Rapid Integration of CPS Security and Safety

Fredrik Asplund, John McDermid, Robert Oates, and Jonathan Roberts

***Abstract*—The security and safety of Cyber-Physical Systems (CPS) often influence each other. Ensuring that this does not have negative implications might require a large and rigorous effort during the development of CPS. However, early in the life-cycle, quick feedback can be valuable helping security and safety engineers to understand how seemingly trivial design choices in their domain may have unacceptable implications in the other.**

We propose the Cyber Risk Assessment Framework (CRAF) for this purpose. The CRAF is based on openly available and widely used taxonomies from the safety and security domains, and a unique mapping of where loss of data security may impact aspects of data with safety implications. This paper represents the first time these different elements have been brought together into a single framework with an associated process. Through examples from within our organisations we show how this framework can be put to good use.

***Index Terms*—Cyber-physical Systems, Co-design, Embedded Systems Security, Safety Critical Systems**

I. Introduction

SECURITY and safety are broad concepts whose mutual boundaries are, at times, difficult to define clearly. At the root of this ambiguity is the fact that both concepts relate to mitigating *risk* [1]. Usually, safety engineering emphasises the accidental triggering of hazards (sources) leading to harm inflicted on people (consequences), while security engineering emphasises the malicious nature of attacks from threats (sources) leading to negative impacts on assets (consequences). The relationship between safety and security comes from the overlap between these perspectives, for instance when harm is inflicted on people as part of an attack or as an accidental side-effect of it. Understanding this relationship is important when engineering Cyber-Physical Systems (CPS), since CPS allow interaction with physical processes through information technology [2]. We designed the Cyber Risk Assessment Framework (CRAF) to facilitate this understanding among engineers using best practice methods. CRAF has been applied by several groups of engineers to safety-critical systems in the marine and defense domains, and has been subject to independent review from both academic and industrial experts [3], [4]. Below we provide two slightly obfuscated scenarios, based on what users have indicated as interesting from their application of CRAF. These scenarios serve to exemplify the link between safety and security, and will also be used to demonstrate how we propose to improve the ability of contemporary engineering to handle this relationship.

F. Asplund is with KTH Royal Institute of Technology, Department of Machine Design, Brinellvägen 83, 10044 Stockholm, Sweden.

J. McDermid is with University of York, Department of Computer Science, Deramore Lane, Heslington, York, YO10 5GH, United Kingdom.

R. Oates and J. Roberts are with Rolls-Royce plc, Software Centre of Excellence, Victory Road, Derby, DE24 8EL, United Kingdom.

Manuscript received XX YY, 20XX; revised XX YY, 2018.

1) Security engineers identified unsecure access to a critical function on a marine vessel. The critical function allowed users to put the vessel's engines in a state which improved their efficiency temporarily, but which would degrade their reliability in the long run. The unsecured access was not acceptable, since all critical functions should be protected and monitored. The security engineers thus suggested that access would be secured by having the system ask for a user name and a password, subject to locking the user out of the system if the wrong credentials were presented repeatedly. When safety engineers analyzed the proposed design change they realized that users would be asked for credentials in highly stressful situations and that being unable to put the engines in their high-efficiency state could jeopardize the safety of the crew. To resolve the situation they suggested that the room in which the system was located would instead be secured with a physical lock, thereby allowing access control at an earlier, less stressful time. However, while this would protect the engines, it would prevent establishing who triggered the mode change after-the-fact. After reviewing the design suggested by the safety engineers, the security engineers deemed the risk associated with it acceptable.
2) Safety engineers identified how cabling could compromise the structural integrity of a fuel tank. They suggested that integrity would be preserved by changing the digital communication from wired to wireless. While the design change was acceptable from the perspective of the safety engineers, they acknowledged that it could increase the risk of data leaking to others than the intended recipients. This triggered an investigation by security engineers into the implications of the change, which eventually deemed the associated risk acceptable due to the encryption employed.

The risk associated with hazards and threats can sometimes be understood by using proven historical data to calculate the probability of undesirable events. However, this is often unrealistic when engineering complex CPS — accidents might occur due to the unexpected aggregation of environmental factors rather than their random combination, and the adversarial nature of security means that attackers innovate, adapt and deploy attacks not seen in historical data. There might also, as in the aforementioned scenarios, be a lack of useful historical data as the risk relates to hitherto unreleased design changes. Therefore, qualitative assessment of security- and safety-related risks fill an important role in best practice in CPS engineering [5], [6]. Assessment in the two disciplines need to be *combined* as the disciplines require two different skill sets, and those able to analyse the one might not fully

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LES.2018.2879631, IEEE Embedded Systems Letters

comprehend the other. In fact, a deep understanding of the cyber part of security does not ensure an understanding of how cyber threats extend out into the physical world. Qualitative approaches can combine safety and security either through *unification* or *integration* [7]. It has been suggested that unification is fundamentally flawed as it opens up the possibility of compromising the techniques in either discipline [8], to which must also be added the extra cost of training engineers outside their field of expertise. While more promising, there are a multitude of integration approaches, such as those modifying the life-cycle [9], processes [10] or methods [11] used. Especially in regard to methods there is a proliferation of research output in support of parts of CPS engineering, such as assurance [12], analysis [13] and architecture design [14].

The majority, if not all, of these approaches can be labelled as *costly* and *intrusive* — they require both a significant effort to introduce and a significant change to development practices. They are also mostly comprehensive, i.e. aiming to help engineers understand all the possible implications of security on safety and vice versa. This implies that the development of a CPS will have to have progressed quite far for these methods to yield useful results. We do not aim to disparage these methods — time will most likely see a number of them becoming best practice in CPS engineering. However, our experience shows that there is room for techniques that instead serve to rapidly integrate CPS security and safety, especially during early parts of the life-cycle. We posit that quick, but guided, feedback to security and safety engineers regarding decisions in each other's discipline is just as important as ensuring complete and correct handling of all cross-disciplinary issues. This type of feedback allows manufacturers to avoid costly rework later in the life-cycle, especially when they are saddled with considerable legacy and in domains where regulation makes it cumbersome to integrate disciplines. Therefore, we developed CRAF to address this gap as a light-weight scalable risk assessment framework for CPS in civil and defence applications in the marine, nuclear, avionics and rail domains.

This standpoint has been explored extensively within our organisations in regard to factors influencing quality, such as the uncertainty of requirements [15], [16]. From these studies we know that software rework in our domain can cost up to 25 times the original cost of deploying functionality, even if an issue is caught in a formal baseline. To thus quantify the implications of CRAF using a recent run-of-the-mill project: 16% of the requirements were security-related, and 5% of these were identified as having unexpected interactions with safety. This suggests a 20% cost saving of using the CRAF to increase the early interactions between security and safety engineers. Even as an organisation learns from past omissions these cost savings may stay substantial, as the complexity of security in CPS is only expected to grow.

Comprehensive state-of-the-art methods aimed at the requirements phase could possibly offer the same savings, but at a larger upfront cost for each new project. However, one should note that if best practice in the future comes to include substantial amounts of formalized design [17], then other approaches for the early integration of safety and security might yield better results than qualitative analysis [18].

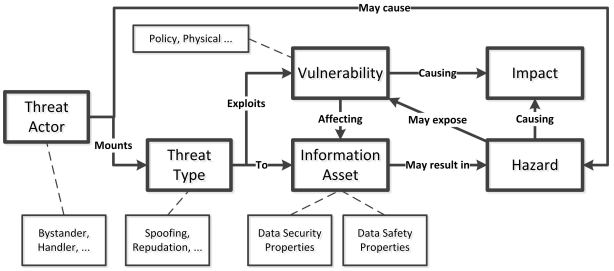

Fig. 1. Linking Security to Safety

II. The Cyber Risk Assessment Framework

The following subsections introduce the publically available taxonomies underlying the CRAF and present the unique mapping we have defined to close the gaps between them.

A. Bringing together Taxonomies

A *Threat Source* is a person or organisation that desires to breach security and ultimately will benefit from the breach in some way. There are taxonomies of threat sources, including specific ones for industrial control systems [19] (see Table I). A *Threat Type* is a type of attack [20] (see Table II). These can be linked to the cyber part of security through the *Data Security Properties* of information assets [21] (see Table III with Integrity expanded as per the definition in the source). Information assets can likewise be connected to *Data Safety Properties* [22] (see Table III). *Vulnerabilities* are inadequacies of an organisation's assets, ranging from its policies to physical hardware [19] (see Table IV), that can be exploited by a threat source.

Although it has been suggested that safety requirements should always be prioritized before security requirements [23], allowing safety responses to compromise security can motivate attackers to jeopardise safety to achieve their goals. Therefore the relationship should be treated as bi-directional. Barring case-specific consequences, there are terminologies available for all identified important concepts. This allows us to model the link between security and safety engineering as shown in Figure 1, using inspiration from older modelling efforts [24]. Aligning the referenced taxonomies along this link constitutes the first novel contribution of the paper.

B. Bridging the Gap to Consequences

To bridge the remaining gap between security and safety we link the data security properties to data safety properties (see Table III). This mapping is based on how the *loss* of data security can lead to an impact on aspects of data which may have safety implications. Developed largely by expert judgment, informed by involvement in [22], this mapping constitutes the second novel contribution of the paper.

In our evaluation of the CRAF we have used HAZOP guidewords made available for linking the data safety properties to hazards related to the physical part of CPS [22]. Presumably other methods found in the academic literature for establishing the impact of security on safety could have been used instead.

TABLE I
THREAT SOURCES [19]

Threat Source	Subtypes
Individual	Outsider, Insider, Trusted Insider, Privileged Insider
Group	Ad hoc, Established
Organisation	Competitor, Supplier, Partner, Customer
Nation State	

TABLE II
THREAT TYPE [20]

Threat Types (STRIDE)
S : Spoofing Identity
T : Tampering with Data
R : Repudiation
I : Information Disclosure
D : Denial of Service
E : Elevation of Privilege

TABLE III
DATA SECURITY TO SAFETY MAPPING [21], [22]

Data Security Property	Data Safety Properties
Confidentiality	Accessibility, Disposability/Deletability, Intended Destination/Usage, Suppression, Traceability
Integrity	Accuracy, Completeness, Consistency, Fidelity/Representation, Format, History, Integrity, Resolution, Sequencing
Availability	Accessibility, Availability, Lifetime, Priority, Sequencing, Timeliness
Non-repudiation	History, Integrity, Traceability, Verifiability
Authorisation / Authentication	Accessibility, Disposability/Deletability, Integrity, Intended Destination/Usage, Lifetime, Suppression

TABLE IV
VULNERABILITY GROUPS [19]

Vulnerability Groups	
Policy and Procedure	Physical
Architecture and Design	Software Development
Configuration and Maintenace	Communication and Network

III. USING THE CRAF

To enable engineers to use the CRAF for rapid integration of security and safety we leverage on exisiting technical risk assessments: best practice in security engineering already scope the security problem and link threat sources to information assets, while on the safety side, engineers identify hazards and link these to the cyber parts of the CPS. Early in the engineering life-cycle, especially when building on legacy, this will result in data available to both sides for quick checks of whether seemingly trivial design decision in one discipline have unacceptable implications according to the other. In the following subsections we use the aforementioned scenarios to exemplify the three steps required to use the CRAF for this purpose.

A. Communicating a Decision

The first *three* columns in Table V and Table VI provide the initial written communication between engineers in Scenarios 1 and 2, respectively. In the first column the identified risk is described. In the the second column the decision advocated by the discipline raising the issue is detailed. The third column notes the affected data properties as found through the use of the CRAF.

B. Raising a Conflict

The fourth column in Table V and Table VI provides the response from the other engineering discipline, where the guidance of the CRAF has led them to identify a conflict.

C. Conflict Resolution

In the fifth column in Table V and Table VI we outline the conflict resolution required to solve Scenarios 1 and 2. Eventually the decision in Scenario 1 came down to one of three alternatives, which are described in Table VII. This is obviously a simplification of the required conflict resolution process given that CPS can be highly complex. However, it serves to show that the CRAF is not meant to be a blind application of guide-words to identify issues that needs to be dealt with. It is meant to highlight decisions and help engineers start thinking in the right direction. Identified issues might require more analysis (as shown in Table V), or a simple acceptance of the decision (as shown in Table VI). Indeed, depending on the application decisions might require complex trade-offs between safety and security.

D. Limitations

The CRAF maps the relationship between safety and security using the implications of how a loss of data security can impact aspects of data that may have safety implications. It is possible that a more appropriate mapping *from* safety *to* security can be identified using other logic. Further research is required to validate the value of the current mapping in both directions, and possibly to identify more useful alternative mappings.

Our risk assessment framework is designed for CPS, which have a cyber component to them. As the above examples indicate, this support extends out into the physical world. However, the vaguer the connection to the cyber part, the more difficult it is to make the connection between disciplines early on. Indeed, our framework might not even be a suitable support if the connection between disciplines is purely in regard to the physical part of a CPS. Further research is required to establish how the CRAF can be modified to overcome this issue.

The Data Safety Properties we used are from a non-exhaustive list [22]. Thus valuable properties for bridging the gaps between the disciplines may still be unidentified.

IV. CONCLUSION

In this letter we propose the Cyber Risk Assessment Framework (CRAF) to support decision-making by enhancing early communication between security and safety engineering during the development of CPS. The framework is constructed based on openly available and widely used taxonomies from the safety and security domains. Through examples from within our organisations we show how the framework can be put to good use. Future work will be aimed at providing a publically accessible evaluation of the framework, improving it in regard to issues solely related to the physical parts of CPS, and validating its mapping between safety and security.

TABLE V
USING THE CRAF FOR SCENARIO 1

Output, Security Risk Assessment	Decision, Security Engineering	Data Security Property	Data Safety Property: Conflict	Conflict Resolution
Identification of unsecured access to [Function X], a critical function which should be protected and monitored.	Access to [Function X] will from now on be protected by credentials.	Authentication	Integrity (Difficult to set data to its true state): Putting the engines in their high-efficiency state is required to protect the safety of the crew. Requiring credentials to be presented in what could be a stressful situation might thus lead to fatalities.	1. Suggestion to secure the room in which the system is located with a physical lock, thereby allowing access control at an earlier, less stressful time. 2. Security risk assessment identifies that this solution does not allow access to [Function X] to be logged. 3. Alternative 3 was chosen as the best acceptable solution using Table VII.

TABLE VI
USING THE CRAF FOR SCENARIO 2

Output, Safety Risk Assessment	Decision, Safety Engineering	Data Safety Property	Data Security Property: Conflict	Conflict Resolution
Cabling compromises the structural integrity of the fuel tank.	Cabling will be replaced with a wireless solution.	Intended Destination/Usage	Confidentiality (Loss of confidentiality): Data regarding the contents of the fuel tank could be exposed.	1. Security risk assessment identifies that the risk associated with the proposed design change is acceptable due to the encryption employed.

TABLE VII
CONFLICT RESOLUTION ALTERNATIVES IN SCENARIO 1

Alternative	Security, Probability of Adverse Event	Security, Impact	Safety, Probability of Adverse Event	Safety Impact
1: Do nothing	High (Unauthorized, unmonitored access)	High (Malicious triggering of [Function X])	Low	Low
2: Require credentials	Low	Low	High (Authorized user denied access)	High (Unable to trigger [Function X] when required)
3: Fit room with physical lock	Medium	Medium	Low	Low

ACKNOWLEDGMENT

Special thanks go to Vicki Derbyshire for her help in proofreading.

REFERENCES

[1] L. Pitre-Cambacds and M. Bouissou, "Cross-fertilization between safety and security engineering," *Reliability Engineering and System Safety*, vol. 110, pp. 110–126, 2013.
[2] M. Törngren, F. Asplund, S. Bensalem, J. McDermid, R. Passerone, H. Pfeifer, A. Sangiovanni-Vincentelli, and B. Schätz, *Chapter 1, Characterization, Analysis, and Recommendations for Exploiting the Opportunities of Cyber-Physical Systems*. Elsevier Inc., 2017.
[3] R. Oates, "A data focussed approach to mapping security failures to safety impacts," in *Presented at the SCSC, Safety and Security Integration seminar*, 2018.
[4] J. Roberts, "Cyber-risk assessment framework encompassing safety and security," in *Presented at SCSSS 2018*, 2018.
[5] IEC, *BS/IEC 61508:2010, Functional Safety of E/E/PE Safety-Related Systems*, International Electrotechnical Commission Std., 2010.
[6] NIST, *Guide for Conducting Risk Assessments*, Std., 2012.
[7] S. Kriaa, L. Pietre-Cambacedes, M. Bouissou, and Y. Halgand, "A survey of approaches combining safety and security for industrial control systems," *Reliability Engineering and System Safety*, vol. 139, pp. 156–178, 2015.
[8] D. Eames and J. Moffett, "The integration of safety and security requirements," in *Computer Safety, Reliability and Security*, vol. 1698. Springer, 1999, pp. 468–480.
[9] T. Novak, A. Treytl, and P. Palensky, "Common approach to functional safety and system security in building automation and control systems," in *Proceedings of ETFA*, 2007.
[10] P. Bieber, J.-P. Blanquart, G. Descargues, M. Dulucq, Y. Fourastier, E. Hazane, M. Julien, L. Lonardon, and G. Sarouille, "Security and safety assurance for aerospace embedded systems," in *Proceedings of ERTS*, 2012.
[11] L. Pitre-Cambacds and M. Bouissou, "Modeling safety and security interdependencies with bdmp (boolean logic driven markov processes)," in *Proceedings of SMC*, 2010.
[12] C. W. Johnson, "Using assurance cases and boolean logic driven markov processes to formalise cyber security concerns for safety-critical interaction with global navigation satellite systems," in *Proceedings of the Fourth International Workshop on Formal Methods for Interactive Systems*, 2011.
[13] I. N. Fovino, M. Masera, and A. D. Cian, "Integrating cyber attacks within fault trees," *Reliability Engineering and System Safety*, vol. 94, no. 9, pp. 1394–1402, 2009.
[14] J. Delange, L. Pautet, and P. H. Feiler, "Validating safety and security requirements for partitioned architectures," in *Proceedings of Ada-Europe*, 2009.
[15] A. J. Nolan, S. Abraho, P. Clements, and A. Pickard, "Managing requirements uncertainty in engine control systems development," in *Proceedings of RE*, 2011.
[16] A. J. Nolan, S. Abraho, P. Clements, J. D. McGregor, and S. Cohen, "Towards the integration of quality attributes into a software product line cost model," in *Proceedings of SPLC*, 2015.
[17] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. S. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyberphysical systems," *Proceedings of the IEEE*, 2012.
[18] M. Sun, S. Mohan, L. Sha, and C. Gunter, "Addressing safety and security contradictions in cyber-physical systems," in *Proceedings of the 1st Workshop on Future Directions in Cyber-Physical Systems Security (CPSSW09)*, 2009.
[19] NIST, *Guide to Industrial Control Systems (ICS) Security*, Std., 2015.
[20] A. Shostack, "Experiences threat modeling at microsoft," in *Proceedings of MODELS*, 2008.
[21] NIST, *Security and Privacy Controls for Federal Information Systems and Organizations*, Std., 2013.
[22] The Data Safety Initiative Working Group, "Data safety guidance 3.0," 2017.
[23] T. Novak, A. Treytl, and A. Gerstinger, "Embedded security in safety critical automation systems," in *Proceedings of ISSC*, 2008.
[24] D. G. Firesmith, "Common concepts underlying safety, security, and survivability engineering," 2003.