Characterizing Cyberattacks against Operational Technology Infrastructures through the Lens of Attack Flows

Sherman Kettner, Caleb Chang, Ekzhin Ear, and Shouhuai Xu

Laboratory for Cybersecurity Dynamics

Department of Computer Science

University of Colorado Colorado Springs

Abstract

Operational Technology (OT) infrastructures play a critical role in modern society and economy. However, their increasing connectivity with public networks such as the Internet has made them vulnerable to cyberattacks, much like traditional Information Technology (IT) systems. In particular, cyberattacks against OT infrastructures remain relatively underexplored and little understood. In this paper, we aim to deepen our understanding of cyberattacks against OT infrastructures. For this purpose, we propose a methodology, including novel cybersecurity metrics to analyze the attack flows of these attacks in an end-to-end fashion, which allows us to draw useful insights. We demonstrate the utility of the methodology by applying it to characterize four real-world cyberattacks against OT infrastructures. This allows us to draw a number of insights, such as: proactively disrupting attackers' reconnaissance; segmenting within OT infrastructures, and from IT infrastructures via zero trust; disrupting attacks before they achieve their intended effects.

Keywords: Cyberattacks, OT infrastructures, attack flows, cybersecurity metrics,

Characterizing Cyberattacks against Operational Technology Infrastructures through the Lens of Attack Flows

Operational Technology (OT) plays a critical role in modern society. Historically, OT infrastructures and systems were protected from cyberattacks by isolating them from the Internet or public networks. However, with the rise of the Internet of Things (IoT) and the relentless pursuit of optimization, OT systems have become increasingly interconnected with conventional Information Technology (IT) networks and the Internet. This growing connectivity has exposed OT infrastructures to a surge in cyberattacks, targeting an area that has traditionally been poorly defended. This paper aims to deepen our understanding of cyberattacks against OT infrastructures to draw insights into hardening these networks against cyberattacks. To achieve this, we propose a methodology to guide our case studies, while noting that the methodology is adaptable for analyzing cyberattacks on OT infrastructures, including those involving proprietary networks with restricted data-sharing policies.

Our Contributions

This paper makes three contributions. First, we propose a methodology for characterizing cyberattacks against OT networks. Second, we conduct a case study to show the utility of the methodology, by applying it to characterize four real-world cyberattacks against OT infrastructures, including the 2009 attack on Natanz Nuclear Facility (Stuxnet), the 2015 attack on Ukraine power grid, the 2017 attack on the Petro Rabigh oil refinery, and the attack on the Maroochy Shire Sewage system. Third, we draw actionable insights, such as: cyberattacks against OT infrastructures can be thwarted by disrupting attacker reconnaissance, which appears to have been little understood or investigated.

Terminology and Methodology

Terminology

Attack Tactics, Techniques, and Procedures (TTP). We adopt the concepts of cyberattack *tactics*, *techniques*, and *procedures* (TTPs) as defined by MITRE in its Adversarial Tactics, Techniques & Common Knowledge (ATT&CK) framework to describe attacks against IT infrastructures [1]. A cyberattack *tactic* is defined as an objective or sub-objective of an attacker. A cyberattack *procedure* is a specific implementation or instantiation of an attack technique. In total, the ATT&CK framework specifies 14 tactics (i.e., "*reconnaissance, resource development, initial access, execution, persistence, privilege escalation, defense evasion, credential access, discovery, lateral movement, collection, command and control, exfiltration, impact*") and 236 techniques [2].

Given that we focus on cyberattacks against OT infrastructures, we also adopt MITRE's ICS TTPs, which specifies 12 tactics (i.e., "*initial access,*

execution, persistence, privilege escalation, evasion, discovery, lateral movement, collection, command and control, inhibit response function, impair process control, impact") and 94 techniques [3].

Cyberattack Flows. This concept is introduced in [4] to describe cyberattacks in an end-to-end fashion. Although it is introduced to describe cyberattacks in the context of IT infrastructures, it can be equally applied to describe cyberattacks in OT infrastructures. We propose improving the concept of *cyberattack flow* with the concept of *attack flow phases*, which provides a higher level of abstraction in describing attacks, meaning that one phase may contain multiple attack steps in a cyberattack flow [4]. Inspired by [5], we define the following three phases:

1. *Setting Conditions:* In this phase, an attacker prepares for the attack by gathering information and developing their capabilities.
2. *Gaining Access:* In this phase, an attacker gains access to the target infrastructure either from their own computer or pivoting from some victim computer that has been compromised by the attacker (i.e., stepping stone).
3. *Deploying Effects:* In this phase, an attacker achieves their objectives—such as data exfiltration, system disruption, or physical impacts—on the target infrastructure.

Methodology

As highlighted in Figure 1, our methodology consists of four steps: (i) defining metrics; (ii) identifying, collecting, and preprocessing cyber threat intelligence; (iii) reconstructing cyberattack flows; (iv) analyzing the reconstructed cyberattack flows to extract insights. Each step is detailed below.

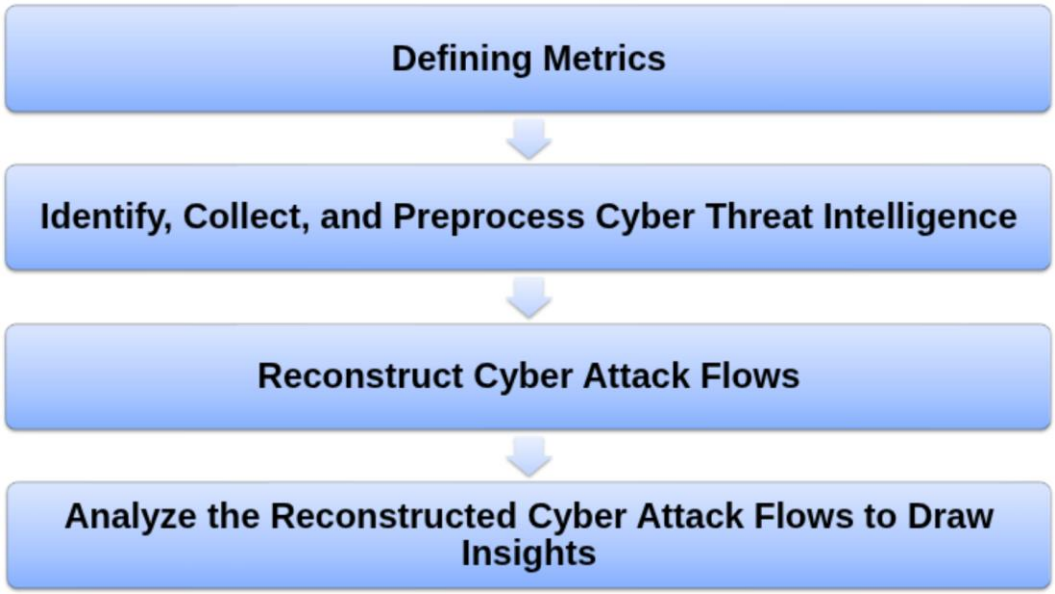


Figure 1: The methodology employed in this study

Step 1: Defining Metrics

To adequately characterize cyberattacks against OT infrastructures, we need to define quantitative cybersecurity metrics. These metrics form the basis for assessing the severity, techniques, and impacts of the attacks. Specifically, we define the following metrics.

Definition 1 (attack complexity). *We define this metric to measure the technical sophistication of a cyberattack against OT infrastructures. More specifically, we measure it via the number of unique* attack techniques *that are used by an attack.*

Definition 2 (attack persistence time). *We define this metric to measure the time an attacker spent in an OT infrastructure without being detected—namely the length of the interval between when the attacker gained initial access to the victim OT infrastructure and when the attack is detected.*

Definition 3 (attack impact). *We define this metric to measure the consequence of an attack, namely the degree of disruption to OT infrastructures, via two submetrics: (i)* recovery time*, namely the time it takes to recover an affected OT infrastructure back to normal service; (ii)* service degradation*, namely the maximum degree to which the OT service is degraded by an attack at any point or interval in time (depending on how the service is measured in the absence of attacks) during the attack, with 100% indicating complete degradation.*

Step 2: Identifying, Collecting, and Preprocessing Cyber Threat Intelligence

Real-world cyberattacks against OT infrastructures are often reported in scattered (possibly non-technical) cyber threat intelligence sources of various forms, such as news reports and (in rare cases) research papers. In this step, we identify and collect any applicable cyber threat intelligence to reconstruct a comprehensive, end-to-end view of each attack. During preprocessing of the collected raw cyber threat intelligence, the key issue is to ensure relevance.

Step 3: Reconstructing Cyberattack Flows

In order to reconstruct cyberattack flows of the attacks described in the preprocessed cyber threat intelligence, we propose extracting the attack techniques from them and then using them to reconstruct cyberattack flows. We observe that attack technique offers the appropriate level of abstraction for this purpose because it disregards the procedure-level implementation details that are not essential for reconstructing cyberattack flows and facilitates the optional reconstruction of tactic-level flows cyberattack flows when desired. For this purpose, we propose to leverage MITRE's attack techniques as reviewed above.

Step 4: Analyzing the Cyberattack Flows to Draw Insights

Having reconstructed the attack flows of a real-world cyberattack against an OT infrastructure, we can leverage the metrics to analyze the attack flows. The resulting characteristics can then inform predictions of future cyberattacks against OT infrastructures, especially designing countermeasures to proactively mitigate them.

Case Studies

To show the usefulness of our methodology, we conduct a case study analyzing four real-world cyberattacks on OT infrastructures to extract actionable insights.

The 2009 Attack on Natanz Nuclear Facility (Stuxnet)

In 2009, the Natanz Nuclear Facility in Iran was the target of a cyberattack, known as Stuxnet, named after the malware it employed. The malware was first reported in 2010 and is the first known malware that targets SCADA systems [6].

Identifying, Collecting, and Preprocessing Cyber Threat Intelligence

To analyze the Stuxnet attack, we used "Stuxnet" as the keyword to conduct a Google search. From the returned results and their cited references, we manually synthesized cyber threat intelligence from a curated set of sources across multiple categories. Technical sources, such as [6], [7], [8], [9], were used to help characterize the attack, including the propagation techniques it used and the exploitation of OT systems. Other reports, such as [10], [11], confirmed that the Stuxnet malware targeted the WinCC database server and were leveraged to validate other details about the malware's evolution. In addition, scholarly articles such as [13], [14], and [15] were referenced to analyze broader implications of the malware and its place within cyberwarfare discourse. Journalistic and intelligence reporting such as [16] and analytical reports like [17] were also incorporated to corroborate operational impact claims and attribute the attack to likely state-sponsored actors.

Reconstructing the Stuxnet Attack Flow

At a high level, Stuxnet exploited four zero-day vulnerabilities [7] to spread the malware, achieve privilege escalation, and bypass defenses autonomously [6], [7], [9] against a wide range of IT systems, including Windows 2000, Windows Server 2000, Windows XP, Windows Server 2003, Windows Vista, Windows 7, Windows Server 2008, and SIMATIC WinCC versions 6.2 and 7.0 [7], [10], [11]. Figure 2 illustrates the Stuxnet attack flow, including three phases as described in our methodology, namely *setting conditions*, *gaining access*, and *deploying effects*, which are elaborated below.

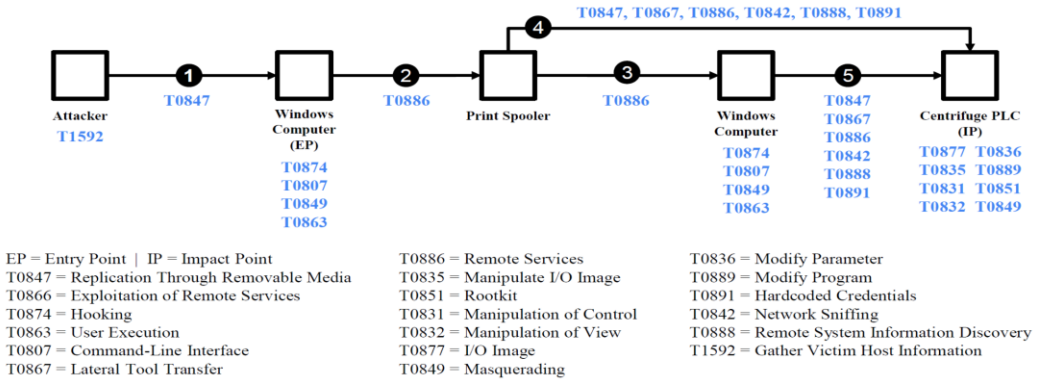


Figure 2: The Stuxnet attack flow

Setting Conditions. The attacker's objective was to compromise the Programmable Logic Controllers (PLCs) that run the Siemens WinCC/PCS7 SCADA software to control the centrifuges at the Natanz nuclear facility. To enable this, the attacker conducted an in-depth analysis of the target environment, acquiring knowledge of both the hardware and firmware used in the facility. This reconnaissance phase included using Gather Victim Host Information (T1592) to enumerate hosts, configurations, and environmental parameters. In the absence of Internet connectivity to the target system, the attacker designed Stuxnet as a highly autonomous, modular, and intelligent malware, capable of self-propagation It was equipped with advanced capabilities, including goal-oriented execution, stealth mechanisms, and self-updating functionality. To make the malware appear legitimate, the attacker equipped the malware with two digital signatures using two stolen private signing keys corresponding to two legitimate digital certificates that were acquired from trusted vendors; there is evidence suggesting that these private signing keys were stolen from legitimate users [7], [10], [12].

Gaining Access. Corresponding to the arc annotated with ① in Figure 2, the widely accepted theory is that Stuxnet was introduced into the Natanz environment via a compromised USB flash drive, manifesting a classic case of the attack technique Replication Through Removable Media (T0847). The initial infection vector likely involved either a double agent or an unsuspecting employee inserting an infected USB device. Once the malware entered a Windows-based workstation (i.e., the *entry point*), it executed via User Execution (T0863) and Command-Line Interface (T0807), while employing Hooking (T0874) to monitor and manipulate API calls. After establishing a foothold, Stuxnet leveraged multiple propagation techniques:

1. peer-to-peer propagation and update mechanisms via technique Lateral Tool Transfer (T0867);
2. exploiting a hardcoded password to access Siemens WinCC database servers via technique Hardcoded Credentials (T0891);
3. leveraging open network shares via technique Remote Services (T0886), corresponding to the arc annotated ②;

4. exploiting the MS10-061 Print Spooler vulnerability via technique Exploitation of Remote Services (T0866), corresponding to the arc annotated ③;
5. exploiting the MS08-067 Windows Server Service vulnerability via technique Exploitation of Remote Services (T0866) ([7]).

These mechanisms enabled technique Lateral Tool Transfer (T0867) and access to Remote Services (T0886) across the infrastructure. To evade detection, the malware used a Rootkit (T0851) technique and gathered additional system details using Remote System Information Discovery (T0888) and Network Sniffing (T0842). Each successive hop mirrored the initial access phase, allowing Stuxnet to move laterally across IT systems to other IT systems and finally the air-gapped OT assets, represented by arcs annotated ④ and ⑤.

Deploying Effects. Once it reached the PLCs governing centrifuge operation, Stuxnet executed a payload designed to manipulate both the system logic and the operator perception. It employed Manipulate I/O Image (T0835) and I/O Image (T0877) to intercept sensor values and feed misleading operational data to human operators. At the same time, it issued destabilizing commands via Manipulation of Control (T0831), causing the centrifuges to spin at destabilizing frequencies and physically degrade the equipment, while simultaneously reporting normal operational values to operators, via exploiting Manipulation of View (T0832) [7], [13]. Parameters were further modified using Modify Parameter (T0836), to fine-tune spin cycles, while the use of Modify Program (T0889) ensured persistent reprogramming of PLC logic.

The malware also maintained stealth through Masquerading (T0849) and Rootkit (T0851) techniques, which prevented alerts to system administrators. Although Stuxnet infected an estimated 100,000 hosts globally, with 60% of infections occurring in Iran, its effects were highly targeted. The malware was engineered to activate only under highly specific configuration conditions that match those of the Natanz facility [7], [9], [13]. Consequently, the impact on non-targeted systems was negligible [14], [15]. According to [16], approximately one-sixth of the centrifuges at Natanz were rendered inoperable, though the true operational disruption may have been more extensive when accounting for diagnostics, recovery time, and firmware remediation.

Analyzing the Stuxnet Attack Flow

The *attack complexity* of Stuxnet is 19 as it employed 19 attack techniques in total, as shown in Figure 2. To determine the *attack persistence time* of Stuxnet, we observe that an early version of the Stuxnet malware was compiled as far back as 6/22/2009 [7], [11], that a later version was compiled on 4/14/2010 [7], that Stuxnet was detected on 6/17/2010 for the first time [8], [11], and that the earliest known compromise occurred on 6/23/2009 [7]. This leads us to measure the attack persistence time as the interval between 6/22/2009 and 7/17/2010, for a maximum of 390 days, or an interval between 4/14/2010 and 7/17/2010 for a more

conservative estimate of approximately 94 days.

Precise *recovery time* is difficult to determine, but some sources suggest recovery and firmware cleanup efforts continued for up to two years [9], [13]. For *service degradation*, the attack was reported to have destroyed roughly one in six centrifuges at the Natanz nuclear facility, corresponding to a reduction in system effectiveness of approximately 16% [16]. However, the actual operational impact was possibly more severe than 16% owing to hidden faults, cascading control logic corruption, and offline maintenance time needed to restore non-destroyed systems. There are reports arguing that the plant's efficiency may have dropped by 50% during peak disruption, which could still be a conservative estimate [7], [9], [17].

The preceding analysis shows that Stuxnet was designed for stealthy, long-term degradation rather than immediate shutdown. Defenders must therefore prioritize early-stage containment if not prevention, such as user training against USB/phishing vectors, isolating critical systems, enforcing MFA or zero-trust principles, limiting attacker impact by disabling unauthorized firmware updates and hiding sensitive PLC configurations. This leads to:

Insight 1. *Preventive defense efforts should focus on training users to reduce their susceptibility to phishing and USB-based threats, isolating critical systems, and enforcing multi-factor authentication, if not zero-trust principles.*

The 2015 Attack on Ukraine Power Grid

In 2015, three regional power distribution companies in Ukraine were targeted in a coordinated cyberattack. The incident marked the most successful cyberattack to that point in time in disrupting an electrical grid. The attacker leveraged the BlackEnergy malware to gain initial access and move laterally within the power grid infrastructure, ultimately allowing the attacker to remotely control the circuit breakers and cause widespread power outages [18], [19].

Identifying, Collecting, and Preprocessing Cyber Threat Intelligence

To reconstruct the 2015 Ukraine power grid attack, we used Google to search for relevant open-source cyber threat intelligence. We manually examined the returned cyber threat intelligence reports, consisting of academic studies, non-academic technical analyses, and institutional reports, including the following. Technical sources [19], [20], [21], [22], [23], provided detailed accounts of the attack vectors, system impacts, and response challenges. A defense-oriented analysis, [24], offered mitigation strategies and insights into the technical environment. [18] contextualized the operational scope, malware deployment, and disruption timeline. These sources collectively enabled a structured reconstruction of the attack flow and a precise mapping to the MITRE ATT&CK techniques.

Reconstructing the Ukraine Power Grid Attack Flow

The attack pivoted from an IT infrastructure to the OT infrastructure. Figure 3

illustrates the attack flow, across three phases described in our methodology.

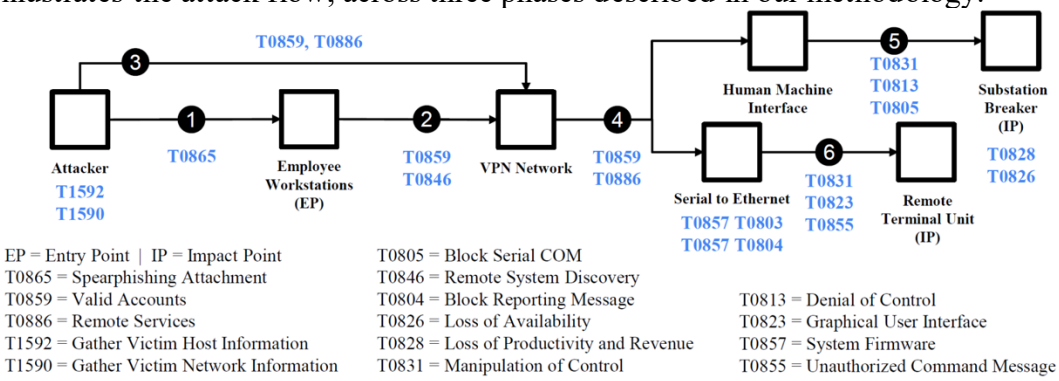


Figure 3: The 2015 Ukraine attack flow

Setting Conditions. The attacker gathered detailed information about the IT infrastructure structure and deployed systems using three attack techniques, namely Gather Victim Host Information (T1592), Gather Victim Network Information (T1590), and Remote System Discovery (T0846). This enabled the tailoring of malware to specific configurations and security postures. Furthermore, the attacker acquired the BlackEnergy malware and prepared malicious Microsoft Word and Excel files, which contain embedded VBA scripts, to deploy the malware upon execution [18], [19], [20].

Gaining Access. The attack commenced with a targeted Spearphishing Attachment (T0865) to aim at employee workstations, marked as ① in Figure 3. Employees inadvertently executed these malicious Microsoft Word and Excel documents, resulting in the deployment of the embedded BlackEnergy malware onto internal systems. Upon gaining a foothold, the attacker used BlackEnergy to harvest user credentials, gaining Valid Accounts (T0859) and expanding its presence within the infrastructure via the VPN, as shown by annotation ②.
Through plugins, such as VS.dll, the attacker conducted Remote System Discovery (T0846) and moved laterally by using the techniques Lateral Tool Transfer (T0867) and Remote Services (T0886). The attacker also leveraged External Remote Services (T0822) to bypass the perimeter defenses, indicated by annotation ③.

Deploying Effects. The attacker obtained extensive access to the OT infrastructure's internal components, enabling the attacker to conduct detailed monitoring and precise targeting. With access to the Graphical User Interface (T0823) on human-machine interfaces, and the serial-to-Ethernet connection highlighted by ④ the attacker used Manipulation of Control (T0831) to remotely trip substation breakers and interfere with industrial operations, indicated as ⑤. This resulted in a Denial of Control (T0813) for operators. Control of System Firmware (T0857) and the use of Block Reporting Message (T0804) as well as Block Command Message (T0803) compromised data integrity and system reporting functions employed by the OT infrastructure. To further complicate

recovery, the attacker sent Unauthorized Command Messages (T0855) and employed Block Serial COM (T0805) techniques to obstruct remote resets, indicated by ⑥.

Analyzing the Ukraine Power Grid Attack Flow

The *attack complexity* of the 2015 Ukraine power grid attack is 15 as it involved 15 attack techniques as highlighted in Figure 3. To measure the *attack persistence time*, we observe that the attacker's access was first gained in either May 2014 [19] or March 2015 [23], with the attack launched on 12/23/2015. This yields a dwell time of approximately 9 to 19 months, indicating extensive preparation prior to execution. In terms of *attack impact*, we estimate the *recovery time* sub-metric as approximately three hours, based on reports that restoration activities took three hours to complete once initiated [20], [23]. We estimate the *service degradation* sub-metric as 100% because multiple substations across three regional distribution operators were disabled, resulting in a 100% temporary loss of power delivery in affected regions [18], [22].

These measurements show that the attack was a long-dwell, coordinated campaign designed to exploit centralized control systems and procedural gaps. This suggests that defenders must prioritize early-stage detection and containment, such as segmentation, user training, and credential monitoring, while limiting the blast radius by restricting VPN access to OT systems and disabling remote operator controls [18], [19]. This leads to:

Insight 2 *Preventive efforts should focus on detecting and disrupting long-term persistence by implementing network segmentation, access monitoring, and user training to prevent credential misuse.*

The 2017 Attack on the Petro Rabigh Oil Refinery (Triton)

In 2017, the Petro Rabigh oil refinery in Saudi Arabia was targeted by a cyberattack involving the Triton malware. The malware was specifically designed to compromise industrial Safety Instrumented Systems (SIS), marking the first known cyberattack to directly target safety-critical infrastructures. The incident was publicly reported later in 2017 and raised significant concerns due to its potential to cause physical harm and loss of life [25], [26], [27].

Identifying, Collecting, and Preprocessing Cyber Threat Intelligence

To reconstruct the Triton attack flow, we consulted detailed technical and contextual sources, which were also obtained via Google searches with keywords "Triton Malware" and "Triton Attack". [28] examined the attack's targeting of Schneider Electric's SIS controllers, while [29], [30], [31], [32] described the techniques for lateral movement, privilege escalation, and safety logic compromise. [33] contributed comparative insights on advanced persistent threats in industrial

control environments. [25] helped confirm the attack's objectives and impacts. News articles by [26], [27], and [34], provided information about the incident timeline, downtime, and broader safety implications. These sources collectively provided sufficient information to enable us to reconstruct the Triton attack flow.

Reconstructing the Triton Attack Flow

At a high level, the Triton attack leveraged sophisticated malware targeting industrial SIS, which are specifically designed to monitor critical industrial processes and intervene automatically when hazardous conditions arise [26], [27], [28]. Figure 4 illustrates the Triton attack flow, structured according to the three phases outlined in our methodology.

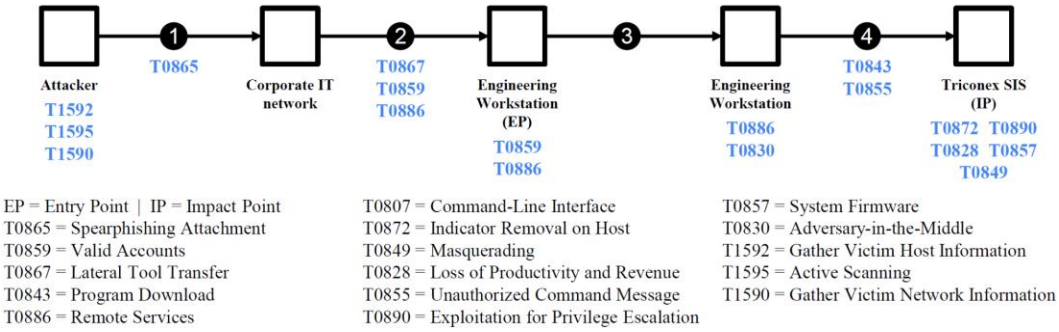


Figure 4: The Triton attack flow

Setting Conditions. The attacker conducted detailed reconnaissance on Petro Rabigh's Industrial Control Systems (ICS), specifically targeting Schneider Electric Triconex SIS controllers [29], [30], [31]. Specifically, the attacker used Gather Victim Host Information (T1592), Active Scanning (T1595), and Gather Victim Network Information (T1590) to identify the exact models of these SIS controllers and corresponding firmware versions. This reconnaissance led to the discovery of a zero-day vulnerability and the development of custom malware tailored to the SIS devices [29], [30], [31].

Gaining Access. While the precise method of initial access into Petro Rabigh's IT infrastructure is unknown, many experts suspect it involved a Spearphishing Attachment (T0865) [33] and [28], indicated by ① in Figure 4. Once inside the corporate IT infrastructure, the attacker established persistent access, potentially using Valid Accounts (T0859) and Remote Services (T0886) to maintain control and prepare for lateral movement ②. Eventually, the attacker pivoted from the IT infrastructure to the OT infrastructure by accessing engineering workstations that served as an entry point into the industrial environment, indicated by ③. This transition was enabled by Lateral Tool Transfer (T0867), repeated use of Valid Accounts (T0859), and continued exploitation of Remote Services (T0886). Once within the OT infrastructure, the attacker employed Adversary-in-the-Middle (T0830) to intercept communications and remain covert [28], [33].

Deploying Effects. Upon reaching the targeted Triconex SIS controller, as

indicated by ④, the attacker used Program Download (T0843) to deliver the final payload. This was accompanied by Unauthorized Command Messages (T0855) to disrupt safety logic execution. To maintain stealth, and possibly complicate post-incident forensics, the attacker employed both Masquerading (T0849) and Indicator Removal on Host (T0872). The attacker also employed Exploitation for Privilege Escalation (T0890) to bypass security boundaries and ultimately tamper with System Firmware (T0857), enabling direct manipulation of safety functions[28], [31], [32]. Although the attack was interrupted before catastrophic effects could be realized, it still caused a Loss of Productivity and Revenue (T0828) and necessitated the shutdown of the affected systems. Although the lasting physical damage was limited, the attack demonstrated the real-world potential of cyber operations in targeting and disrupting systems designed to protect human life. The attack's unprecedented targeting of safety-critical infrastructure marked a turning point in the evolution of cyber physical threats [25], [26], [27].

Analyzing the Triton Attack Flow

The *attack complexity* of the Triton attack is 16 because it used 16 attack techniques as highlighted in Figure 4. The *attack persistence time* is estimated at three years, with initial access reported in 2014 and the final attack execution occurring in 2017 [26]. This extended dwell time suggests that the adversary conducted extensive reconnaissance and preparation, particularly of SIS. In terms of *attack impact*, the attack reportedly led to at least one week of downtime [25], with an earlier disruption occurring in June 2017 [26], suggesting that the total operational impact may have ranged from one to four weeks. During this time window, plant output halted and critical processes were disabled, indicating an approximate 100% *service degradation* [26], [27]. The preceding measurements suggest that effective defenses should leverage defense in depth and segmentation. This leads to:

Insight 3 *Preventive efforts should combine ingress-focused defenses (e.g., IT/OT segmentation) to prevent tailored attacks that pivot from IT to OT infrastructures.*

The 2000 Maroochy Shire Sewage Attack

In 2000, the Maroochy Shire sewage control system in Queensland, Australia, was targeted by a deliberate cyberattack. A disgruntled former contractor gained unauthorized wireless access and used it to manipulate the facility's ICS. This incident is one of the earliest documented cyberattacks on OT infrastructure.

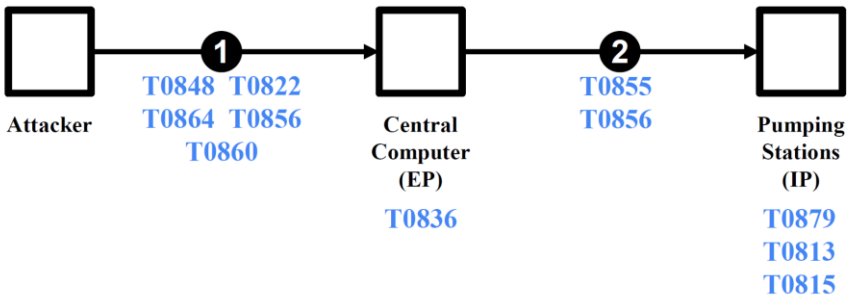
Identifying, Collecting, and Preprocessing Cyber Threat Intelligence

To reconstruct the Maroochy Shire Sewage Attack, we examined cyber threat intelligence from technical reports, academic research, and media coverage

returned by a Google search using the keyword "Maroochy Water Breach." Analyses by [35], [36], and [37] detailed the attacker's privileged access as a former employee, the use of radio-based control gear, and tactics aligned with multiple MITRE ATT&CK techniques. These sources also provided insight into the attacker's use of spoofed messages, rogue master emulation, and alarm suppression. [38] corroborated the financial and environmental consequences. Together, these sources enabled a high-fidelity reconstruction of the insider threat's behavior, access path, and real-world impact.

Reconstructing the Maroochy Shire Sewage Attack Flow

To reconstruct the Maroochy Shire Sewage attack flow, we leveraged the available forensic and investigative reports referenced above. At a high level, the attacker, namely, a former contractor with knowledge of the system, used a stolen wireless controller and reverse-engineered the protocols to access the supervisory control and data acquisition (SCADA) network remotely. Over the course of several weeks, the attacker issued unauthorized commands to manipulate sewage valves and pumps, ultimately causing over 800,000 liters of untreated sewage to spill into public waterways and facilities [37], [38]. Figure 5 highlights the Maroochy Shire Sewage attack flow, including the three phases in *setting conditions*, *gaining access*, and *deploying effects*, which are elaborated below.



EP = Entry Point | IP = Impact Point
T0848 = Rogue Master
T0864 = Transient Cyber Asset
T0860 = Wireless Compromise
T0822 = External Remote Services
T0878 = Alarm Suppression
T0838 = Modify Alarm Settings
T0879 = Damage to Property
T0813 = Denial of Control
T0815 = Denial of View
T0855 = Unauthorized Command Message
T0836 = Modify Parameter
T0856 = Spoof Reporting Message

Figure 5: The Maroochy Shire Sewage attack flow

Setting Conditions. Due to the nature of insider threats, traditional external reconnaissance was unnecessary. The attacker, a former employee, was already intimately familiar with the OT infrastructure, including its control systems. The attacker also retained prior access to a Transient Cyber Asset (T0864) [35], [37], providing the attacker privileged insight without the need for typical reconnaissance. The attacker's knowledge of the communication infrastructure,

combined with access to legitimate hardware, enabled the attacker to emulate a Rogue Master (T0848) device, allowing seamless interaction with operational assets [35], [37].

Gaining Access. In this case, the attacker's access diverged significantly from that of traditional cyberattacks. As indicated by the arc annotated with ① in Figure 5, the attacker used Wireless Compromise (T0860) by leveraging stolen radio-based communication gear and reconnected to the control network remotely via External Remote Services (T0822). These actions allowed the attacker to reinsert themselves back into the system without breaching perimeter defenses. Additionally, the attacker sent falsified messages consistent with the Spoof Reporting Message (T0856) technique to appear as a legitimate device on the network [35], [37].

Deploying Effects. After successfully establishing access to the central control computer (the *entry point*), the attacker used the Modify Parameter technique (T0836) to alter operational functions and sent Unauthorized Command Messages (T0855) to downstream pumping stations, as indicated by the arc annotated with ②. These malicious commands caused the systems to misbehave while disguising the attacker's true intent [35], [37]. The effects were profound. The attacker leveraged Modify Alarm Settings (T0838) and Alarm Suppression (T0878) to prevent detection by operations staff, while the pumping stations suffered from Denial of View (T0815) and Denial of Control (T0813), making it difficult for operators to observe or correct system behavior in real time. The deliberate sabotage culminated in Damage to Property (T0879), with 800,000 liters of untreated sewage discharged into the environment. Financial consequences included at least $176,000 in cleanup costs for the local city council and an additional $500,000 in damages to the contractor, Hunter Watertech Pty Ltd [35], [36], [37], [38]. This incident clearly demonstrates the significant physical and financial impacts possible from insider threats in critical infrastructure.

Analyzing the Maroochy Shire Sewage Attack Flow

The *attack complexity* of the Maroochy Shire Sewage attack is quantified as 12 because it used 12 attack techniques, as highlighted in Figure 5. The *attack persistence time* is measured as the time interval between when unauthorized access started, which is reportedly between Late January 2000 [35] and 02/28/2000 [37], and when the attack was recognized in March 2000. This leads us to estimate the *attack persistence time* as between one and two months. In terms of *attack impact*, we measure the *recovery time* as multiple days [38], which is the time it takes the company to clean up the sewage discharged into the environment. We measure the *service degradation* sub-metric as 0% because the service remained sufficient to operational demands, despite the sewage discharge.

The preceding analysis shows that even low-complexity insider threats can bypass conventional defenses and cause long-term disruptions. Preventive efforts should therefore focus on identity and trust management, including automatic

revocation of credentials upon termination to invalidate residual access. It also demonstrates that even an attack that does not reduce system output can have other negative consequences. This leads to:

Insight 4 *Preventive defense should be employed to mitigate insider threats, including immediate credential revocation.*

Discussion

Contrasting Analysis via Metrics

Now we analyze the four OT attacks by leveraging the measurement results of their metrics. First, we observe that although Stuxnet exhibited the highest *attack complexity*, or the use of 19 attack techniques, its ultimate physical impact was relatively contained when compared with that of the Maroochy Shire Sewage attack and that of the Triton attack. This suggests that using more attack techniques does not necessarily result in greater damage.

Second, we observe that the *attack persistence time* underscores the importance of early detection, while noting that extended dwell times often enable attackers to refine their strategies. However, direct comparisons can be misleading due to the differing nature of attacks. For example, The Maroochy Shire Sewage attack required minimal time in the target infrastructure, as the attackers had extensive preattack intelligence. By contrast, the Ukraine power grid attacker spent considerable time within the infrastructure analyzing and planning their actions. The Triton attacker was in the Petro Rabigh infrastructure for approximately three years before activating, illustrating how long-term access can support customized attack. These differing contexts highlight the need for nuanced interpretations of this metric.

Third, we observe that the *recovery time* sub-metric reflects a system's ability to "bounce back" from a compromised state to make everything normal. We also observe that the *service degradation* sub-metric measures only the immediate functional impact on system operations. In the Maroochy Shire Sewage attack, this impact was minimal, as the core control systems continued to operate despite unauthorized manipulation. By contrast, the Stuxnet attack introduced measurable technical degradation—potentially reducing centrifuge efficiency by up to 50% due to corrupted control logic and the need for offline maintenance. However, it remains unclear whether this degradation translated into long-term operational disruption or strategic failure [7], [9], [17].

Contrasting Analysis via Attack Phases

First, we observe that in all four attacks discussed above, attackers obtained near-complete knowledge of their target environment before launching their operations. This pattern highlights the importance of the *setting conditions* phase, specifically, that detailed reconnaissance is a prerequisite for successful cyberattacks against OT

systems. This suggests that *proactively* thwarting adversary reconnaissance is critical to defending cyberattacks against OT infrastructures, such as intelligently obfuscating or isolating critical system components (such as device types, firmware versions, or network topology). If attackers cannot identify what to target, they may be unable to launch an effective attack at all. This leads to:

Insight 5 *Proactively disrupting adversary reconnaissance may be an effective approach to thwarting sophisticated attacks, making it a promising direction for future research.*

Second, we make two observations on the *gaining access* phase. (i) We observe that human involvement is a pivotal factor in this phase. Both the Ukraine power grid attack and the Petro Rabigh attack relied on the Spearphishing Attachment technique (T0865) to initiate compromise, while the Stuxnet attack required physical delivery via Replication Through Removable Media (T0847). The Maroochy Shire Sewage attack bypassed technical defenses entirely through insider threats, demonstrating that not all threats originate outside the perimeter. These examples underscore the persistent vulnerability of human factors, despite improvements in technical defenses. Mitigating this risk requires a layered approach: robust training and awareness, strict access controls, and safeguards such as multi-factor authentication (MFA) and asset tracking. (ii) We observe that lateral movement in the OT infrastructures was central to the success of attacks. The Stuxnet attack leveraged multiple propagation techniques, such as Lateral Tool Transfer (T0867) and Remote Services (T0886), to move between IT and OT infrastructures. The Ukraine power grid attacker exploited VPN access to cross the IT-OT boundary. The Triton attacker used Valid Accounts (T0859) and administrative tools to blend into normal operations. By contrast, the Maroochy attacker required no lateral movement because the entry point was already the impact point. These distinctions reinforce the need for network segmentation especially zero-trust, behavioral monitoring, and the principle of least privilege as defenses against lateral movements. These lead to:

Insight 6 *Reducing human susceptibility to cyber social engineering attacks and segmenting OT infrastructure using zero-trust principles can effectively thwart attacks.*

Third, in the *deploying effects* phase, we observe that the four attacks achieved a 100% attack success rate with respect to their intended goals, whether that meant sabotage, system manipulation, or physical damage. This highlights a sobering truth: once attackers reach a critical control point, they are highly likely to cause impact. Prevention and early detection must therefore be prioritized, since once effects are deployed, it may already be too late to avert harm. This leads to:

Insight 7 *It would be ideal to prevent all attacks from penetrating into OT infrastructures, or thwart them before they achieve their intended effects.*

Conclusion

As OT infrastructures continue to play a critical role in modern society, securing them against cyberattacks has become more urgent than ever due to their growing exposure to public networks. Through case studies of four attacks, we have drawn useful insights that can guide the design of future defenses, including: (i) proactively thwarting reconnaissance; (ii) reducing human's susceptibility to cyber social engineering attacks and segmenting OT infrastructures using zero-trust principles; (iii) preventing attacks before they achieve their intended effects. These insights can help shape the next generation cyber defense strategies.

Acknowledgement. The research was supported in part by the DoD VICEROY and UC2 programs. The views are that of the authors and do not reflect the policies of the funding agencies or the government.

References

[1] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "Mitre att&ck: Design and philosophy," in *Technical report*, The MITRE Corporation, 2018.

[2] MITRE Corporation, "MITRE ATT&CK for Enterprise," 2024. [Online]. Available: https://attack.mitre.org/matrices/enterprise/

[3] MITRE Corporation, "MITRE ATT&CK for Industrial Control Systems (ICS)," 2024. [Online]. Available: https://attack.mitre.org/matrices/ics/

[4] L. Zeien, C. Chang, L. T. C. Ear, and D. S. Xu, "Characterizing Advanced Persistent Threats Through the Lens of Cyber Attack Flows," *Military Cyber Affairs*, vol. 7, no. 1, p. 5, 2024.

[5] E. Ear *et al.*, "Characterizing Russia's Cyber Operations in Ukraine through the Lenses of Cyber Attack TTPs," *USCYBERCOM CyberRecon'2024*, 2024.

[6] D. Kushner, "The real story of stuxnet," *IEEE Spectr*, vol. 50, no. 3, pp. 48–53, 2013, doi: 10.1109/MSPEC.2013.6471059.

[7] N. Falliere, L. O Murchu, and E. Chien, "W32.Stuxnet Dossier: Version 1.4," Feb. 2011. [Online]. Available: https://pax0r.com/hh/stuxnet/Symantec-Stuxnet-Update-Feb-2011.pdf

[8] A. Matrosov, E. Rodionov, D. Harley, and J. Malcho, "Stuxnet Under the Microscope: Revision 1.31," 2011. [Online]. Available: https://www.rpac.in/image/ITR%201.pdf

[9] R. Langner, "Stuxnet: Dissecting a Cyberwarfare Weapon," *IEEE Secur Priv*, vol. 9, no. 3, pp. 49–51, 2011, doi: 10.1109/MSP.2011.67.

[10] Antiy, "Report on the Worm Stuxnet's Attack," Oct. 2010. [Online]. Available: https://www.antiy.net/media/reports/stuxnet_analysis.pdf

[11] M. De Falco, "Stuxnet Facts Report: A Technical and Strategic Analysis," Tallinn, Estonia, 2012. [Online]. Available: https://ccdcoe.org/uploads/2018/10/Falco2012_StuxnetFactsReport.pdf

[12] R. Abrams, “Why Steal Digital Certificates?,” Jul. 2010. [Online]. Available: https://www.welivesecurity.com/2010/07/22/why-steal-digital-certificates/
[13] J. R. Lindsay, “Stuxnet and the Limits of Cyber Warfare,” *Security Studies*, vol. 22, no. 3, pp. 365–404, 2013, doi: 10.1080/09636412.2013.816122.
[14] S. Collins and S. McCombie, “Stuxnet: the emergence of a new cyber weapon and its implications,” *Journal of Policing, Intelligence and Counter Terrorism*, vol. 7, no. 1, pp. 80–91, 2012, doi: 10.1080/18335330.2012.653198.
[15] J. P. Farwell and R. Rohozinski, “Stuxnet and the Future of Cyber War,” *Survival (Lond)*, vol. 53, no. 1, pp. 23–40, 2011, doi: 10.1080/00396338.2011.555586.
[16] E. Nakashima and J. Warrick, “Stuxnet was work of U.S. and Israeli experts, officials say,” *Washington Post*, Jun. 2012, [Online]. Available: https://www.washingtonpost.com/world/national-security/stuxnet-was-work-of-us-and-israeli-experts-officials-say/2012/06/01/gJQAlnEy6U_story.html
[17] M. Baezner and P. Robin, “Stuxnet,” May 2018. [Online]. Available: https://www.researchgate.net/publication/323199431_Stuxnet
[18] CISA, “Cyber-Attack Against Ukrainian Critical Infrastructure,” Jul. 2021. [Online]. Available: https://www.cisa.gov/news-events/ics-alerts/ir-alert-h-16-056-01
[19] R. M. Lee, M. J. Assante, and T. Conway, “Analysis of the Cyber Attack on the Ukrainian Power Grid – Kaspersky,” 2016, *E-ISAC and SANS*. [Online]. Available: https://assets.contentstack.io/v3/assets/blt36c2e63521272fdc/blt6a77276749b76a40/607f235992f0063e5c070fff/E-ISAC_SANS_Ukraine_DUC_5[73].pdf
[20] A. Shehod, “Ukraine Power Grid Cyberattack and US Susceptibility: Cybersecurity Implications of Smart Grid Advancements in the US,” Dec. 2016. [Online]. Available: https://web.mit.edu/smadnick/www/wp/2016-22.pdf
[21] Y. Xiang, L. Wang, and N. Liu, “Coordinated attacks on electric power systems in a cyber-physical environment,” *Electric Power Systems Research*, vol. 149, pp. 156–168, 2017, doi: https://doi.org/10.1016/j.epsr.2017.04.023.
[22] D. E. Whitehead, K. Owens, D. Gammel, and J. Smith, “Ukraine cyber-induced power outage: Analysis and practical mitigation strategies,” in *2017 70th Annual Conference for Protective Relay Engineers (CPRE)*, 2017, pp. 1–8. doi: 10.1109/CPRE.2017.8090056.
[23] J. Styczynski and N. Beach-Westmoreland, “When the Lights Went Out,” 2016, *Booz Allen Hamilton*. [Online]. Available: https://www.boozallen.com/content/dam/boozallen/documents/2016/09/ukraine-report-when-the-lights-went-out.pdf
[24] J.-P. Hauet, P. Bock, R. Foley, and R. Françoise, “Ukrainian Power Grids

Cyberattack," *InTech*, Mar. 2017, [Online]. Available: https://www.isa.org/intech-home/2017/march-april/features/ukrainian-power-grids-cyberattack

[25] CISA, "MAR-17-352-01 HatMan—Safety System Targeted Malware (Update B)," Feb. 2019. [Online]. Available: https://www.cisa.gov/sites/default/files/documents/MAR-17-352-01%20HatMan%20-%20Safety%20System%20Targeted%20Malware%20%28Update%20B%29.pdf

[26] M. Giles, "Triton is the World's Most Murderous Malware, and It's Spreading," Mar. 2019. [Online]. Available: https://www.technologyreview.com/2019/03/05/103328/cybersecurity-critical-infrastructure-triton-malware/

[27] B. Sobczak, "The Inside Story of the World's Most Dangerous Malware," Mar. 2019. [Online]. Available: https://www.eenews.net/articles/the-inside-story-of-the-worlds-most-dangerous-malware/

[28] A. Di Pinto, Y. Dragoni, and A. Carcano, "TRITON: The First ICS Cyber Attack on Safety Instrument Systems," in *Black Hat USA 2018 - Research Paper*, 2018. [Online]. Available: https://scadahacker.com/library/Documents/Cyber_Events/Nozomi%20-%20TRITON%20-%20The%20First%20SIS%20Cyberattack.pdf

[29] B. Johnson, D. Caban, M. Krotofil, D. Scali, N. Brubaker, and C. Glyer, "Attackers Deploy New ICS Attack Framework 'TRITON' and Cause Operational Disruption to Critical Infrastructure," Dec. 2017. [Online]. Available: https://cloud.google.com/blog/topics/threat-intelligence/attackers-deploy-new-ics-attack-framework-triton/

[30] F. Intelligence, "TRITON Attribution: Russian Government-Owned Lab Most Likely Built Custom Intrusion Tools for TRITON Attackers," Oct. 2018. [Online]. Available: https://cloud.google.com/blog/topics/threat-intelligence/triton-attribution-russian-government-owned-lab-most-likely-built-tools/

[31] B. Jeffries, S. Saravia, C. Carter, and Z. Ankuda, "Cyber Risk to Mission Case Study: Triton," Bedford, MA, Oct. 2022. [Online]. Available: https://apps.dtic.mil/sti/trecms/pdf/AD1183008.pdf

[32] F. B. of Investigation, "Private Industry Notification: Cyber Threat Advisory," Mar. 2022. [Online]. Available: https://www.ic3.gov/CSA/2022/220325.pdf

[33] R. Kumar, R. Kela, S. Singh, and R. Trujillo-Rasua, "APT attacks on industrial control systems: A tale of three incidents," *International Journal of Critical Infrastructure Protection*, vol. 37, p. 100521, 2022, doi: https://doi.org/10.1016/j.ijcip.2022.100521.

[34] F. Bajak and M. Levy, "Breaches by Iran-Affiliated Hackers Spanned Multiple U.S. States, Federal Agencies Say," Dec. 2023. [Online]. Available: https://apnews.com/article/hackers-iran-israel-water-utilities-critical-infrastructure-cisa-554b2aa969c8220016ab2ef94bd7635b

[35] N. Sayfayn and S. Madnick, “Cybersafety Analysis of the Maroochy Shire Sewage Spill,” Cambridge, MA, May 2017. [Online]. Available: https://web.mit.edu/smadnick/www/wp/2017-09.pdf

[36] J. Slay and M. Miller, “Lessons Learned from the Maroochy Water Breach,” in *International Federation for Information Processing Digital Library; Critical Infrastructure Protection;*, May 2007, pp. 73–82. doi: 10.1007/978-0-387-75462-8_6.

[37] M. Abrams and J. Weiss, “Malicious Control System Cyber Security Attack Case Study - Maroochy Water Services, Australia,” 2008. [Online]. Available: https://www.mitre.org/sites/default/files/pdf/08_1145.pdf

[38] G. Cohen, “Throwback Attack: An Insider Releases 265,000 Gallons of Sewage on the Maroochy Shire,” Nov. 2021. [Online]. Available: https://www.controleng.com/throwback-attack-an-insider-releases-265000-gallons-of-sewage-on-the-maroochy-shire/