Cybersecurity in Prognostics and Health Management

Kai Goebel[1,2]

[1]*SRI International, Palo Alto, CA, 94304, USA*
[2]*Lulea Technical University, Lulea, Sweden*
kai.goebel@sri.com

ABSTRACT

PHM continues to show its value by improving operational efficiencies, increasing safety, reducing downtime, and decreasing cost of operations. PHM technologies are therefore not only being deployed as retrofit solutions but are being integrated into new systems as standard practice. Deployment covers areas such as medical equipment, nuclear power plants, aeronautics applications, oil and gas, mining, and many others. As the impact of PHM increases, it is imperative to also consider the potential vulnerabilities that are being exposed. Hackers have famously used Supervisory Control and Data Acquisition (SCADA) and Programmable Logic Controller (PLC) systems to sabotage industrial facilities. As such, it is important to understand the exposure to malfeasance to ensure that PHM does not end up being the enabling mechanism for unauthorized access to the system it is meant to keep in running order. It is also important to understand the measures that need to be taken to avoid or respond to an attack. These range from extensive penetration testing to conducting extensive counter-social engineering training, setting up a PHM-specific Computer Emergency Response Team (CERT) plan and team in place. This paper discusses various threats that are emerging and that may have to be considered when designing a PHM solution. Additionally, the National Institute of Standards and Technology (NIST) cybersecurity framework is discussed in the context of PHM. Finally, this paper looks at the diagnostic capabilities of PHM systems to detect cyber security attacks and to contain these threats.

1. INTRODUCTION

Cybersecurity and PHM overlap in a number of ways because PHM technology may increase a system's attack surface. There is an obvious need to protect the integrity of critical system information to avoid loss of proprietary information or to protect from unauthorized access (Kwon, Hodkiewicz, Fan, Shibutani and Pecht, 2016). Systems and accessed information, if tempered with, may lead to economic loss if they result in taking systems off-line, if set-points are changed, if the system fails to react to a safety margin threshold being breached, or if needed maintenance is not performed. If incorrect decisions are communicated back to the system - assuming that the system has the capability to act upon them – it is possible that the system could be driven into an undesirable state. A well-known example is the stuxnet virus that targeted SCADA/PLC systems (specifically Siemens control software Step 7) and caused centrifuges at the Natanz plant (see Figure 1) to self-destroy (Koch & Kuehn, 2017). In general, there are two major vulnerabilities of SCADA systems: unauthorized access to software (virus infections, intentionally induced changes, or other problems that can affect the control host machine); and vulnerability to packet access to network segments that host SCADA devices where little or no security of actual packet control protocol exist. Theoretically, anyone sending packets to a SCADA device could be in a position to control it.

Figure 1: Centrifuges at Natanz (Iran Times, 2016)

https://doi.org/10.36001/IJPHM.2024.v15i2.4063

The Repository of Industrial Security Incidents (risidata) is a database of incidents of a cyber security nature that have (or could have) affected process control, industrial automation or SCADA systems (Byres & Fabro, 2015). risidata recorded more than 240 incidents affecting control systems between 1997 and 2015. Many PLCs that are in use today were deployed several decades ago and typically don't have security features such as user authentication, antivirus software, or endpoint detection and response agent tools to catch malware or suspicious activity. “Once an attacker gets into the PLC, it's easy and he can stay there a long time” (Higgins, 2017)."

Figure 2: Power plant like the one targeted by “Triton” (Gibbs, 2017)

In 2017, hackers used malware dubbed “Triton” to take control of a safety workstation at an industrial power plant in the middle East. In particular, Triconex 3008 safety controllers made by Schneider Electric were impacted. They are part of the safety instrument system used to monitor industrial processes and automatically take actions or shut things down if unsafe conditions are detected. Triton infected the Triconex systems by modifying the code to disable certain safety mechanisms. It took advantage of security weaknesses in the safety systems’ programming environment and infrastructure. By attacking the Triconex safety controllers, the malware could disable automated emergency shutdown functions meant to prevent physical damage. Operators noticed the attack when some controllers inadvertently entered a failsafe mode and caused related processes to shut down (Gibbs, 2017). This attack breached the safety system (which is at the heart of some PHM activities) and as such indicates the potential for other parts of any industrial plant being compromised - while operators may not even initially detect the attack.

CyberX compiled data from more than 850 production industrial controls system networks across all industrial sectors. This data show that vulnerabilities in industrial control systems continue to be soft targets in key areas such as plain-text passwords (69%), direct connections to the internet (40%), weak anti-virus protections (57%), and wireless application protocols (WAP) (16%) (CyberX, 2019).

Figure 3: Entergy's Chiltonville Training Center is an identical twin mock-up of the control room at the Pilgrim nuclear power plant. (R. Lubbock, 2019)

Similarly, in 2015 hackers gained access to the SCADA systems at three regional electric power distribution companies in Ukraine resulting in loss of power for 80,000 customers (Zetter, 2016). The hackers were able to take the SCADA systems offline and overload communication processors, preventing operators from reestablishing connectivity for 6 hours. This prevented operators from regaining remote control capabilities, forcing them to go to substations manually to restore power. Existence of the manual backup systems was fortunate, because a fully automated system might have had a harder time recovering from such an attack. Figure 3 shows a mockup of a nuclear power plant’s control room, having supervisory controllers, individual power plant controllers, and balance-of-plant controllers

PHM systems may also have to think about protection against ransomware. In 2018, the Boeing plant in North Charleston, S.C., was hit by a ransomware attack with the WannaCry virus which resulted in equipment being locked down and demands for ransom payment in exchange for release of the computer systems. Boeing feared at the time that the virus might hit equipment used in functional airplane tests, which could lead to it spreading even to airplane software (Gates, 2018). The virus exploited Windows operating system vulnerabilities, which is relevant to us in the current context, because many ground-based elements of a PHM system run on the Windows operating system. Typically, airborne software is not as vulnerable because these are developed using much stricter safety assurance standards (DO-178C, e.g.).

For in-air inspection, operators will have to worry about unauthorized access of communication links. An example for this (albeit not in the PHM domain) was the jamming and

reprogramming of GPS signals of a bat-wing RQ-170 Sentinel drone over Afghanistan in 2012. Communications links between drone and operator were cut and the drone was then commandeered by the hackers to land at a location of their choice. This was accomplished by simply injecting noise on the communications channel which prompted the drone to switch into autopilot mode. Then the drone's GPS coordinates were reconfigured (Peterson & Faramarzi, 2011) and subsequent "spoofing" of fake GPS signals made the drone "land […], without having to crack the remote-control signals and communications" from the US control center (Peterson & Faramarzi, 2011). Beyond drone inspection applications, the ramifications extent to autonomous terrestrial vehicles and other mobility applications.

Figure 4: Downed US RQ-170 Sentinel drone (Peterson & Faramarzi, 2011)

In that context, it will not surprise that prognostic methods can be vulnerable to adversarial cyber attacks. Mode & Hoque (Mode & Hoque, 2020) found that a host of deep learning models had a number of cyber weaknesses that would lead, in case of an attack, to serious deficiencies in the remaining useful life estimation. While a disciplined process to improving the security posture will be examined in section 5 in this paper we will mentioned here the methods that Karandikar et al. (Karandikar, Knutsen, Wang & Løvoll, 2022) have explored that address some aspects of security of PHM systems. These include using a blockchain-enabled approach to federated learning aimed at promoting standardization for wider adoption. Such as federated learning approach would leverage a Docker-based infrastructure for data collection, storage, and analysis. This, combined with a methodology for ensuring tamper proofing of PHM data, can serve as a robust foundation for enhancing standardization, trust, and transparency in federated learning implementations of PHM algorithms.

An area of overlap between PHM and cybersecurity is to use PHM **for** Cybersecurity. The latter has been examined by Evans, Mishra, Yan, & Bouqata (2016) who describe how security related protections could be integrated fully with Monitoring and Diagnostics systems that assess the health of complex assets and systems and in particular combining system parameters already in use for PHM with security parameters to detect complex cyber threats. Indeed, the idea of using PHM principles for intrusion detection is not new. As summarized in Samrin and Vasumati (2017), the gamut of anomaly detection and classification tools used commonly in PHM (e.g., Naïve Bayes classifier, Neural Networks, Fuzzy clustering, k-means, k nearest neighbors, SVM, random forest, and decision trees, often in combination with some other technique) can be employed for network intrusion detection as well. Cassandro et al. (Cassandro, Wu, Wang & Li, 2024) take a resilience perspective where PHM serves as the backbone bridging security and reliability. The foundation of the approach is a resilience performance curve with which one can quantify resilience that considers both reliability and security. Murthy (2023) argues that for realizing the benefits of using PHM systems in detecting cybersecurity threats they need standardized guidelines for architects and designers. Emerging technologies like blockchains, digital twins, cognitive infrastructures/computing, quantum computing/quantum key distribution, and Secure DevOps offer promising advances for improving the resilience of PHM systems and services. Enterprise models for ICS need to account for these emerging technologies.

While work on the use of PHM techniques to detect cybersecurity threats is more than a decade old more work has been done in a related field of detecting counterfeit components, a perennial problem in the electronics industry. DARPA funded a program a couple of decades ago called IRIS (Integrity and Reliability of Integrated Circuits) to help detect and respond to such threats. The Center for Advanced Lifecycle Engineering (CALCE) at the University of Maryland has done extensive work in this area, including the use of PHM techniques to detect hardware intrusions and counterfeit components (Khemani, Azarian & Pecht, 2021). While the focus is often on software threats, authors discuss the concept of a hardware trojan (HT) and the need to implement PHM to ensure hardware security. Wang et al. (Wang, Tehranipoor & Plusquellic, 2008) provide a high-level survey of techniques to detect HTs. The University of Florida has an institute for national security that has a number of researchers working in this area (see e.g., the compilation volume by Shi, Q., et al. (2018)). Finally, Cranfield University's IVHM Centre has done research specifically on the use of PHM in detecting and mitigating such cybersecurity risks (e.g. Aslam et al. (Aslam, Jennions, Samie, Perinpanayagam & Fang, 2020)).

In summary, PHM community is gradually awakening to the critical importance of cybersecurity. As high-profile incidents come to light, the issue is steadily gaining prominence on the field's agenda, though it has yet to reach the forefront of collective concern. The following sections will examine in more detail which systems can be

compromised, how they can be compromised, and what on can do about the problem.

2. HOW SYSTEMS ARE COMPROMISED

Fundamentally, attacks seek to disrupt, deny, degrade, deceive, or destroy. Some of the common mechanisms to impair an industrial system include (Stouffer, Falco & Scarfone, 2011):

- Impeded or delayed transmission of data across Industrial Control Systems (ICS) networks, potentially causing disruptions in ICS operations.
- Unsanctioned alterations to instructions, commands, or alarm thresholds, with the potential to harm, disable, or deactivate equipment, create environmental consequences, or pose risks to human safety.
- Dissemination of inaccurate information to system operators, either to conceal unauthorized changes or to induce operators to take inappropriate actions, resulting in various adverse outcomes.
- Tampering with ICS software or configuration settings or infecting ICS software with malware, leading to a range of detrimental consequences.
- Interference with the proper functioning of safety systems, placing human lives in jeopardy. (NIST, 2018)

Some of these attack mechanisms are further explored in the next sections.

2.1. Direct Control

The aforementioned Stuxnet attack on Iran's uranium-enrichment facilities showed not only how SCADA devices might be hacked to cause real-world effects but how sensors can be fooled by faked logs so staff are left in the dark about the unfolding attack. "stuxnet manipulated the code blocks sent from the control computer, executed dangerous commands on the PLC and made the centrifuges spin at a higher frequency than originally programmed. The attacks on the PLC were only executed approximately every 27 days to make the attack stealthy and difficult to detect, which indeed is a central part of an Advanced Persistent Threat (APT). Stuxnet also took over the control computer and displayed false output on the STEP 7 software. This attack step was a core part of the attack and known as deception. In this case, the engineers located at the nuclear plant did not receive any indication of errors, assuming the centrifuges were spinning at the correct frequency. By receiving false output in STEP 7, the engineers would assume the meltdown was caused by human error, rather than malware, and acted accordingly" (Leyden, 2018).

Not every direct control attack is as sophisticated as stuxnet. It took only a disgruntled former employer to remotely access the actuators of a sewage system to discharge raw sewage in Melbourne (Gonda, 2014, Cherdantseva, Burnap, Blyth, Eden, Jones, Soulsby & Stoddart, 2016) (aka the "Maroochy Shire incident")

2.2. Spoofing

Altering (or "spoofing") sensor data in a specific way can result in changed response of the system as intended by an attacker. The above-mentioned redirection of a drone was accomplished in part through spoofing. Essentially, false data are injected which causes the mobility system to think it is at a location that it is not actually at (Tippenhauer, Poepper, Rasmussen & Capkun, 2011). Satellites have also been found to be vulnerable to cyberattacks (Lemos, 2019). In particular during war times, certain nation states have been tampering with the global navigation system GNSS, presumably to prevent drone attacks. While not a direct attack on a PHM service, a service that relies on GPS position information, such as an inspection drone, would be susceptible to the attack.

2.3. Denial of Service

Denial of Service (DoS) attacks are one of the easiest attacks on IoT systems. Any PHM system that communicates via wireless network might be susceptible to "jamming" the channel with an interrupting signal. Similarly, "flooding" (multiple packet transmission) or "collision" (timed flooding) are effective attacks.

Attacks are possible at the various layers of the network (physical, data link, transport, application). To jam a network at the physical layer, all a hacker needs to do is to broadcast a radio signal on the same frequency as the network, thereby overpowering the original signal. A jamming attack can be either intermittent or constant, but both have a detrimental impact upon the network. Attacks are also possible at the data link layer where collision attacks are more common. An adversary will intentionally violate the communications protocol which requires the retransmission of any packet affected; Further, the transport layer is susceptible to flooding DoS attacks where multiple connection requests are sent to a device, agent, or server. Because resources must be allocated to handle requests, overloading with malicious requests will quickly deplete resources. Finally, the application layer can be susceptible to a path-based DoS attack, whereby the attacker may insert spurious or replayed packets into the network. As the packet is forwarded to the destination, energy and bandwidth are consumed by forwarding nodes. This attack may starve the network of authentic data transmission, as resources along the path to a base station or application server are consumed (Tsiatsis, Karnouskos, Hoeller, Boyle & Mulligan, 2019).

In industrial control systems, equipment may have a low tolerance to bogus traffic, or it may be connected via low bandwidth links that are easily saturated. Also, while DoS

events can be caused by system overload, there are other ways that an attacker may deny access to service:

- Physical Equipment Destruction: This type of attack inflicts physical damage through digital means, as demonstrated by the Saudi Aramco incident where tens of thousands of user workstations were disabled. Again, the Stuxnet malware subtly altered an industrial control process with the goal to ultimately destroy centrifuges.
- Ability to Fix Denial: Attackers executing denial of service attacks intentionally disrupt the affected system's ability to resolve the issue. It may overwrite firmware, delete accounts, or block administrative access. For instance, remote process control equipment can be turned off and its firmware damaged, or network routes in infrastructure can be subtly updated to prevent administrator access.
- Public Identifier Theft: Attackers may hijack essential online accounts like domain names or Twitter accounts. In the case of cloud services, gaining access allows them to disable infrastructure while preventing the victim from reverting changes.
- RF Interference: Attackers may employ radio jamming to disrupt local WiFi or long-range wireless connections to remote sites, such as sensor installations. This type of attack involves some risk to the attacker because they need to be in physical proximity to the target location.

(National Cyber Security Centre, 2016)

As an example, the DoS attacks on the Ukraine power grid in 2015 and 2016 targeted the communication systems used by the ICS and SCADA systems. By overwhelming the communication networks, the attackers disrupted the ability of operators and the central control system to send and receive commands to and from remote substations and facilities. This made it difficult to control and monitor the power distribution network effectively. The DoS attacks had the effect of isolating remote substations from the central control system. Substation operators lost the ability to receive real-time data from the central control center, affecting their ability to respond to grid conditions and perform critical tasks. It should be noted that the DoS attacks were part of a larger coordinated cyberattack, which also included malware attacks against the ICS systems (Park & Walstrom, 2017).

2.4. Ransomware

Hackers may engage in attacks for personal gain through extortion/blackmail, typically by launching ransomware. This is a special type of malware that secretly installs on a computer and then either holds data hostage, or it is a sophisticated leakware that threatens to publish the data. It works by locking the system or encrypting critical files until a ransom is paid. This form of attack has not been reported yet through PHM related systems - but the potential for that may exist. IoT provides a large attack surface.

2.5. Wiping Data

Computers in Saudi Arabia's civil aviation agency and other Gulf State organizations have been wiped by the Shamoon malware after it resurfaced some four years after wiping thousands of Saudi Aramco workstations. Saudi Arabia's General Authority of Civil Aviation lost "critical data" in attacks that brought operations to a halt for several days (Pauli, 2016).

3. MOTIVATIONS

It may be helpful in preventing attacks by understanding what the motivation for hacking into PHM devices is. The reasons range from the thrill of making headlines by amateur hackers to sophisticated espionage schemes by nation actors. As such, the effort and level of sophistication associated with an attack varies widely.

3.1. Enrichment

Recent news on cybersecurity has been dominated by ransomware attacks where bad actors extort companies to pay a certain amount of money (often in cryptocurrency) to release a lock on the system that they disabled in the attack.

3.2. Sabotage

Sabotage is the act of deliberately destroying, damaging, or obstructing operations. It is done to gain a commercial, military, or political advantage. Depending on the hacker's intent, sabotage will target an organization to disrupt operations, create chaos or just be a nuisance. Where nation actors are involved, the sophistication and effort to cause harm can be considerable, as exemplified by the stuxnet attack (which was presumed to be caused by nation actors). Considerable effort is required by the hackers to infiltrate the system. In the case of stuxnet for example, the attacks used several Zero-day vulnerabilities to attack the system (Zetter, 2014). Zero-day vulnerabilities are those for which a defense does not yet exist. Typically, a hacker would use one of those vulnerabilities, if the objective were, say, "simple" theft. Here, though, the dedication shown is symptomatic for a nation-actor when resources play not a primary role. In addition to using the Zero-day vulnerabilities, the stuxnet attackers had very detailed knowledge of a particular PLC which again is testimony to the directed dedication only found when a sponsored attack is being carried out.

3.3. Industrial Espionage

In industrial espionage the target of a hack might be a trade secret, for example a proprietary product specification or formula, or information about business plans. It may also just be set points, best practices, efficiencies at which operations

are run, operational settings, and similar. In all cases, industrial spies are seeking data their own organization can exploit to its advantage. Those victimized by hacking incidents are not always forthcoming with that information (or details on which information has been compromised), because the harm to the victim's reputation if it's revealed that they were negligent on their security due diligence may outweigh the benefit of taking legal action against their attacker.

3.4. Headlines

Sometimes hacking groups are trying to make a statement in an attempt to make headlines. Hacktivists are hackers who breach systems to make political or ideological points or to access information that they can use for these purposes. And some hackers may feel that they need to prove something to their peers or friends, and hack something only for the challenge.

3.5. Boredom

And then there is a crop of hackers that seem to engage in their practices for no good apparent reason at all. They just do it because they can, or maybe for the thrill of it.

4. INDUSTRIAL CONTROL SYSTEMS

To understand how PHM interacts with Industrial Controls and what vulnerabilities arise, it is helpful to understand the various components of an Industrial Controls Systems.

Enterprise Network
Logistics and Business Planning
Site Operations
Supervisory Controls
Monitoring and Low-Level Controls
Sensors and Actuators

Table 1: Industrial Control System Reference Model (adapted from Didier et al., 2011)

Table 1 shows a reference model for an Industrial Controls Systems which is comprised of six levels. The lowest level includes actuators and sensors that interact in the physical world. The next layer performs monitoring functions using information from sensors and issuing control commands. The Supervisory Controls level handles operational functions and interacts with both the low-level controls as well as with site operations. This layer frequently communicates with the logistics layer which handles inventory management, communication between different division, and where databases of time-stamped events such as process output and alarms are kept. The highest layer is comprised of the centralized information technology services which includes business to customer interaction (Knowles, Prince, Hutchison, Disso & Jones, 2015).

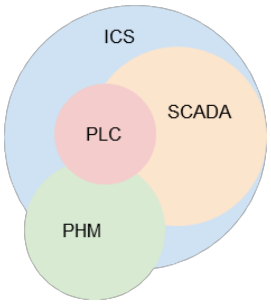

Figure 5: Notional overlap of ICS, PLC, SCADA and PHM

Figure 5 shows a notional overlap of ICS, PLC, SCADA and PHM. ICS is the overarching umbrella that covers also PLC and SCADA subsystems. PLCs and SCADA overlap as PLCs can be part of SCADA systems. PHM touches all systems but has also additional coverage outside of the ICS world as it deals with machinery and server systems that are not under the purview of ICS.

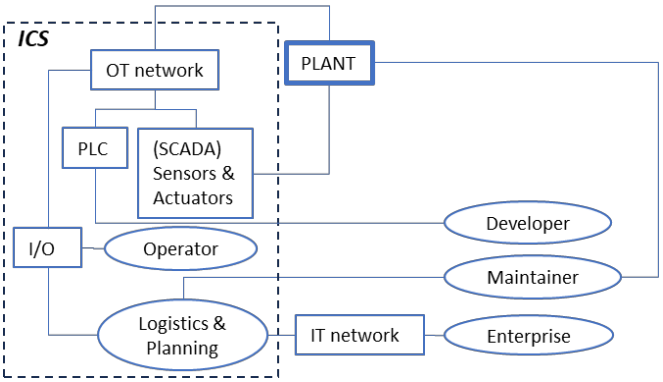

Figure 6: Plant reference model

As shown in Figure 6, a plant (as a generic term for a set of equipment that benefits from PHM) is controlled by some form of industrial control system (ICS). Operators, maintainers, and developers have access to the ICS. Communication between ICS and the plant is typically realized with a hardwired communications bus. Communication between operators and the ICS can be realized in several ways. All information is being transferred between different physical locations through some sort of databus. These can be hard-wired solutions such as Ethernet (for production system), WAN, RS-232, RS-485, or RS-422. They can also be tapping into wifi or other private radio link (bluetooth, infrared and laser communications that communicate sensor data) for a distributed operation. Where distance prevents these solutions in remote area telemetry can be realized using an appropriate network or some other form of radio signal.

Programming Logic Controllers (PLCs) were first developed in the 1960s to replace hardwired relay control logic on the factory floor in the automotive industry (Zetter, 2014). These PLCs allow discrete (bit-form) input and output in an easily extensible manner, and they permit its operation to be monitored. Early PLCs were programmed on custom terminals in "ladder logic", which strongly resembled a schematic diagram of relay logic and typically adhere to the IEC 61131/3 control systems programming standard. More modern PLCs are programmed using application software on personal computers. Prior to the discovery of the Stuxnet virus, security of PLCs received little attention. Modern PLCs generally contain a real-time operating system such as OS-9 or VxWorks, and exploits for these systems exist (much as they do for desktop computer operating systems).

Originally, a PLC was designed solely for autonomous operation within an industrial control system, with no intention of external connectivity or internet accessibility. PLCs relied on isolated, air-gapped networks and stringent physical access controls as a security strategy.

However, the evolution of ICS design has led to an increased exposure of PLCs to the internet (Schaefer, 2023). Air-gapped networks have proven to be an inadequate design choice and offer no compelling security argument in the context of modern ICS. As an illustrative example, researchers developed ladder logic code for the Siemens S7-1200 PLC, which generated frequency-modulated RF signals just below the AM radio band to encode stolen data. This purloined data could range from network topology information to sensitive blueprints, and it can be decrypted using a Software Defined Radio and a PC connected to the targeted site via an antenna. The researchers explained that an attacker could either fly a drone over the facility to capture the stolen intelligence or establish a nearby setup with a Software Defined Radio and a PC for data collection (Higgins, 2017)

Naturally, direct network infiltration might not always be essential. Reflecting on the Stuxnet incident, for instance, the malware was introduced through a USB device. Although its intended target was a specific site, it inadvertently propagated to over 115 countries, infecting critical infrastructure globally, despite the fact that most control systems were theoretically designed to be air-gapped (Igure, Laughter & Williams, 2006). It should be noted that ICS solutions that do not use PLCs are just as susceptible to attacks as PLCs as they have the same access points and vulnerabilities as PLCs do. Personnel that have access privileges to PLCs (or other controllers) include developers, operators, and possibly maintainers.

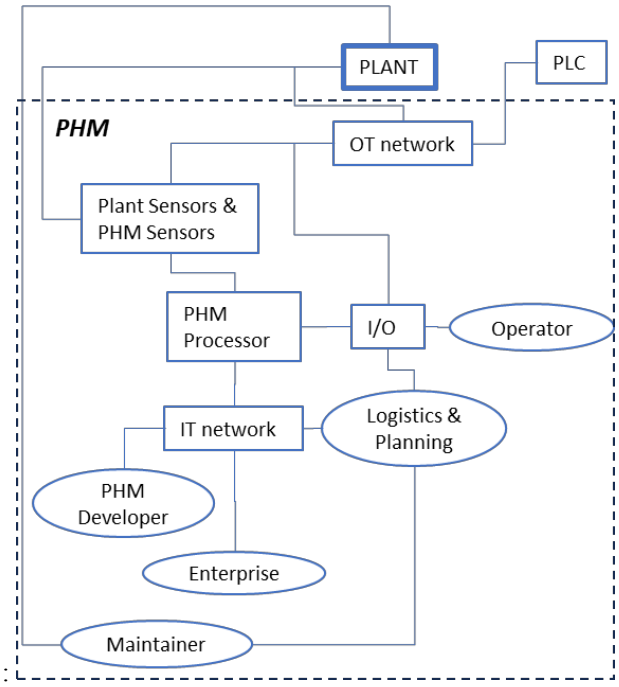

Figure 7: PHM System within a Plant Environment

Figure 7 shows the boundary of a PHM System within a Plant Environment. PHM systems are typically made up of sensors, processors, and I/O devices. While realizations of PHM systems vary based on policy and desired functionality, PHM systems often benefit from cloud-based access and therefore would have access to the IT network. That said, some enterprises insist on air-gapping the PHM implementations which requires on-premises solutions.

PHM also accesses both plant sensor information as well as additional information from additionally deployed sensors. PHM processors are typically edge devices as well as servers that centrally provide PHM analytics.

4.1. Sensors

Sensors play a pivotal role in OT systems as they inform about operational status and are used to control equipment. They are of course also used for PHM purposes. Therefore, if their integrity is impaired through outside influence, the integrity of the whole operation is at risk. Indeed, it is possible for sensors to be thrown off remotely even if there is no direct electrical connection. For example, strong electromagnetic interference may cause malfunction or erratic behavior. In 1999, the radar system of a Navy ship operating 2 miles off the coast of San Diego interfered with the wireless network of a local water company (San Diego County Water Authority) preventing workers from opening and closing valves so field engineers had to be dispatched to operate them manually to prevent a dam from overflowing (Byres & Fabro, 1999). The same incident resulted in

problems for the San Diego Gas and Electric (SDGE) Companies which were unable to remotely actuate critical value openings and closings as a result.

4.2. SCADA

The need to interact with sensors and control systems has long been recognized. In the mid-70s, the term SCADA (short for "Supervisory Control and Data Acquisition") emerged (Antón, Fraunholz, Lipps, Pohl, Zimmermann & Schotten, 2017) SCADA is an architectural framework for control systems that leverages computers, networked data communication, and graphical user interfaces to oversee and manage high-level processes. It also interfaces with the process plant or machinery using additional peripheral devices like programmable logic controllers (PLC) and discrete PID controllers.

SCADA systems are vulnerable through unauthorized access to software (virus infections, intentionally induced changes, or other problems that can affect the control host machine) and through packet access to network segments that host SCADA devices. Theoretically, anyone sending packets to a SCADA device could be in a position to control it. Often, supervisory control and data acquisition (SCADA) or industrial control systems (ICS) sit outside of traditional security walls (Storm, 2014). Indeed, Pauli reported (2014) that researchers found more than 60,000 exposed control systems in energy, chemical and transportation systems. And at the 2013 Chaos Communication Congress, a couple of security researchers asserted that they could demonstrate "how to get full control of industrial infrastructure" to the energy, oil and gas, chemical and transportation sectors (Storm, 2014).

The software's Project database had vulnerabilities related to the encryption and storage of passwords. These vulnerabilities enabled attackers to easily and dangerously obtain full access to PLCs (Pauli, 2014). They probed and found holes in "popular and high-end ICS and supervisory control and data acquisition (SCADA) systems used to control everything from home solar panel installations to critical national infrastructure."

SCADA Strangelove (a group of information security researchers founded that focus on security assessment of industrial control systems) has identified more than 150 zero-day vulnerabilities in SCADA, ICS and PLCs, with five percent of those being "dangerous remote code execution holes" (Storm, 2014). Amongst others, the project also released a password-cracking tool that targeted the vulnerability in Siemens PLC S-300 devices (Wuesst, 2014).

4.3. Communication Protocols

Industrial control systems (ICS) and PLCs can make use of multiple different communication protocols. Typically one distinguishes between the OT network and the IT network. OT stands for "Operational Technology". It is less well understood and will be briefly described here. Fundamentally, it refers to the hardware and software in industrial equipment that is needed to cause it to change operational settings. The OT network is a hardwired data exchange environment that uses industrial protocols such as Profibus and Modbus. It runs on standalone network and it was historically considered relatively safe not by conscious design choices but due to the "security by obscurity" concept, i.e., reliance on the idea that attackers could not think of possible vulnerabilities. Many ICS protocols were initially developed without inherent security features, leaving them susceptible to remote code execution, packet sniffing, and replay attacks due to the absence of authentication and encryption.

In the 1990s, congressional mandates necessitated companies to oversee and manage their industrial emissions (US Congress, 1990a, 1990b). This led to a transition to commercial operating systems. However, security has historically not been a priority in the design of industrial control equipment, and this trend persists even as IoT technology integrates into the realm of SCADA systems. Consequently, common problems like the use of default hard-coded credentials and the absence of encryption persist (Leyden, 2018). This results in inadvertently exposing these networks to viral threats.

Two of the more common communications protocols are briefly described below.

4.3.1. Profibus

Profibus is an international fieldbus communication standard. It is used to link several devices together and allows bi-directional communication. Profinet enables bi-directional communication. It uses the traditional Ethernet hardware, which makes it compatible with most equipment. Profinet is widely used in the automation industry, and its design is based on the Open Systems Interconnection (OSI) model. and is the preferred communication protocol for the Siemens Simatic PLCs.

4.3.2. Modbus

Modbus, established and published by Modicon (Schneider Electric) in 1979, stands as a serial communications protocol. Often referred to as master-slave communication, it allows one master to control up to 247 slave devices. Typically, the control computer takes on the role of the master, while automation devices or PLCs function as the slaves. Originally conceived as a communication protocol for PLCs, it later evolved into an international standard for linking various industrial devices. Modbus is known for its ease of deployment, cost-effectiveness, and its suitability for SCADA systems. The Modbus protocol exists in three variations: American Standard Code for Information

Interchange (ASCII), remote terminal unit (RTU), and transmission control protocol/Internet protocol (TCP/IP).

Fundamentally, both Profibus and Modbus lack basic security controls like authentication and encryption. This makes them vulnerable to man-in-the-middle, eavesdropping, and spoofing attacks. While Modbus has traditionally operated on isolated serial connections, deployment on TCP/IP networks and the internet has increased its exposure to cyberattacks. Lack of visibility into Modbus networks hinders detection of attacks. Security monitoring of Modbus traffic would help identify unusual traffic patterns.

5. REMEDIATION

An obvious question is what can be done about these threats. Regrettably, there is no quick fix that solves all these issues. Many solutions are known but are often treated as unwelcome barriers to deployment and easy operations. It is argued that a thoughtful, methodical approach to Cybersecurity in PHM will yield the best results. Ideally there is an internal Computer Emergency Response Team (CERT) team that is cognizant of PHM aspects of cybersecurity and that can help with assessment and responses. Whether or not such a team exists, cybersecurity policy is a key element together with following a disciplined approach as described in the NIST Cybersecurity Framework. Additional means such as Canaries are described as well.

5.1. Cybersecurity Policy

Cybersecurity policy is an area of increasing importance for PHM (Goebel, Smith & Bajwa, 2019). Cybersecurity policy is meant to provide guidance about the protection mechanism of an organization's crucial physical and information assets. At the minimum, it will specify intentions and conditions that aid to protect assets along with instructions to carry out these intentions. The cybersecurity policy is the mechanism that directs users to build, install, and maintain systems to assure the confidentiality, integrity, and availability of both the PHM system as well as the system that it connects to.

5.2. Assessing Vulnerabilities

Sadly, it turns out that there is no free lunch in PHM cybersecurity and securing an ICS environment, including PLCs, presents a significant challenge, primarily because these systems lack inherent cyber-resilience. Consequently, it is imperative to integrate cybersecurity resilience measures into the PHM environment. To that end, it is necessary to understand what the additional vulnerabilities are that result from a PHM system. Figure 8 shows the possible vulnerabilities that should be considered during a PHM systems integration.

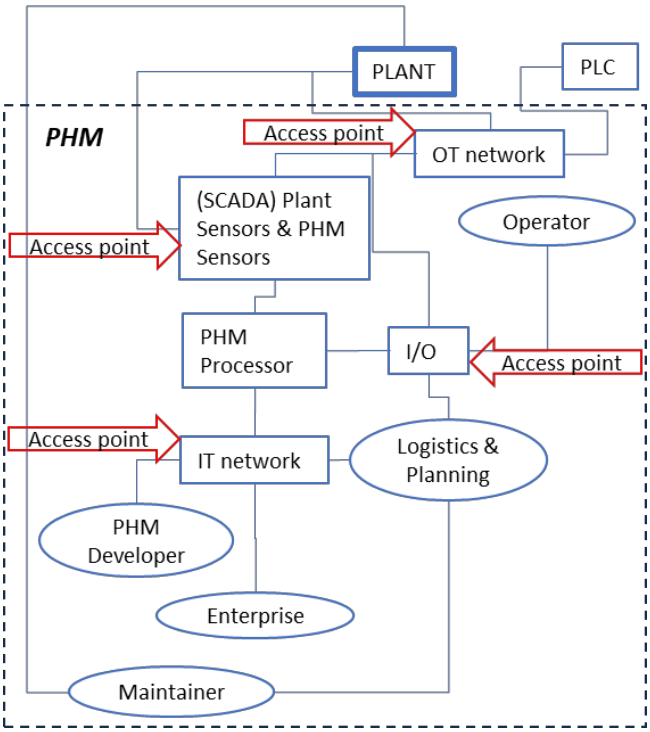

Figure 8: Potential cybersecurity vulnerabilities in a PHM system

These vulnerabilities result from access points that may be introduced when implementing the PHM solutions. They can include the IT, I/O, the OT network, and the sensors. An integral part of this process is the identification and comprehensive understanding of potential risks. Equally important is the continuous evaluation of the likelihood of impending attacks (Houmb & Martin, 2018). An illustrative threat analysis that includes the attack vectors mentioned above is presented in section 8.

Remedial measures encompass at the minimum strategies like implementing perimeter defense mechanisms, such as firewalls, to mitigate the risk of unwanted network traffic and to avoid that PHM systems increase the vulnerability of a plant to cyberattacks. More comprehensive measures may have to include network monitoring, ideally employing non-intrusive, ICS-specific, anomaly-based network monitoring, designed not to add extra strain on ICS networks. Additionally, endpoint protection and monitoring are crucial to reduce the vulnerability of PLCs to remote attacks and to promptly detect any signs of intrusion. This necessitates the use of ICS-specific endpoint protection and monitoring tools. It becomes apparent that a methodical approach towards assessing the cybersecurity threat is needed to avoid ad-hoc solutions that do not provide the needed protection. A solution is the NIST cybersecurity framework that is introduced in the next section.

5.3. NIST Cybersecurity Framework

The NIST Cybersecurity Framework (NIST, 2023) was developed in part in response to the Cybersecurity Enhancement Act of 2014 (15 U.S.C. § 272(e)(1)(A)(i), 2014). The purpose of the Cybersecurity Framework is to manage cybersecurity risks not just for PHM solutions but for any conceivable application. It is the de-facto standard that defines best practices for a large number of enterprises. PHM operations may benefit from this framework if it is adopted for its purposes. The framework postulates that an understanding of the organization's business drivers and security considerations specific to its use of PHM technology need to be available. There are five concurrent and continuous functions that define the framework. They are: Identify, Protect, Detect, Respond, and Recover. Figure 9 shows the functions and several elements within that function.

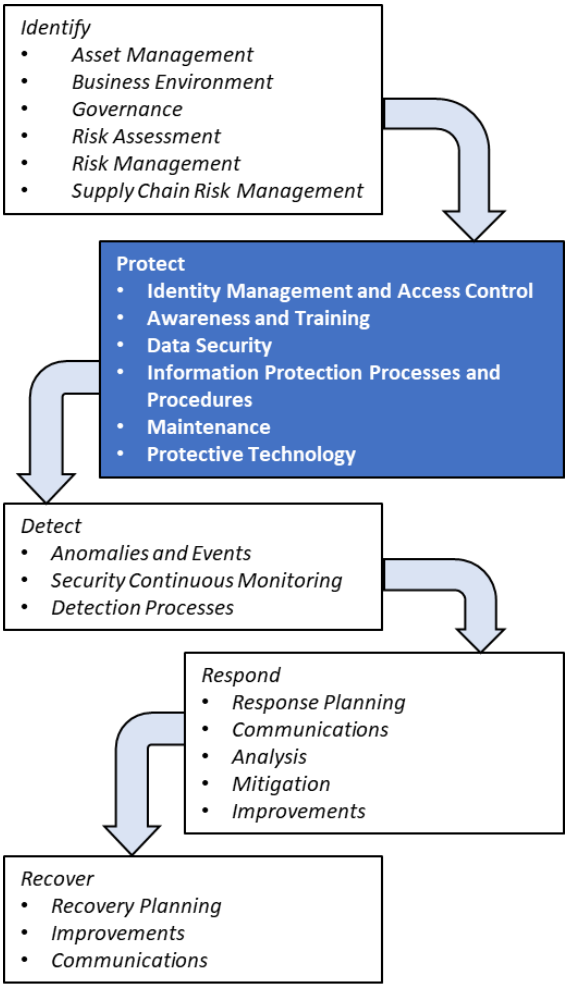

Figure 9: NIST Cybersecurity Functions (NIST 2023)

The following sections will go through each function and will spend particular attention to the "Protect" function because it impacts the design and operation of PHM technology the most.

5.3.1. Identify

The **Identify** function suggests that an understanding of the business context and the resources that support critical functions will help in recognizing the related cybersecurity risks. The overall purpose is to develop an organizational appreciation that helps to manage cybersecurity risk to systems, people, assets, data, and capabilities. It is meant to enable the organization to focus and prioritize its cybersecurity efforts. Examples of categories that need attention include: Asset Management; Business Environment; Governance; Risk Assessment; and Risk Management Strategy. These are further explained in the following sections

In **asset management** for PHM (Prognostics and Health Management) objectives, all essential hardware and software assets are identified, including devices, software, data, and personnel. This involves cataloging physical devices and systems, documenting software applications, mapping data and communication flows, and cataloging external information systems. Prioritization of resources is done based on their importance and potential impact on the system. Additionally, cybersecurity roles and responsibilities are established to ensure accountability and access control for personnel and stakeholders.

Securing logical access to the ICS network involves implementing a DMZ network architecture with firewalls to block direct communication between corporate and ICS networks. Separate authentication methods and credentials are essential. Additionally, a multi-layered network topology ensures the highest security for critical communications.

Securing physical access to the ICS network and devices is vital to prevent potential disruptions. Various physical access controls like locks, card readers, and guards can be used to accomplish that. Protecting individual ICS components involves swift security patch deployment, disabling unused ports and services, minimizing user privileges to necessity, monitoring audit trails, and implementing security controls, such as antivirus software and file integrity checks, to prevent, deter, detect, and mitigate malware when feasible.

Next, the organization's mission, objectives, stakeholders, and governance must be grasped and prioritized, informing cybersecurity roles, responsibilities, and risk management as part of understanding the **business environment.** Similarly, organizational policies, procedures, and processes for regulatory, legal, risk, environmental, and operational requirements must be understood, shaping cybersecurity risk management.

As the last part in this phase, **risk assessment** provides insight into cybersecurity risks for organizational operations, evaluating the impact on mission, functions, image, and

reputation. It establishes the organization's priorities, constraints, risk tolerances, and assumptions for supporting operational risk decisions within a Risk Management Strategy. This assessment should also encompass supply chain risks. For an example on threat modeling, please refer to Section 8.

5.3.2. Protect

The Protect function is at the heart of cybersecurity safeguarding activities. Here, the development and implementation of appropriate measures is prescribed. The Protect Function supports the ability to limit or contain the impact of a potential cybersecurity event. Examples of categories within this function include: Identity Management and Access Control; Awareness and Training; Data Security; Information Protection Processes and Procedures; Maintenance; and Protective Technology.

As part of **identity management, authentication and access control**, access to physical and logical assets and associated facilities is restricted to authorized users and processes, managed in accordance with assessed risks. However, securing assets in applications with PHM systems can be challenging if the organization deploying the PHM solution lacks control. Nevertheless, it is crucial to limit access to authorized users. The level of vulnerability varies depending on the application, with mass-produced assets like autonomous vehicles being more susceptible to tampering. Best practices include embedding sensors to reduce tampering, using integrated sensor validation, and protecting sensor-to-device communication through physical and non-physical means, such as encryption. The IoT field offers valuable examples and lessons for approaching this issue.

Personnel is provided with cybersecurity **awareness education and training** to ensure that they will perform their cybersecurity related duties and responsibilities consistent with related policies, procedures, and agreements. Cybersecurity-related information needs to be included into PHM product documentation that is released to the customer. This includes installation of upgrades, change of passwords, and the like.

5.3.3. Detect

Critically, if a cybersecurity event happens, it is important to detect it in a timely manner to minimize the impact. To that end, it is necessary to develop and implement appropriate activities to identify the occurrence of a cybersecurity event. The Detect Function enables timely discovery of cybersecurity events. Examples of outcome Categories within this function include: Anomalies and Events; Security Continuous Monitoring; and Detection Processes.

Detection assumes that the information system and assets are monitored to identify cybersecurity **anomalies and events** and verify the effectiveness of protective measures. Moreover, the detection processes and procedures are maintained and tested to ensure continued awareness of novel anomalous events.

5.3.4. Respond

Incidents are inevitable and an incident response plan is essential. A major characteristic of a good security program is how quickly a system can be recovered after an incident has occurred. (NIST, 2018). Once an attack has been detected, it is necessary to develop and implement appropriate activities to take action regarding a detected cybersecurity incident. The Respond Function supports the ability to contain the impact of a potential cybersecurity incident. Examples of outcome Categories within this Function include: Response Planning; Communications; Analysis; Mitigation; and Improvements.

Response processes and procedures are meant to respond to the cybersecurity event in an effective manner. There are a number of options ranging from shutting the system down to shoring up IT defenses. Response activities should be coordinated with internal and external stakeholders (e.g. external support from law enforcement agencies). As a last step, an analysis should ensure effective response and support recovery activities.

A device planted on a factory floor can identify and list networks, and trigger controllers to stop processes or production lines.

5.3.5. Mitigate

Similarly important is the prevention of expansion of an event, mitigate its effects, and resolve the incident. Here it is important to ensure that the system can actually recover. What has been learned from the utility outage incident in Ukraine in 2017 is that the ability to restore operations must not be hamstrung by the persistence of the attack. For example, if a software attack was the cause of the event, then an automated reboot of the system may not, in fact, resolve the issue. What helped to restore operations in Ukraine was the fact that the system could be switched back on by hand. While this may seem arcance, the importance of being able to isolate certain system functions from the attack is key.

Maintaining functionality during adverse conditions involves designing the ICS so that each critical component has a redundant counterpart. Additionally, if a component fails, it should fail in a manner that does not generate unnecessary traffic on the ICS or other networks, or does not cause another problem elsewhere, such as a cascading event.

5.3.6. Recover

The NIST framework also lays out steps to recover from an event. Specifically, one should think about developing and implementing appropriate activities to maintain plans for resilience and to restore any capabilities or services that were

impaired due to a cybersecurity incident. The Recover Function supports timely recovery to normal operations to reduce the impact from a cybersecurity incident. Example categories within this function include: Recovery Planning; Improvements; and Communications.

Recovery planning involves that processes and procedures are executed and maintained to ensure restoration of systems or assets affected by cybersecurity incidents. Organizational response activities are **improved** by incorporating lessons learned from current and previous detection/response activities. Restoration activities are coordinated with internal and external parties through appropriate **communications**, for example through coordinating centers, Internet Service Providers, owners of attacking systems, victims, other CSIRTs, and vendors.

5.4. Canaries

A "canary" is a concept that is more specifically associated with early warning systems (Wang & Pecht, 2011). It is a bait or decoy file, account, or system that is placed intentionally within a network or system. If an attacker accesses or modifies the canary, an alert is triggered. This serves as an early warning system, indicating that an intrusion or unauthorized access is likely underway. Canaries are often used to identify and respond to threats before the latter can cause significant damage.

Canaries are part of a proactive defense strategy, often used to detect and deter potential attackers by creating uncertainty and making it more challenging for them to proceed undetected. These measures are designed to improve the security and resilience of PHM systems against cyber threats. These could be realized for example by creating simulated control systems, data servers, or communication channels that appear to be part of the legitimate infrastructure. It could also be realized by implementing so-called honeypots or decoy systems that simulate real assets in the target network. These systems should look like genuine control systems, or communication devices. Additionally, one might consider setting up robust network monitoring tools and intrusion detection systems to track network traffic and activity across the system, then configuring alert mechanisms to notify security personnel or administrators when suspicious activity is detected within the deceptive components. This can include monitoring for unauthorized access attempts, unusual data transfers, or unexpected system interactions. Ensuring that all activities and access attempts within the deceptive components are logged securely with subsequent analysis can provide valuable information in case of a security breach.

Implementing canaries in a system can enhance cybersecurity defenses by acting as an early warning system, providing valuable insights into potential threats, and helping to respond to security incidents. However, there is a trade-off to balance security measures with the need to maintain the safety and integrity as well as operational cost of the systems.

6. OTHER CYBERSECURITY STANDARDS

Besides the NIST Cybersecurity framework, SAE, IEEE, and ISO have developed several standards for cybersecurity. ISO/IEC 27001, part of the ISO/IEC 27000 family of standards, is an information security management system standard (jointly by International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC)) to formally specify a management system that is intended to bring information security under explicit management control. It is one of the more popular information security standards and accredited certification to the Standard is recognized worldwide. Certification demonstrates the commitment to data security and provides a valuable credential when tendering for new business. Similar to other standards, the list of possible controls is rather long (ISO 27001 lists 114 of them in Annex A). However, to achieve certification, not all of those controls need to be implemented. Based on a risk assessment, the necessary controls are identified in conjunction with justification to why other controls are excluded from the ISMS.

IEEE works on standards such as C37.240-2014 - IEEE Standard Cybersecurity Requirements for Substation Automation, Protection, and Control Systems. To that end, risks expected to be present at a substation are being addressed in such a manner that access and operation to legitimate activities is not impeded, particularly during times of emergency or restoration activity.

SAE J3061 Cybersecurity Guidebook for Cyber-Physical Vehicle Systems defines a lifecycle process framework. It provides information on common existing tools and methods used when designing, verifying and validating cyber-physical vehicle systems.

SAE Information Report J2931/7 establishes the security requirements for digital communication between Plug-In Electric Vehicles (PEV), the Electric Vehicle Supply Equipment (EVSE) and the utility, ESI, Advanced Metering Infrastructure (AMI) and/or Home Area Network (HAN). JA7496 addresses Cyber-Physical Systems Security Engineering Plan (CPSSEP). This standard, which deals with aerospace and automotive systems, was developed by SAE's G-32 (Cyber Physical Systems Security) committee, which was established in 2019 to develop standards for the entire mobility sector. Other technical committees within SAE have also been working on standards related to cyber security. For example, the HM-1 technical committee (on IVHM) is currently working on AIR7381 "Integrated Vehicle Health Management (IVHM) and Cybersecurity," aimed at the aerospace industry.

Other organizations such as North American Electric Reliability Corporation (NERC), ARINC, Information Assurance standard of Small and Medium-sized Enterprises (IASME), European Telecommunications Standards Institute (ETSI), and others all have established committees to meet

the growing demand guidance to protect the various assets under their charge.

7. LLMs

This discussion would be incomplete without acknowledging the potential role Large Language Models (LLMs) play in cybersecurity. One of the most significant impacts they can have on the security landscape is the ability to scale, expedite, and amplify existing threats, not only for phishing attacks, but also in generating code that can interface with Internet of Things (IoT) systems, conceivably posing a significant risk to cybersecurity.

7.1. Potential threats originating from LLMs

Specific concerns regarding LLMs relate to PHM in the following ways:

7.1.1. Code Generation and Vulnerabilities

LLMs can be used to generate code for PHM applications, including software for monitoring equipment health or controlling industrial systems. While this can speed up the development process, it also opens the door to vulnerabilities. As shown in several studies, LLMs may generate code that contains security flaws, such as insecure communication protocols, inadequate authentication mechanisms, or improper handling of sensitive data. For PHM, which often deals with critical infrastructure, any vulnerability in code generated by LLMs can lead to system compromise, data breaches, or even catastrophic equipment failures.

For instance, if an LLM generates faulty code to manage communication between PHM sensors and central servers, attackers could exploit this weakness to spoof sensor data, leading to incorrect diagnostics or even sabotaging operations by injecting malicious commands.

7.1.2. Automation of Cyberattacks

One of the most concerning aspects of LLMs is their ability to automate cyberattacks. While not explicitly specific to PHM, malicious actors could use LLMs to rapidly generate phishing emails, malware, or ransomware specifically tailored to actually target PHM systems. For example, an LLM could be trained to impersonate trusted vendors or internal personnel, tricking operators into installing malware on PHM systems or providing unauthorized access to critical components. This is a particular concern because PHM systems have accountability beyond their own domain if they are integrated with Industrial Control Systems (ICS).

7.1.3. Insufficient Domain Expertise

While LLMs excel in generating human-like responses and processing general data, they may lack the specific domain expertise required for critical PHM applications. For example, PHM systems must ensure accurate diagnostics and real-time response to equipment health issues. However, if LLMs are used to make real-time decisions or generate critical system alerts, their lack of deep domain-specific understanding might lead to incorrect interpretations of sensor data or poor decision-making in high-stakes environments. Ordinarily, the lack of domain knowledge can be overcome by a process called Retrieval Augmented Generation (RAG), where an outside authoritative knowledge base is referenced to help provide the missing domain information. However, this is a custom process that needs to be carefully curated to achieve the desired results. Where this process is omitted or where it falls short, vulnerabilities may arise where an LLM fails to recognize a developing cyberattack, leaving the PHM system exposed to threats.

7.1.4. Scalability of Threats

LLMs can scale up cyberattacks, because they can automate probing for vulnerabilities as well as creating malware and phishing campaigns much faster than human attackers. For example, an LLM-generated attack could target multiple edge devices in a PHM system, disabling sensors or corrupting the data sent to the central monitoring station.

7.2. LLMs used to defend Threats

On the flip side, these models can also be leveraged for the detection and mitigation of phishing attempts. Cybersecurity leaders recognize multiple ways in which generative AI can support organizations in bolstering their defenses, such as reviewing code for efficiency, identifying potential security vulnerabilities, and exploring novel tactics that malicious actors may employ. It can also automate repetitive tasks like report writing. Nonetheless, it is crucial to proceed with caution. Of note is a Stanford study (Clark, 2023) which revealed that AI assistants produced code with more vulnerabilities than code generated by an experienced software engineer. In an experiment (Alney, 2023) that involved instructing ChatGPT to identify flaws in security code of varying quality, researchers identified several limitations. Despite the responses sounding authoritative and relevant, they offered limited value in terms of security review. Furthermore, they consumed the time of human reviewers as they had to ascertain that each point raised was incorrect, despite the well-structured and persuasive language used in the AI-generated output. As individuals who attempt to work with LLMs will discover, these models are all too willing to generate code that can interface with Programmable Logic Controllers (PLCs) or Supervisory Control and Data Acquisition (SCADA) systems, a capability that is not widespread and poses additional security concerns.

In conclusion, while LLMs hold significant potential for enhancing PHM systems through automation and advanced analytics, their integration must be carefully managed to mitigate cybersecurity risks. Addressing these concerns

through proactive measures can help ensure the security and reliability of PHM in critical industries.

8. THREAT MODELING

In threat modeling vulnerabilities are identified, enumerated, and prioritized from a hypothetical attacker's point of view. The purpose of threat modeling is to provide defenders with a systematic analysis of the most likely attack vectors, and the target that an attacker seeks.

One method used is a threat tree which graphically represent how a potential threat to a cyberphysical system can be exploited. STRIDE is a model of threats used in conjunction with a model of the target system that includes a full breakdown of processes, data stores, data flows and trust boundaries. It assesses the functions spoofing, tampering with data, repudiation, information disclosure, denial of service, and elevation of privilege (Donovan, 2021). Typical Threat and Risk Analysis (NIST, 2012, VDI/VDE, 2011) involves these steps:

- Determine Scope (identify what systems and applications need to be protected, the sensitivity of what is being protected
- Collect Information and Data. These include policies and procedures (both those that are already in place and those that are missing) as well as pertinent system data such as physical location of the system, operating system type, services running, service pack levels, network applications running, network surveying, port scanning, wireless leakage, intrusion detection testing, firewall testing, access control permissions.
- Identify potential vulnerabilities. This is typically done using a vulnerability tool. It also includes penetration testing.
- Analyze threats uncovered and assign a rating.
- Perform threat analysis. Threats are described as anything that would contribute to the tampering, destruction or interruption of any service or item of value. The analysis will look at all conceivable risk elements, both from humans (Hackers, maintainers, technicians, backup operators, non-technical staff, inadequately trained IT staff, etc.). A complete threat analysis would also consider non-cybersecurity threats such as lightning strikes, contamination, EMI events, fire, etc. One would then go through the findings to determine anything that may contribute to tampering, destruction or interruption of any service or item of value.
- Develop a strategy to remove these threats with measures that include installation of new software, implementing additional access controls, shielding, etc.

8.1. Example

Below, an example of a preliminary threat assessment for a PHM system that uses edge devices, additional sensors, and a web-based server is examined. The specific use case considered is for a chemical plant where the PHM system monitors the health and performance of various critical components, such as mixers and compressors. The system employs edge devices installed on several components to collect sensor data and perform initial data processing. These are sensors in addition to the sensors that are installed for control purposes and have been added to enhance data collection capabilities. Data from edge devices are transmitted to a centralized web-based server for further analysis, reporting, and remote monitoring.

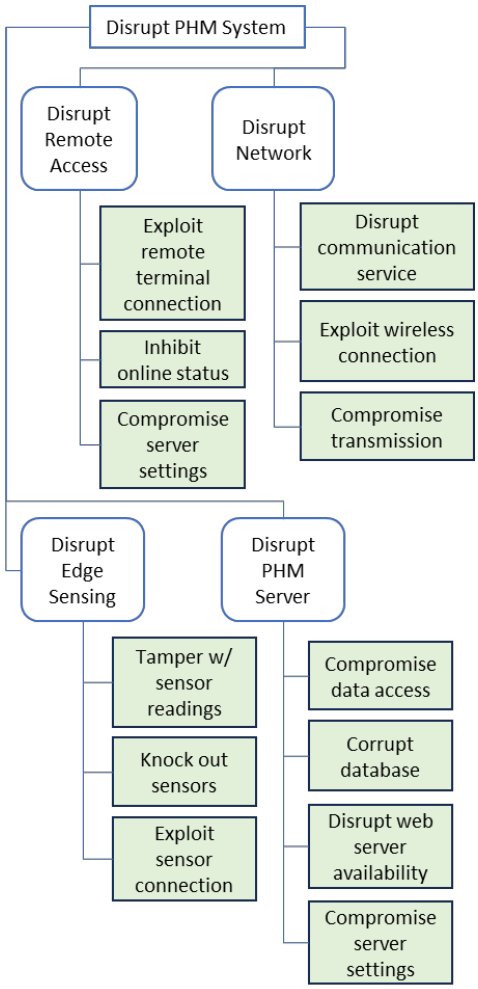

Figure 10: Attack Tree for PHM system

Recall the plant reference model as encapsulated in Figure 6. A simple version of an attack tree is shown in Figure 10. Special attention should be directed towards identifying and addressing the specific access points that serve as potential entry points for various threats. Each threat needs to be addressed by at least one defense countermeasure. While the

comprehensive coverage of all threats and countermeasures exceeds the scope of this paper, a few of them are listed here for illustrative purposes.

Threat: Exploit Remote Terminal Connection

- Vulnerabilities: Weak or default passwords, unpatched software, insufficient access controls.
- Risks: Unauthorized access can lead to data breaches, system disruption, or sabotage.
- Countermeasures: Implement strong authentication, access controls, and regular security updates.

Threat: Disrupt Edge Sensing

- Vulnerabilities: Lack of physical security measures on the equipment, inadequate tamper detection.
- Risks: Physical access can lead to the compromise of edge devices and data.
- Countermeasures: Implement physical security measures and tamper-evident seals on edge devices.
- Vulnerabilities: Inadequate data integrity checks, lack of encryption, insecure data transmission.
- Risks: Tampered data can lead to inaccurate health assessments and, in the aviation context, potential safety risks.
- Countermeasures: Implement data encryption, digital signatures, and secure data transmission protocols.

Threat: Compromise PHM Server

- Vulnerabilities: Insufficient data protection, misconfigured data access permissions.
- Risks: Data leakage can lead to the exposure of sensitive information, impacting safety and competitiveness.
- Countermeasures: Encrypt sensitive data, control access permissions, and monitor data access.

Threat: Disrupt Network

- Vulnerabilities: Limited server capacity, lack of DoS mitigation measures.
- Risks: DoS attacks can disrupt system availability and impact real-time monitoring and analysis.
- Countermeasures: Implement DoS protection, load balancing, and system monitoring.

This threat assessment provides only a partial overview of the potential risks and vulnerabilities associated with the PHM system that uses edge devices, additional sensors, and a web-based server. It is essential to continually assess and update the threat landscape and associated countermeasures to ensure the ongoing security of the system, especially in critical industries like aviation. A more complete vulnerability assessment of a SCADA system is shown in (Ten, Liu, & Govindarasu, 2007).

9. CONCLUSIONS

This paper reviewed the key cybersecurity challenges and solutions in the context of Prognostics and Health Management (PHM). PHM, while improving operational efficiency, introduces new vulnerabilities due to increased system connectivity and the integration of additional sensors. These vulnerabilities expose critical systems to potential threats, such as unauthorized access, data breaches, and sabotage, particularly in industries like aviation and energy.

To summarize, these challenges can be broadly categorized into the following areas:

1. Increased Attack Surface
2. Legacy Systems and Compatibility
3. Data Integrity and Availability
4. Network Vulnerabilities
5. Complexity of Securing Distributed Systems
6. Resource Constraints on Edge Devices
7. Human Factor and Insider Threats
8. Balancing Safety and Security
9. Lack of Standardization
10. Emerging Cyber Threats

The integration of cybersecurity measures into PHM systems is essential to mitigate these risks. The NIST Cybersecurity Framework provides a robust foundation, offering guidelines for identifying, protecting, detecting, responding to, and recovering from cybersecurity threats. In particular, its focus on continuous monitoring and endpoint protection proves invaluable for defending industrial control systems. One promising area involves using PHM principles themselves to enhance cybersecurity, such as integrating anomaly detection tools with PHM systems to detect cyberattacks early. Future work should explore how PHM diagnostic capabilities can be expanded to detect and mitigate cyber threats, thereby transforming PHM from a vulnerability to a security asset.

In conclusion, the intersection of PHM and cybersecurity presents both risks and opportunities. As PHM technologies become more widespread, proactive cybersecurity measures will be crucial to safeguarding these systems and ensuring their reliability in critical applications.

ACKNOWLEDGMENTS

The author wishes to thank Ravi Rajamani for his thorough review and helpful suggestions.

REFERENCES

15 U.S.C. § 272(e)(1)(A)(i). (2014). The Cybersecurity Enhancement Act of 2014. (S.1353), US Congress

Alney, C., (2023). Security Code Review with ChatGPT. nccgroup, https://research.nccgroup.com/2023/02/09/

security-code-review-with-chatgpt/, last accessed 10/14/23

Antón, S., Fraunholz, D., Lipps, C., Pohl, F., Zimmermann, & M., Schotten, H. (2017). Two decades of SCADA exploitation: A brief history. Proceedings IEEE Conference on Application, Information and Network Security (AINS), pp. 97-104.

Aslam, S., Jennions, I.K., Samie, M., Perinpanayagam S., & Fang, Y., "Ingress of Threshold Voltage-Triggered Hardware Trojan in the Modern FPGA Fabric–Detection Methodology and Mitigation," in *IEEE Access*, vol. 8, pp. 31371-31397, 2020, doi: 10.1109/ACCESS. 2020. 2973260.

Byres, E., & Fabro, M. (2015). The Repository of Industrial Security Incidents. risidata. https://www.risidata.com/, last accessed 10/14/23.

Brenner, B., (2011). SCADA Hacking madness. cso online, https://www.csoonline.com/article/2134949/scada-hacking-madness.html, last accessed 3/19/19

Cassandro, R., Wu, G., Wang, H., Li, Z.S. (2024). Prognostics and Health Management for Cyber-Physical System Resilience: A Security and Reliability Perspective. In: Karanki, D.R. (eds) Frontiers of Performability Engineering. Risk, Reliability and Safety Engineering. Springer, Singapore

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K., (2016). A review of cyber security risk assessment methods for SCADA systems. Computers & Security, Volume 56, pp. 1-27.

Clark, A. (2023), New AI wave will find uses and abuses in cybersecurity. Axios, https://www.axios.com/2023/02/17/cybersecurity-ai-tech-chatgpt-bing, last accessed 10/14/23.

CyberX, (2019) Global ICS & IIOT Risk Report.

Dieier, P., Macias, F., Harstad, J., Antholine, R., Johnston, S. Piyevsky, S. Schillace, M., Wilcox, G., Zaniewski, D., & Zuponcic, S. (2011). Converged Plantwide Ethernet (CPwE) Design and Implementation Guide. ENET-TD001E-EN-P, Cisco Systems, San Jose, CA and Rockwell Automation, Milwaukee, WI.

Donovan, F. (2021). What is STRIDE and how does it anticipate cyberattacks? Security Intelligence, https://securityintelligence.com/articles/what-is-stride-threat-modeling-anticipate-cyberattacks/, last accessed 1/28/2024.

Evans, S., Mishra, P., Yan, W., & Bouqata, B. (2016). Security Prognostics: Cyber meets PHM. Proceedings of IEEE PHM Conference.

Finkle, J., (2018). Schneider Electric says bug in its technology exploited in hack. Reuters, https://www.reuters.com/article/us-schneider-cyber-attack/schneider-electric-says-bug-in-its-technology-exploited-in-hack-idUSKBN1F7228 last accessed 3/19/19.

Gates, D., (2018). Boeing hit by WannaCry virus, but says attack caused little damage. Seattle Times, https://www.seattletimes.com/business/boeing-aerospace/boeing-hit-by-wannacry-virus-fears-it-could-cripple-some-jet-production/, last accessed 4/10/2018.

Gibbs, S. (2017). Triton: hackers take out safety systems in 'watershed' attack on energy plant. The Guardian, https://www.theguardian.com/technology/2017/dec/15/triton-hackers-malware-attack-safety-systems-energy-plant, last accessed 4/10/2018.

Goebel, K., Smith, B., & Bajwa, A. (2019). Ethics in Prognostics Health Management. International Journal of PHM, 012.

Gonda, O. (2014). Understanding the threat to SCADA networks. Network Security, Volume 2014, Issue 9, pp. 17-18.

Higgins, K. (2017). Stealthy New PLC Hack Jumps the Air Gap. Dark Reading, https://www.darkreading.com/threat-intelligence/stealthy-new-plc-hack-jumps-the-air-gap-/d/d-id/1330381, last accessed 1/28/24.

Houmb, S., & Martin, E., (2018). More exploits: the great PLC hack. control design, Oct 25, https:// www.controldesign.com/articles/2018/more-exploits-the-great-plc-hack/, last accessed, 3/26/19.

Igure, V., Laughter, S., & Williams, R., (2006). Security issues in SCADA networks. Computers & Security, 25, 498-506.

Iran Times (2018). Image source http://iran-times.com/ap-says-iran-will-be-able-to-enrich-more-uranium-sooner/, last accessed 4/7/2018.

Karandika, N, Knutsen, K., Wang, S., Løvoll, G. (2022). Federated Learning on Trusted Data for Distributed PHM Data Analysis, Proceedings of the PHM Conference

Khemani, V., Azarian, M. H., & Pecht, M. G. (2021). Prognostics and Secure Health Management of Electronic Systems in a Zero-Trust Environment. Annual Conference of the PHM Society, 13(1). https://doi.org/10.36001/phmconf.2021.v13i1.3006

Knowles, W., Prince, D., Hutchison, D., Disso, J., & Jones, K. (2015). A Survey of Cyber Security Management in Industrial Control Systems. International Journal on Critical Infrastructure Protection, pp. 52-80.

Koch, R., & Kuehn, T. (2017). Defending the Grid: Backfitting Non-Expandable Control Systems. Proceedings 9th International Conference on Cyber Conflict

Kwon, D., Hodkiewicz, M., Fan, J., Shibutani, T., & Pecht, M. (2016). IoT-Based Prognostics and Systems Health Management for Industrial Applications. Special Section on Trends and Advances for Ambient Intelligence with Internet of Things (IoT) Systems, IEEE Access.

Lemos, R. (2019). Cybersecurity Experts Worry About Satellite & Space Systems. InformationWeek, darkReading, https://www.darkreading.com/attacks-breaches/cybersecurity-experts-worry-about-satellite-and-space-systems/d/d-id/1335131, last accessed 7/3/2019

Leyden, J. (2018). Pwned with '4 lines of code': Researchers warn SCADA systems are still hopelessly insecure. The Register, https://www.theregister.co.uk/2018/06/18/physically_hacking_scada_infosec/ last accessed 10/17/2023

Lubbock, R. (2019) Entergy's Chiltonville Training Center is an identical twin mock-up of the control room at the Pilgrim nuclear power plant. WBUR, https://www.wbur.org/news/2019/05/31/plymouth-reactor-training-center., last accessed 12/22/2023

Mehta, A., & Gruss, M., (2019). Pentagon hopes to have new cybersecurity standards for contractors in 2020. Fifth Domain, https://www.fifthdomain.com/dod/2019/03/26/pentagon-hopes-to-have-new-cybersecurity-standards-for-contractors-in-2020/ last accessed 3/28/19.

Mode, G, Hoque , K. (2020). Crafting Adversarial Examples for Deep Learning Based Prognostics, 19th IEEE International Conference on Machine Learning and Applications.

Murthy, R. (2023) Exploring the Use of PHM for Software System Security and Resilience, IEEE International Conference on Prognostics and Health Management

National Cyber Security Centre. Denial of Service Guidance. https://www.ncsc.gov.uk/collection/denial-service-dos-guidance-collection, last accessed 10/16/2023

NIST, (2012), SP 800-30 Guide for Conducting Risk Assessments, NIST report, Revision 1.

NIST, (2018). Framework for Improving Critical Infrastructure Cybersecurity. NIST public report, version 1.1, April 16 https://doi.org/10.6028/NIST.CSWP.04162018 8

NIST, (2023). CSWP 29 (Initial Public Draft), The NIST Cybersecurity Framework 2.0

Park, D., & Walstrom, M. (2017). Cyberattack on Critical Infrastructure: Russia and the Ukrainian Power Grid Attacks. Jackson School of International Studies, U. Washington, https://jsis.washington.edu/news/cyberattack-critical-infrastructure-russia-ukrainian-power-grid-attacks/, last accessed 10/13/2023

Pauli, D., (2014). Hackers gain 'full control' of critical SCADA systems. itn news, https://www.itnews.com.au/news/hackers-gain-full-control-of-critical-scada-systems-369200 last accessed 3/19/19

Pauli D., (2016). Shamoon malware returns to again wipe Saudi-owned computers. The Register, https://www.theregister.co.uk/2016/12/02/accused_iranian_disk_wiper_returns_to_destroy_saudi_orgs_agencies/, last accessed 3/19/19

Peterson, S., & Faramarzi, P. (2011). Iran hijacked US drone, says Iranian engineer. Christian Science Monitor, https://www.csmonitor.com/World/Middle-East/2011/1215/Exclusive-Iran-hijacked-US-drone-says-Iranian-engineer , last accessed 4/10/2018

Saleh, S., Prateek, M., & Poor, V., (2018). 27th USENIX Security Symposium. August 15–17, 2018, Baltimore, MD.

Samrin, R., & Vasumati, D. (2017). Review on Anomaly based Network Intrusion Detection System. Proceedings of 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT).

SCADA Strangelove http://www.scada.sl/

Schaefer, W. (2023) The Rising Importance of PLC Cybersecurity: An Essential Look into Industrial Vulnerability. engineering.com, https://www.engineering.com/the-rising-importance-of-plc-cybersecurity-an-essential-look-into-industrial-vulnerability/, last accessed 9/16/24

Shi, Q., Forte, D., Tehranipoor, M.M. (2018). Deterrent Approaches Against Hardware Trojan Insertion. In: Bhunia, S., Tehranipoor, M. (eds) The Hardware Trojan War. Springer, Cham. https://doi.org/10.1007/978-3-319-68511-3_13.

Sikder, A., Petracca, G., Aksu, H., Jaeger, T., & Uluagac, S. (2018). A Survey on Sensor-based Threats to Internet-of-Things (IoT) Devices and Applications. arXiv:1802.02041v1 [cs.CR] 6 Feb 2018.

Son, Y., Shin, H., Kim, D., Park, Y. Noh, J. Choi, K., Choi, J., & Kim, Y., (2015). Rocking Drones with Intentional Sound Noise on Gyroscopic Sensors. 24th USENIX Security Symposium, pp. 881-996.

Storm, D. (2014). Hackers exploit SCADA holes to take full control of critical infrastructure. Computerworld, Jan.15, 2014, https://www.computerworld.com/article/2475789/hackers-exploit-scada-holes-to-take-full-control-of-critical-infrastructure.html last accessed 10/16/2023

Stouffer, K., Falco, J., Scarfone, K., (2011). Guide to industrial control systems (ICS) security. NIST special publication, 800, 16-16.

Ten, C., Liu, C., & Govindarasu, M., (2007). Vulnerability Assessment of Cybersecurity for SCADA Systems Using Attack Trees. Proceedings of Power Engineering Society General Meeting.

Tippenhauer, N., Poepper, C., Rasmussen, K., & Capkun, S. (2011). On the Requirements for Successful GPS Spoofing Attacks. Proceedings of 18th ACM Computer and Communication Security, pp. 75-86.

Tsiatsis, V., Karnouskos, S., Hoeller, J., Boyle, D., & Mulligan, C. (2019). Internet of Things. Academic Press.

US Congress. (1990a). H.R.3030, Clean Air Act.

US Congress. (1990b). H.R.5931, Pollution Prevention Act.

VDI (2011). IT Security for Industrial Automation. VDI/VDE Guideline 2183.

Wang, W., Pecht, M. (2011) Economic Analysis of Canary-Based Prognostics and Health Management, IEEE Transactions on Industrial Electronics, Vol. 58, Issue 7.

Wang, X., Tehranipoor M., & Plusquellic, J., "Detecting malicious inclusions in secure hardware: Challenges and

solutions," *2008 IEEE International Workshop on Hardware-Oriented Security and Trust*, Anaheim, CA, USA, 2008, pp. 15-19, doi: 10.1109/HST.2008.4559039.

Wuesst, C. (2014). Targeted Attacks Against the Energy Sector. Symantec Report.

Zetter, K. (2014). Countdown to Zero Day. Crown Publishers, New York.

Zetter, K. (2016). Everything we know about Ukraine's Power Plant Hack. Wired, https://www.wired.com/2016/01/everything-we-know-about-ukraines-power-s

BIOGRAPHY

Kai Goebel is Director of the Intelligent Systems Lab at SRI/PARC. Prior to working at SRI/PARC he was at NASA Ames Research Center where he directed the Prognostics Center of Excellence. His research interest is in the areas of machine learning, real time monitoring for safety, diagnostics, and prognostics. He has fielded numerous applications for manufacturing systems, aircraft engines, unmanned aerial systems, space systems, transportation systems, energy applications, and medical systems. He holds 21 patents and has published more than 400 papers. He received the degree of Diplom-Ingenieur from Technische Universitaet Muenchen in 1990 and the Ph.D. from the University of California at Berkeley in 1996. Dr. Goebel worked between 1997 and 2006 at General Electric's Corporate Research Center in upstate New York where he was also an adjunct professor at Rensselaer Polytechnic Institute. Dr. Goebel is now an adjunct professor at Lulea Technical University. He is a co-founder of the Prognostics and Health Management Society and was associate editor of the International Journal of PHM. While pursuing the Ph.D. at UC Berkeley, he was member of the student Pugwash group at Cal where he helped to craft the Engineering Ethics pledge that was listed on commencement brochures.