat and H. Hansson, "Dynamic Reconfiguration of Safety-Critical Production Systems", in: 25th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC '20), Perth, Australia.
5. O. Jaradat, I. Sljivo, R. Hawkins and I. Habli, "Modular Safety Cases for the Assurance of Industry 4.0", in: 28th Safety-Critical Systems Symposium, SCSS 2020, York, UK, February 11-13, 2020.
6. F. Reclus, K. Drouard, "Geofencing for fleet & freight management". In: 9th International Conference on Intelligent Transport Systems Telecommunications, (ITST). pp. 353–356 (2009)

ASSURING
AUTONOMY
INTERNATIONAL PROGRAMME