SAND2020-1299C

APPLICATION OF A SIMPLIFIED PROCESS TO IDENTIFY AND MANAGE SENSITIVE DIGITAL ASSETS

Michael T. Rowland
Sandia National Laboratories
Albuquerque, New Mexico, United States of America
Email: mtrowla@sandia.gov

John Sladek
Canadian Nuclear Safety Commission
Ottawa, Canada

Michael StJohn-Green
Mike StJohn-Green Consulting Ltd
Cheltenham, United Kingdom

Robert Anderson
Idaho National Laboratory
Idaho Falls, Idaho, United States of America

Abstract

The identification of digital assets and their management (i.e. assignment to security levels; specification of computer security requirements) within cyber security programmes at nuclear facilities has historically been a complex process. The current approaches use a system or asset-centric approach with the aim of applying cyber-security retro-actively. An example of such an approach is provided in US NRC Reg Guide 5.71 whereby Licensee systems are classified as critical systems if they have meet one or more of the following criteria: (i) Performs Safety, Security or Emergency Preparedness (SSEP) functions; (ii) Affects critical systems, functions or pathways; or (iii) Supports critical systems.

This paper outlines a simplified approach for the cyber security of nuclear facilities by aligning identification and management of digital assets to security goals using an integrated cyber security management system (CSMS). Different security goals require different processes and procedures to ensure these goals are achieved through the protection of the Licensee's significant functions. This approach avoids the complexity of the current asset-based approach by assigning computer security requirements to a few functions rather than directly to the many assets, in a manner analogous to nuclear safety.

This paper outlines a simplified process to establish a CSMS to identify and manage sensitive digital assets. Key to this process are (1) identifying security goals; (2) determination of the inherent value of functions and their SDAs; (3) determination of the compromise value of functions and their SDAs; and (4) Assignment of a final security level to SDAs;

No matter how capable the team performing the analysis, or how accurate the results are, compromise of digital assets can lead to indeterminate effects. Indeterminate effects reduce the confidence in the functional analysis that dominates elements (1) and (2), and necessitates element (3). The process for element (3) is to limit the potential for compromise resulting in indeterminate effects to those indeterminate effects that are linked to an adversary profile and to a credible scenario. This process will never be as accurate as the results of analysis of (1) and (2) since there are uncertainties adversary profile and the credible scenarios will not have high confidence. When the results of (3) are taken into account in the assignment of the security level as per element (4) it is effective.

1. INTRODUCTION

Protection of computer-based systems (including digital I&C systems) is recommended by the Nuclear Security Recommendations on Physical Protection of Nuclear Material and Nuclear Facilities (INFCIRC/225/Revision 5) [1], paragraph 4.10, which states that "computer based systems used for physical

protection, nuclear safety and nuclear material accountancy and control should be protected against compromise (e.g. cyber-attack, manipulation or falsification) consistent with the threat assessment or design basis threat". This same requirement is often used as a basis for national computer security regulation for nuclear facilities.

The current focus on 'systems' and 'assets' leads to tactical approaches (i.e. bottom-up) to computer security, which complicates the application of a strategy. The identification and management of sensitive digital assets (SDAs) is a key process within a strategy to ensure that security requirements are appropriately specified and imposed as well as ensuring that effective measures are put in place to protect either the significant function and/or the sensitive information. However, the strategy must focus on efficacy and efficiency that applies a graded approach that accounts for the limited resources that are available to protect these assets.

The process outlined in this paper demands as its primary objective that the inherent value of an SDA is captured completely (i.e. security goal, function contribution, effects of mal-operation). This necessitates the owner/operator of the SDA to undertake significant effort to identify attributes and elements that contribute to the value of the SDA.

Cyber-attacks have provided adversaries with a potential capability to place I&C systems in indeterminate states, where the function of the system can no longer be guaranteed [2]. This, coupled with the potential that an owner/operator may not correctly determine the value the SDA (i.e. the adversary might be able to discover its true value; arbitrage[1]), there is a need for a final adversary-focused step to confirm or raise confidence in the determined SDA value.

This paper details three categories of security goals: (1) nuclear safety; (2) physical protection; and (3) information security. Security levels will be used to an approximation of and a proxy for 'inherent value' of the SDA [3]. Security level '1' represents the highest or most stringent protection requirements while security level '5' represents the lowest or least stringent protection requirements. This simplified approach is based upon the Facility and System Computer Security Risk Management processes detailed within [3].

2. CURRENT PRACTICES

2.1. Nuclear Safety Paradigm

For Nuclear I&C systems, a mature process is used to categorize safety functions and classify safety systems. I&C systems important to safety are identified on the basis of their necessary I&C safety functions and the definition of systems that perform certain combinations of these functions [4]. The systems important to safety are based on the following fundamental safety functions that are required for all plant states [5]:

- Control of reactivity;
- Removal of heat from the reactor and from the fuel store;
- Confinement of radioactive materials, shielding against radiation and control of planned radioactive releases, as well as limitation of accidental radioactive releases.

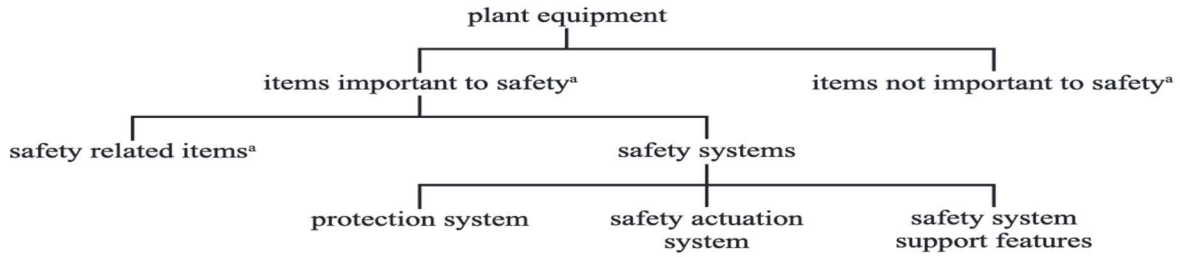
Para 5.34 of [5] specifies that the method for classifying the safety significance of items important to safety shall be based primarily on deterministic methods complemented, where appropriate, by probabilistic methods, with due account taken of factors such as:

(a) The safety function(s) to be performed by the item.
(b) The consequences arising from failure of the item to perform its safety function.
(c) The frequency with which the item will be called upon to perform a safety function
(d) The time following a postulated initiating event at which, or the period for which, the item will be called upon to perform a safety function

The identification of the functions performed by an item is a critical element in determining its safety classification. This approach is generalized by [6] as follows:

---

[1] the adversary may discover that the value of an asset is greater than that assigned by the owner, and therefore the protection given by the owner is not proportionate to its value to the adversary



[a] In this context, an 'item' is a *structure, system or component.*

Fig. 1 Plant Equipment in Terms of safety function [6]

Currently international consensus categorizes functions of Safety Systems as Category A (with minor exemptions), while those of important to safety are categories B and C. Based upon these categories the correspondence with safety classes is as shown in the table below.

Categories of I&C Functions important to safety			Corresponding classes of I&C systems important to safety
A	(B)	(C)	1
	B	(C)	2
		C	3

Table 1: Correlation between classes of I&C systems and Categories of I&C functions [4][2]

The categorization of I&C control functions and classification of systems have been captured in the Level 1 international standard IEC 61513:2011 [7], and further detailed in the Level 2 international standard IEC 61226:2009 [8]. The resulting classification determines the relevant design criteria and is applied to all the information and command functions, and the I&C systems that provide those functions [8].

The IAEA provides guidance on categorization and classification. Fig. 1. from [9] highlights the overall approach.

2.1.1. US NRC [10]

US NRC Reg Guide 5.71 provides a framework to aid in the identification of those digital assets that must be protected from cyber-attacks. These identified digital assets are referred to as "critical digital assets" (CDAs). Licensees should address the potential cyber security risks arising from compromise of CDAs by applying the defensive architecture and the collection of security controls identified in this regulatory guide [10].

This approach focuses on CDAs and their controls. The US NRC also provides for a categorization of CDAs as direct or indirect, which provides for a graded approach to the application of controls.

2.1.2. IEC 62645 [11]

The IEC approach to assigning security levels[3] uses the safety class as the primary factor, but allows for an adjustment based on other factors. The function's categorization for safety is used to establish a minimum security degree and the adjustment would allow for an upwards (more stringent) increase in security degree. For example, a Category C/Class 3 system would be assigned a minimum security degree of '3' but could be adjusted upwards to security degree '2' or '1'.

The IEC standard does not extend beyond I&C or Electrical Systems and does not address Emergency Preparedness (EP), Nuclear Material Accounting and Control (NMAC), and Physical Protection Systems (PPS).

---

[2] Parenthesis indicates that the category can be elevated to a higher corresponding class of I&C system, but this is not mandatory.

[3] The IEC uses the term 'Security Degrees'. This is equivalent to 'Security Levels'

It may be extended to other systems based on the competency of the assessor, and the security culture of the organization.

2.1.3. Canadian Approach, CSA N290.7 [12]

The Canadian approach is captured within Canadian Standards Association CSA N290.7:2014 *Cyber security for nuclear power plants and small reactor facilities*. . CSA N290.7 requires that facilities identify Cyber Essential Assets (CEAs), the Canadian equivalent of SDAs, and determine their significance (i.e. high, moderate or low), which is the equivalent of the computer security level of NST047. The scope of N290.7 covers I&C systems used for safety and also includes EP and PPS. I&C systems which perform NMAC functions are included within the scope of the standard, but NMAC systems which are information technology systems are not.

For I&C systems performing safety functions, the primary factor for determining significance is the safety categorization which is determined using a similar process to IEC 62645 [10].]. EP systems are typically assigned a low significance. For PPS the significance is determined based upon the security function performed by the asset, and the assigned significance can be low, moderate or high.

CSA N290.7 requires that a computer security architecture be based upon defense-in-depth. This differs from IAEA guidance which recommends that the DCSA implementation establishes defense in depth. IAEA recommends that the requirements for DCSA be based upon a trust model with the objective of reducing adversary access to attack pathways. The DCSA is established based upon the arrangement of zones and measures that meet these requirements (i.e. to enforce access controls and restrict information flows to those authorized and compliant with the trust model). However, "CSA N290.7-14 does not require the implementation of a zone model with graded protective principles." [13]. Furthermore, CSA N290.7 does not require that the security architecture be developed based upon a trust model.

2.1.4. UK approach

The UK approach strongly aligns with that used for safety class with three additional elements that may raise the level of the SDA or system. These are:

- System provides main displays within the main control room;
- System has a large network;
- System if compromised may result in an extended plant shutdown.

If a category B system has any of these elements, the system is considered a 'significant category B' system and is assessed to meet the category A requirements [12]. This approach considers the potential for compromise (greater opportunity for systems with large networks) and consequences of loss of view (displays in main control room) or loss of control (extended plant shutdown)[4]. This approach is strongly based on the IEC 62645 approach (enhanced by providing additional guidance for category B functions) and has similar limitations.

2.2. Comparison of Safety and the Proposed Approach

The IAEA safety approach [9] mirrors strongly with the proposed process for SDAs. The key steps of both approaches are summarized below:

Process	Safety [9]	Proposed Approach (CSMS)
Establish Design Basis (Management System)	Basic understanding of plant design, its safety analysis and how main safety functions will be achieved.	Understand security goals, their security analysis, and how the main security goals will be achieved.

[4] The IAEA does not consider 'plant shutdown' as a consensus regulatory concern for safety, but Member States can determine other criteria upon which to regulate.

Process	Safety [9]	Proposed Approach (CSMS)
Inherent Value of Functions	Identification of all functions necessary to achieve main safety functions	Identification of functions that implement, support, or assist in realizing the security goals;
	Categorization of all Functions	Assign a security level to all functions based upon the potential consequence (i.e. the harm that could occur if the function is not provided when needed, or if the function has been maliciously modified).
	Identification and classification of systems, structures, and components	Identify the digital assets (or systems) that perform or support these functions and assign a security level to the digital assets based upon the contribution that the digital asset provides to provision of the function (i.e. directly performs function, directly supports the function, or indirectly supports the function);
Fault Tree Compromise Analysis of functions	Identification of design provisions to prevent accidents.	Evaluate the effects of compromise of functions using an adversary profile and characterization.
	Identification and classification of SSCs implemented as design provisions	Identify and assign a security level to the digital assets based on effects of compromise.
Decision	Is the classification correct and complete? i.e. The Inherent Value and the Fault Tree Value are identical	Is the security level assignment correct and complete? i.e. The inherent and the compromise value are identical. If not, assign the level requiring the highest protection.
Selection of Measures	Select applicable engineering design rules	Select applicable cyber security control methods[5]

Table 2: Safety and Security Processes

From the table above, the two approaches are analogous; SDAs and SSCs are identified and assigned (or classified) using strategic processes similar to Fig. 1 of [9]. The key difference is the proposed approach requires different methods be used at each process step to ensure the achievement of Safety, Security, and Information Security goals. That is to ensure that functions are performed as required, and that the nuclear facility is protected from malicious compromise.

Safety approaches align strongly with the current IAEA guidance and with the process for identification and management of SDAs outlined in this paper. Safety approaches are associated with mature and established processes (e.g. Quality Assurance, Document Control, and Procurement) that provide and raise confidence in the achievement of Nuclear Safety.

These mature processes conflict with a strategic approach for cyber security since these processes did not consider computer security. Achievement of the security goals requires different processes and a cyber security management system (CSMS) that maintains and achieves those goals. The CSMS is subordinate to the Facility's Integrated Management System.

For example, programmatic Security Level 2 requirements for Safety must not be identical to programmatic Security Level 2 requirements for Information Security or for Physical Security as the underlying Trust Models and demands in the delivery of security are considerable different. Similarly, the processes for configuration control, and quality assurance must accommodate the differences between these domains to ensure the attainment of the security goal. For example, a technical control measure assigned to security level 2, may

[5] Selection of methods is done in a graded manner and based upon the requirements for the security level, but are outside of the scope of this paper.

require an urgent security patch or configuration update achieve the security goal, which would be difficult to achieve in a timely manner if the quality assurance requirements for safety category B were applied to the security measure.

Additionally, Safety identifies key attributes such as (1) function, (2) consequence of failure, (3) frequency of performance of the function, and (4) places time constraints upon the performance of the function. This fails to meet security objectives since it does not consider the effects of compromise on the system function [2] such as indeterminate states, unexpected behaviours or actions (e.g. loss of control), or malicious acts that impede actions taken in response to postulated initiating events (e.g. loss of view). Addressing these security attributes require s additional analysis, tools, and procedures to ensure these requirements are addressed.

This paper proposes that while the strategic processes for achieving safety and security goals align, the processes to accomplish those goals are different. The application of safety processes to computer security needs to be applied to only those SDAs for which the primary security goal is also safety. Application of safety processes to SDAs that have other primary security goals (e.g., security, information protection) are highly likely to impede or severely limit the ability of these SDAs to provide their function in achievement of their goal throughout their lifecycle.

3. COMPLEXITY OF 'SYSTEM' AND 'ASSET' APPROACHES

The asset or system-based approaches for classification contravenes not only strategic computer security approaches, but also those outlined for safety above. While there are a limited number of potential functions, there is an immensely large number for potential 'systems' and 'assets' that implement each function.

These implementations could be similar to or significantly different from one another. This increases both entropy (e.g. an identical function has many possible system implementations) and increases complexity.

3.1. Simple Analogy using Passwords

The focus on 'assets' can be highlighted by an analogy of a password. In this case, a specific password shall be considered as an 'asset' providing a function. The function being to verify the identity of a user during authentication. The primary goal of the password in this case will be to provide 'Information Security'.

Application of an Asset-centric approach demands an investigation into the specific implementation of the asset, in this case, the password. When considering a specific implementation of a password, the strength must be determined. Generally, this considers both length and the number of potential characters to choose from at each position. Many organizations require passwords of a certain length and the use of different 'categories' (symbol, lowercase, uppercase, number) to be included in a password to ensure that the chosen characters are diverse, thereby decreasing the likelihood of successfully guessing a password, and somewhat increasing the effort required to 'crack' the password.

When comparing two passwords, the relative strength can be determined, but the rationale for the specific sequence of characters or the 'design' decision would be extremely difficult to determine since it is based on personal and environmental factors (unless randomly generated passwords are used). For example, a lowercase character may be used at position 2 in password #1, while password#2 utilizes a number at the same position. These implementation differences have little direct importance to the strength of the password.

Regardless of the strength of the passwords used, the implementation of the user authentication function may be deficient, allowing an adversary to bypass it without the need to guess or 'crack' the password. If the implementation is deficient, the strength of the password is irrelevant.

Consequently, the goal of information security needs to prioritize protection of the authentication process for all users over those for a single user/password. This goal and function require processes that are customized and efficient for both the user and organization.

4. SIMPLIFIED APPROACH

The simplified approach establishes a framework (e.g. CSMS) that develops, implements, maintains, monitors, and coordinates subordinate processes that are customized for each of the categories of security goals.

This simplified approach is outlined through a 4-step process: (1) identifying security goals; (2) determination of the inherent value of functions and their SDAs; (3) determination of the compromise value of functions and their SDAs; and (4) Assignment of a final security level to SDAs.

Each process step will be outlined below and then illustrated with three case examples.

4.1. Identify and enumerate the nuclear security goals

The nuclear security goals for a nuclear power plant at a high level are:

- *Physical Security* – protect against the unauthorized removal of nuclear material
- *Physical Security* – protect against sabotage resulting in unacceptable radiological consequences (URC)
- *Information Security* – protect against the unauthorized disclosure, alteration, modification, destruction or denial or use of sensitive information
- *Safety* – Control of reactivity;
- *Safety* – Removal of heat from the reactor and from the fuel store;
- *Safety* – Confinement of radioactive materials, shielding against radiation and control of planned radioactive releases, as well as limitation of accidental radioactive releases.

These goals are graded based upon the associated consequences. For security, IAEA NSS No. 13 [1] provides graded requirements for protection against unauthorized removal and sabotage. For information security, IAEA NSS 23-G [15] provides graded requirements for protection of sensitive information. For safety, SSG-30 [9] and the aforementioned IEC standards [7][8] provide a categorization that may be used to impose graded requirements for systems providing these functions.

Distinct, customized processes are developed based upon the category of security goals. For example, Safety goals will likely adopt and adhere to the safety processes with some adaptation to allow for CSMS integration. Physical Security will need to adapt to new design processes to account for computer security. Information Security will need to adopt processes that allow for integration within the overall Facility Management System.

4.2. Identify the functions that provide, support, or assist in realizing the security goals and assign a security level to each of the functions and their SDAs.

Once the security goals have been identified and enumerated and a framework for grading has been determined, the functions that provide, support, or assist in the realizing of these goals need to be identified.

NST047 describes these as “facility functions” with relationships shown below [3]. Facility functions differ from the component or system function as they exist independent of implementation. The facility function is the primary consideration in the application of a graded approach using the concept of security levels. The facility function is assigned a security level based on the potential consequence associated with compromise of the function.

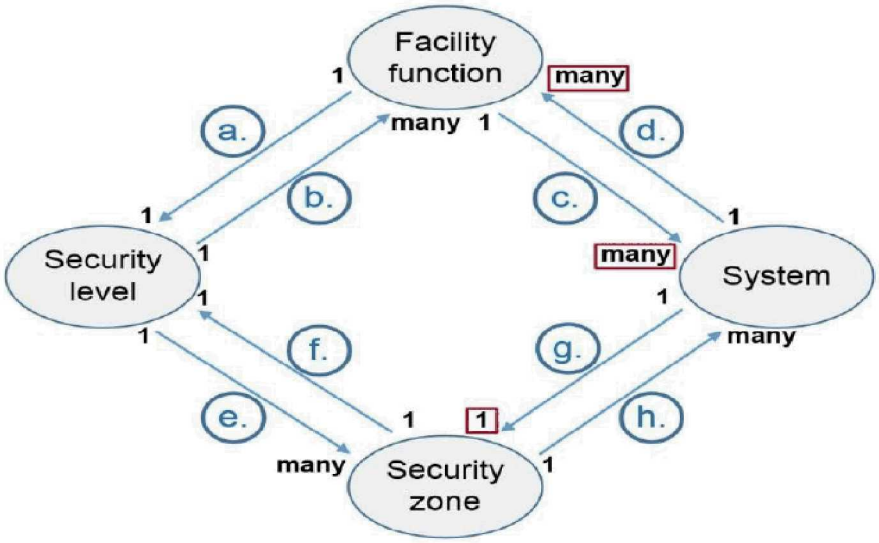


Fig. 3 – Relationships between Computer Security Concepts [3]

This step looks at the potential consequences of non-performance of the facility function. The objective of this step is to determine the consequences on the function which are listed in para 2.21 of [16]. To determine the extent of the consequences, a graded effort to identify and analyze potential indeterminate states and unexpected behaviours or actions needs to be undertaken.

This step should not consider the adversary but should be performed from a designer's perspective to enumerate specific asset states that would fail, degrade, or alter, a function or that would provide an opportunity to affect other functions. It is critical that the designer enumerate fully all credible unexpected behaviours or actions and potential altered states (i.e. to bound potential or likely indeterminate states).

Once the impact to functions is known and bounded, the digital assets that perform or support these functions should be identified. In new designs, the assets, configuration, and architecture can be restricted (through design constraints) to ensure that security requirements are met. However, in legacy designs the assets are already in service and security needs to be retroactively applied to existing assets that may have been designed without explicit security requirements.

The identification of assets must follow the approach outlined in the System Computer Security Risk Management Process detailed within para 5.5 of [3]. This approach analyses necessary information flows that exist within the implementation of the system to determine the 'connectivity[6]' and 'trust' of assets. Connected and trusted assets are assigned the same level as the function.

4.3. Evaluate the effects of compromise using an adversary profile and characterization

This step raises confidence in the evaluation by applying an independent process that evaluates actions or sequences of actions taken of the adversary with a specific focus on confirmation of the security level arrived at in step 2. Adversary actions should be based upon the adversary characterization contained in the Design Basis Threat, augmented by open source threat intelligence, and other information sources that provide information on adversaries having cyber skills. The most capable adversary profile should be used to generate attack trees (or credible scenarios) that have the potential to lead to unacceptable consequences. The number of credible scenarios must be sufficient to provide for a high-quality assessment of the effects of compromise upon the system function. A security level is assigned based upon the consequences arising from these scenarios.

[6] Assets are considered connected if there is any mandatory periodic information flow between them with no technical control measure in place to restrict or control this flow of information.

4.4. Assignment of a final security level to SDAs

If the security level assigned in steps 2 and 3 differ, then the most conservative security level should be assigned to ensure appropriate protection of the function and its SDAs. This step can only raise the security level (i.e. increase requirements) and not lower them.

5. CASE EXAMPLES

For all case examples below, Step 1 is captured in a sample subset of security goals in section 4.1 above.

5.1. Case Example 1: Safety Function – Control of Reactivity – Control of neutron flux during normal operation within the core via increase or subtraction of neutron absorption capability

5.1.1. Step 2

The security goal is *Safety – control of reactivity*; while the function being identified is the control of neutron flux within the reactor core during normal operation via addition or subtraction of neutron absorption of capability. This can be completed many ways (1) insertion and removal of control rods, in most conventional designs; (2) adding light water to or removing it from the liquid zone control compartments in CANDU designs; or by other means. The implementation elements will be detailed below.

For this step, the function is analyzed similarly to [8] and [9] which considers:

(a) The consequences of failure to perform a safety function.
(b) The frequency with which the item will be called upon to perform a safety function
(c) The significance of the contribution of the function to the security goal

This results in a categorization of the safety function, would be Category B, as the consequences of failure to perform the function would result in transition from normal operation to abnormal operation. The function is required to reach and maintain a safe state during normal operation. Using a conservative approach, the failure to reach and maintain a safe state is high (i.e. unsafe reactivity) and therefore, using table 1 of [9], the safety category assigned is 2. This would imply an initial security level of '2'

Based on the example, the Supervisory control function that if not provided or altered could lead to failure to reach or maintain a safe state. However, this function is not accredited for reaching controlled states after anticipated operational occurrences or after design basis accidents. Nevertheless, the constraint of this function needs to be understood and documented to ensure that the correct requirements are specified and enforced.

The function was assigned security Level 2. This step needs to consider the implementation of the function via digital systems. For example, in a system that uses both light water and control rods (graphite) to absorb neutrons. In a hypothetical system, containing a single network to control both the addition/removal of light water or control rods, consists of 'x' assets. It is assumed that there is one controller, one network, and a number of sensors and actuators (all digital).

All 'x' assets would then be assigned Security Level 2.

5.1.2. Step 3

The output of step 4, was to assign all 'x' assets to security level 2. However, given certain adversary capabilities, it is likely that large networks provide greater opportunity for them to access sensitive digital assets and affect their significant function. Additionally, large networks may use commercial of the shelf equipment or network communication protocols or services that have publicly known and critical vulnerabilities that reduce the level of effort required on the adversary to exploit these devices. These attributes would result in not only an increase in opportunity, but also an increase in the probability of success for an adversary that was able to access these devices or network.

Given the potential for indeterminate states (e.g. mal-operation) [2] resulting in amplification of the severity of consequences associated with failure to perform the safety function, or assuming a credible scenario whereby the adversary can disable the Category A systems or degrade them to prevent them from recognizing

certain conditions (e.g. mal-operation of the Category B system; 'x' assets); or become aware of latent design deficiencies not considered or known. This would necessitate, based on conservative decision making, in an increase of the security level from 2 to 1.

5.1.3. Step 4

The two processes arrive at different security levels (i.e. '2' and '1'). As security level 1 is the most conservative decision, the assigned security level is '1'. This aligns with the UK ONR approach, with the 'significant category B' being an example of increasing the security requirements above that considered using safety assessment alone.

5.2. Case Example 2: Information Security - Enforcement of access controls for read and write operations across a zone boundary to protect integrity of sensitive information within a Biba trust model.

5.2.1. Step 2

The security goal of a Zone Boundary Technical Control Measure such as a Firewall (e.g. Security Gateway/Appliance), is *Information Security* – protect against the unauthorized disclosure, alteration, modification, destruction or denial or use of sensitive information.

Just as in Safety, which mandates consideration of the function with respect to plant state, the information security goal must consider the security level (sensitivity of information; trust model). In this case, the function is to enforce access control rules for read and write operations across the boundary to protect the integrity of sensitive information of SDAs within the computer security zone being protected.

The failure to enforce access controls could result in the compromise of SDAs within the zone, but it is important to consider that the boundary technical control measure only gains significance based on the assets that are within the zone.

This function's significance is informed by the harm that may result from its failure to contribute to the security goal (in this case the protection of sensitive information/SDAs). Functions providing decoupling of zones must be considered separately from other functions that can exist at a single level.

In this example, the function will provide boundary protection (i.e. decoupling) between zones containing assets at security level 2 and security level 3. This function would be initially assigned a security level of 2/3. It is critical that the assignment to a security level considers the functions/assets that are being protected. For example, a decoupling of a zone containing zero assets would result in no security level being assigned to this boundary protection function.

However, the implementation and configuration will likely require a function to provide for remote connectivity and 'write-access' from the remote location to several level 2 functions, which needs to consider the corresponding trust model.

Failure of this remote connectivity function could degrade or alter the function to provide boundary protection thereby exposing the protected digital assets to compromise. However, this is a pre-cursor or support event, and although having significance, is not accredited with a design basis function that would result in failure to reach a secure or safe state given other malicious actions that must be directed against those SDAs that perform safety or security functions and deliver either a Safety or Security goal.

This step illustrates why functions that require technical control measures demand specific requirements that impose different measures. For example, the access control rules for protection of Safety SDAs are based upon the Biba trust model, while the protection of the confidentiality of the configuration of the firewall needs to be based upon the Bell-LaPadula model (with separation of duties/Brewer Nash also to be considered).

In this example, the function(s) could exist (or have) three distinct levels based on processes: (1) Security level 2 for the protected zone (inferred from the security level of the assets being protected); (2) Security level 3 (or higher) for the interface with the unprotected zone; and (3) Security level 2 for the administration of the asset (inferred from (1) and based on the remote connectivity function).

The function of the enforcement of access controls at the level 2/3 boundary can be most readily implemented via one or more firewalls. However, if implemented by more than one firewall, it would be important to ensure that an equivalent level of protection is provided to each firewall.

In this example, all firewalls would be assigned a level 2/3, the side of the firewall that connects to the protected area is level 2, with the outward facing side being assigned level 3. Any privileged access would also be assigned to level 2. The access controls detailed in NSS 17 [6] recommended for communications across a level 2/3 zone boundary are for write down, with limited handshaking allowed (e.g. TCP/IP). If the Firewall failed to provide the decoupling, the security requirements would not be met.

The decoupling assets (Firewall(s)) would therefore be assigned to security level '2'. But this does not take into account other functions that may be assigned to the asset such as zone access control administration and interface with unprotected side.

5.2.2. Step 3

The output of step 2, resulted in three security level assignments based upon implementation of the control. However, an adversary could compromise the administrative console or a specific vendor (supply chain attack) to make ineffective the boundary protections. This degrades defence in depth because of the loss of independence, but does not degrade the ability of the SDAs to provide the function.

This step is therefore important to reinforce the demand for segregation of duties, but also the need for independence and diversity for boundary controls.

Also critical is that technical control measures adopt configuration management and quality assurance practices that are complementary to the assets that they protect but are not identical. The need for agile deployment of countermeasures at the boundary control device is necessary to adapt to ever-changing adversary TTP.

The requirements for the security levels and management are further refined by this step, but in this example, the security levels remain unchanged from step 2. This is based on the attribute that no credible scenario exists whereby the decoupling assets (i.e. Firewall(s)) can directly[7] lead to an impact on a Safety or Security goal. Additionally, it cannot directly impact the information security goal as the Firewall is not an SDA so long as it does not store, process, control or transmit sensitive information[8].

5.2.3. Step 4

The assigned security levels determined in Step 2 and Step 3 align, so no change is necessary. The firewall shall have security level '2' requirements for the protected side and management port and security level requirements '3' for unprotected side. The requirements and procedures for these SDAs need to be customized to allow for an SDA to be assigned more than one security level (due to its location at the boundaries of zones).

5.3. Case Example 3: Security – Protect against unauthorized removal of categorized nuclear material via control of access to inner areas

5.3.1. Step 2

The security goal is *Security – unauthorized removal*; while the function being identified is the control of access to inner areas. This requires (1) identifying authorized personnel; (2) authentication of identity; and (3) provision of access. This is only one function that would be part of an integrated physical protection system that provides prevent, detection, delay, and response functions.

---

[7] 'directly' implies that no additional step or action is to be completed by the adversary before the consequences of unauthorized removal of nuclear material or sabotage resulting in unacceptable radiological consequences is achieved.

[8] Definition of sensitive information assets from IAEA NSS 20 [17]; the firewall is likely to have information that has importance for security, but if care is taken to compartmentalize information, as well as to implement independent and diverse controls within the overall DCSA, the consequences will not solely or directly compromise of nuclear security.

The function does not detect malicious acts by insiders, but minimizes the number of potential malicious actors. In this instance the analysis is similar to that in Case Example 1, which would imply an assignment of Security Level 2 based on the determination that the function is required to reach and maintain a secure state and a failure to perform the function would be a high consequence. The functions are not protective (i.e. detect, delay, response) and are required only to maintains a secure state.

Contemporary security designs apply a high degree of integration into all functions, including detection. Therefore, most PPS designs utilize a single 'flat' network to provide all functions. This results in a single zone assigned security level 2 encompassing hundreds of assets.

It becomes clear that a graded approach to provision of computer security to protect the functions is not possible, and without a gradation of requirements, an implementation of defence in depth becomes difficult. Based upon this example, all PPS assets would be assigned security level 2 based upon association with this function.

5.3.2. Step 3

Due to the current designs of PPS, the integrated flat network, would imply that mal-operation of access control function would degrade or alter the detection function. Para 2.42 of [3] identifies 'intrusion detection (including assessment) at the critical detection point' as a function important to security.

It is then conceivable that a capable adversary with opportunity to connect to the access control networks would be able to compromise some or all of the PPS SDAs. The flat network allows for direct trusted access from one asset to other assets (including those performing different functions). Assets providing intrusion detection could be potentially compromised from any point on the network, and compromise of the intrusion detection function is likely to result in failure to reach, maintain, or return to a secure state. This would result in a high likelihood that the malicious act would be completed prior to interception by response forces.

Given that intrusion detection protects against unauthorized removal and sabotage especially against insiders (i.e. undetected access to inner areas or vital areas), it would be prudent to assign Security Level 1 to the entire PPS to be coherent with Safety processes.

5.3.3. Step 4

The output of step 3, resulted in assignment of security level 1 to all PPS SDAs. This result is based upon flat network PPS design, and will incur significant effort to ensure security. The required effort may exceed available resources, stress the computer security team, and may lead to fatigue of the computer security. Step 4 therefore needs to consider and prioritize requirements to transition to a segmented architecture that is function based.

The segmentation would need to iteratively follow this simplified process to implement defence-in-depth and a graded approach to computer security. This step does not change the security level from '1', as it is constrained to only elevate, but may provide insight into future designs that can segregate functions along a secure architecture while allowing for the necessary inter-functional information flows necessary for security.

6. CONCLUSION

This simplified approach relies heavily on the Facility and System Computer Security Risk Management processes detailed within [3] as well as the Safety Categorization and Classification processes outlined in the IAEA Safety Standards Series [5] [9]and aforementioned IEC standards [7][8].

This paper also attempts to show how the process consists of three steps based on a secure by design approach with a fourth verification step as to confirm the results of the previous steps. The goals, functions, design, and assets are all specified by the designer. The difficulty is that mal-operation or indeterminate states of system functions are inherently unbounded. The competency of the designer (or evaluator) becomes critical in step 2 to identify the most likely and significant unbounded events and analyse them. Given that this process is knowledge-based it is necessary to add a step 3 to use an adversary-centric approach to confirm or discover additional (i.e. unanalyzed) events. Requirements to ensure the independence of steps 2 and 3 would lead to better outcomes associated with the overall process.

Fundamentally and strategically, goals must be achieved, and functions must be performed with no impediment from the programmatic requirements. The sole focus must be on the overriding objective to protect the correct performance of the function and undertaking design and analysis activities to raise confidence that both the protection and resilience of the assets guarantee that the function will be performed in the period of time for which it is required. This demands a CSMS that develops, manages, monitors, and coordinates different customized processes that efficiently provide protection to SDAs.

ACKNOWLEDGEMENTS

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

REFERENCES

[1] INTERNATIONAL ATOMIC ENERGY AGENCY, Nuclear Security Series, Recommendations, Nuclear Security Recommendations on Physical Protection of Nuclear Material and Nuclear Facilities (INFCIRC/225/Revision 5), NSS No. 13, Vienna, 2011.

[2] ROWLAND, M. T. et al. "Computer Security for I&C Systems at Nuclear Facilities", 116-19945, paper presented at 10th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies, San Francisco, United States of America, 2017

[3] INTERNATIONAL ATOMIC ENERGY AGENCY, draft publication Computer Security Techniques for Nuclear Facilities, To be published.

[4] WORLD NUCLEAR ASSOCIATION, Safety Classification for I&C Systems in Nuclear Power Plants – Current Status & Difficulties – CORDEL Digital Instrumentation & Control Task Force Report No. 2015/008, WNA, England, Sept 2015, http://www.world-nuclear.org/uploadedFiles/org/WNA/Publications/Working_Group_Reports/safety-classification-for-iandc-systems-in-npps.pdf

[5] INTERNATIONAL ATOMIC ENERGY AGENCY, Safety Standards, Safety of Nuclear Power Plants: Design, Specific Safety Requirements No. SSR-2/1 (Rev. 1), IAEA, Vienna, 2016

[6] INTERNATIONAL ATOMIC ENERGY AGENCY, Nuclear Security Series, Technical Guidance, Reference Manual, Computer Security at Nuclear Facilities, NSS No. 17, Vienna, 2011.

[7] INTERNATIONAL ELECTROTECHNICAL COMMISSION, Nuclear power plants - Instrumentation and control important to safety - General requirements for systems and standard, IEC 61513:2011, IEC, Geneva, 2011

[8] INTERNATIONAL ELECTROTECHNICAL COMMISSION, Nuclear power plants - Instrumentation and control important to safety - Classification of instrumentation and control functions, IEC 61226:2009, IEC, Geneva, 2009

[9] INTERNATIONAL ATOMIC ENERGY AGENCY, Safety Standards, Safety Classification of Structures, Systems, and Components in Nuclear Power Plants, Specific Safety Guide No. SSG-30, IAEA, Vienna, 2014

[10] UNITED STATES NUCLEAR REGULATORY COMMISSION, Regulatory Guide 5.71, Cyber Security Programs for Nuclear Facilities, January 2010

[11] INTERNATIONAL ELECTROTECHNICAL COMMISSION, Nuclear power plants - Instrumentation and control systems – Requirements for security programmes for computer-based systems , IEC 62645:2014, IEC, Geneva, 2014

[12] CANADIAN STANDARDS ASSOCIATION, Cyber security for nuclear power plants and small reactor facilities, CSA N290.7, CSA, Toronto, 2014

[13] INTERNATIONAL ATOMIC ENERGY AGENCY, International Physical Protection Advisory Service, Mission Report: Canada, 19-30 October, 2015, http://www.nuclearsafety.gc.ca/eng/pdfs//IPPAS/Canadas-IPPAS-Mission-Report-2015-eng.pdf.

[14] DYER, P. "Cyber Security on Nuclear Plant in the UK", IAEA Technical Meeting on Engineering and Design Aspects of Computer Security in Instrumentation and Control Systems for Nuclear Plants, Gloucester, UK, May 2017

[15] INTERNATIONAL ATOMIC ENERGY AGENCY, Nuclear Security Series, Implementing Guide, Security of Nuclear Information, NSS No. 23-G, Vienna, 2015.

[16] INTERNATIONAL ATOMIC ENERGY AGENCY, Nuclear Security Series, Technical Guidance, Reference Manual, Computer Security of Instrumentation and Control Systems at Nuclear Facilities, NSS No. 33-T, Vienna, 2018.

[17] INTERNATIONAL ATOMIC ENERGY AGENCY, Nuclear Security Series, Nuclear Security Fundamentals, Objectives and Essential Elements of a State's Nuclear Security Regime, NSS No. 20, Vienna, 2013.