SAND2017-10655C

Potential Weaknesses in the Cyber Systems of High-Security Physical Protection Systems

J. Clem[1], W. Atkins[1], R. Baker[1], J. Daley[1], V. Urias[1]

[1]Sandia National Laboratories (SNL), Albuquerque, NM, USA

E-mail contact of main author: jfclem@sandia.gov

Abstract

This paper represents an update of one previously presented by the main author at the International Conference on Computer Security in a Nuclear World: Expert Discussion and Exchange at IAEA Headquarters, Vienna 2015. This updated work describes important results from a multiyear research and development effort undertaken by researchers at Sandia National Labs who investigated potential weaknesses in the cyber systems of representative high-security physical protection systems (PPS). It first discusses general cyber threats to PPS, and then provides a detailed description and analysis of three cyber-enabled attacks including attack vector, relative difficulty, and consequences.

Key Words: Cybersecurity, Physical Security, Physical Protection Systems (PPS)

1. Introduction[1]

Clem et al., argue [1,2] the nuclear security community depends on physical protection systems (PPS) assuming they are isolated and therefore secure, work as intended, and are not especially prioritized for compromise or misuse by a determined adversary. Important concerns conveyed by an expert committee reporting to the United States Congress include (cyber) "interactions and dependencies among security countermeasures," and "the adequacy of attack scenarios used to design, update, and test the security systems" [3]. To address risks, the same experts recommended that the National Nuclear Security Administration adopt a total systems approach to "characterize the interactions and dependencies of security countermeasures…" [3]. The committee concluded it is critically important to understand the adversary, i.e., their objectives and perspectives on the security system itself [3].

[1] Section 1. includes both paraphrased language and excerpts from a paper previously delivered by the main author to the International Conference on Computer Security in a Nuclear World: Expert Discussion and Exchange at IAEA Headquarters, Vienna (2015). Minor edits have been made to support this paper.

Given the preponderance of digital information technologies in most modern PPS, and given the global availability of cyber attack tools (including custom exploit development), attack scenarios targeting theft or sabotage of nuclear assets now must consider cyber exploitation. In the authors' experience, a troubling number of stakeholders understand their PPS and associated network(s) to be isolated and therefore not reachable by adversaries through cyber means. Yet attackers have compromised cyber-based components in similar systems, including important cyber systems that govern processes and hardware in the physical world (e.g., such as those that manage critical infrastructure and other critical operations) [4]. Even when stakeholders acknowledge the general threat posed by cybersecurity vulnerabilities, they still do not know what *credible* cyber attack pathways might be available to an adversary (where vulnerabilities are discoverable, reachable, and exploitable). Unfortunately, current methods and tools used by the nuclear security community (e.g., physical security simulations) cannot validate the achievability of hypothesized cyber-enabled physical attacks available to adversaries. Specifically, current methods fail to enumerate and convincingly demonstrate exploitability of technical cyber vulnerabilities in the PPS. Hence decision makers do not fully comprehend potential cyber threat impacts on PPS performance.

This paper represents an update of one previously presented by the main author at the International Conference on Computer Security in a Nuclear World: Expert Discussion and Exchange at IAEA Headquarters, Vienna 2015. This updated work describes important results from a multiyear research and development effort undertaken by researchers at Sandia National Labs who investigated potential weaknesses in the cyber systems of representative high-security physical protection systems (PPS). It first discusses general cyber threats to PPS, and then provides a detailed description and analysis of three cyber-enabled attacks including attack vector, relative difficulty, and consequences.

2. A Brief Description of Cyber Threats to Physical Protection Systems[2]

Clem et al., [1,2] explain that many PPS subsystems and components communicate using modern Internet Protocol (IP) networks—the kind used throughout the world in enterprise environments. The evolution of our protection systems to include and rely upon digital technologies unintentionally has introduced an entirely new category of threats: cyber attacks targeting the digital portions of a PPS to improve an adversary's chances of success for one or more future physical attacks. These are called *cyber-enabled physical attacks* or *cyber/physical attacks* interchangeably.

At the field device level, controllers commonly run a commodity embedded operating system, such as VxWorks from Wind River or Windows CE from Microsoft, on top of which custom programs are executed to perform specific required functions. These commodity embedded operating systems provide tremendous benefits to vendors. All core aspects of an

[2] Section 2. includes excerpts from a paper previously delivered by the main author to the International Conference on Computer Security in a Nuclear World: Expert Discussion and Exchange at IAEA Headquarters, Vienna (2015). Minor edits and reductions have been made to support this paper.

embedded computing environment—kernel, application programming interface, hardware drivers, networking functionality, etc.—are available immediately and maintained by the operating system manufacturer.

Backend systems might be particularly attractive targets for attackers because of their central role in the operation of modern PPS. An attacker with control over one or more backend systems could significantly degrade the effectiveness or availability of a PPS by respectively reporting false information to operators or destroying the contents of permanent data storage. For the moment, ignoring potential vulnerabilities in the PPS vendor component firmware and software, it can be concluded that attackers familiar with only core commodity operating systems will see a large and familiar attack surface when targeting a PPS. Attackers, perhaps with little specialized knowledge, also could exploit flaws in these operating systems made public by vendors, security researchers and fellow attackers to obtain access to and negatively affect the performance and/or availability of a targeted PPS.

The authors are aware of the pervasiveness of built-in "backdoor" accounts and remote access mechanisms that many ICS vendors include in their products to support debugging and remote troubleshooting. In the vast majority of cases vendors do not provide any means to disable such backdoor mechanisms. In some cases, they require remote access as part of their conditions with the owner/operator prior to signing support contracts. These interfaces, unfortunately, are especially attractive to attackers because of their historically poor security design and implementation. Although the authors did not investigate instances of such backdoor mechanisms in PPS, future cybersecurity assessments could reveal their existence.

Modern PPS have adopted the approach of using Ethernet and IP networks as a communications backbone. It would not be unexpected to discover that a common physical infrastructure is used for both enterprise information technology (IT) and PPS networking at low- and medium-security installations. In such installations, logical—not physical—separation is the strongest form of isolation possible for PPS subsystems and components. Such isolation techniques (e.g., firewalling, virtual local area networks [VLANs], encryption, virtualization, and the use of "sneakernet" data transfer mechanisms) may not sufficiently prevent cyber attacks against the PPS that make use of the enterprise IT network as a hopping point.[3]

Attackers able to blend both cyber and physical attack means may use both wired and wireless PPS interfaces as avenues for gaining initial footholds and mounting an assault against digital PPS subsystems and components prior to conducting a physical attack. These attackers would rely on unprotected or improperly protected communications interfaces, i.e., those that permit communication with networked PPS components existing physically outside the boundaries of the area(s) protected by the PPS. In some cases, an attacker may surreptitiously implant their own equipment to access these interfaces remotely to decrease the probability of detection, attribution, and/or apprehension during a future physical attack.

[3] Techniques for defeating these forms of logical isolation are well understood but outside the scope of this R&D.

3. Three Demonstrations of Cyber-Enabled Physical Attacks[4]

Clem et al., [2] report that a multi-disciplinary R&D team of physical security system engineers, technologists, and cybersecurity experts was formed to identify and validate credible cyber threats to SNL's representative PPS testbed using red team techniques. SNL's Integrated Security Facility (ISF) is a physical security research, development, testing, and training area with a fully functional and integrated PPS. The ISF was created from the decommissioned security systems surrounding a former Category I nuclear material site. The ISF includes a Perimeter Intrusion Detection and Assessment System (PIDAS) featuring a fully sensored exclusion area with inner and outer security fences, entry control portals, high security lighting, vehicle barriers, video surveillance systems, personnel access control systems, and many additional capabilities. In accordance with best practices for real-world high-security sites, the ISF PPS is architected in sectors around its protected areas. Multiple vendor AC&D systems are installed, each assigned monitoring and control for a given set of sectors. The ISF testbed extends to a remote storage bunker that is sensored, alarmed, and monitored from the CAS in the main location. The bunker's PPS elements are interconnected to the CAS via a terrestrial wireless communications link. The ISF's PPS is interconnected via a TCP/IP communications network and instantiated in commodity IT servers and workstations running commercial AC&D and other supporting software.

3.1. Demonstration One: Hacking the Access Control System from a Remote Bunker

The first cyber-enabled physical attack scenario identified and implemented by the R&D team is characterized as an outside-inside attack. The attack begins on the outside of the area protected by the PPS and then moves to networked PPS systems that enable adversary access to the area protected inside the security perimeter.

The R&D team's cyber attacker discovered a vulnerable communications box at the base of a wireless communications tower at the testbed remote bunker site. With this discovery, the attacker inserted low-cost, off-the-shelf, consumer-grade wireless networking gear into the communications box.[5]

Next, the attacker moved approximately one kilometer from the remote bunker site to conduct the actual penetration (using wireless communications), demonstrating the ability to avoid unwanted attention from security personnel. Next, the attacker connected wirelessly to the gear he inserted in the communications box at the remote bunker. Once connected, he accessed the PPS network. The attacker discovered the access control server and found it

[4] Section 3. through section 3.4. includes excerpts from a final report of research written by the authors of this paper, Emulytics for Cyber-Enabled Physical Attack Scenarios, SAND2017-1603, 2017. The text includes both minor edits and reductions to support this paper.

[5] The attacker effectively performed a Man-in-the-Middle attack by rerouting the PPS communications from the bunker through his gear and back to the PPS wireless communications tower. In addition, the gear the attacker inserted served as an unauthorized wireless access point to which he could connect, and thus he could then connect to the PPS subnet that serves the remote bunker.

configured with weak administrator credentials, then guessed the system administrator login. The attacker accessed the system's access control software used to manage access authorization at the testbed site. He enrolled a blank access card that can be obtained easily and set a personal identification number (PIN) for the card. He gave himself site-wide access privileges. Later, the attacker walked up to the facility and entered the protected area uncontested, using his recently enrolled access credential.

To complete the attack scenario, the cyber attacker used a laptop computer, publicly available no-cost attack software (Metasploit), and less than $400 of additional equipment.[6] It is important to note that no changes to any system configuration were made to facilitate attack success, and that vulnerabilities in the PPS architecture and implementation enabled the attacker to easily avoid detection at the remote bunker.

In summary, the R&D team used an adversarial attack methodology to successfully:

1. Discover testbed PPS vulnerabilities;
2. Leverage vulnerabilities in the PPS architecture and implementation to connect to the PPS network;
3. Exploit weaknesses in the PPS network and backend servers to access the access control software;
4. Enroll a blank physical access card in the access control database; and then
5. Walk into the protected area undetected, using the attacker's unauthorized credential.

3.2. Demonstration Two: Implantation of Cyber Technology to Overcome Network Isolation

A second attack scenario identified and implemented by the R&D team is characterized as an insider-enabled external attack. The attack begins with the implantation of low-cost commercially available hardware tools inside the ISF CAS and results in the adversary's remote access to the PPS network and systems. As with the first attack, the adversary compromises the access control system to enroll an unauthorized credential, resulting in undetected access by a physical attacker.

The portrayed adversary is assumed to have gained physical access to the PPS network from inside the CAS. The attacker in this first stage implanted two inexpensive Ethernet over Powerline adapters (also referred to as powerline network adapters and available from most home electronics retailers) to extend the PPS network using the electrical distribution cabling at the CAS. One was implanted near the PPS server rack and the other on the outside of the

[6] The additional equipment consisted of commercial off-the-shelf (COTS) consumer-grade wireless network router and small switch, fiber to Ethernet adapters, and a wireless range-extending antenna.

facility. The attacker connected the first adapter to an electrical socket under the raised floor in the server room, and connected an Ethernet cable to the adapter from the PPS network switch in the server rack. Placed smartly under the raised floor, the newly introduced hardware remained out of sight of any CAS personnel. Then the attacker deployed the second adapter and connected it neatly to an electrical socket on the outside of the CAS building. To avoid arousing suspicion, the adapter was hidden in a commonly used field distribution box. The attacker also placed a low-cost commercially available penetration testing tool known as a Pwn Plug inside the enclosure.[7]

The Pwn Plug features a myriad of communications interfaces, including Ethernet, Bluetooth, and cellular. For this scenario, the tool's Ethernet interface was interconnected with the PPS network via the second powerline network adapter, then connected to the Internet via its cellular interface, which connected to the nearest cellular communications tower. The Pwn Plug was configured – upon initialization – to establish an encrypted connection to an Internet-located server owned by the attacker. Once the hardware was in place and powered on, the attacker, working from a location more than 10 miles away (it could have been thousands of miles), remotely connected to his Internet-hosted server from his Windows laptop (also using an encrypted connection).

Because the Pwn Plug was successfully connected to the PPS network via the powerline network adapters, and because it had successfully connected to his Internet server, the attacker used the Pwn Plug to target vulnerable systems in the PPS network. Repeating the objectives of the first scenario, and using a nearly identical set of steps, the attacker successfully compromised a vulnerable Windows server on the PPS network. The results of the attack are the same: the attacker successfully enrolled an unauthorized badge/credential in the access control system, then later used that credential to gain undetected, unauthorized entry in to the testbed's secure area. However, it is noted that the attack leveraged a different vendor's AC&D access control software than that used in the first attack.

To complete the attack scenario, the attacker used a laptop computer, publicly available no-cost attack software (Metasploit), and less than $3000 of additional equipment.[8] It is important to note that no changes to any PPS system configuration were made to facilitate attack success. This attack required a relatively short period of physical, on-site access in which the attacker or accomplice implanted simple-to-install equipment, though some of it was configured prior to implantation.

It is important to note that the vulnerabilities exploited in both the first and second attack demonstrations were not related to any discovered vulnerability in either vendor's AC&D software. Instead, the attacker gained unauthorized and undetected access to the PPS

[7] The Pwn Plug is typically used by system defenders to support cybersecurity assessments. But, like so many cybersecurity tools, it can be used to conduct cyber exploitation just as easily.

[8] The additional equipment consisted of a commercial field distribution box, COTS consumer-grade Ethernet over Powerline adaptors, Pwn Plug, USB cellular modem, and dedicated IP space on an Internet-connected server.

network, and then attacked vulnerable servers that had not been sufficiently secured to protect against targeted exploitation by an adversary.

In summary, the R&D team used an adversarial attack methodology to successfully:

1. Use insider access to easily implant and hide low-cost, small-footprint cyber technology;
2. Interconnect attacker technology to the PPS network;
3. Remotely connect to the PPS network and discover vulnerable systems;
4. Exploit weaknesses in the backend servers to gain access to the access control software;
5. Enroll a blank physical access card in the access control database; and then
6. Walk into the protected area undetected, using the attacker's unauthorized credential.

3.3. Demonstration Three: AC&D Software Supply Chain Corruption

This attack combined a data breach, software exploitation, and social engineering. It ended with a physical breach of the "secured area" protected by the testbed PIDAS. The attack carefully modeled the real-world deployment of software updates by an actual vendor of AC&D software. Suppression of sensor alerts to operator workstations is the key feature.

The demonstrated attack accounted for the method used by the vendor to supply software updates to its customers. In this case, the vendor sends an email notifying their customers that an update or new version of the software is available, and directs them to an Internet-connected FTP[9] server to download the update. However, high-security PPS are not expected to have direct connections to the Internet. Therefore, the customer must download updates from a non-PPS system, save the update to removable media (e.g., thumb drive, or compact disc) transfer the update to the AC&D server, and install it during a planned maintenance cycle.

The following core steps outline the demonstrated attack:

1. Obtain a copy of the AC&D software used by the site;
2. Reverse engineer portions of the software to identify critical functions;
3. Modify the software with malicious changes;

[9] FTP: File Transfer Protocol. FTP is used to support the exchange of electronic files between remote computers over a network, including the Internet.

4. Clone the vendor's software update FTP site;
5. Upload the modified AC&D software to the attacker's FTP site;
6. Perform a spear-phishing email attack against site personnel;
7. Confirm the email attack was successful (i.e., the site downloaded the attacker's AC&D update);
8. <optional> Probe the target's PPS to determine if it has been degraded; then if it has,
9. Launch a physical attack crossing the PIDAS with confidence AC&D sensor events will not be transmitted to CAS operator workstations.

Each phase of the attack required different skills and knowledge. For example, modifying the software to the attacker's benefit while keeping it functional in all other aspects required a high level of ability (including knowledge and sophistication) within the software reverse engineering realm. The social engineering portion of the attack demanded perfect or near-perfect replication of vendor emails to its customers, and the ability to perfectly, or nearly perfectly clone the vendor's FTP server. This required a different set of knowledge and experience, but it is quite common to see these kinds of attacks in the wild, demonstrated by a range of adversaries. Some steps, particularly the spear-phishing campaign and physical attack were either not fully implemented or not implemented as doing so would have served no purpose to advance the R&D. Instead, the R&D team made appropriate assumptions to satisfactorily demonstrate the critical parts of the attack scenario and incorporate other elements of the R&D not discussed in this paper.

In summary, the R&D team demonstrated an adversarial attack methodology to successfully:

1. Gain access to the vendor's AC&D software;
2. Modify the behavior of the software;
3. Spoof the vendor's software update channel;
4. Induce the targeted site to download and install the modified software; then
5. Physically cross the PIDAS without alerting CAS personnel.

3.4. Analysis

Clem et al., concluded [2]:

Each of the three demonstrated attacks produced devastating impacts on the performance and reliability of the testbed protection system. The demonstrated attacks do not guarantee that operational systems deployed in the real world are susceptible to the same attacks. However, the weaknesses and vulnerabilities discovered and then exploited by the R&D team are typical of systems-of-systems environments including enterprise IT and industrial control systems that have been victimized by adversaries around the world. The first attack is viewed as one enabled by "low-hanging fruit" where inadequate and/or incomplete protections were *applied to critical PPS IT components. But the demonstrated attack should help put to rest the notion that isolation, as a technique to insulate PPS from cyber attackers, provides adequate defense. This attack demonstrated that isolation is, at best, a layer that prevents accidental unwanted connections and, at worst, a dangerous assumption of a system's security. The second attack is viewed as one that is more difficult to defend against because it leveraged opportunistic access by an insider to implant adversary communications and hacking equipment. The final attack consisted of relatively common hacker methods as well as more sophisticated methods to corrupt the site's AC&D software. Still, there are identifiable mitigations that can reduce the susceptibility of site owners, but it requires PPS vendor cooperation and effort. For example, the vendor's software could benefit from improved protections (details are intentionally withheld from this report to protect the vendor's intellectual property and reputation).*

Additionally, the last attack did not incorporate video surveillance systems or potential presence of guard patrols that a real-world attacker would be expected to encounter when crossing a secured boundary such as a PIDAS. Still, any adversary clever enough to be able to compromise the AC&D systems is expected to have the forethought to also target video surveillance systems, and conduct decoy operations directed at human security personnel.

High-security PPS are designed to, in order, detect, delay, and provide response functions that enable interruption and neutralization of human attackers. Each demonstrated attack targeted what might be fairly characterized as the crown jewel of the PPS: the AC&D used by security operators in the CAS. *Detection* was the key PPS performance requirement compromised in each attack. In the first two attack scenarios, detection of an unauthorized person entering the secured area could not occur because their entrance was made with a valid credential, properly enrolled in the access control system element of the PPS. In the third scenario, the sensors in the PIDAS worked as expected, but detection is presumed **not** to have occurred because no alert from the compromised AC&D system was transmitted to operator workstations. Thus, operators would **not** have enabled any active delay elements, and response personnel would **not** have been notified of an intruder.

R&D personnel did not directly test the PPS delay and response functions. However, because delay and response logically follow detection, and because modern PPS include information and communication technologies (ICT) that implement them, the R&D test outcomes on the detection function support inferences of adverse, consequential impacts from cyber exploitation on delay and response, minimally in the scenarios that were tested. In fact, given the wide-ranging configuration options that modern AC&D software provide to site operators, many more options existed for the R&D team's cyber adversary to conduct malicious operations given their achievement of full access/permissions for the AC&D software in the first two scenarios.

Overall, SNL subject matter experts judged the realism of the PPS testbed environment to be representative of modern PPS found throughout the world. Other SNL experts judged the cyber vulnerabilities discovered and exploited to be representative of those found in similar ICT environments. The method and techniques used by SNL researchers to discover and then exploit cyber weaknesses in the PPS testbed were reflective of processes used by both cybersecurity red teams and real-world cyber attackers.

4. Conclusions

The R&D discussed in this paper supports recommendations received by the U.S. Congress that a total systems approach be taken to characterize the interactions and dependencies between security countermeasures in high-security PPS [3]. Results support a recommendation that attack scenarios used to design and engineer PPS be adjusted to include adversary intent to disrupt and defeat the PPS itself through cyber means. Indeed, cyber vulnerabilities posited by SNL researchers to be discoverable, reachable, and exploitable enabled several successful attack vectors against the R&D testbed PPS.

A red team (adversary-based) methodology was used by SNL researchers to identify, analyze, and exploit testbed PPS systems to gain unauthorized, undetected access using legitimate system functionality. The R&D team further used common adversary techniques to obtain, modify, and reinsert critical software into the PPS environment; crucial functionality of the PPS software was defeated, resulting in the failure of sensor events to be communicated to operator workstations.

Certain limitations in the testbed environment, namely the lack of trained operators, security patrols, and video surveillance during testing of the third demonstration, leave open the opportunity to more rigorously test certain cyber-enabled physical attack scenarios in the future. However, these elements should not be presumed to represent adequate mitigations to defend against cyber exploitation of the AC&D and larger PPS. Facility vulnerability to cyber exploitation depends on many factors, including human performance, processes and procedures used to manage technology, and the security system's design and features.

REFERENCES

[1] CLEM, J., et al., "Investigation of Cyber-Enabled Physical Attack Scenarios", 2015 (Proc. International Conference on Computer Security in a Nuclear World: Expert Discussion and Exchange Vienna, 2015, IAEA, CN-228/4D3/060).

[2] CLEM, J., et al., Emulytics for Cyber-Enabled Physical Attack Scenarios, SAND2017-1603, Sandia National Laboratories, Albuquerque, NM (2017)

[3] National Research Council, Committee on Risk-Based Approaches for Securing the DOE Nuclear Weapons Complex, Understanding and Managing Risk Security Systems for the DOE Nuclear Weapons Complex–Abbreviated Version, The National Academies Press, Washington, DC (2011).

[4] KUSHNER, D., "The Real Story of Stuxnet", IEEE Spectrum (2013), https://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet