Logical Foundations for Protecting Materials and Facilities from those with Malicious Intent—Proposed 1st Principles for Security Systems

Sue A. Caskey, Adam D. Williams, Lauren Crabtree, and John "JR" Russell
Sandia National Laboratories*, Albuquerque, NM, USA,
[sacaske; adwilli; lcrabtr; jlrusse] @sandia.gov

Introduction

The implementation of processes to protect us and our resources from those with malicious intent has been witnessed within every known civilization. This long history of developing protective solutions that meet the operational, environmental, technological, and intellectual constraints of a given time provides useful lessons learned and insights. Additionally, significant efforts in the U.S. has been put forth in recent decades to leverage this observations and insights to protect critical pieces of infrastructure. Yet, these lessons and insights still tend to be applied in an ad hoc fashion. This paper will explore a collection of proposed 1st principles, aiming to demonstrate that these principles are the fundamental concepts of security.

To define the 1st principles, the team worked to build a paradigm – or an outline of definitions – used to bound the security system conversation. This paradigm then supported the development of higher-level concepts to derive 1st principles. The team also worked backwards from existing security system heuristics to define the 1st principles and theories for security systems. The figure below reflects on the relationships between the 1st principles, theories, and heuristics.

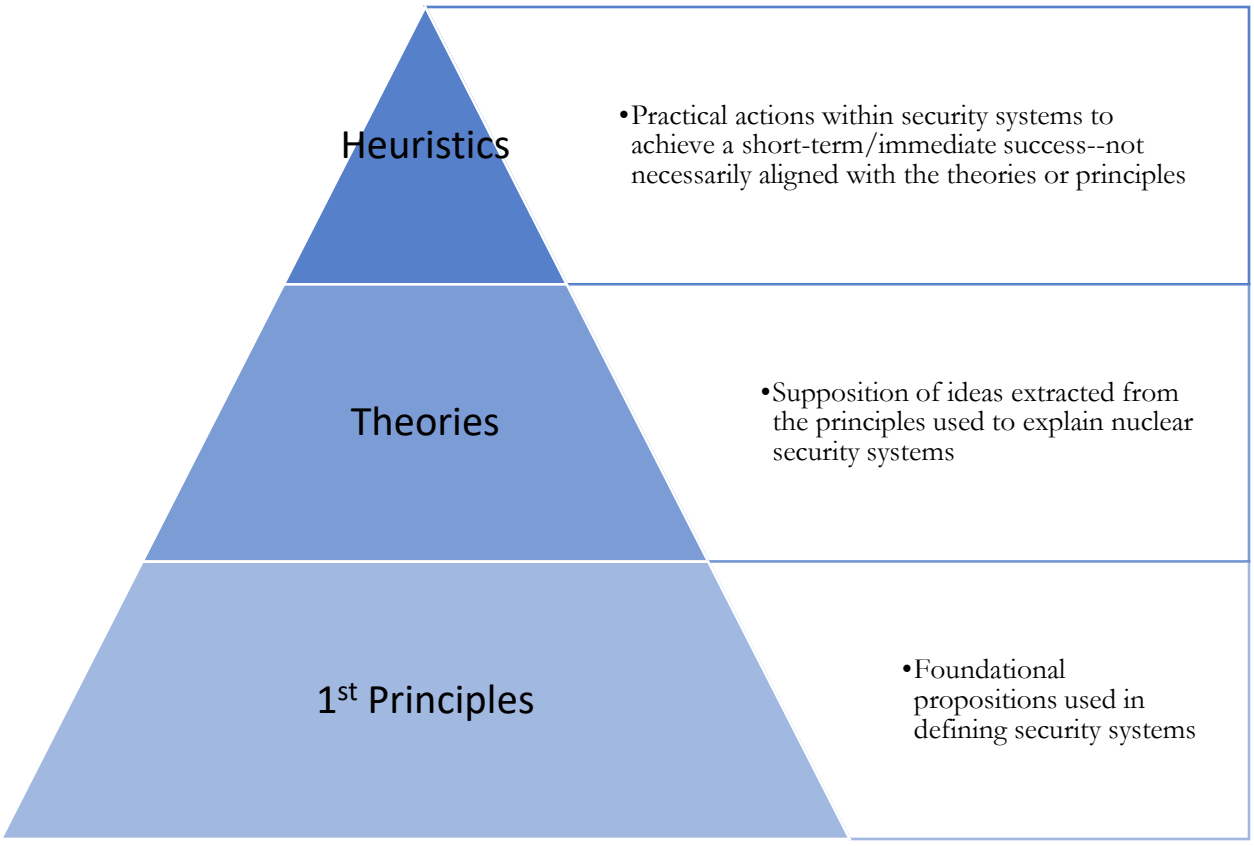

Figure 1 Relationships and definitions used in this paper for 1st principles, theories, and heuristics as defined and used by the team in this report

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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
SAND2021-9308 C

1st Principles

The 1st principles presented in this paper are based on a dynamic, systems theoretic paradigm of nuclear security. Specific characteristics related to this paradigm of security include:

- To be secure is to be in a state free from threat, driven by the intent of a threat and not just the absence of attractiveness. As threats can be considered *dynamic, complex systems* existing within our environment, this secure state also becomes dynamic and can be impacted by external fluctuation in the environment or from the threat.
- Security systems are also *dynamic, complex systems* whose performance directs movement related to a secure state—suggesting that any internal or external perturbations (e.g. component behavior, weather, threat actor capabilities, etc.) can move the system closer or further from this state.
- Security *risks* are the gaps between current state and secure state.

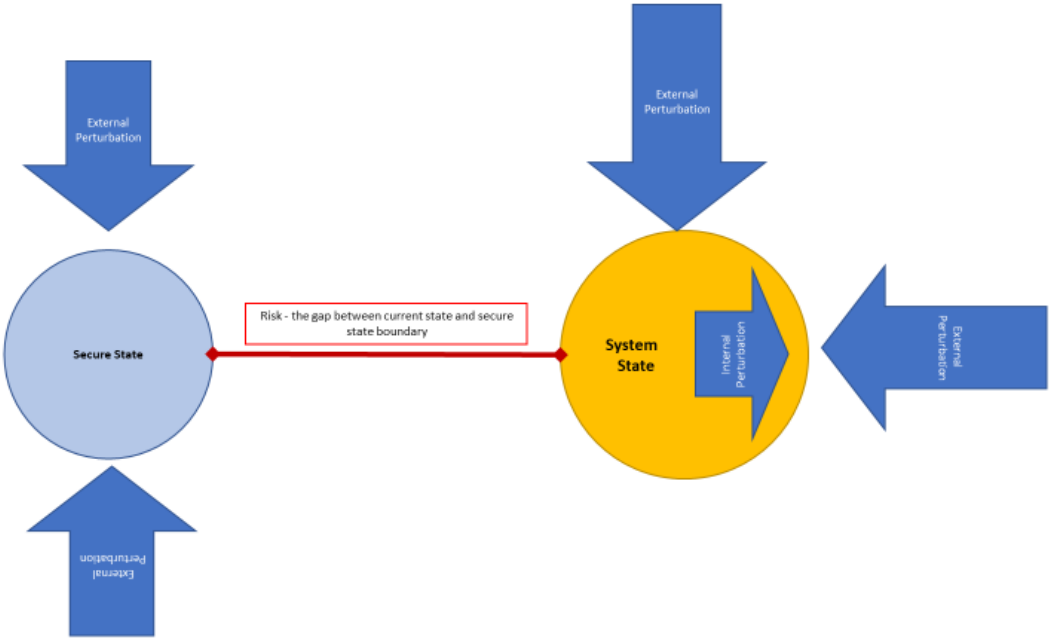

Figure 2 Representation of the systems theoretical paradigm of nuclear security. In this diagram the perturbations (internal or external) are reflected as arrows that could move the secure state and the current state of the system closer or further apart. The red line reflects the security risk or the gap between the current system state and the defined secure state.

This security paradigm was outlined based on discussions with security experts (Gunda, 2021-submitted) and leveraged concepts presented in several classic security system texts (Garcia, 2008), (Biringer, 2007). From this paradigm, we have identified three 1st principles of security systems, theories, and several heuristics that can be used to support security system design. Each principle will be outlined in detail along with systems theories and examples. The defined principles are interdependent –there are complex cause and effect relationships between them. While each serves an independent foundation for security systems, there is an overlap between them conceptually.

The 1st principles of security systems are defined as:

1. Security risk will never be zero,
2. Security risk is dynamic; and,
3. Threats are adaptive.

The diagram below is a simple Venn diagram highlighting the overlapping elements of each principle.

Figure 3 Venn diagram reflecting the three 1st principles of security systems and how they are not independent but rather overlap

First Principle #1 is based on the concept that no security system will ever be perfect, that is there will always be some level of risk. Risk can be reduced or shifted, but never fully removed, specifically when considering the protection of materials and facilities defined as high consequence facilities[1] (HCFs) (U.S. Department of Homeland Security, 2020). Since we have defined a security risk as the gap between the current system state and the secure state, this principle equates to a secure state always being just out of reach of our current system state. Garcia (Garcia, 2008) defines security risks based on the intentionality of the threat actor, the ability of the threat actor to achieve the intended goal, and the consequences of achieving this goal—simplified mathematically as:

$$Risk = P_A \times (1 - P_E) \times C$$

P_A = probability of attack (for a specified timeframe) – this reflects the intentionality of the threat actor
P_E = probability that system will be effective against attack – this reflects the security system's effectiveness at changing or reducing the ability of the threat actor to achieve their goal
C = Consequence of attack – the consequences of the threat actor achieving their goal.

Based upon the above equation, risk could be zero if P_A or C is 0, or if P_E is 1. However, the reality is there will always be a potential for an attack, and there will always be some level of consequence of the attack (especially when considering HFCs) unless a HFC is no longer operational. And while in a perfect world a security system could prevent successful attacks from an adversary, the reality is that no system is perfect. The goal of the security system is to perform at a level sufficient to

[1] Defined as those whose incapacitation would have a devastating impact on national security, economic prosperity, and/or public health

balance out the potentiality of an attack and the consequences of an attack – creating a low or acceptable risk level. This mirrors the concept in radiation safety of "as low as (is) reasonably achievable," (U.S. NRC, 2021), where some level of risk is considered acceptable but it is recognized that the risk will never be zero.

In thinking more specifically regarding the security system and its overall effectiveness, a security system is a complex and multi-dimensional system. As such it is susceptible to perturbations – including from those within the system (internal perturbation) as well as by those external to the system (external perturbation), that can impact emergent behaviors. For example, in thinking about internal perturbation, a security system is directly dependent on humans supporting and functioning within the system (e.g. guards, operators); humans are far from 100% predictable or reliable. Similar, a security system is often dependent on external infrastructure such as power or communication, which again are far from 100% reliable (external perturbation).

Finally, in modeling a system behavior, it is not conceivable to understand every defining element - the darkness principle of systems theory (Whitney, 2015). This suggests that for approaches like Garcia's, (Garcia, 2008) to define a risk as zero likely reflects a lack of fully characterizing all the element needed to define the risk. Consider the recent pandemic of the novel coronavirus (Covid-19) -security guards are a key element in security for most HCF, but due to positive cases and quarantine requirements, the world witnessed a security guard shortage. In an article from the UK (Joshi, 2020), security guards had one of the highest reported deaths from Covid-19, by profession. This is a tragic yet effective example of how unforeseen perturbations in the system can dramatically change the system's overall effectiveness.

First Principle #2 is that security risk is dynamic. This principle also influences the first 1st principle in that since security risk is dynamic, as the security system moves closer to the ideal secure state, the dynamic, complex nature of both will cause any direct overlapping to likely be short lived . While it may be possible to calculate a risk value, common approaches do so specific to static perceptions regarding threat, system state, and even the consequences. Considering again the risk equations defined by Garcia, each of the three variables defined is subject to the perceptions of those defining the problem space. From the systems theory principle of complementarity, we know that the broader the perceptions the more the representation will reveal about the system (Whitney, 2015), but we also know our understanding of the system is subject to the darkness principle so we will never have complete knowledge. As such, while a calculated risk value may aid in understanding the system, this value is only reflective of a moment in time and a specific set of perspectives.

Considering the dynamic nature of the system (creating internal perturbations), the emergent behavior of attackers (external perturbations), and changing environmental conditions (external perturbations) – there exist perturbations which can alter the intentionality of the threat, the effectiveness of the system, or the consequences. The cyber-attack that occurred against an Indian nuclear power plant in 2019 is a clear example of how emergent behavior of attackers can impact the system's effectiveness altering the risk in a rapid and dynamic manner (Singh, 2019) – this example also illustrates the relationship between the dynamic risk and the adaptive nature of threats (First Principle #3) in that the adaptive behavior of the threat becomes an external perturbation pushing the secure state and the security system further apart. An interesting example of how internal perturbation that can occur by external situations could be reflected in studies looking at the impact of communication reliability across coaxial cables based on rapid temperature changes (Sobolewski, 2003). Consider locations like in Russia where coaxial cables are used to communicate information

regarding the state of security – in 2020, there was a heat wave across Siberia that created a rapid change in temperature– while not specifically reported on, this rapid temperature change had the potential to impede security system's alarm and assessment communication. Similarly, sociopolitical environment can also create external perturbations pushing the system further (or potential nearer) to the secure state. For example, a rapidly changing political climate as was witnessed in the winter of 2021 when a riot in Washington D.C. stormed the US Capital Complex (Barrett, 2021) - this riot was a situation the security system of the US Capital Complex was unprepared to address therefor the risk (distance between the secure state and the system's current state) were increased.

First Principle #3 the team defined is the adaptive nature of threats. Threat actors are adaptive in both the intent of their actions as well as the methods of their actions. Specifically, they may alter in their intent based upon opportunities or challenges. In defining threats, most often we are focusing on the actions or perceptions regarding intent from past actions. While prediction of future actions is not actually possible, recognition of the adaptive nature of threats is critical in considering security risk. An example of the emerging and adaptive behavior of threat actors was demonstrated in the 2008 Lashkar-e-Taiba attacks in Mumbai. During this attack, the actions of the actors within the Taj Mahal Palace could be defined as opportunistic as they exploited the real-time media reports to alter their attack strategies (CNN Editotorial Research, 2020).

While the security paradigm states that risk is defined based on the intentionality of the threat as compared to the attractiveness of the material or facility, attractiveness can influence emergent and adaptive behaviors of the threat actor. Consider the reports regarding threat actors' interest, specifically that of the Islamic State, in chemical weapons following the use by Syria against its civilian population (NTI, 2020). This successful misuse altered the intent of threat actors and created a potentially new potential of attack from these actors toward chemical HCFs.

Security Systems Theory

These three 1st Principles have created a foundation upon which to build security systems a theory. We propose the following security system theory - adequate security performance emerges from actively observing and proactively responding to security risk. By extension, a security system should not be evaluated/analyzed as static. A security system must also be implemented to support and align with the operational objectives of the HCF to include other systems such as the safety system.

Observation can be conducted by use of detection measures. Detection is broadly defined as an action or process to identify the presence of something, specifically with the ability to differentiate between information-bearing patterns (e.g. a person) as compared to random patterns (Wilmshurst, 1990). In many domains this would be referred to as the sensitivity and the specificity of the measure. Observation should also include environmental scanning, broadly defined to mirror concepts from Beers Viable Systems Model (Espejo, 1998) – in general, this would include watching for any signaling behaviors that could define emerging threat behaviors or other perturbations that could impact the system.

In considering why *active* observation is important, security systems have been recognized as far back as the Mycenean Age (Collins, 2015). Based on historical documentation: walls, moats, hedge rows, etc. were created to help protect people and important materials, in this case often food stores, from adversaries. These barriers were created in concentric layers (protection layers) with the most

important materials in the inner most layer and identified areas requiring increasing stringency of *active* observation. Historical records also discuss the success of the patient thief, the adversary who was able to get through the protection layers as no one was watching. More recently this issue was witnessed in Brazil (Lehman, 2005), where thieves worked for months to dig a tunnel under a city street and broke into a vault acquiring nearly $70 million USD. As such a security system without *active* observation (or detection) can be considered a deterrent but does not comply with the proposed security systems theory.

As defined as part of detection, the security system must be able to specify it has detected something and have the sensitivity to recognize a response is needed, then can implement a response. The response is the reaction stated in the security systems theory.

A concept that must also be defined as part of a security system is the access processes. A HCF would not be able to meet its operational objectives without allowing entry by those requiring access – e.g. workers needing to perform specific tasks within the HCF and specifically may need access to materials or equipment that could be considered a target for a threat actor. Access control processes can be generalized to be the processes that allow an authorized individual to bypass the detection system (Russell, 2020). This individual must be authenticated as part of the process and accountable for both their access and their actions within the HCF.

Heuristics

If 1st principles are the foundation for theories, then theories are the basis for analysis methodologies, be they formal analysis techniques or heuristics. For this paper, we leave the formal analysis techniques for subsequent discussion and define a set of heuristics related to the proposed 1st principles and theory.

The heuristics based on our presented theory could be stated as:

- Without detection, physical security barriers are only a deterrence,
- Without assessment and response to the detected threat, there is no detection, and
- Without resilience security risk will grow over time.

The following are security design parameters should be considered in order to ensure the systems followed the presented heuristics. The systems should include a continuous line of detection, that is no gaps within the detection perimeter. The system should be designed to ensure the specificity of the detection measures align so there are no weak links within the continuous line of detection. And the detection measures should be implemented in depth, which is the use of layers of detection in concert with protection layers (Russell, 2020). In considering response, the response following detection and assessment must be sufficient in prevention of the threat's intent regarding the HCF. This includes timeliness and adaptiveness in the behaviors of the responders (Sandia National Laboratories, 2018). A security system must also have metasystems that support the system's ability to absorb, recover, and maintain itself in the wake of internal perturbation and external turbulence. These concepts mirror those from complex systems governance and resilience theories (Gunda, 2021-submitted). The concepts include (Whitney, 2015), (Gunda, 2021-submitted):

- A security system must have a process to ensure the system's objective is supported in the event of changing conditions,
- The security systems must have the ability to regulate its internal environment to maintain stable operations,

- The system much be able to make internal adjustments to maintain stable operations,
- The system needs sufficient redundancy to ensure stable operation,
- The system needs sufficient diversity to ensure stable operation, and
- The system's elements need to be sustainable and maintainable as well as the systems as a hole.

Summary

This paper has proposed a collection of 1st principles, systems theories, and heuristics that can be used in considering security risk and used in designing and the evolution of security systems. The aim of this paper is as a kick starter in furthering discussions regarding these topics and allow for refinement or solidification of these concepts.

References

Barrett, T. R. (2021, January 7). US Capitol secured, 4 dead after rioters stormed the halls of Congress to block Biden's win. *CNN Politics*, pp. https://www.cnn.com/2021/01/06/politics/us-capitol-lockdown/index.html.

Biringer, B. M. (2007). *Security Risk Assessment and Management: A Professional Practice Guide for Protecting Buildings and Infrastructures.* Hoboken: John Wiley & Sons.

CNN Editotorial Research. (2020, December 3). Mumbai Terror Attacks Fast Facts. *CNN World*, pp. https://www.cnn.com/2013/09/18/world/asia/mumbai-terror-attacks/index.html.

Collins, P. R. (2015). *Principles of Security and Crime Prevention.* Oxforshire: Routledge.

Espejo, R. H. (1998). *The Viable System Model: Interpretations and Applications of Stafford Beer's VSM.* Hoboken: Wiley.

Garcia, M. L. (2008). *The Design and Evaluation of Physical Protection Systems, Second Edition.* Boston: Butterworth-Heinemann.

Gunda, T. C. (2021-submitted). Revisiting Current Paradigms: A Systems Approach for Nuclear Facility Security Assessments. *Journal of Nuclear Materials Management*, Currently in Review.

Joshi, A. (2020, June 26). *Coronavirus: Security guards are most at risk of dying with COVID-19, figures show.* Retrieved from Sky News: https://news.sky.com/story/coronavirus-security-guards-are-most-at-risk-of-dying-with-covid-19-figures-show-12015241

Lehman, S. (2005, August 10). Tunnel leads to Brazil's biggest bank heist ever. *Associated Press.*

NTI. (2020, October 23). *The Chemical Threat.* Retrieved from Nuclear Threat Initative: https://www.nti.org/learn/chemical/

Russell, J. (2020). Security system principles. (S. Caskey, Interviewer)

Sandia National Laboratories. (2015). *Pandemic Response Plan.* Albuquerque: Unpublished.

Sandia National Laboratories. (2018). *Response.* Retrieved from International Training Course on the Physical Protection of Nuclear Facilities and Materials : https://share-ng.sandia.gov/itc/course-materials.html

Singh, A. C. (2019). Lessons from the cyberattack on India's largest nuclear power plant. *Bulletin of the Atomic Scientists*, pp. https://thebulletin.org/2019/11/lessons-from-the-cyberattack-on-indias-largestnuclear-power-plant/.

Sobolewski, J. (2003). Coaxial Cable. In R. Meyers, *Encyclopedia of Physical Science and Technology (Third Edition)* (pp. 277-303). Tarzana: Elsevier Science Ltd.

U.S. Department of Homeland Security. (2020, July 10). *Critical Infrastructure Sectors.* Retrieved from Cyber Infrastructure Security Agency: https://www.cisa.gov/critical-infrastructure-sectors>

U.S. NRC. (2021). *ALARA*. Retrieved from NRC Library - Glossary: https://www.nrc.gov/reading-rm/basic-ref/glossary/alara.html

Whitney, K. B. (2015). Systems theory as a foundation for governance of complex systems. *International Journal of Systems Engineering*, 15-32.

Wilmshurst, T. (1990). *Signal Recovery from Noise in Electronic Instrumentation.* Boca Raton: CRC Press.