A Systematic Way to Incorporate Security in Safety Analysis

Elena Lisova, Aida Čaušević, Kaj Hänninen, Henrik Thane, Hans Hansson
Mälardalen Real-Time Research Centre, Mälardalen University,
Västerås, Sweden
{elena.lisova, aida.causevic, kaj.hanninen, henrik.thane, hans.hansson}@mdh.se

Abstract—Today’s systems are being built to connect to public or semi-public networks, are able to communicate with other systems, e.g., in the context of Internet-of-Things (IoT), involve multiple stakeholders, have dynamic system reconfigurations, and operate in increasingly unpredictable environments. In such complex systems, assuring safety and security in a continuous and joint effort is a major challenge, not the least due to the increasing number of attack surfaces arising from the increased connectivity. In this paper we present an approach that aims to bridge the gap between safety and security engineering. The potential of the approach is illustrated on the example of E-gas system, discussing the cases when unintentional faults as well as malicious attacks are taken into consideration when assuring safety of the described system.

I. INTRODUCTION

For complex, software intensive safety-critical systems there are well-established approaches to ensure their safety. Safety practices are dictated by safety-standards that prescribe how systems should be developed, verified and maintained to minimize risks of accidents during their lifetime. Traditionally, such systems used to be closed, but nowadays are becoming more and more open as they include interaction via modern communication infrastructures and cloud services. Systems are no longer separate units, but part of larger cooperating systems. They are connected to public or semi-public networks, where information errors can propagate throughout the system in many, sometimes unpredictable, ways. Systems no longer communicate only with human operators, but also with other systems, have dynamic reconfigurations, and unpredictable operating environments.

Security is an increasingly important aspect in safety assurance, as the open interconnected nature and increased reliance on software-based solutions in emerging systems makes them susceptible to security threats at a much higher degree than existing more confined products. Despite the academic efforts to identify interdependencies and to propose combined approaches [1–5] for safety and security, there is still a lack of integration between safety and security practices in the industrial context. One of the main reasons for this is the fact that the disciplines have separate standards, different techniques and processes to assure risk reductions. Moreover, security concerns are generally not covered in any details in safety standards potentially resulting in successfully safety-certified systems that still are open for security threats. To the best of our knowledge, there is no safety or security standard that directly and fully addresses these properties in an joint harmonized effort. However, SAE International has provided SAE J3061 [6], a guidebook on cyber-security, that covers the vehicular domain. This document cannot be seen as a standard itself, since it provides only a guide on how to include cybersecurity when developing complex systems, without any details on which methods, and techniques are the most applicable and should be used. At the moment, International Organization for Standardization (ISO) is driving the development of the ISO 21434 standard that will be aligned with ISO 26262 and will provide details on cybersecurity engineering.

The work presented in this paper is motivated by a study performed by Hänninen et al. [5] that has demonstrated the importance of an integrated approach for safety and security, in order to keep risks of accidents and incidents in complex systems and products at acceptable levels. In this paper we go one step further and present an approach to provide safety and security assurance of software intensive systems where we have identified ways to extend safety work to include aspects of security during the context establishment and initial risk assessment procedures. The ambition of our proposed approach is to improve safety and increase efficiency and effectiveness of the safety work within the frames of the current safety standards. In this paper we focus mostly on ISO 26262 [7]. For a joint safety and security approach to be accepted in an industrial context, the proposed approach must comply with current standards for assessment purposes. An increased understanding of the industrially used standards and current state of practice is thus needed for further research improvements in the area.

The proposed approach is illustrated on the example of E-gas system — an Electronic Glow Adjustable Switch that is replacing a mechanical accelerator cable used to connect the accelerator pedal and the throttle in a vehicle. We provide a simplified hazard analysis in two cases: i) only non-intentional faults considered, ii) malicious attacks from outside taken into consideration. We then compare and discuss results.

The reminder of the paper is structured as follows. In Section II we provide details on our approach for a systematic security assurance in safety-critical systems. Next, Section III provides a hazard analysis of an illustrative example of the E-gas system described in Section III-A. Section IV provides an overview of relevant related work. Finally, conclusions with future work directions are presented in Section V.

II. A SYSTEMATIC WAY TO ADDRESS SECURITY

In this section we go step by step through the safety process and discuss how to incorporate security considerations in a
systemic way in each of the steps.

A. System Definition and Interface Classification

The system definition is the basis for all safety work since it defines the functionality, the environment, the interfaces and the boundaries of the system. It is also a foundation for the risk analysis. The hazards identification process starts from the system definition by identifying hazards that can lead to accidents, incidents, damage or significant financial losses, thus it is important that the system definition is complete and correct.

Today’s systems are characterized by a number of complex interconnections, distributed control centers and service providers. Therefore the safety system definition must be extended to comply with these needs including both failures from within the product itself and intentional misuse and sabotage. The traditional reasoning about sources of hazards (failures and foreseeable misuse) must be extended to also include intentional misuse. Given this, the failure model of the environment and interface parts of the system definition has to be extended with actors and assets that are part of the system or interface with it.

Due to the issues described above, one can conclude that the establishment of a system definition is not trivial. Based on our experiences and common security threats, we have identified the need to identify all assets of the system to be developed, analyze the ways how these assets can be attacked, and learn about possible malicious adversaries interested in remote or local tampering with the system.

We have identified that the following threats and interfaces have to be considered when extending a typical safety system definition: i) people (internal and external personnel, subcontractors, competitors, litigants, press, hackers, criminals, terrorist etc.); ii) nature and accidents (e.g., fires, storms, floods, transportation accidents etc.); iii) interfaces and assets (e.g. fieldbuses and input/output (I/O) for system functionality, internal product buses and interfaces, sensors, actuators, configuration interfaces, control interfaces, monitoring interfaces and diagnostics interfaces, maintenance interfaces, testing interfaces and upgrading interfaces, infotainment interfaces, external product interfaces (e.g., authentication and authorization interfaces, session management interfaces, Universal Serial Bus (USB) interfaces etc.), cellular interfaces and additional assets (such as mobile-enabled devices, printers, USB devices, control centers, cloud services, computers etc.).

Our main intention is to identify the ways to tamper with communication links and to identify all possible personnel that might make harm to the system. Only provided that we have this information, the traditional system definition can be extended. Moreover, we cannot expect anymore the safety organizations to be responsible and have sufficient knowledge for the system definition establishment, as we see a need for different organizational teams to contribute with their respective knowledge in an effort to establish a complete and correct definition.

B. Risk Assessment in Security informed Safety Reasoning

Risk assessment and risk management enable analysis and control of the risks imposed by systems on humans, the environment, or on operations. Risk assessment provides identification, analysis and classification of potential sources of harm and their possible effects to the system. Risk management defines measures to eliminate or reduce the identified risks to acceptable levels. In doing a risk management the measures to be implemented need to be balanced with the cost of reducing the risk. On the other hand, residual safety and security risks will always remain in the final system because of the difficulty to identify all potential sources of harm under all situations and circumstances.

Safety and security techniques share similarities given that the assurance and establishing confidence in the systems are based on assessment of the arguments and the evidence of risk reduction provided for the systems. Structured development methods to produce arguments and evidence are required in normative standards in an effort to provide means for product, process and environment assurance from a safety as well as from a security perspective. There might however occasionally be an unclear border between the responsibilities and the risks managed by safety measures and security measures [8]. Both safety standards and security standards address effects on humans, environment and operations.

C. Hazards and Risk Identification

The goal of a risk and hazard analysis is to identify, quantify, rank, and list hazards that might cause accidents or losses during the lifetime of the product. Various techniques can be used to do so at different stages of the life-cycle. A preliminary hazard analysis is usually performed in the concept phase before any development has been initiated, to be further refined when more details of the system design emerges, and repeated whenever performing maintenance. It is often guided by experiences from similar projects and different analysis techniques such as a fault tree analysis (FTA), a failure mode and effects analysis (FMEA), an event tree analysis (ETA), etc.

The security domain has similar guidance from e.g., threat models, an attack tree analysis. In our extension of the hazard analysis, we include security threats in the safety analysis, assuming that failures are not only coming from the system itself but also from people with malicious intent. The extended scope of the system definition allows identification of previously unforeseen safety hazards, and additional ways in which a system might enter a hazardous state. This results in a more security-aware safety management process. When including security threats as potential sources of hazards, additional and previously unforeseen hazards may be identified. One of the main reasons for this is the fact that the scope of reasoning is extended and that intentional- and accidental misuse of the constituents of the system definition (i.e., the product, the interfaces etc.) must be accounted for. Thus, it is recommended that the hazards identification process, which may be a structured brainstorming meeting with representatives from various disciplines, apply the new system definition (as
D. Root Cause Analysis

Assuming that a system definition is correct and complete, we have been interested in finding out root causes for risks originating from malicious intents. A typical root cause analysis (RCA) includes a number of approaches, tools, and techniques used to uncover causes of problems [9], and in our case potential vulnerabilities. In our work we have focused on sabotage, with the main purpose to identify system vulnerabilities that are both remotely (we assume a communication network being present) and physically, i.e., locally, exploitable.

In the first step we have identified ways how faults can be introduced in the system, given that either or both local or remote interfaces exist. We have provided an interface classification as follows:

1) no tampering possible;
2) physical tampering is possible;
3) local interfaces, where one has to be physically present to utilize the interfaces (only one to one effect mapping);
4) local area network, wired that might provide impact on several items in the system, one to \( N \), where \( N_{\text{max}} = 254 \) (e.g., assuming class C IPv4 network (mask 255.255.255.0) on Ethernet);
5) local area network, wireless, same effect as previous, one to \( N \), where \( N_{\text{max}} = 254 \) (e.g., assuming a class C IPv4 network (mask 255.255.255.0) on Ethernet);
6) wide area network, wired, same effect as previous, one to \( M \);
7) wide area network, wireless, same effect as previous, one to \( M \).

Please note that numbers \( N \) and \( M \) denote the number of possible systems to be affected. Our approach with respect to RCA includes the following steps:

- to identify involved interfaces based on an initial architecture description or system definition;
- provided that interfaces are identified, we have all the information needed to understand the technology used to realize the interface;
- interfaces need to be classified according to the classification schemes proposed in the safety standards. At the same time, new security related hazards that have been found using the extended hazard analysis have therefore to be classified according to the same scheme. Note that this does not imply that the risk classification proposed by the security standards should be ignored for security risks. The reason to classify the related safety risks according to a safety scheme serves two main purposes: i) to assure compliance with the safety standard being used and ii) to assure that all safety risks have been classified according to the same scheme. Note also that a risk classification originating from any reused sources of already identified hazards may have to be reassessed [5] given that the scope of the system definition has changed. Both new as well as old hazards must be mitigated with safety measures to be able to claim that the risk is tolerable. The functional safety standards mandate different levels of rigor for the development and maintenance process, including techniques and measures to be applied depending on the identified risk level (i.e., safety integrity level (SIL), automotive safety integrity level (ASIL), performance level (PL), etc.). A consequence of our extended analysis is that proper mitigations may not be found in the safety standards, but they have to be taken from the security standards. Here it is necessary to translate the rigor required between the different domains and standards.

E. Risk Classification

All identified risks and hazards must be classified according to the classification schemes proposed in the safety standards. However, along with the extended analysis new intolerable risks, as well as end up in situations where we need to re-classify already existing risks.
cases, the risk reduction process needs to consider whether the risk can be reduced according to a safety standard or according to a security standard, or with a combination of both safety and security standards. It is important to stress out that to be able to certify the product, the development and maintenance process steps required in the safety standard must be followed in all cases when implementing mitigations even if the mitigation comes from a security standard [5].

III. SAFETY ANALYSIS VS SECURITY INFORMED SAFETY ANALYSIS

In this section we introduce an E-gas system [11] for which we provide safety analysis and point out differences arising from applying the approach proposed in this paper that incorporates security analysis, compared to a pure safety analysis.

A. Use Case — E-gas

E-gas stands for Electronic Glow Adjustable Switch that is replacing a mechanical accelerator cable used to connect the accelerator pedal and the throttle in a vehicle. The purpose of E-gas is to continue increasing the safety and comfort of vehicles, while at the same time reduce the emission of pollutants. Based on the accelerator pedal, E-gas detects a drivers intention and informs engine control unit that adjusts the opening and closing of throttle.

We assume that our version of the E-gas system is composed of an input that corresponds to an accelerator pedal, a logic implemented through the engine control unit, and an output in the form of an e-throttle as depicted in Figure 1. For this example we provide hazard analysis in two cases, considering (i) only non-intentional faults and (ii) malicious attack from outside of the system taking into consideration the approach presented in this paper. Finally, we compare and discuss obtained results.

B. Hazard Analysis

1) Hazard Analysis considering only non-intentional faults: The system is defined as presented in Subsection III-A. Given that definition, the assumed safe state is turn off the engine or stop the propulsion. The following hazards for the system are of interest for this work: i) unintended acceleration; ii) missing acceleration; iii) unintended deceleration iv) missing deceleration. Note that the list of possible hazards presented above is not complete, as the aim is not to perform complete safety analysis, but to point out differences between the classical hazard analysis and the approach proposed in this work. Thus, for illustrative purposes we will focus only on the hazardous event: unintended acceleration in a highway scenario.

For such an event, according to ISO 26262 [7] the following properties can be derived: i) the severity level $S = S3$ implying life-threatening or fatal injuries ii) the exposure level $E = E4$ that means with a high probability; iii) the controllability level $C = C1$ meaning simply controllable. We have identified the possible actions to take over control that are: i) use brake, ii) slow down by using engine through the clutch, iii) turn off the engine. Based on these parameters, the considered hazardous event can be assigned with ASIL B. The ASIL risk classification provided by ISO 26262 describes levels A-D and a specific level is assigned based on hazardous event severity, exposure and controllability.

The considered hazard raises the following safety goal – to prevent unintended acceleration. Thus, the corresponding safety concept is if unintended acceleration then turn off the engine. The next step in the analysis is derivation of related safety requirements. For example, plausibility checks can be imposed for a sensor data, a logic outcome and an actuator input. Overall, based on the formulated requirements for ASIL B reliability/availability of E-gas shall be no more than $10^{-7}$ failures per hour.

2) Hazard Analysis considering malicious attack coming from outside the system: According to the approach proposed in this work, the system definition of E-gas shall be complemented with consideration of e.g. internal and external personnel, nature of accidents and extended system interfaces descriptions as described in Section II-A. Thus we introduce consideration of intentional faults. For example, we assume that a vehicle maintainer (locally in a workshop or remotely via an Internet update) is able to flash data or code. Given such a system definition the safety analysis provided in the previous subsection needs to be enriched with security relevant information.

As mentioned in Section II, consideration of intentional faults can lead to formulation of new hazards, however the E-gas example is a system with quite limited functionalities, thus we choose to continue with hazards identified in the previous subsection as they cover the main physical processes involved.

Focusing on the same hazardous event, re-evaluation of its levels of severity, exposure and controllability is required according to the proposed approach. Considering severity, remote malicious influence is classified as type 7 in Section II-D, which opens up possibility to influence several vehicles simultaneously and thus requires a different classification than $S = S3$, which we had when we consider non-intentional faults only. We refer to such case as $S = S3+$. Generally speaking, introduction of security considerations requires re-evaluation of the existing classification of severity, as malicious intent due to its persistence may bring new dimensions of harm. Exposure does not change within the proposed approach for this hazardous event as it has already physically maximum possible coverage, $E = E4$ that is equivalent to “always”. However, controllability in its turn within new settings becomes a challenging property to assess. As we consider malicious attacks towards the system, an adversary might tamper mechanisms with the aim to take control over the system during its failure.
and try to deactivate them before deploying the attack. For example, an adversary can make it impossible to brake when trying to counteract to an unintended acceleration. Thus, we say that controllability can vary from $C = C1$ to $C = C3$ depending on the access to other sub-systems given a system intrusion. The combination of identified parameters gives us ASIL D that results in reliability/availability required to be more than $10^{-5}$ failures per hour.

Our approach that follows the classical analysis’ flow but incorporates security considerations results in a higher ASIL. Therefore, for connected systems, where a possibility for a malicious attack exists, safety cannot be guaranteed without considering security, as security may impact the integrity level and thus results in a different set of requirements. For example, to enhance controllability partitioning can be used, as the probability that an adversary can influence mitigation mechanisms might be lower.

Getting a higher ASIL level means not only a necessity to enable lower failure rate per hour for hardware (especially, as this does not help against malicious faults), but implies a systematic process changes. Security needs to be threatened in a systematic way during the development process in case the developing system is supposed to have external connection, dependencies or cooperate with other systems, i.e., unless it is an isolated system. Thus, ASILs’ requirements need to incorporate corresponding security measures. ASIL determines the required safe failure fraction, determines the processes, methods, techniques for the hardware and software. Considering software, methods and techniques for its development may depend on an ASIL classification that incorporates both safety and security. For example, such systematic techniques may include architecture development (e.g., considering its partitioning), protection of the public point of access (e.g., defense in depth approach with incorporating firewalls and levels of authorization), corresponding testing methods (e.g., to complement classical safety tests with penetration tests). In case of hardware, we suggest that its failure rate can rely on the pure safety analysis, however, the processes, methods and techniques for its development may have to change, e.g., increasing reliability by applying redundancy is not enough if vulnerabilities exist in the hardware (HW) that has been used for applying redundancy. The systematic part of developing hardware may be affected similar to the one for software, i.e., one have to assure that insiders cannot assign things in the hardware that can be access from outside etc.

**IV. RELATED WORK**

In the following we describe the most relevant related work and correspondingly position our work among them.

Security-Aware Hazard and Risk Analysis (SAHARA) [1] is an approach proposed by Macher et al. that combines two already well known approaches HARA [7] coming from automotive domain and STRIDE [12] that focuses on threat modeling to review system design in a methodical way usually used in security domain. An underlying method in this approach results in quantified security impact on the safety-critical system development. Provided safety analysis relies on ISO 26262 and utilizes HARA analysis. Security analysis is done based on the STRIDE method independently. Gathered findings form security analysis are further used in an ASIL quantification concept to determine security level. The approach provides an information on resource limits allocated for risk management for security threats. The approach proposed in this paper provides a systematic way to incorporate security in safety process, however it stays quite abstract and does not refer to a particular security technique. Also security incorporation starts already with the system definition, i.e., from the very beginning of the safety process.

Young and Leveson propose a method called STPA - Sec [2], based on already existing top-down safety hazard analysis method System-Theoretic Process Analysis (STPA). The approach is intended for the concept phase of the system development and requires a multidisciplinary team consisting of security, operations, and domain experts to identify and constrain the system from entering vulnerable states that lead to losses. Identified hazards are presented as control problems, where each control action is reviewed under set of a different conditions and guide words, in order to identify loss scenarios, marked as insufficient control or safety constraints. The approach allows to focus on vulnerable states in order to avoid threats to exploit them and create disruptions, and eventual losses, and is not aligned with any standard. The approach is suitable for parallel consideration of both safety and security properties, but enables a single property analysis, as well. Similar to STPA-Sec our approach is applicable to the concept phase, however in contrast to STPA-Sec it aims to cover the whole life-cycle. We also consider interfaces of control logic elements and in this way point out one of the main attack surfaces.

Failure Modes, Vulnerabilities and Effect Analysis (FMVEA) [4], is a method based on already existing approach from the safety domain FMEA, described in IEC 60812 [13] and enables hazards and threats identification. The method incorporates both failure mode and failure effect model for safety and security cause-effect analysis. It is a high level approach suitable for design and verification phase in a system development and for a analysis of only single causes of an effect. Threats are quantified using threat agents that represent attackers, while threat modes are extracted using a STRIDE model resulting in threat effects and attack probabilities. Given that the analysis depends on the accuracy of a system model, a benefit of FMVEA is the possibility to reuse previously acquired results and redo the analysis in case a new threat or vulnerability is identified [14]. However, this is not done in a continuous manner. FMVEA can be used as a technique withing our approach for hazards identification.

Raspotning et al. describe Combined Harm Assessment of Safety and Security for Information Systems (CHASSIS) as a high level approach to combine safety and security methods in order to provide a joint safety and security assessments approach, usually suitable for early phases of system development [3]. The approach is based on modeling misuse cases and misuse sequence diagrams within a UML behavior diagram. Analysis results in a security and safety requirements specification, with the idea to unify analysis by providing a trade-off analysis to check whether mutually dependent or independent features exist. The approach specifies requirements based on ISO 26262 [7] and Hazard and Operability Study (HAZOP) tables combined with Boolean logic Driven
Markov Processes (BDMP) [15]. Given this, a high level of details is required as well as good expert knowledge for the analysis. The approach depends on expert knowledge and reusability in repeated analysis is not applicable since the level of experiences might be different in different teams, most likely affecting the results [14]. This method also can be used within the approach presented in this paper for initial assessment and identification of dependencies between vulnerabilities and failures.

Beside the above described approaches there are a number of projects that focus on combining safety analysis with security such as Sesamo [16], HEAVENS [17], FIA [18], CloSS [19], SAFSEC-CPS [20], SCOTT [21].

V. CONCLUSIONS

Safety of complex safety-critical systems with interconnections to other systems and existing dependencies outside of the system cannot be guaranteed anymore without considering security, since security breaches may jeopardize safety i.e., it is essential to system assets, system vulnerabilities and potential threats towards it. Thus, in this work we propose an approach for a systematic security consideration within the safety process. We start with extending the system definition to introduce intentional faults. Next, we investigate how hazard and risk identification process can be influenced by the security involvement. The interface classification is proposed to handle the risk assessment. Furthermore, a methodology for the root case analysis incorporating security consideration is presented, as well. Based on these steps the risk classification challenge and mitigation techniques are discussed. To illustrate our findings we consider an example of the E-gas system. We illustrate a hazard analysis of such system (only one hazard studied) in a case when no intentional faults are considered and in a case when a malicious attack is taken into consideration. The results show that these two cases lead to different ASIL levels, meaning that the hazard leads to a different set of safety requirements and imposes more strict requirements on the system development process, when the malicious attack is considered. For the future work we plan to extend this approach to cater for other phases of the system life-cycle and aim for its evaluation on a larger safety-critical system.

ACKNOWLEDGMENTS

This work is performed within the CloSS project funded by Software Center Sweden, the SAFSEC-CPS project funded by The Knowledge Foundation and the Serendipity project funded by The Swedish Foundation for Strategic Research.

REFERENCES


