Why We Cannot (Yet) Ensure the Cyber-Security of Safety-Critical Systems

Chris Johnson[1]

University of Glasgow

Glasgow, UK

Abstract *There is a growing threat to the cyber-security of safety-critical systems. The introduction of Commercial Off The Shelf (COTS) software, including Linux, specialist VOIP applications and Satellite Based Augmentation Systems across the aviation, maritime, rail and power-generation infrastructures has created common, vulnerabilities. In consequence, more people now possess the technical skills required to identify and exploit vulnerabilities in safety-critical systems. Arguably for the first time there is the potential for cross-modal attacks leading to future 'cyber storms'. This situation is compounded by the failure of public-private partnerships to establish the cyber-security of safety critical applications. The fiscal crisis has prevented governments from attracting and retaining competent regulators at the intersection of safety and cyber-security. In particular, we argue that superficial similarities between safety and security have led to security policies that cannot be implemented in safety-critical systems. Existing office-based security standards, such as the ISO27k series, cannot easily be integrated with standards such as IEC61508 or ISO26262. Hybrid standards such as IEC 62443 lack credible validation. There is an urgent need to move beyond high-level policies and address the more detailed engineering challenges that threaten the cyber-security of safety-critical systems. In particular, we consider the ways in which cyber-security concerns undermine traditional forms of safety engineering, for example by invalidating conventional forms of risk assessment. We also summarise the ways in which safety concerns frustrate the deployment of conventional mechanisms for cyber-security, including intrusion detection systems.*

[1] Contact: School of Computing Science, University of Glasgow, Scotland, G12 8RZ. http://www.dcs.gla.ac.uk/~johnson, Johnson@dcs.gla.ac.uk

1 Introduction

There is a growing threat to the cyber-security of safety-critical systems. This is, in part, due to the integration of a small number of Commercial off the Shelf (COTS) products across the supply chain of national critical infrastructures. In previous generations, critical infrastructures tended to rely on bespoke systems that were not reused across different industries (Johnson, 2015). Now COTS components in safety-related applications include, but are not limited to, Linux, VOIP and Satellite/Ground Based Augmentation Systems (SBAS) such as WAAS in North America and EGNOS in Europe. A growing number of potential attackers have the technical knowledge to undermine safety-related applications across the transportation, energy distribution, food and water industries, etc.

At the same time, we have seen the rise of a new generation of terrorist threats, based around semi-stable regimes that resemble nation states. These regimes have access to trained engineers and process equipment within their borders. State-like terrorist regimes have become skilled in cyber-security, partly as a consequence of the policies implemented by Western governments. Police and intelligence agencies have denied them access to conventional social media. These regimes have responded by developing cryptographic skills and peer-to-peer networking using techniques originating with the deep or dark web. This has connected terrorist groups with strong political, religious and ideological motivation to the semi-commercial hackers who already sell zero-day exploits, malware libraries and root kits.

A key theme in this paper is that neither government nor private industry has moved at the rate required to maintain our defences against the growing threats from cyber-criminals and terrorist states. Public-private partnerships have failed to deliver regulatory guidance and appropriate audit mechanisms. The fiscal crisis prevented many safety regulators from recruiting and retaining staff with sufficient expertise in both cyber-security and safety-applications. Cultural and organizational barriers between safety and security have compounded this. It takes time before someone with a deep knowledge of conventional cyber-security can also gain an understanding of the concerns that arise, for instance in the nuclear or aviation industries. This is important when cyber-security techniques cannot simply be transferred from more conventional office based systems to safety-critical environments.

Political and organizational barriers also help to explain our limited progress in securing national critical infrastructures. These barriers arise because different regulators are responsible for the cyber-security of national data networks and for the safety of particular industries. In the UK, this is illustrated by the distinction between OFCOM, the Centre for the Protection of National Infrastructures and the Civil Aviation Authority or the Office for Nuclear Regulation. In the United States, similar distinctions arise between the Federal Communications Commis-

sion, the Department for Homeland Security and the Federal Aviation Administration or the Nuclear Regulatory Commission, not to forget the National Institute of Standards and Technology (NIST) as the body responsible for the Federal cyber-security provisions within the Federal Information Security Management Act (FISMA), or the host of local and State organizations that may also be included as stakeholders. These organizational and political distinctions create significant practical consequences. Companies often do not understand their reporting obligations for cyber incidents across national and international agencies, including Computer Emergency Response Teams (CERTS), police, intelligence and critical infrastructure organizations, as well as telecoms and industry regulators. This situation is compounded when superficial similarities between safety and security have led to the development of inappropriate policies that cannot be sustained using existing engineering practice. The following pages focus on two classes of concern. Firstly, there are situations in which cyber-security concerns undermine existing safety practices:

- Conventional safety risk assessments cannot be sustained when systems might be exposed to coordinated and malicious attacks;

- Existing safety-management systems offer limited support for cyber-security – especially given differences in incident reporting and root cause analysis between these two areas;

- Cyber-security concerns challenge many existing safety-related software engineering techniques, for instance, the use of software diversity and N-version programming lead to extended supply chains that are difficult if not impossible to secure.

The second, inverse set of concerns arise when safety issues complicate the application of existing cyber-security techniques:

- The limitations of conventional intrusion detection systems. White list enumerations of permitted processes cannot easily be applied to complex legacy systems and there are dangers when valid, safety-related processes are denied necessary resource. In contrast, black list enumerations of malware do not work because of the failure of cyber incident reporting in safety-related systems, noted in the previous section;

- The limitations of conventional forensic techniques. Existing support tends to focus on Internet Protocol related systems rather than on industrial Supervisory Control and Data Acquisition (SCADA) infrastructures using very different Programmable Logic Controllers (PLCs) and protocols. Further concerns stem from the competing risks that arise when deciding to either immediate isolate a compromised system leaving ap-

plication processes in a potentially unsafe state or in continuing to operate until shut-down but with the risk of over-writing critical evidence;

- The limitations of conventional cyber-security policies for air-gapped SCADA systems. For control systems, where many devices are not networked the implementation of conventional security patching policies may arguably increase rather than decrease potential vulnerabilities.

2 The Failure of Safety Critical Techniques in Cyber Security

The introduction has argued that we are ill-prepared to face a growing range of threats against the COTS infrastructures that support many safety-critical industries. Later sections will explain why a range of existing cyber-security techniques cannot easily be applied in safety-related domains. In contrast, this section explains why safety-techniques are often compromised by cyber-security concerns.

2.1 Cyber-threats Undermine Safety Risk Assessments

The European Network and Information Security Agency (ENISA, 2006) and the European Air Traffic Management Organisation (EUROCONTROL, 2006) advocate the use of safety management concepts to support the cyber-security of critical infrastructures (Johnson, 2015). In conventional applications, safety management systems use incident reporting and other forms of operational monitoring to determine whether an implemented system meets the safety requirements derived form an initial risk assessment. If new forms of hazard emerge, or if the system failed to adequately mitigate an identified risk, then further development is required. In other words, risk assessment, design and operation, monitoring and incident reporting form a virtuous circle.

These components of safety management systems also provide the foundations of information security management systems. Risk assessment helps to identify threats and vulnerabilities. Appropriate design and operating procedures help to ensure that the threats are mitigated. Incident reporting and audit provide the feedback necessary to revise the initial risk assessments when new threats emerge and to identify any situations in which operations fail to meet the requirements derived from an initial risk assessment. In theory, the use of similar concepts should support the integration of safety and information security management systems. Unfortunately, these superficial similarities hide a host of differences that undermine attempts to transfer the benefits of safety management systems into the security domain, revealing the lack of engineering expertise and operational experience that has informed much previous guidance (Johnson, 2015a). For example, the presence of an intelligent adversary undermines independence

assumptions in conventional safety assessments. Blended attacks are timed to coincide with routine component failures. Similarly, if the symptoms of one form of cyber attack are identified then it is very likely that a system may have been compromised in other ways.

Cyber-security concerns undermine existing safety engineering practices in other ways. Not only do they challenge the probabilistic components of risk assessment, cyber-threats also undermine safety-related consequence assessments. One reason is that we have limited experience of the new forms of advanced persistent threat, such as the state machine that varied the behaviour of Stuxnet or the use of Command and Control servers to hide Duqu from intrusion detection systems. This makes it very dangerous to predict the potential outcome from future modes of attack. The growing interconnection of critical infrastructure leads to hidden interdependencies. There are also concerns over 'cyber-storms' where a single attack brings down many different infrastructures – for instance when critical systems run under the same variant of Linux or where multiple services depend on timing information from the same satellite infrastructures.

None of these caveats would be significant if we had a range of tools and techniques that could be used to combine conventional safety risk assessments and cyber-threat analysis. Fault trees have superficial similarities to attack trees but the underlying semantics are different. In consequence, many organisations end up with parallel systems that are incapable of transferring lessons between safety and security. There are some notable exceptions (Pièrre-Cambacédès and Bouissou 2010, Johnson, 2015). However, there are few published case studies in integrated approaches to safety and security and even less agreement over the general utility of these tools across different industries.

2.2 Cyber-threats Challenge Safety Incident Reporting

It has taken many years to establish strong incident reporting cultures across safety-critical industries. In contrast, very few companies have the same security reporting culture. One reason for this is that employees reporting safety concerns are typically protected by a 'no blame' or 'proportionate blame' environment. In contrast, security violations can trigger disciplinary or legal action against other employees. A mismatch between security policy and practices can undermine a nascent reporting culture. Management implicitly approve of many security violations, for example the use of USB devices by sub-contractors, because they are anxious to maintain operations. Such incidents are seldom reported.

Many companies are reluctant to report incidents to Computer Emergency Response Teams (CERTs), regulatory agencies or industry associations. The loss of control and the reliance on external agencies can also compromise intellectual property where investigators must be familiar with commercially sensitive information in order to diagnose the causes of an attack. There are also political con-

cerns over the exchange of information about cyber-security incidents across national borders, even to otherwise friendly states.

Further barriers prevent the use of incident reporting to support safety and cyber-security. Lessons learned applications ensure that safety recommendations are disseminated as widely as possible. The aim is to avoid any recurrence of potential accidents. However, the disclosure of information about a cyber-incident might encourage future attacks. It can undermine market confidence; it can trigger regulatory action and litigation.

While there are well-established reporting mechanisms for safety concerns, cyber-incident reporting has been undermined by a series of 'turf wars' across Europe and North America. For many companies, it can be unclear whether reports should be sent to an industry regulator, such as the US Federal Aviation Administration or Nuclear Regulatory Commission, to a security agency, such as the Department of Homeland Security, or the US CERT, or to telecoms regulators who have responsibility for collating information about wider cyber-security concerns, such as the Federal Communications Commission. In some cases, a single incident must be reported to more than one agency. For example, in the UK a cyber-attack with safety related consequences must be reported to the national industry safety regulator and potentially also to a subset of the National Crime Agency, the National Cyber Crime Unit, GOVCERT, the UK Information Commissioner as well as the CESG/Centre for the Protection of National Critical Infrastructure via providers registered under the Cyber Incident Response (CIR) or the Cyber Security Incident Response Scheme (CSIR).

2.3 Cyber-security Undermines Safety-Critical Development

The tensions between safety and security extend across the engineering lifecycle, from risk assessment to detailed development practices. For instance, redundancy is typically used to increase the dependability of critical systems. If one component fails then a backup can maintain operation. However, this provides few benefits in software related systems without some level of diversity. Two redundant versions of the same code are likely to contain the same bugs and hence will fail in the same way, even if they are provided with slightly different data. In consequence, N-version programming techniques rely on using two or more contractors to develop multiple versions of the same program. In the event that one fails, it is intended that the other will not. The use of a diverse supply chain helps to ensure that both programs do not share common bugs, assuming that their requirements are correct.

Unfortunately, software diversity creates immediate problems for security management. The end user must secure two or more different supply chains – in other words, redundant diversity opens up multiple routes through which compromised code might be integrated into a safety-critical system. The customer must audit multiple sub-contractors to ensure that they meet agreed cyber-security

requirements. These security concerns are seldom considered in safety-critical software development. Customers have few guarantees that suppliers have vetted their staff, have prevented the introduction of code from untrusted sources, etc.

The meta-level point is that integrating safety and security reveals a host of tensions, which can only be addressed through integration. Without this, we cannot assume that the two isolated communities will deliver viable solutions to the problems identified in this paper. For example, a small number of safety-critical companies are now offering diverse supply chains from within their own organization. In other words, they will provide customers with two pieces of software each performing very similar functions but with assurances that they were implemented by different teams of employees using diverse development methods. This simplifies the supply chain but relies upon a range of innovative software management and development practices. It remains to be seen whether such practices are strong enough to address the natural safety concerns that arise when redundant software comes from the same supplier.

3 The Failures of Safety Techniques in Cyber-Security

The previous section has argued that there is an urgent need for integrated tools and expertise because safety-techniques are often compromised by the introduction of cyber-security concerns. In contrast, the following paragraphs argue that existing cyber-security policies cannot easily be applied in safety-related domains. We focus on three examples:

1. Conventional intrusion detection systems undermine the safety of complex applications;
2. Secondly, existing forensic guidance for conventional office based systems would lead to loss of life in safety-critical systems;
3. Finally, the air-gapped architectures of many SCADA environments undermine existing principles of security management.

3.1 Safety Concerns Limit Cyber-Intrusion Detection Systems

NIST (2012) advocate the use of several different intrusion detection systems (IDS) within critical applications. This raises significant concerns when an IDS might erroneously block critical processes in safety-related applications. There are two main approaches to intrusion detection. Blacklisting relies on detecting the characteristics of malware. Whitelisting is discussed in subsequent sections and relies on recognizing approved code.

Most blacklist IDS are designed to protect office-based systems. The signatures and symptoms of malware are compiled from evidence about incidents re-

ported through a range of mechanisms, including honey pots (Spitzner, 2002) but also through confidential reporting to the major security companies. Section 2.2 summarized the barriers that prevent the exchange of information about cyber incidents in safety-critical systems. Unless we can identify the malware signatures that characterize the growing threat to industrial control systems then attempts to develop blacklist IDS will provide very limited protection for SCADA applications (Naedele, 2007).

In order to protect a system, it is important to update a blacklist as soon as a malware signature has been identified. However, this creates problems in safety-related applications where there are requirements to conduct exhaustive tests prior to any software modifications. Uploading a corrupted blacklist could also cause the failure of a detection system with knock-on consequences for safety-related processes. Safety engineers would face a difficult decision between the competing requirement to update blacklists as soon as a new signature was identified and the requirement to ensure that the new signature did not undermine the safety of application processes.

In contrast to blacklisting, whitelist IDS ensure that only approved programs can be executed. They profile and report any deviations from 'normal behavior'. This can be implemented by creating a hash digest of all software applications. If the hash of an executable does not match anything in the list, it will trigger a security event. It is also important to prevent unauthorized users from changing the lists indicating which files can be executed.

Whitelisting offers benefits for safety-critical applications. The focus on identifying 'normal processes' eliminates the need to continually update malware signatures. Whitelists provide some protection against zero-day exploits – even if the signature of an attack is unknown, the malicious code will not be included on the approved hash list. However, the application of this approach raises a number of concerns. For instance, the same attack across multiple instances of a control system will simultaneously lead to a large number of distributed security events. These can overwhelm an organization's ability to respond in a timely fashion and may also be triggered by non-malicious causes. Software updates that are not reflected by changes in the whitelist can lead to a large number of false positives. Safety-critical processes could be denied computational resources. It, therefore, becomes imperative that staff and sub-contractors follow agreed security update procedures during all software installations. In some safety-critical applications this is relatively straightforward – for instance in long-lived SCADA systems where software updates on PLCs are relatively rare. In other contexts, such as air traffic management, where tens of sub-contractors each have intellectual property concerns, it can be very hard to determine what is and what is not a 'normal process'.

Data diodes ensure that information can only travel in one direction; for instance, by removing the send and receive transceivers from one direction of a fiber-optic cable. They can be used so that process data only flows from an operational zone to business systems but not vice versa. These devices can also isolate IDS from critical processes. The uni-directional flow of data reduces concerns that

the detection system will have an adverse effect on application safety. Unfortunately, greater levels of monitoring lead to an increasing number of false alarms. This can undermine cyber situation awareness and can lead to denial of service when operators incorrectly halt an application that they fear has been compromised. In contrast, raising IDS tolerance thresholds increase the potential for missed positives. In safety-critical systems this leads to the possibility that the over-tolerant configuration of an IDS allows companies to continue operating with malware inside critical applications. In conventional office-based systems, machine-learning techniques have been successfully deployed with threat visualization to integrate automated intrusion detection with human decision-making. Further work is required to determine whether these approaches might also be adapted to address the false-positive/false-negative concerns that undermine cyber-situation awareness in safety-critical systems.

3.2 Safety in Air Gapped Systems Undermines Cyber Policies

Most existing cyber-security tools and techniques focus on distributed architectures that are based around the conventional IP-stack. In contrast, many safety-related applications rely on computational devices such as PLCs that are isolated even from local area networks. The behavior of monitoring and control applications may not change for over a decade. The 'air gap' between the device and any network improves the cyber-security of safety-critical applications because it limits the opportunities for remote attacks. However, the 'air gap' also limits opportunities to use blacklist IDS. There is no easy way for system administrators to automatically update nodes with malware signatures. Operators must manually install any updates on each isolated device across the plant. This leads to a paradox. It is hard for any attacker to compromise a stand-alone PLC unless it is hooked to another device – for example to install a patch or update the IDS. The more often these updates occur then the greater the risk of cross-contamination. Systems managers of safety-critical systems, therefore, often deliberately ignore conventional cyber-security advice, preferring to leave isolated devices unpatched. Other problems limit the application of whitelisting in air-gapped systems. Without network access, it may be weeks or months before an engineer can examine the logs in sufficient detail to note an infection on a remote device.

3.3 Safety Concerns Undermine Conventional Cyber-Forensics

Detailed guidelines cover the forensic analysis of cyber-security incidents. For example, the US Department of Justice (2004, 2008) suggest that forensic investigators must preserve the 'chain of evidence':

- "Immediately secure all electronic devices, including personal or portable devices.
- Ensure that no unauthorized person has access to any electronic devices at the crime scene.
- Refuse offers of help or technical assistance from any unauthorized persons.
- Remove all persons from the crime scene or the immediate area from which evidence is to be collected.
- Ensure that the condition of any electronic device is not altered.
- STOP! Leave a computer or electronic device off if it is already turned off".

These principles support cyber forensics in office-based systems. They also illustrate the problems of integrating existing security practices into safety-critical systems. It is hard to envisage how any investigatory agency could immediately secure "all electronic devices" distributed across a compromised process control system, or indeed how to enumerate all of the devices connected to a modern, national air traffic management system. Many safety-critical companies have minimal access control policies, so that it is not always clear to external agencies who exactly has authorized access to the devices at a crime scene. Typically, there are strong forms of perimeter access – where only staff and authorized sub-contractors can gain access to a facility or machine room but once inside they have wide-ranging access to racks and network components. This is a strong contrast with financial institutions and even web service providers where it is normal to have fine grained access control policies that prohibit software engineers from accessing a machine room. Removing "all persons from the crime scene" could be catastrophic in a crowded Air Traffic Control centre or nuclear control room where operations, engineering and safety management must cooperate during contingency operations, including the aftermath of a cyber-attack. To "leave a computer or electronic device off if it is already turned off" would prevent the use of redundant protection systems.

The Department of Justice guidelines aim to preserve evidence by urging investigators to turn off any compromised systems. Continued operation may overwrite valuable data or enable attackers to disguise the manner in which a system was compromised. However immediately isolating a safety-critical system might endanger the lives of the public and of operators. Starting a fallback system can reduce this risk until an application reaches a safe state. However, this increases the risk of cross-contamination. In other words, halting a primary application to preserve forensic data can lead to the infection of the secondary system at a time when engineers and investigators are unlikely to know the mechanisms by which an attack was originally propagated. Without some form of integration, it is impossible for operational staff, senior management and investigatory agencies to balance the risks between the safety of application processes, the potential for cross-contamination and the legal requirements to preserve the evidence necessary for prosecutions in the aftermath of an attack.

4 Conclusions and Further Work

We face a growing range of cyber-threats to safety-critical systems. State-like terrorist groups have access to significant finance and engineering resources. The threats to safety-critical systems also stem from commercial markets in malware through the peer-to-peer networks of the Dark/Deep web where zero day exploits can be bought by those lacking the technical skills necessary to develop them. At the same time our vulnerabilities are increasing – through the integration of COTS applications including Linux, Voice over IP (VOIP) and Satellite Based Augmentation Systems into safety critical applications. The public-private partnerships established to enhance cyber-security have done little to address these concerns partly because the fiscal crisis and organizational barriers have left us without regulators who are competent in the cyber-security of safety-critical applications.

Superficial similarities between safety and security have led to the development of policies that cannot be sustained using existing engineering techniques. There are unique concerns for safety that prevent us from simply re-using existing guidance from office based systems in the aftermath of cyber-attacks. We cannot immediately isolate safety-related processes during forensic investigations without risking the lives of those who depend on critical infrastructures. Similarly, we cannot reuse convention Intrusion Detection Systems if these applications could block critical processes or if updates to malware signatures inadvertently bring down safety-related systems.

There are many areas for further work, including the causal analysis of cyber-incidents in safety-critical systems. Systemic factors help create the context in which an incident or accident is likely to undermine the safety of application processes. In contrast, security investigations tend to focus more on deliberate or unwitting violations. This focus on the direct human causes of a security incident is similar to the 'perfective approach' that characterized safety-related reporting more than a decade ago (Johnson, 2003). We might, therefore, expect that the focus of security investigations to shift towards systemic factors in the future. For this to happen it seems likely that we will need a new generation of root cause analysis techniques. Most existing approaches use counter-factual reasoning in the aftermath of safety-related incidents. Recommendations are derived by identifying causes, which had they been prevented then the incident would not have occurred. Such reasoning cannot easily be applied to cyber-attacks. It is hard to argue that a security incident would have been prevented given that adversaries launch multiple, simultaneous attacks, some of which go undetected.

References

ENISA, (2006), Risk Management: Implementation principles and Inventories for Risk Management/Risk Assessment methods and tools, Conducted by the Technical Department of ENISA, Section Risk Management, Heraklion, Greece, June 2006.

EUROCONTROL (2006), Safety Case Development Manual, Technical report DAP/SSH/091, Brussels, Belgium.

Naedele, M. (2007), Addressing IT Security for Critical Control Systems, Proceedings of the 40th Hawaii International Conference on System Sciences, IEEE Computer Society.

Johnson, C.W., (2003), Failure in Safety-Critical Systems: A Handbook of Accident and Incident Reporting, University of Glasgow Press, Glasgow, Scotland, ISBN 0-85261-784-4, available in in electronic form.

Johnson, C.W., (2015), Contrasting Approaches to Incident Reporting in the Development of Security and Safety-Critical Software. In F. Koorneef and C. van Gulijk (eds.), SAFECOMP 2015, Springer Verlag, Heidelberg, Germany, 400-409, LNCS 9337, ISBN 978-3-319-24254-5.

Johnson, C.W., (2015a), Barriers to the Use of Intrusion Detection Systems in Safety-Critical Applications. In F. Koorneef and C. van Gulijk (eds.), SAFECOMP 2015, Springer Verlag, Heidelberg, Germany, 375-384, LNCS 9337, ISBN 978-3-319-24254-5.

Piètre-Cambacédès, L. and Bouissou, M., (2010), Modelling Safety and Security Interdependencies with BDMP (Boolean logic Driven Markov Processes). IEEE International Conference on Systems Man and Cybernetics (SMC), 10-13 Oct. 2010, 2852 – 2861.

Spitzner, L. (2002), Honeypots tracking hackers. Addison-Wesley. pp. 68–70. ISBN 0-321-10895-7.

U.S. Department of Justice (2004), Forensic Examination of Digital Evidence: A Guide for Law Enforcement.

U.S. Department of Justice (2008), Office of Justice Programs, Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition, Washington DC, 2008. http://www.nij.gov/publications/ecrime-guide-219941/

U.S. National Institute of Standards and Technology (NIST, 2012), Computer Security Incident Handling Guide (Draft), Special Publication 800-61 Revision 2 (Draft), Gaithersburg, Maryland.