SECURITY-INFORMED SAFETY: INTEGRATING SECURITY WITHIN THE SAFETY DEMONSTRATION OF A SMART DEVICE

Robin Bloomfield, Eoin Butler, Sofia Guerra and Kate Netkachova
Adelard LLP
24 Waterside, 44-48 Wharf Road, London N1 7UX, UK
{reb,eb,aslg,kn}@adelard.com

Robin Bloomfield and Kate Netkachova
City University of London
Northampton Square
London EC1V 0HB, UK
{R.E.Bloomfield, Kateryna.Netkachova.2}@city.ac.uk

ABSTRACT

Safety and security engineering have, over the years, developed their own regulations, standards, cultures, and practices. However, there's a growing realisation that security is closely connected to safety. Safety must be security-informed: if a safety-critical system isn't secure, it isn't safe. A safety demonstration is incomplete and unconvincing unless it considers security. In our work for government and industry, we have used the Claims, Arguments, Evidence (CAE) framework to analyse the impact of security on a safety justification or safety case and identified the significant changes needed to address security explicitly. This will impact the design and implementation process as well as the assurance and V&V approach.

In this paper we discuss the impact of integrating security when developing a safety demonstration of a smart device. A smart device is an instrument, device or component that contains a microprocessor (and therefore contains both hardware and software) and is programmed to provide specialised capabilities, often measuring or controlling a process variable. Examples of smart devices include radiation monitors, relays, turbine governors, uninterruptible power supplies and heating ventilation, and air conditioning controllers.

Key Words: Smart (embedded) devices, safety assessment, security-informed safety, cyber

1 INTRODUCTION

Safety and security engineering have, over the years, developed their own regulations, standards, cultures, and practices. However, there is a growing acceptance that security is closely connected to safety: it can no longer be assumed that a safety-critical system is immune from malware because it is built using bespoke hardware and software, or that the system cannot be attacked because it is separated from the outside world by a so-called "air gap". Safety must be security-informed: if a safety-critical system isn't secure, it isn't safe. There are also business drivers for security-informed safety: stakeholders do not want to pay twice for assurance, or worse, discover conflicts between safety and security that significantly impact project timescales and require considerable system re-architecting. In addition, reputational risks from actual or period security incidents could be very costly. In the UK, the government has recently published a Civil Nuclear Cyber Security Strategy that further motivates the need and urgency of tackling security issues on both legacy and new systems [1].

In this paper we discuss the impact of integrating security when developing a safety demonstration of a smart device. A smart device is an instrument, device or component that contains a microprocessor (and

therefore contains both hardware and software) and is programmed to provide specialised capabilities, often measuring or controlling a process variable. One essential property is that it cannot be programmed by the end-user. That is, although the end-user may be able to perform limited configuration of the device, they cannot add new functionality and cannot modify the existing functionality in a fundamental way. Examples of smart devices include radiation monitors, relays, turbine governors, uninterruptible power supplies and heating ventilation, and air conditioning controllers.

2 ASSESSMENT FRAMEWORK

2.1 Overall Assessment Strategy

We are concerned with the assessment of safety of a smart device within a particular I&C context. A smart device is not safe as such ("safety is a system property") but will have functional and performance properties defined that enable the safety and security of the wider system. Some of these properties will enable the device to deliver a trustworthy service and others enable it to be a "good citizen" and support other parts of the I&C system components.

A device assessment, at its core, involves presenting evidence about behaviour to support claims about dependability properties. A systematic evaluation of potential weaknesses and vulnerabilities in the system is also a key part of the assessment. Standards are important in defining design constraints that need to be satisfied and have very significant role in the overall licensing of the system. The overall assessment approach is therefore described as property-based, vulnerability-aware, and standards-informed and is illustrated by the strategy triangle in Figure 1 [2].

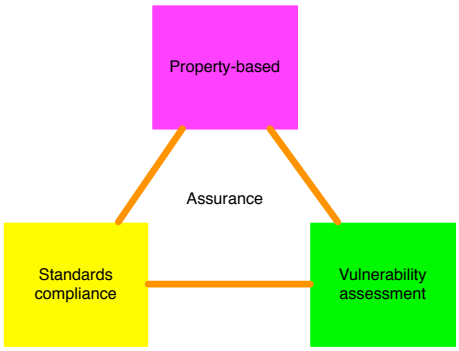

Figure 1. The strategy triangle

2.1.1 Property-based Approach

A property-based approach focuses directly on the behaviour of and constraints on the device being assessed. It explores claims about the satisfaction of the safety requirements and the mitigation of the hazards to the system. For example, for a smart sensor, this is likely to include accuracy and timing, in addition to functionality, operability, etc. A selection of techniques can be made to show that the properties are met, each of which supports one or more of the properties.

2.1.2 Vulnerability-based Approach

Vulnerabilities are weaknesses in a system, e.g., if a divide by zero is not caught by error handling, they could lead to hazardous situations but are not strictly hazards. Vulnerabilities can exist at different levels of abstraction (e.g. lack of diversity at the architecture level, or buffer overflow within the implementation). There are several methods and techniques that can be employed to perform a vulnerability analysis for a smart device and its parent system. At a component level, these approaches

aim to identify both generic failure modes and their causes, or to provide evidence of their absence. At a system level, the device failure modes should be considered in terms of the system/application vulnerabilities, and whether mitigations are adequate.

Once vulnerabilities have been identified, the assessment will seek to show that this are absent or mitigated. Some techniques, such as static integrity analysis, are particularly suited to identifying vulnerabilities. There may be an overlap between techniques used to support the property-based approach and techniques used to support the vulnerability assessment.

2.1.3 Standards Compliance

The third part of the strategy triangle is concerned with compliance with standards. Adequate compliance with standards will need to be demonstrated as part of the overall licensing process. In this paper we do not consider the standards aspects as we focus on the impact of security in the safety assessment of the device. The issues we raise should be addressed within the standards making process.

2.2 Incorporating Security Issues

A safety assessment is incomplete and unconvincing unless it considers security. In previous work, we have used the Claims, Arguments, Evidence (CAE) framework [3, 4] to analyse the impact of security on safety assessment or safety cases and identified the significant changes needed to address security explicitly [5, 6]. We investigated the impact that security might have on a generic case by considering the three aspects and deciding whether we need to: change the claims, augment the arguments or change how we deal with evidence. In terms of methodology, the steps we took were:

- Express the safety case about system behaviour in terms of Claims-Arguments-Evidence.
- Review how the claims might be impacted by security, including the top-level claims.
- Review security controls to see if these can be used to provide an argument and evidence for satisfying the claim.
- Review architecture and implementation impact of deploying controls and iterate the process.

Incorporating security into the safety assessment impacts the design and implementation process as well as the assurance, and the approach to verification and validation. In particular, we found that the following are some of the most significant considerations from a security perspective:

- Integration of requirements, e.g., of safety, with security and resilience.
- Supply-chain integrity, e.g., mitigating the risks of devices being supplied compromised or having egregious vulnerabilities.
- Post-deployment malicious events that will change in nature and scope as the threat environment changes and a corresponding need to consider prevention (e.g. implementing a risk based patching policy) but also recovery and resilience.
- Reduced lifetime of installed equipment as there is a weakening of security controls as attackers' capabilities and technologies change.
- Threats to the effectiveness and independence of safety barriers and defence in depth.
- Design changes to address user interactions, training, configuration, and software vulnerabilities and patching. These might lead to additional functional requirements for security controls.
- Possible exploitation of the device/service to attack itself or other systems.

This paper discusses how these and similar security considerations are addressed when applied to a safety assessment of smart devices. An overview of some of the issues discussed in this work is provided

in Figure 2. It shows how a device implements some system level aspects and also supports other parts of the systems (human and technical components).

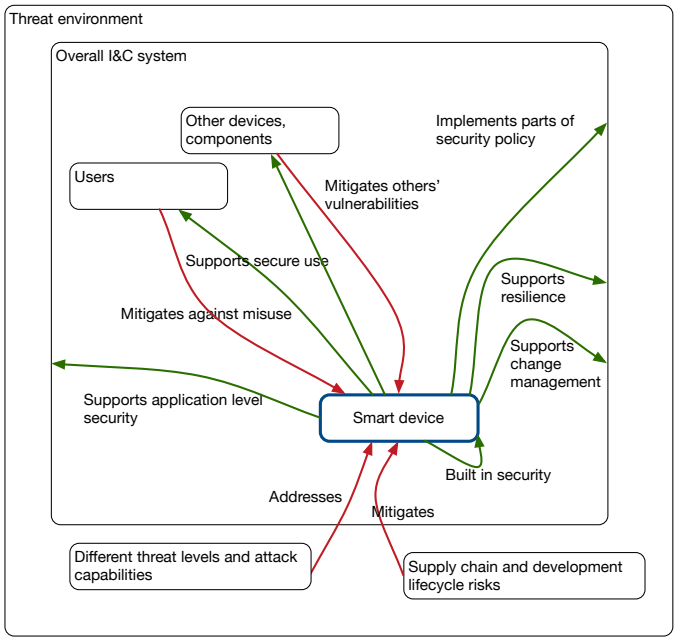

Figure 2. Overview of security considerations for smart devices

The rest of this paper elaborates these issues and explores how they might impact the assessment of a smart device. We consider how security will impact the

- Assumptions of the assessment
- Required properties (as in the property-based approach – see Section 2.1.1)
- Vulnerabilities to be addressed (Section 2.1.2)
- Trust in the assessment

3 IMPACT OF SECURITY CONSIDERATIONS

3.1 Environment and Threat Assumptions

In order to define the properties of interest for the safety demonstration (as in Section 2.1.1), the safety assessment will need to consider the environment of the system and of the device manufacturer. However, often the threat assumptions are left implicit or factored out. For security-informed cases this would not be adequate, as we need to define what assumptions we are making about the threats to the system. For example, the nature of attackers, their resources and motivations as well as any additional claims about perimeter security would need to be defined. In previous assessments, in common with many security risk assessment approaches, we have used a graded approach to describing the capabilities that an attacker would need in order to achieve a particular impact failure. An example of these capability levels is shown in Table I.

We would expect government agencies to provide advice on what is appropriate, i.e., they would provide the interface to threat intelligence and justification of the Design Basis Threats.

Table I. Attack capability levels

Capability level	Interpretation for smart devices
E	An expert in security engineering who can: • use tools specific to the domain, which may be customized for the attacks • develop novel equipment and tools specific to the attack • use publicly available and proprietary information on how the device and surrounding system works and what mitigations are in place against attacks • develop large test beds and trials for the attack • coordinate timing of several attacks • influence expert insiders
D	An expert in security engineering who can: • use tools specific to the domain, which may be customized for the attacks • access equipment for trials and attack development • use publicly available and proprietary information on how the device and surrounding system works and what mitigations are in place against attacks • influence knowledgeable insiders
C	Someone with a basic understanding of security engineering who can: • use tools specific to the domain but without customization • use publicly available information on how the system works and what mitigations are in place against attacks • influence insiders (but at routine skill level)
B	Someone with physical access to the system, for example: • an engineer who is able to plug a maintenance console into the equipment but has no specific training or authorization to access the system in this way • an unwitting participant, using a compromised machine or device
A	Someone without access to the system, for example: • unskilled individuals using scripts or programs developed by others to attack computer systems and networks • someone who has been co-opted into scaling a distributed denial of service attack • an enterprise IT user

3.2 Property-based Approach – Required Properties

Security will impact the properties required for the application and for the smart device itself. The properties that would need to be implemented and demonstrated are likely to increase in scope to include functionality arising from additional security controls and from addressing security attributes such as confidentiality. There may also be design changes to support increased security of user interactions, training and configuration. Security is classically thought of as encompassing the attributes of availability, integrity and confidentiality. The requirements for integrity and availability would already be considered intrinsically as part of a safety assessment but may need modifying in the light of security aspects, threat assumptions and requirements trade-offs. Security will have a major impact on the risks to integrity and availability and this will inform the security informed hazard analyses.

3.2.1 Including Confidentiality Issues

In terms of confidentiality, there is a need to consider it in more detail for two reasons:

- Assets in the system could have value and become targets for attack, e.g. control algorithms, set points or “recipes”.
- Information such as product details, project management information and tool chain details could be acquired and used to escalate or enable an attack.

So there are issues of confidentiality of the deployment and development processes as well as that of the device itself. The overall assessment should involve new claims about confidentiality:

- The device does not leak information that leads to unacceptable increase in risk of successful attack.
- The device protects confidentiality of assets that have direct information value.

For the smart device the overall deployment should address confidentiality. Some of this may be in the safety assessment where safety related risks are concerned. For example, the claim that the smart device does not leak information may be related to information about the process that the device is monitoring or controlling but equally it might be information about the device itself which might allow the device to be identified and specific vulnerabilities to be discovered, e.g., by using Shodan search engine if it was connected to the internet.

In addition, the smart device should protect confidentiality of assets that have direct informational value, e.g., characteristics of plant processes, materials and algorithms. This may also apply to parts of the safety assessment as knowledge of, for example, the techniques used to assess a device can be of assistance to a capable attacker.

For a smart device that is only available "off the shelf", the design may include intrinsic mitigations that provide sufficient security. However, typically, additional mitigations will need to be addressed by the configuration and location of device to restrict access as well as by management and technical processes to monitor information flow. These application level assumptions would need to be captured and communicated, e.g., as extra conditions for the device deployment.

3.2.2 Integrated Requirements

Security considerations will have an impact on the properties that would need to be demonstrated for the smart device. For example, there might be increase in scope and complexity, e.g., due to design changes to support increased security of user interactions, training and configuration, supply chain integrity, authentication or access control. There will need to be careful design to ensure that these additional requirements do not conflict with the safety properties and their demonstration, and in practice some trade-offs or compromises may be necessary.

Figure 3, taken and generalized from [5], shows different aspects of the safety requirements and security policy interactions. In the bottom left corner we have an area of maximum operational benefit, where with low levels of threat and no significant safety challenge it is relatively straightforward to satisfy both aspects. The other areas indicate how at certain threat level security concerns might dominate (e.g., with a need to restrict access to the device). In this case, the safety analysis must show that these constraints are acceptably safe even if they do cause higher workload or operational complexities. There is a corresponding zone where safety issues dominate and the security policy is the same or weakened. In this case, the security analysis must show that identified security threats are satisfactorily mitigated by other means during this time. Finally, the top right hand corner is a very uncertain area where some special capabilities might be needed, e.g., in the form of a manual override to security policy. The interactions between a security policy and the safety requirements need to be assessed and any trade-offs identified. In some circumstances, increased security may reduce safety so it is essential to consider these holistically.

As well as the functional trade-offs that might be needed, the security perspective brings with it the need to consider defence in depth from a security perspective and also resilience (resilience is the ability to adapt and recover). It is more credible that security issues will defeat defence in depth and lead to the system being shut down or some worst credible case incident occurring. There needs to be a stronger focus on resilience to emphasise the need for adaptation and recovery and the possibility of security-

induced failures. A balanced system level view should be taken of resilience – balancing the role of the device itself in being resilient and the resilience of the overall service it is part of.

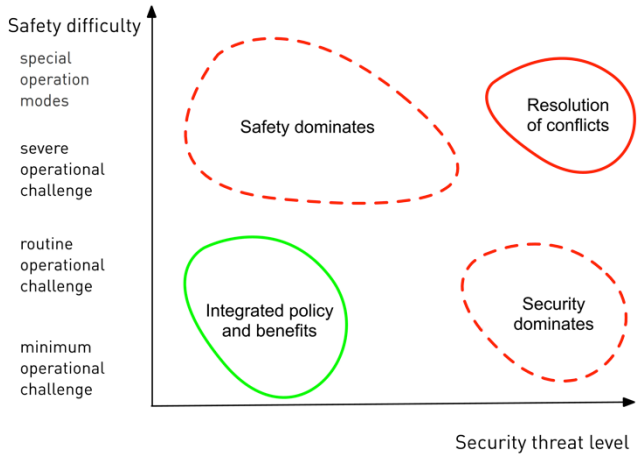

Figure 3: Defining an integrated security and safety policy

3.3 Vulnerability-based Approach

The importance of identifying vulnerabilities in the software and design can be greatly impacted by the security design basis threats. While product vulnerabilities may have already been addressed in the safety assessment, the claims will need to be increased in depth and also in scope as issues of lifecycle threats and malicious threats to evidence need to be included. For example, although safety standards already require the trustworthiness of tools to be justified, the inclusion of security concerns means that the possible malicious inclusion of code by tools or the deliberate non-reporting of findings will also need to be considered. To assess what is needed in a systematic way will need security-aware hazard analysis techniques. We have been using security- (or cyber-) informed Hazop to assess industrial systems [7] and have adapted this well-known approach with additional security guidewords and an enhanced multidisciplinary team.

In undertaking static-code analysis, both security and safety perspectives are needed to assess the likelihood of vulnerabilities being exploited, their mitigations' effectiveness and consequences. The security perspective will increase the scope of the static analysis and impact the sentencing of vulnerabilities that are discovered.

There may also need to be increased verification effort to show independence of critical functionality from failures of other software or components and to address wider classes of software vulnerabilities. The safety assessment of the smart devices in the UK includes a number of verification activities that could easily be extended to take into account security vulnerabilities as part of the vulnerability assessment. See, for example, the analyses described in [8].

3.4 Post Deployment – Assessed Properties Hold in the Future

The discussion so far has considered the properties and vulnerabilities of the smart device at the time of the assessment (i.e., pre deployment). It is also necessary to show that these properties will continue to hold during the operating lifetime of the device. There are two important changes to the safety assessment when we consider the claim that the properties continue to hold in the future. The first is that we need to ensure that the future system is robust to malicious threats and associated changes (e.g., new version of

the system to address newly discovered vulnerabilities) as well as to the safety related set of changes that are normally considered. Second, we need to address the change in nature and intensity of the threat environment and the weakening of security controls as the capability of the attacker and technology changes.

Although safety systems are already designed to support operational changes for calibration and maintenance, the ease of recovery principle, which states that the security of the system should not depend on anything that cannot be easily changed, could have far reaching impact on the architecture of safety systems.

Moreover, changes to threats over the lifetime of the system will probably mean that controls that were adequate initially will need to be reconsidered. This is illustrated in **Figure 4**, which shows the lifecycle of cryptographic hashes and how their strength decays over time [9].

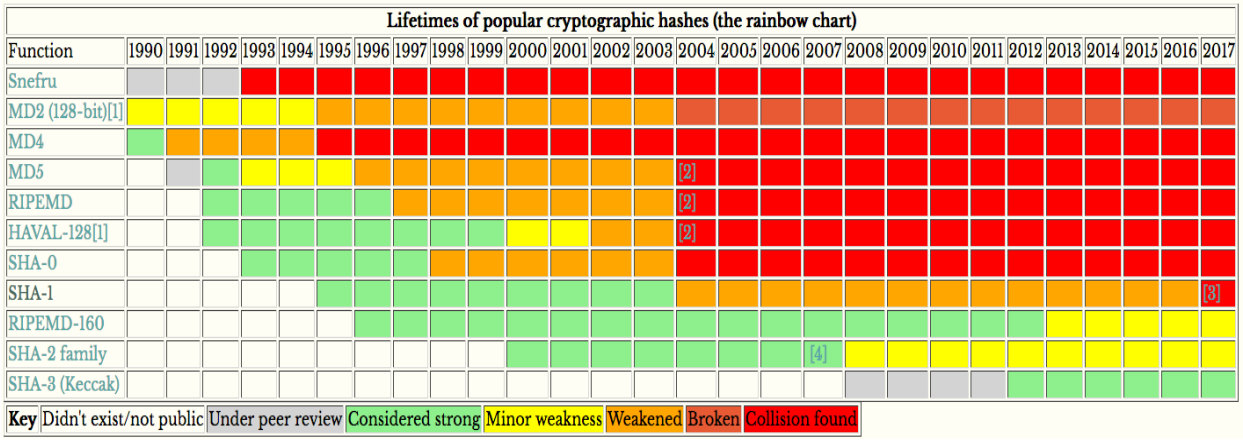

Figure 4: Lifetime of cryptographic hashes

This decay of security controls arising from the changing threats, vulnerability discovery and attack sophistication time have implications for the architecture and lifecycle of embedded safety systems where design life may be decades. It may have major impact on the proposed lifetime of the installed equipment and require a design philosophy that enables refurbishment and change.

From a smart devices perspective this is a problematic area for a number of reasons, e.g.:

- Addressing refurbishment and obsolescence can be difficult due to the specialized nature of the equipment or the location on plant where it installed.
- Patching and upgrading devices can be problematic: difficult trade-offs have to be made between the increased security and the risks and safety risks of the update.
- The devices may use third party libraries and components that become unsupported (e.g., in the past, use of Windows XP in advisory systems).
- The approach to safety justification may not match the tempo needed to address security concerns: this is both a technical and institutional challenge.
- The issue needs to be addressed at product selection stage and also at deployment. Use of diverse populations of devices where feasible and progressive and staged maintenance may provide some useful mitigations.

As smart devices cover a wide range of products (e.g., from chart recorders, gas analyses, temperature transmitters) the approach needs to be creative and flexible.

3.5 Trust in the Assessment

The impact of security aspects will also need to be addressed as part of the assessment process. There will be a need to ensure that personnel security is appropriately managed and there are the correct processes and systems in place for the handling of information. There are two particular issues that need to be emphasized:

- Whether the device that will be delivered to plant has the same properties as the device that has been assessed
- Can the evidence that is provided to the assessor be sufficiently trusted

3.5.1 Supply Chain Integrity

Security will have an impact on how we address supply chain integrity that is how we can be sure that the supplied device is as trustworthy as the one that has been assessed. Supply chain integrity could be very significant if the safety critical system being delivered needs to be secure. A variety of intelligence, design, and procurement approaches will be necessary and these should be captured within the design basis threats and the safety and security policy.

For example for smart devices:

- There could be randomness in the procurement process or an anonymised process so that suppliers do not know the destination of the devices they are selling.
- The assessment and evaluation could be on the actual devices that are supplied to the plant, not just on "typical" devices. This would need some innovation in the assurance to make it more cost effectively repeatable and that would need corresponding changes to design, e.g., so authenticity of firmware could be established.
- There would need to be an additional means to validate identity and ensure authenticity of the source code of the devices (e.g. via a digital signature). This will mitigate the risk that the code might be replaced or modified anywhere within the supply chain.

The assessment process for smart devices in the UK includes a detailed audit of the supplier's processes and facilities [8], which gives a privileged access to information that could be used to assess supply chain integrity including quality assurance.

3.5.2 Explicit Discussion of Trustworthiness of Evidence

Changing the threat assumptions will have an impact on how we address the evidence that is fundamental to the safety case. We need to make an explicit claim that the evidence is trustworthy and we may need to factor this by the different organisations that provide it. For example, the feasibility and risks from the deliberate tampering with evidence and the non-reporting or falsification of findings should be addressed. Although safety standards already require the trustworthiness of tools to be justified, the inclusion of security concerns means that the different threats become credible, so that, for example, the possible malicious inclusion of code by tools needs consideration.

4 CONCLUSIONS AND FURTHER WORK

In this paper we discuss the impact of integrating security when carrying out a safety assessment of a smart device. We have elaborated and explored the impact of a number of generic issues on the safety assessment of smart devices by exploring who security would be integrated within the strategy triangle. We consider how security will impact the threat assumptions of the assessment, the required properties of the device arising both directly from its role in the safety system as well as taking into account security issues and assisting other components to mitigate the security risks. We have discussed how the nature and depth of vulnerability analysis will be increased and how to ensure that the required properties will

continue to be met over time. Lastly, we consider the need to explicitly address the trustworthiness of the assessment process itself and the evidence used.

A security informed safety strategy has to address the wide range of possible devices as well as the legacy aspects of the applications, the system level safety justifications and the devices themselves. We have shown the depth and breadth of the impact of security on smart device assessment, and also some of the issues the supplier and designer of these devices face. Although the current assessment framework in the UK has the necessary components to consider security (e.g., it includes detailed consideration of vulnerabilities including static analysis as well as thorough and comprehensive audit of the supplier), there will need to be significant changes to the assessment, regulation and design of devices to address security. To achieve a technically sound and cost-effective approach to security informed smart device assessment will need:

- Engagement with suppliers to increase security of design and build security and support for application level security.
- Architecture level approaches that develop integrated requirements for smarts devices.
- Innovation at the application level to mitigate security risks and implement security engineering principles.
- Security informed hazard analysis techniques to support the design and assessment.
- Improvement to static analysis to deal with more complex devices and a greater range of vulnerabilities.
- Developments in standards and guidelines to support licensing as well as assessment.
- Integration of security assessments during the analysis to support the safety demonstration.

There is also the need for a security informed safety assessment framework. In future work, we hope to develop technical approaches based on a claims, argument evidence framework and CAE Blocks [10] to demonstrate how security and safety of smart devices can be addressed in an integrated manner building on our research and assessment work in nuclear [11] and other sectors.

5 ACKNOWLEDGMENTS

This work has been partially supported by the UK EPSRC project “Communicating and Evaluating Cyber Risk and Dependencies” (CEDRICS, EP/M002802/1), which is part of the UK Research Institute in Trustworthy Industrial Control Systems (RiTICS).

6 REFERENCES

1. UK Department for Business, Energy and Industrial Strategy, Civil Nuclear Cyber Security Strategy, February 2017, available for download from www.gov.uk/government/publications.
2. Bishop, P.G., Bloomfield, R.E. and S Guerra, S., The future of goal-based assurance cases. In Proceedings of Workshop on Assurance Cases. Supplemental Volume of the 2004 International Conference on Dependable Systems and Networks, pp. 390-395, Florence, Italy, June 2004.
3. Adelard Safety Case Development Manual, © Adelard, ISBN 0 9533771 0 5, 1998.
4. Bishop, P. G., Bloomfield, R. E., Guerra, S. & Thuy, N., Safety justification frameworks: Integrating rule-based, goal-based and risk-informed approaches, in: 8th International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies 2012, NPIC and HMIT 2012: Enabling the Future of Nuclear Energy. (pp. 1283-1290. ISBN 9781627480154.
5. Netkachova, K., Müller, K., Paulitsch, M. & Bloomfield, R. E., Investigation into a Layered Approach to Architecting Security-Informed Safety Cases, IEEE/AIAA 34th Digital Avionics

Systems Conference (DASC), Sept 2015, Prague, Czech Republic, DOI: 10.1109/DASC.2015.7311447.

6. Bloomfield, R.E., Netkachova, K. and Stroud, R., Security-Informed Safety: If it's not secure, it's not safe, 5th International Workshop on Software Engineering for Resilient Systems (SERENE 2013) 3-4 October, Kiev, Ukraine.
7. Bloomfield, R. E., Bendele, M., Bishop, P. G., Stroud, R. & Tonks, S. (2016). The risk assessment of ERTMS-based railway systems from a cyber security perspective: Methodology and lessons learned. Paper presented at the First International Conference, RSSRail 2016, 28-30 Jun 2016, Paris, France.
8. Guerra, S., Butler, E. and George, S.,Safety Demonstration of a Class 1 Smart Device. NPIC 2017.
9. Lifetimes of cryptographic hash functions http://valerieaurora.org/hash.html. Accessed: 10 March 2017.
10. Bloomfield, R.E., Netkachova, K.: Building Blocks for Assurance Cases. In: IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW) 2014, pp. 186-191, doi:10.1109/ISSREW.2014.72.
11. Guerra, S, Sheridan, D, Chozos, N, Justifying Digital Cots Components When Compliance Cannot Be Demonstrated – The Cogs Approach, International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies: NPIC 2015, Charlotte, NC.