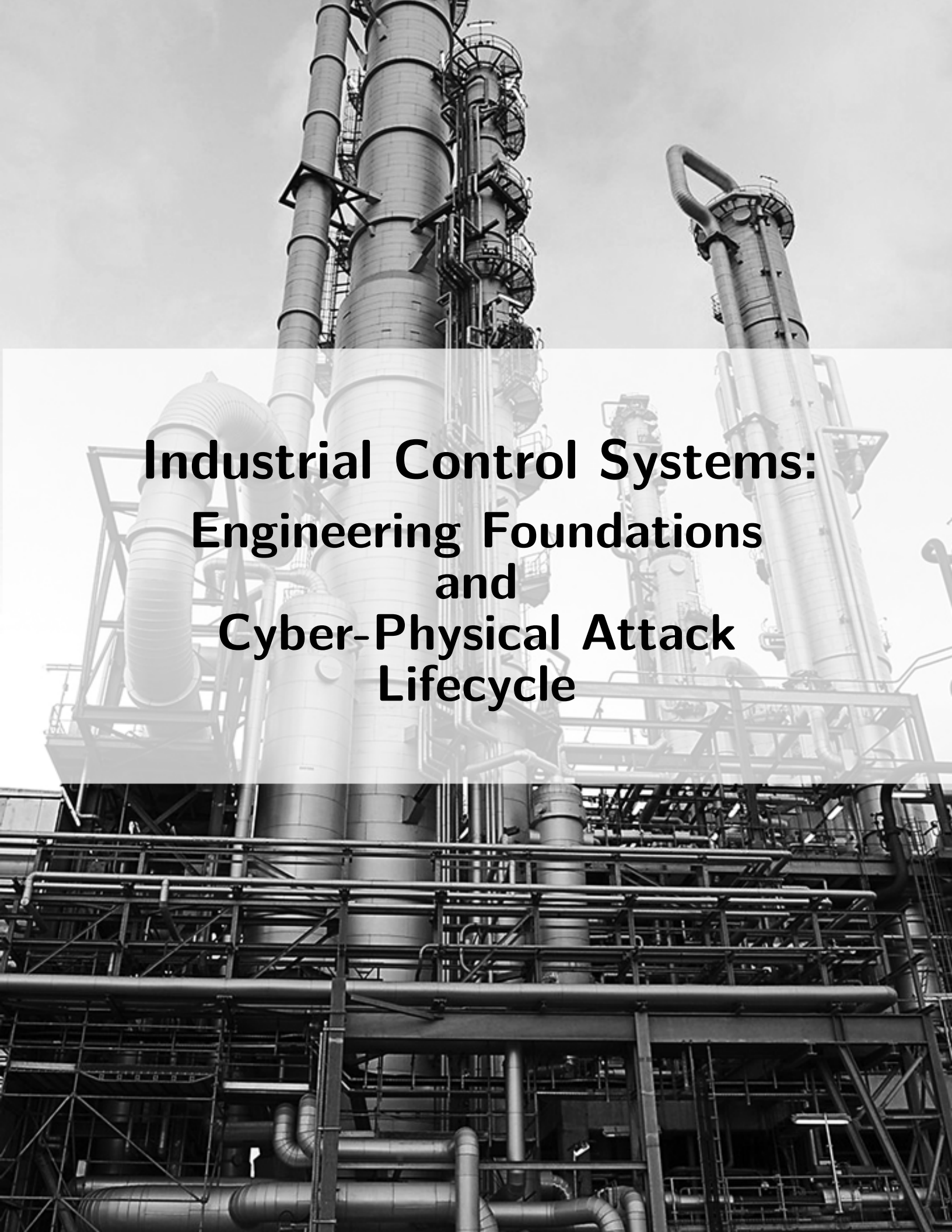


Industrial Control Systems:
Engineering Foundations and Cyber-Physical Attack Lifecycle

Technical white paper

Marina Krotofil
ICS Security Engineer
@marmusha

May 2023

Contents

Executive Summary

Industrial control system (ICS) is a collective term used to describe different types of control systems and associated instrumentation, which include end-point devices, systems, networks, applications and controls used to operate and/or automate industrial processes.

Understanding attack methods and tools allows defenders to conduct informed threat assessment and proactively implement adequate security controls and monitoring tactics. An "attack lifecycle" or "kill chain" are common methods to describe the process of conducting cyber attacks. In the final stage, the attackers act upon their motivation and take the required action to achieve their planned mission. In the ICS domain, the attacker aims at achieving a desired outcome in the physical world. This technical white paper introduces a Cyber-Physical Attack Lifecycle and illustrates its attack stages on the example of a targeted attack on a chemical plant.

This paper consist of four parts:

- Introduction to cyber-physical systems security (Chapter 1);
- ICS Engineering Foundations, based on own experiences, the current body of knowledge on ICS infrastructures and OT cyber security, including progress achieved in Industry 4.0 concept (Chapter 2 and Appendix A);
- Introduction to a Vinyl Acetate chemical process and its control model (Chapter 3 and Appendix B);
- Cyber-Physical Attack Lifecycle and an illustration of designing a targeted cyber-physical attack on a chemical process (Chapter 4).

Note to the readers. ICS/CPS security is a complex field, it would not be possible to include all related knowledge base, nuances and critical discussions without overloading the paper and turning it into a long, overwhelming read. To the best of our abilities, we kept this white paper as concise and as comprehensive as possible.

Acknowledgment. The author expresses her deep appreciation to Jason Larsen for his pioneering work in the field of "physical damage" and for enabling the development of a new body of knowledge in cyber-physical attacks field such as presented in this whitepaper.

1. Introduction

Advances in computing and networking have added new capabilities to physical systems that could not be feasibly added before. This has led to the emergence of engineered systems called cyber-physical systems (CPS): systems where the events in the physical world are managed with the help of modern advances in computation and control. Complex machines such as aircraft or robots, building automation systems, smart cities and smart grids, railways and agricultural systems, medical devices and industrial infrastructures, in general, are examples of cyber-physical systems. Some of these industrial processes are categorized as critical infrastructures (CI) because societal well-being depends on their reliability (e.g., water and power utilities). On the other side of the spectrum of the CPS definition are smaller appliances and gadgets such as consumer electronics and wearables widely known as the Internet of Things (IoT) devices. Instead of being an inseverable part of a larger cyber-physical ecosystem, they are stand-alone devices that are programmed for certain clearly defined applications and rely on Internet connectivity for continuous transmission of data for analysis and feedback. In the white paper, we consider larger cyber-physical systems such as industrial infrastructures due to their higher complexity.

Cyber-physical systems, as a new type/kind of systems, were mapped as a novel research area in a series of National Science Foundation (USA) workshops starting in 2006. This is also when the term *cyber-physical systems* first emerged. In cyber-physical systems, physical processes affect computations and vice versa, with computations and physical processes being so tightly integrated that it is not possible to identify whether behavioral attributes are the result of computations, physical laws, or both working together. In the words of Edward A. Lee, one of the pioneers of the cyber-physical systems discipline [134]:

> As an intellectual challenge, CPS is about the intersection, not the union, of the physical and the cyber. It combines engineering models and methods from mechanical, environmental, civil, electrical, biomedical, chemical, aeronautical and industrial engineering with the models and methods of computer science.

The design of cyber-physical systems is subject to a wide range of physical requirements, such as dynamics, power and physical size as well as to systems-level requirements, such as safety, security and fault tolerance. While security is part of the requirements, it is deeply embedded into engineering specifications.

> In CPS, "cyber security" is not seen as a property that concerns the confidentiality, integrity and availability of information systems and has no intrinsic connection with physical processes, but as a novel discipline of cyber-physical security – an integration of physical/engineering and information sciences.

1.1 Cyber-Physical Systems Security

Perceived and real security threats affecting cyber-physical systems have been attracting considerable attention in the media, among decision makers, regulators, and the research community. Cyber-physical systems, by their very nature, cause effects in the physical world, which in some cases can have disastrous physical consequences. On one hand, this is an issue that had to be dealt with already before physical systems were connected to cyberspace. Well-designed systems would have been deployed with appropriate safety measures in place. Conceivably, those measures can restrain cyber-physical attacks once they have transited from cyberspace into the physical domain. On the other hand, those countermeasures were designed under certain assumptions, e.g., physical security protecting access to premises or independence of component failures. Conceivably, those assumptions get invalidated once modern Information Technology (IT) systems get integrated with existing physical plants.

The concern for physical consequences puts cyber-physical systems security apart from the science of *information security*. Integrating modern IT systems with existing physical systems exposed these installations to new security threats. Some of these threats are well-known in IT security and countermeasures have been studied at length. Those threats are new only because of a new application area. Other threats may indeed be specific to cyber-physical systems. The white paper focuses on new security aspects intrinsic to cyber-physical systems that establish cyber-physical security as an object of research in its own right.

1.2 Motivation

For a long time, the primary focus of industry and academia has been on securing the communication infrastructure and hardening control systems. There is a large body of literature on how to adapt existing IT security methods to the characteristic features of the control domain. However modern malware for persistent attacks may now be equipped with "trusted" certificates, travel in USB sticks and laptops of "trusted" parties, carry zero-days exploits and rootkits, and propagate through "trusted" security updates. It is becoming increasingly difficult to prevent, detect and halt these attacks based solely on technical security measures deployed in the cyber layer.

It is often claimed that "once communications security is compromised the attacker can do whatever she wants". These are presumptuous claims. The attacker may well be able to inject any input she wants but this does not necessarily amount to being able to influence processes in the physical world at will. The processes and their actuators have to

be properly understood. Process physics and built-in safety measures might get in the way of the attacker. To address the limitations of defending a system using only IT methods, a new line of research has focused on understanding the *adversary's interactions with the physical system.* Analyzing the effects of attacks in the process control domain is an active research area. Figure 1.1 shows the logical layers of a cyber-physical system.

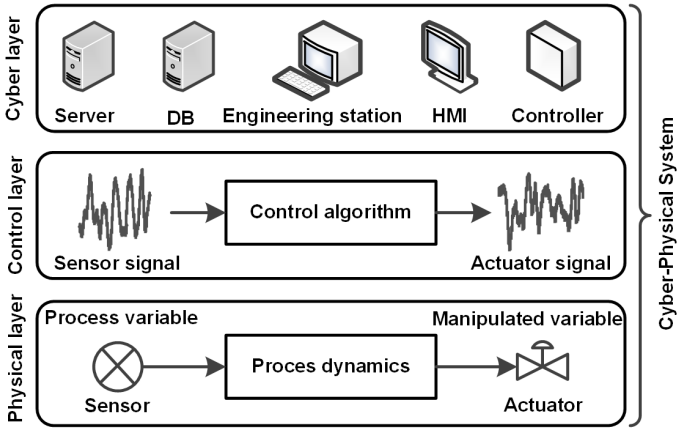


Figure 1.1: Logical layers of a cyber-physical system

While compromising the cyber layer is essential, the attacker needs to interact with the control system to achieve the desired outcome in the physical world. Exploration of such interactions from the attacker's perspective and a better understanding of needed defenses is the research angle explored in this white paper.

2. ICS: Engineering Foundations

Industrial control system (ICS) is a collective term used to describe different types of control systems and associated instrumentation, which includes end-point devices, systems, networks, applications and controls used to operate and/or automate industrial processes. Process Control Systems (PCS) is a special case of Industrial Control Systems, which refers to monitoring and managing continuous or batch processes such as chemical plants or water utilities in order to ensure the desired output.

Depending on the complexity of a process and the spread of its supporting infrastructure, the process control function may be implemented with various types of control systems. The distinction between the major types of control systems is visualized in Figure 2.1 and can be summarized as follows:

- Programmable Logic Controller (PLC): Typically used to control a small-size/low complexity process or a unit of a larger process. The control function of a PLC may or may not be monitored by a human operator;
- Distributed Control System (DCS): Typically used for a large-scale, complex process in a single location. DCS includes an integrated control center for supervisory process control by human operators;
- Supervisory Control and Data Acquisition (SCADA): Typically used to control a complex process distributed via a large geographic area and may include several manned and unmanned control rooms/centers.

In practice, SCADA systems have evolved over the years and become very similar to DCS in functionality. It is not uncommon that both types of control systems are used in the same industrial site, supplemented with PLC-based controls to administer small supporting functions. This white paper focuses on the architectures and terminology closely related to distributed control systems due to their prevalent usage in the (petro)chemical sector, power generating, pharmaceutical, manufacturing, and other large-scale continuous and batch processes.

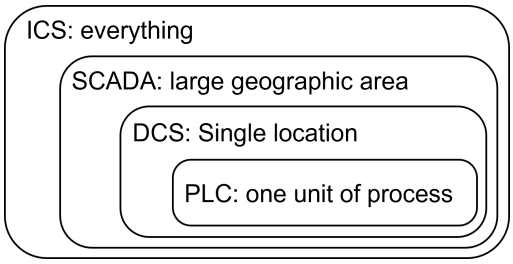


Figure 2.1: Distinction between various types of control systems

2.1 Fundamentals of Process Control

Process control is an umbrella term for both an engineering discipline and engineered infrastructures which include various systems and software that exert control over production processes. Control systems include sensors and data processing electronic units, actuators, networks to connect equipment, and algorithms to relate process variables to product attributes.

In the process industry *process* refers to the methods of changing or refining raw materials to create an end product. *Control* refers to the methods that are used to control process variables when manufacturing a product. This is done for three major reasons: (1) Reduce variability, (2) Increase efficiency, (3) Ensure safety. The first two points are important for the plant economy. Reduced variability lowers operational costs and ensures consistent quality of the end product. Efficiency refers to the accurate maintenance of optimal production conditions to decrease the production bill. Precise control is important for preventing runaway processes and ensuring safe operations.

The complexity of modern production processes is usually simplified by dividing the control load into subsystems containing separate control loops. The control loop is a fundamental building block of industrial control systems. A block diagram of a basic feedback control loop is shown in Figure 2.2.

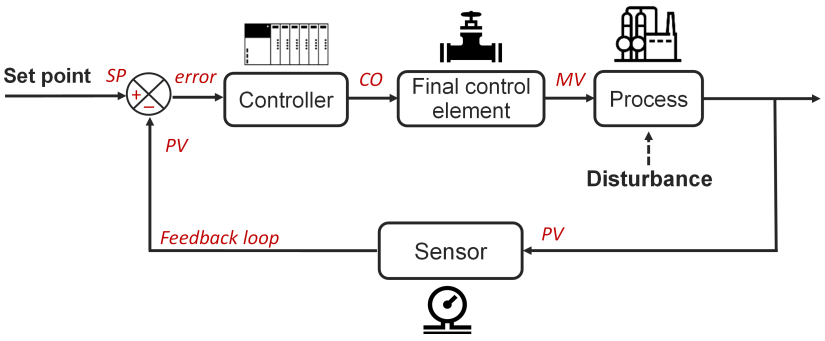


Figure 2.2: The components of a control loop

A control loop monitors the production process via sensors deployed around the production infrastructure and interacts with the process through actuators also called final control elements (FCE). The autonomous control function over a process is achieved through a

specialized computing element called a controller. Configuration of the control loop starts with a decision on a *set point* (SP) which is the desired value of a certain process parameter, e.g., a tank level L. Level L is called *process variable* or *process value* (PV) and must be kept as close to the setpoint as possible using control methods (minimization of *error*). Level L can be measured directly in which case the PV may be called *measured variable*, or indirectly by measuring two process variables, e.g., in- and out-flows. Process variables are fed into a controller containing a control algorithm based on a complex set of equations. The controller calculates the offset between SP and PV and generates an actionable controller output (CO) signal to the FCE. The actuator then adjusts the *manipulated value* (MV), e.g., flow F to bring the process closer to the SP.

In practice, control loops can be complex. More common are multivariable or advanced control loops in which each MV depends on two or more of the process variables as shown in Figure 2.3. The strategies for holding basic and multivariable control loops at setpoints are not trivial, and the interactions of numerous setpoints in the overall process control scheme can be subtle and complex. Process interactions may cause loop interactions via hidden feedback control loops and may result in control loops instability. For this reason, heavy control loop couplings among subsystems are avoided.

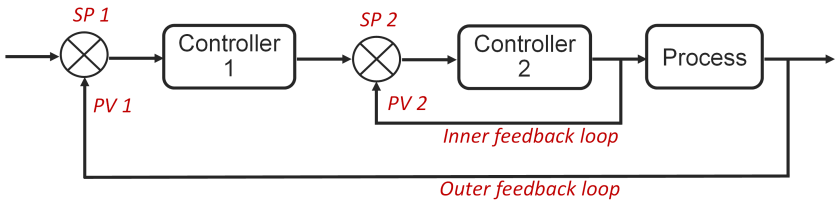


Figure 2.3: Multivariable control loop

A variety of process controllers are used to manage processes, e.g., On-Off controllers. However, the most common control algorithm used in industrial applications is a so called PID controller.

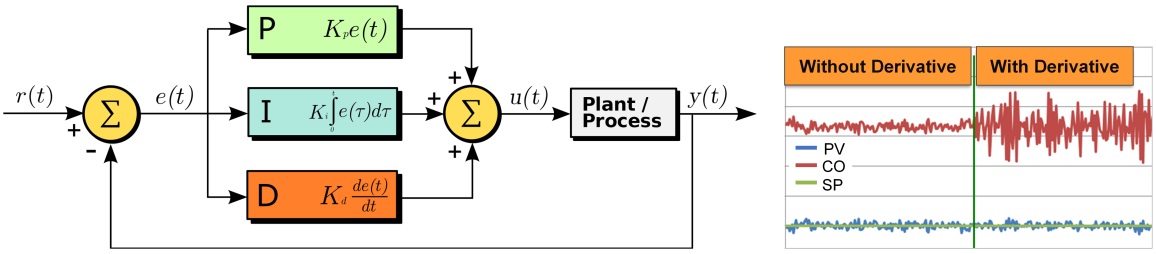
Definition: A Proportional–Integral–Derivative (PID) controller is a control loop mechanism employing feedback that is calculated according to the following formula:

$$u(t) = K_p e(t) + K_i \int_0^t e(t')\,dt' + K_d \frac{de(t)}{dt}, \tag{2.1}$$

where K_p, K_i and K_d are non-negative coefficients for the Proportional, Integral, and Derivative terms respectively.

The block diagram of the PID controller described with Equation 2.1 is shown in Figure 2.4a. The controller continuously calculates an error value $e(t) = r(t) - y(t)$ as the difference between a desired setpoint $SP = r(t)$ and a measured process variable $PV = y(t)$ and applies a correction based on proportional, integral, and derivative terms. The controller attempts to minimize the error $e(t)$ over time by adjustment of a control variable or controller output $u(t)$ to a new value determined by a weighted sum of the control terms.

The PID coefficients are tuned (or weighted) to adjust their effect on the process. Controller tuning allows for the optimization of a process and minimizes the error between



(a) Block diagram of a PID controller [226]

(b) Impact of signal noise [204]

Figure 2.4: PID controller and impact of a derivative component on controller output

the variable of the process and its set point. Along formal approaches for controller tuning such as the Ziegler-Nichols and Cohen-Coon methods [204], a "trial and error" approach is often used due to its simplicity. In this method, the proportional action is the main control, while the integral and derivative actions subsequently refine it. First, the proportional gain is increased until the output of the loop oscillates around the setpoint. Once P has been set to obtain a desired fast response, the integral term is increased to stop the oscillations.

Even though the full PID controller in theory delivers the most accurate controller output, most practical control systems use very small or zero derivative coefficients K_d. The derivative response is proportional to the rate of change of the process variable and causes the controller output to decrease if the process variable is increasing rapidly. Because of its inherent properties, the derivative response is highly sensitive to noise in the process variable signal as shown in Figure 2.4b. If the sensor feedback signal is noisy, the derivative response can make the control system unstable. For this reason, PI controllers are the most frequently used in the industry.

An industrial plant may have thousands of measurements and control loops. *Plantwide process control* involves systems and strategies required to control an entire plant consisting of many control loops and interconnected unit operations with an emphasis on structural decisions. The structural decisions include the selection/placement of actuators and measurements as well as the decomposition of the overall problem into smaller sub-problems [148, 133]. There are two main approaches to the problem: a *mathematically-oriented* approach and a *process-oriented* approach. It was acknowledged in [133] that the mathematically-oriented approach to plant-wide control configuration is difficult to implement practically, both because of the size of the problem and the large cost involved in making a precise problem definition, which would include a detailed dynamic and steady state modeling (see also Chapter 3). An alternative to the mathematical approach is a design procedure based on heuristic rules founded on experience and understanding of the process (process-oriented approach). One of the prominent heuristic approaches was proposed by Luyben et al. [148]. The approach is composed of nine steps and centers around the fundamental principles of plantwide control: energy management; production rate; product quality; operational, environmental and safety constraints; liquid-level and gas-pressure inventories; the makeup of reactants; component balances; and economic or process optimization.

While basic process controls are designed and built with the process itself to facilitate

fundamental operational requirements, Advanced Process Controls (APC) are typically added subsequently, often incrementally over the course of many years, to address particular performance or economic improvement opportunities in the process. Figure 2.5 illustrates the economic benefits of advanced process control.

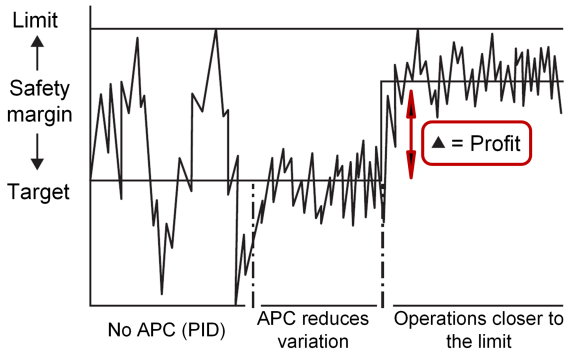


Figure 2.5: Benefits of Advanced Process Control [208]

One of the frequently employed APC methods is Model Predictive Control (MPC), a class of advanced process controllers capable of utilizing system information through a well-developed model and real-time process measurements to predict the future trajectory of the process. This is achieved through complex predictive modeling, which then allows the control system to take corrective action in advance, to ensure the process remains within the optimal trajectory in the future. A key feature of MPC is that future process behavior is predicted using a dynamic model and available measurements. Due to the requirement of having a high-precision process model, MPC is predominately used in control loops critical to plant economy and safety.

2.2 Industrial Control Systems Architecture

The first generation of process control systems used to be pneumatic and based on manually operated relay systems. Later control systems evolved into analog electronic systems which allowed to automate the majority of labor-intensive tasks. In the 1970s a new generation of computerized control systems emerged resulting in a new discipline called computer-aided manufacturing. Over the past decades, these systems evolved in parallel with general-purpose computing technologies to become a highly sophisticated distributed ecosystem of applications and hardware systems. Because computers allowed to automate manual tasks, modern process controls are also called *automation systems* or simply automation.

For many years control systems used to be air-gapped with process management being executed locally within the boundaries of individual production sites. With more data becoming available and increased complexity of equipment and process control approaches the need for interconnection of process control and enterprise (IT) networks as well as for third-party remote access has emerged. In the 1990s, T. J. Williams, a member of the Purdue University Consortium for Computer Integrated Manufacturing, published the Purdue Enterprise Reference Architecture (PERA) that provided guidance on the interface design between process control functions and enterprise functions [231]. On the premise of

the PERA model, the International Society of Automation (ISA) released a layered network model which described common vendor-independent information flows which can be applied in any manufacturing or process industry production architecture known also as the ISA-95 international standard [48] for operational technology environments (OT). It quickly became a de-facto standard for OT professionals to think about, design and implement industrial control systems within OT environments.

To address cyber security concerns and ensure secure integration of OT and IT environments ISA developed ISA-99 [50] which is currently known as the IEC-62443 series of standards on the cyber security of Industrial Automation and Control Systems (IACS) [49]. The collection of IEC-62443 standards and technical reports is organized into four general categories called General, Policies and Procedures, System and Component. Part IEC-62443-1-1 of the standard specifies how IACS assets should be organized into network layers based on device function and requirements as shown in Figure 2.6. The industrial network architecture also involves provisioning of the OT Demilitarized Zone (DMZ) zone to securely segregate the OT environment from the untrusted enterprise IT network and the Internet.

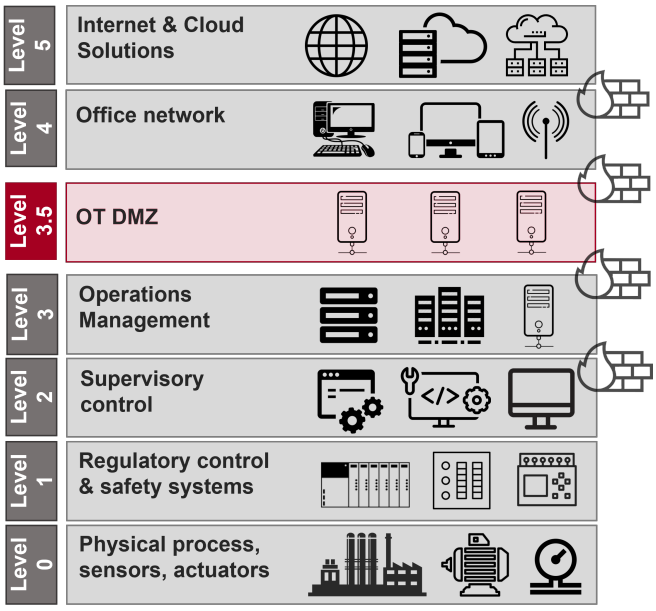


Figure 2.6: Purdue reference architecture for process automation

The Purdue model and the IEC-62443 reference architecture are often referenced synonymically due to their direct relation. They are regarded as industry best practices and are a widely adopted concept model for the logical network segmentation of industrial control systems. In the following we provide a summarization of each network level, including example devices and systems, and their functions. Note that the below descriptions are indicative. Real world infrastructures can deviate slightly or significantly. The largest deviations may begin from Layer 2 and layers above.

Level 0: Physical Process, Sensors and Actuators

The lowest level of the automation hierarchy includes the physical process and supporting physical infrastructure and defines activities involved in sensing and manipulating the physical processes. This level is also called *Field Level*. Correspondingly, sensors and actuators are often collectively called *field instrumentation*. Figure 2.7 gives a view of a physical level infrastructure from a water treatment environment.

Process equipment (static). This group of process equipment includes static process infrastructure such as pipes, various vessels and storage tanks, reactors, distillation columns, filtering and absorption equipment, ovens, etc.

Process equipment (dynamic). This group of equipment includes equipment which if actuated/powered. is capable of executing dynamic/kinetic behaviors. Examples include pumps, compressors, furnaces, fans, conveyors, robots, etc.

Sensing equipment. In broad terms, this group of equipment captures various parameters of the physical world. Examples include various types of sensors such as flow, temperature or pressure; proximity, fire and gas detectors; chemical/quality analyzers, etc.

Actuating equipment. An actuator converts a control signal into a mechanical action. Examples include valves, electric motors, steam or gas turbines, internal combustion engines, etc. Strictly speaking, although valves are frequently seen as monolithic elements, practically they consist of two elements – (1) a mechanical part of the valve, (2) devices designed to automatically control/actuate and monitor the position of valves in relation to their open or closed positions (valve actuators and positioners). Similarly, pumps are often seen as actuating equipment despite their actuating function being dependent on the accompanied motor.

In the distant past field instruments were $3-15$ *psi* "dumb" pneumatic-based devices that performed control via local single-loop controllers. Information was shown locally with gauges and recorded manually with pen and paper by a technician. In special cases, local chart recorders were used when data needed to be saved for analysis. A small degree of intelligence and ability to communicate with remote controllers was next added to field instruments in the form of a $4-20$ mA analog signal that represents 0 to 100% of process variable or controller output. Currents less than 3.8 mA or more than 20.5 mA are seen

Figure 2.7: Physical process infrastructure and field equipment [180]

as indicators of a fault and with that provide rudimentary diagnostic functionality. With the introduction of the HART protocol, it became possible to transmit some data and settings over the $4-20$ mA analog signal with up to 1200 *bps* speed in half-duplex mode. A major step forward occurred when microprocessors became small and robust enough for installation in field instruments, enabling local conversion of an analog signal into its digital representation suitable for transmission over a network. Added computing capabilities allowed instruments to manage other tasks such as calibration and complex self-diagnostics and with that transformed them into *smart instruments*.

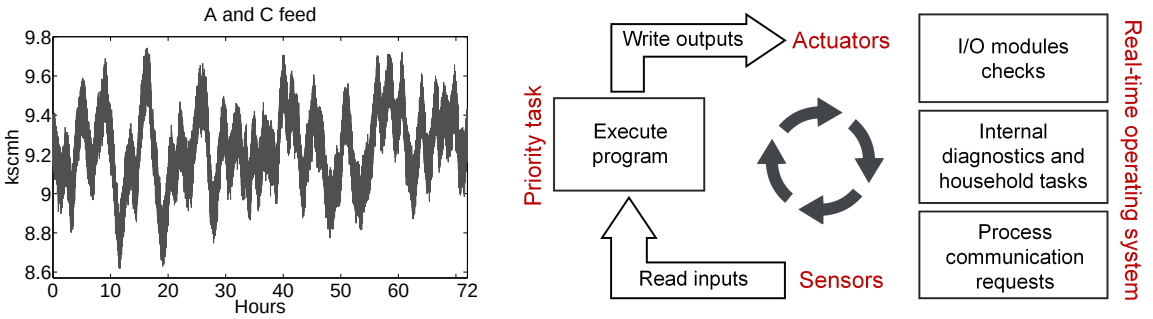
Modern smart sensors (also called transmitters) can send back multiple readings along with alarm conditions. Valves now have computing and data storage capabilities to calculate and retain in the local data history a current valve signature of pressure vs. stem travel, compared with the signature when the valve was installed, and provide diagnostic information or alarms based on the detected differences. Despite the popularity of smart sensors, cheap simpler sensors are still often used for monitoring major equipment. For instance, a standard electric motor may now have over a dozen sensors providing real-time measurements to determine the health of equipment. This data is sent to an upper layer of the control network architecture where it is fed into an application for predictive maintenance and other purposes. Note that the utilization of a large number of sensors increases the required investment and maintenance costs, and optimal sensor placement is an active research area.

As the extent of field data consumption increased, field data started to be called telemetry data. Telemetry engineers are responsible for the accuracy of data during their acquisition in the field, in transit and in storage. It is worth mentioning that raw sensory data rarely can be used directly. The electrical output of a sensing element is usually small in value and has non-idealities such as offset, sensitivity errors, non-linearities, noise, etc. Therefore, the raw sensor or, more precisely, transducer output is subjected to signal conditioning such as amplification, filtering, range matching, etc. Acquired signal conditioning may happen directly in the sensor or in external dedicated signal processing modules. The example of a conditioned and digitized signal is shown in Figure 2.8a. Sensor signals are combined together, and aggregated with other sensor signals to extract additional new information about process performance and the state of the plant. To drive efficiencies, sensors are increasingly communicating directly with their vendor monitoring software in the cloud via cellular networks.

Level 1: Regulatory Control

The level immediately adjacent to Field Level hosts controlling equipment responsible for monitoring operating parameters and controlling functions of the physical process and equipment. This level is called Regulatory or Basic Control. The types of control equipment include PLCs, Variable Frequency Drives (VFD), dedicated PID controllers, Remote Terminal Units (RTU), etc. Controllers interface field instruments via analog or digital Input/Output (I/O) modules. We will use the example of a Programmable Logic Controller to describe the major concepts related to regulatory control.

The main workflow of the programmable controller consists of two parts: the process of self-diagnosis and communication response and the execution process of the *user program* also called *control logic* as shown in Figure 2.8b. This workflow is repeated several times a second and is called *scan cycle*. In relation to process control, the scan cycle consists of



(a) Sensor signal

(b) PLC scan cycle

Figure 2.8: Sensor signal and PLC scan cycle

three steps: reading inputs (process variables), solving control logic and writing outputs (manipulated variables). The PLC queries each input card and saves the readings in a variable table (VT). VT contains all the variables needed by the control logic: setpoints, counters, timers, inputs and outputs. During the program scan cycle, every change in the I/O of the PLC is ignored until the next program scan cycle. The PLC executes the user program one instruction at a time and when the program execution completes, the outputs are updated using the temporary values from VT. The PLC updates the status of the outputs and restarts the scan cycle by starting a self-check for faults. The PLC scan cycle is an infinite loop that runs until reboot or power off.

Execution of the scan cycle takes some amount of time. This time is called the *scan time* and often amounts to milliseconds (ms). Usually, industrial environments mandate real-time control over an industrial process. Failing to execute control operations in a timely manner may result in the failure of an industrial process, which leads to unacceptable consequences. To overcome this problem, the majority of PLCs are equipped with Real Time Operating Systems (RTOS) to execute their tasks in a predictable manner and within strict time constraints. User program or control logic, as the name suggests, defines what operations should happen, when and under which conditions. It also contains so called interlocks that define mutually exclusive conditions to prevent undesired (harmful) states of the process. Industrial controllers are programmed by control engineers. The code can be developed in-house but most frequently it is developed by external engineering companies (third-party subcontractors). The source codes of control logic are stored on the Engineering Workstation and/or on data servers or data shares. The IEC standard 61131-3:2013 [46] outlines five PLC programming languages. These are split into two categories: (1) Graphical (Ladder Diagram, Function Block Diagram, and Sequential Function Chart) and (2) Text Based (Instruction List, Structured Text). Although these languages are vendor and application agnostic, vendor specific language subsets are common as well. An example of a ladder diagram with a function block for a Siemens PLC is shown in Figure 2.9. The control logic is compiled with vendor-proprietary compilers and uploaded to a PLC for execution in the form of bytecode or binary code.

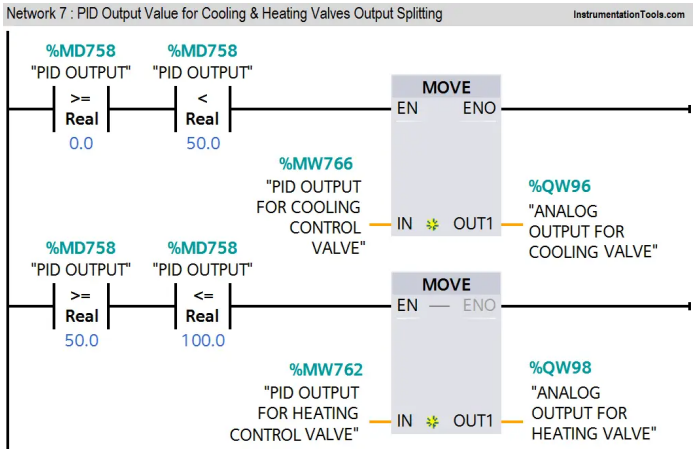


Figure 2.9: Example of a ladder diagram [219]

One unique characteristic seen across industrial control systems is the use of *points* – a concept not applied in conventional IT systems. Points are responsible for all aspects of an ICS, essentially denoting a data source or a controllable function. Each point is assigned a unique ID, e.g., I1.1, DB1.DBX1.1, etc. ID values can be input channels, memory locations or other process specific variables. Point IDs are supplemented with more logical "friendly" names (e.g., Valve 1) called tags. Tag names are descriptive and given to points for operator convenience. In large facilities tag naming/numbering typically follows a site-specific nomenclature in the form of a "code" (e.g., TMHS KQP536NR) and tag description is provided as a separate attribute. Tags typically encode some meaningful data like plant area code, equipment type, equipment tag, unique sequence number and similar. Points can be soft or hard. A hard point denotes a physical input or output to/from sensors and actuators. A soft point denotes results derived from mathematical calculations and control logic actions. From a software perspective, a tag is considered much the same as a variable name. In essence, points and their associated tags link external devices and process parameters with process control applications and can have either Read, Write, or Read/Write permissions. Figure 2.10 shows an example of a tag table in a PLC.

Default tag table

	Name	Data type	Address	Retain	Visible..	Acces...	Comment
1	Emerg-OFF	Bool	%I1.0	☐	☑	☑	Emergency-OFF (nc contact)
2	S3	Bool	%M0.3	☐	☑	☑	pushbutton START S3 (no contact)
3	B1	Bool	%I0.1	☐	☑	☑	sensor safety fence closed (no contact)
4	B2	Bool	%I0.2	☐	☑	☑	sensor cylinder A moved out (no contact)
5	M0	Bool	%Q0.0	☐	☑	☑	move out cylinder A
6	S1	Bool	%M0.1	☐	☑	☑	pushbutton manual mode S1 (no contact)
7	S2	Bool	%M0.2	☐	☑	☑	pushbutton automatic mode S2 (no conta..
8	S4	Bool	%M0.5	☐	☑	☑	pushbutton ON S4 (no contact)
9	S5	Bool	%I0.5	☐	☑	☑	pushbutton OFF S5 (no contact)
10	Motor1	Bool	%Q0.2	☐	☑	☑	motor conveyor belt M01
11	B0	Bool	%I0.3	☐	☑	☑	sensor bottle counting
12	S6	Bool	%I0.6	☐	☑	☑	reset counter / new box

Figure 2.10: Example of a tag table in a PLC [6]

Controllers communicate with field instruments and with each other via real-time communications protocols. Many industrial protocols were initially designed to communicate serially over RS-232/485 physical connections at low speeds, but have since evolved to operate over Ethernet networks using routable protocols such as TCP/IP and UDP/IP. A good example of this evolution would be the Modbus protocol which evolved from its serial version Modbus RTU to the Ethernet version called Modbus TCP. For many years Modbus held the title of the most widely used network protocol in industrial applications. With time a variety of proprietary and functionally richer protocols such as EtherNet/IP, PROFINET and EtherCAT began to prevail.

An interlock functionality mentioned above can be also seen as a basic safety feature that makes the state of two mechanisms or functions mutually dependent. In most applications, an interlock is used to help prevent a machine from harming its operator or damaging itself by preventing one element from changing state due to the state of another element, and vice versa. For instance, the pump should not run if there is no flow or if lubricating material is not supplied. Interlocks can be not only programmable but also consist of any electrical, electronic, or mechanical devices or systems such as the Safety Instrumented Systems discussed below.

Safety Systems

Besides basic process control system (BPCS) Levels 0-1 may also include an independent control system to carry out safety functions. Safety Instrumented System (SIS), also sometimes called Emergency Shutdown Systems (ESD), provides dedicated fail-safe operations of high reliability. It is used to protect against catastrophic consequences such as equipment damage, severe environmental impact and personal injuries/casualties by shutting down the production application. SIS/ESD is part of the Independent Protection Layers (IPL) to handle safety incidents as shown in Figure 2.11.

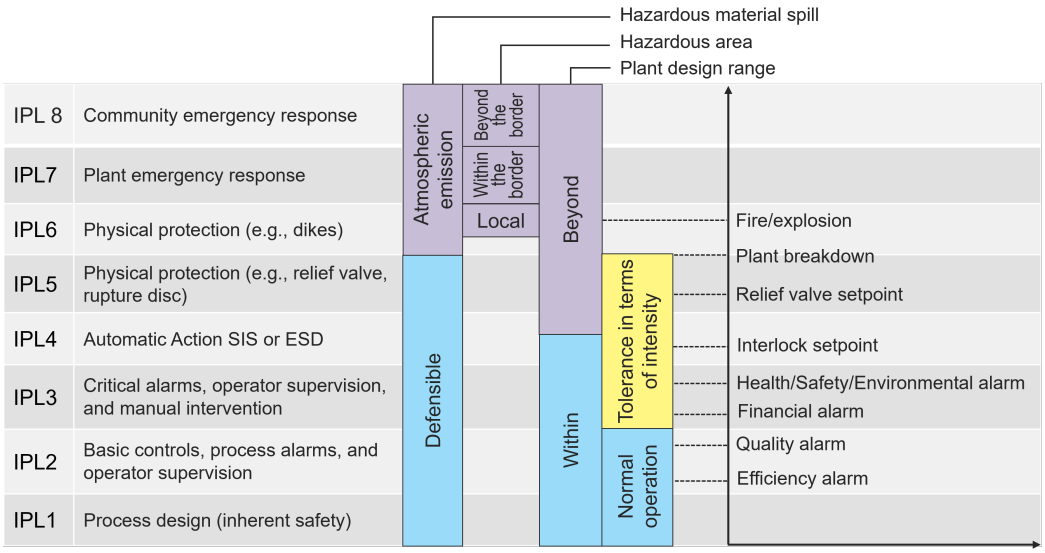


Figure 2.11: Independent Protection Layers as specified in IEC 61508/IEC 61511 [94]

SIS (IPL4) typically consists of sensors and logic functions that detect dangerous conditions and final elements, such as valves, that are operated to achieve a safe state. SIS is categorized by a Safety Integrity Level (SIL) – a performance measure that is related to the probability that the safety instrumented function will not work when challenged (when needed). The required SIL may be determined from Hazard and Operability Studies (HAZOP), Failure Mode and Effect Analysis (FMEA), Layers of Protection Analysis (LOPA), risk graphs, and other methods. The higher the severity of hazards, the higher the SIL level of the safety controller must be.

Historically, safety systems were hardwired and segmented off from the main control system. In recent years SIS evolved to the point where they now include programmable electronic technology and an IP stack, which allows for over-the-network changes to the functionality of these systems. To maintain the high reliability of an SIS, it is advisable to segregate BPCS and SIS functions with the needed communication conduit being tightly configured. As a best practice, the data from safety systems is first communicated to a regulatory controller via a point-to-point communication link (Modbus RTU or Modbus TCP) and then forwarded to the systems in the upper layers of the control architecture. Some vendors offer Integrated Control and Safety System (ICSS) solutions that combine elements of process control and functional safety into a single architecture. Usage of such architectures is typically driven by the cost reduction of plant engineering and subsequent ease of maintenance through integrated tools. A comprehensive overview of the SIS architectures can be found in [113].

Level 2: Supervisory control

Individual controllers contain only a small portion of control logic and therefore can control the process to achieve a local optimum at most. As the title suggests, this level is involved in supervising the control function over the whole plant, typically done by human operators and engineers. The main component of this level is the control room (or control center) where operators monitor and control the plant through Human Machine Interface (HMI) software which contains a graphical representation of the process as shown in Figure 2.12. If several different types of control systems are used in the plant, their data are jointly displayed in the control room. The supervisory control function is performed 24/7 in shifts. Whenever a process deviates too much and requires correction, operators may change a setpoint, manually operate actuators and even overwrite interlock conditions.

HMI display has a layered design, each layer/view representing a different level of detail about plant state. What is displayed in each level is, in general, plant (customer) specific, however, there is a general guidance [56]:

- Level 1 – Plant overview;
- Level 2 – Unit overview;
- Level 3 – Equipment overview;
- Level 4 – Trends/Elements of control logic.

Unit overview and process variables trends are among main displays viewed by the operators as they provide the most relevant and timely (near real-time) information about plant state. During the process upset the operators may switch between different displays to determine root cause of the disturbance and collaborate with the operators who monitor adjacent plant units as well as plant maintenance personnel on the shop floor.

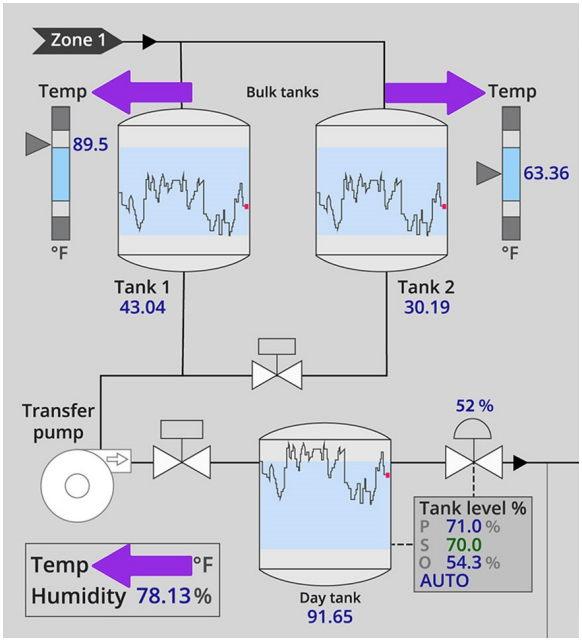


Figure 2.12: Example of a Human Machine Interface view [65]

The operators are notified of abnormal process conditions or equipment malfunctions that require actions to be taken by an application called an alarm system. Alarm management is the most critical function of the supervisory control as the potential consequences of an unresolved abnormal situation can be dramatic, ranging from significant financial losses to casualties as shown in Figure 2.13.

		Available Response Time	PRIORITY CLASS				
Response Class	SHORT	< 5min	L	M	E	E	E
	MEDIUM	5 – 15 min	L	M	M	E	E
	LONG	> 15 min	L	L	M	M	E
Consequence Category	ECONOMICS		No/Slight Effect (< 10k)	Minor Effect (10k – 100k)	Medium Effect (100k – 1M)	Major Effect (1M – 10M)	Extensive (>10M)
	HEALTH & SAFETY		No/Slight Injury	Minor Injury	Major Injury	Single Fatality	Multiple Fatalities
	ENVIRONMENT		No/Slight Effect	Minor Effect	Local Effect	Major Effect	Massive
CONSEQUENCE CLASS			Negligible	Low	Medium	High	Extreme

Figure 2.13: Alarm criticality and consequences cost (the source is intentionally withhold)

The alarm systems can include both the basic process control system and the safety instrumented system, each of which uses measurements of process conditions and logic to generate alarms. Packaged systems such as fire and gas systems or complete packaged units are typically included into overall alarm management systems. Figure 2.14 illustrates the process of alarm generation, visualization and logging in the alarm system.

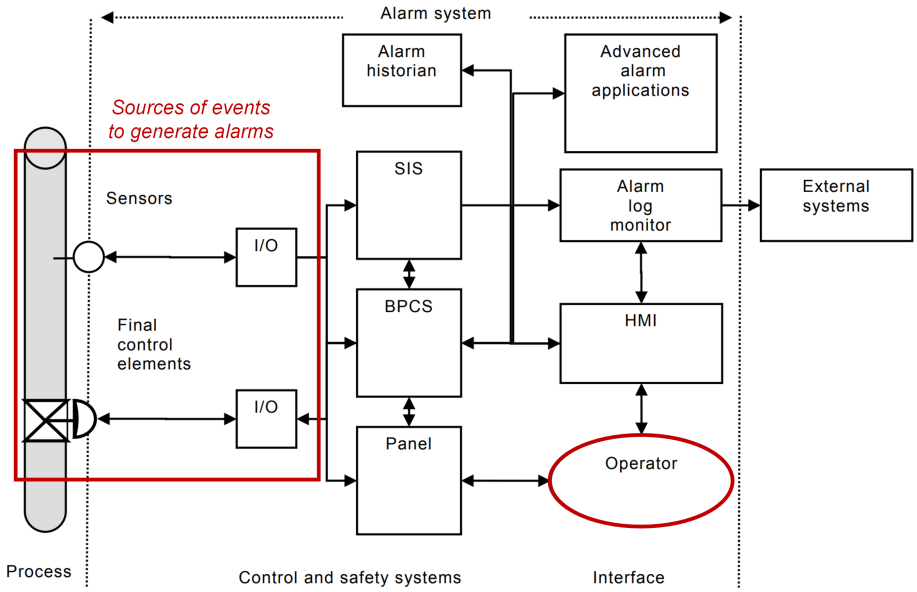


Figure 2.14: Alarm system generation and response workflow [47]

Most frequently alarm information is communicated to the operator via an HMI, usually a computer screen or an annunciator panel. The number of alarms to be triaged per period of time is strictly regulated to maintain appropriately sufficient cognitive function by the operators. In process industries, the acceptable alarm rate is equal to $6-7$ alarms per hour or $1-2$ per 10 min as shown in Table 2.1, with critical alarms not exceeding 10% of the total number of alarms [89].

Average Alarm Rate	Acceptability	Performance and risk
< 10	Very likely unacceptable	Inefficient / High Risk
5 – 10	Likely over-demanding	Medium performance & risk
2 – 5	Possibly over-demanding	
1 – 2	Manageable	
> 1	Very likely acceptable	Efficient / Low Risk

Table 2.1: Impact of alarm rate on operator acceptability and performance, per 10 min in steady state, based on [89]

The other systems found at the supervisory level include typical components of a Distributed Control System: various applications for configuration, maintenance and diagnostics of the control system, control logic and equipment, databases, communication and batch management servers, engineering desktops and others. A key element in the control system infrastructure is an Engineering Workstation (EWS), a high-reliability and sometimes ruggedized computing platform equipped with software applications and tools to program and modify the industrial process. It includes specialized tools needed to directly communicate with, configure, and update the primary control equipment such as PLC, BPCS, SIS, intelligent electronic devices (IED) and others. Due to the nature of its tasks,

an EWS contains significant amounts of sensitive documentation specific to the process and infrastructure design, configuration, and plant operations. It is also frequently used to store historic information related to process configuration changes. Another key element of the control system is a purpose-built real-time database (RTDB) for short-term storage of process data to be consumed by other applications. The process RTDB may also include a local instance of the data historian application, specialized software that collects point values, alarm events, batch records, and other information from industrial devices and systems for long-term storage. Additional systems may include asset management solutions, local backup servers and any other system which is required to support plant management operations.

Level 3: Operations Management

At this level reside systems that support functions involved in managing the workflows to produce the desired end products. Examples include Manufacturing Execution Systems (MES) and Manufacturing Operations Management Systems (MOMS). These systems take orders and business data from the corporate systems to manage production scheduling and dependencies, ensure production capacity optimization and schedule dispatching of the finished goods. This level is also sometimes called "shared services" because it hosts applications and services jointly used by multiple plants such as centralized backup and asset management systems, network management functions, file servers and others.

Among the pillar systems hosted at this level is the Plant Information Management System (PIMS) also known as a historian. Figure 2.15 shows the Graphical User Interface of one of the most widely used historian software, OSIsoft PI.

Figure 2.15: Data historian view [128]

A PIMS collects and integrates information about production processes and infrastructures from a multitude of different sources, including local historian instances from Level 2. Historians use specialized algorithms for lossy and lossless data compression to reduce storage footprint as well as specialized software to enrich data with context and visualize statistical trends.

It might not always be obvious whether a certain system belongs to Level 2 or Level 3. In general, no system or application which is directly involved in or depended upon for process control should be hosted at this level. In other words in case of issues with the L3 infrastructure, it should be possible to disconnect this network segment from the rest of the OT infrastructure without immediately affecting the process control function. For this reason an essential service such as historical storage is provisioned at both L2 and L3. However, because L3 hosts such services as scheduling systems and material flow applications, disruptions at the Operations Management level can lead to hours or days of interferences in production and enterprise processes, with the potential for considerable revenue loss. In some instances, it might be required to shutdown the manufacturing processes.

OT Zone

Levels 0-3 of the network reference architecture are collectively called *OT zone* or *Industrial zone*. The rise of automation leading to higher efficiencies has created an increased need for bidirectional data flows between OT and IT systems. The OT zone is required to be protected from Internet-enabled networks such as the enterprise layer and any other "untrusted" environments like third-parties. This is achieved through the provisioning of a dedicated perimeter network also known as a Demilitarized Zone which acts as a conduit system between the OT zone and other external environments.

Level 3.5: OT DMZ

OT DMZ is a recent addition to the Purdue model. Similar to a conventional IT DMZ, the OT DMZ is a buffer zone where services and data can be shared between two networks belonging to different trust zones. OT DMZ is the only non-functional level and provides an interface where the IT and OT worlds *"converge"*. For this reason, the OT DMZ level is frequently counted as 3.5 to highlight its extension to L3. Among the key functions of the DMZ is to enable communication between OT and enterprise applications in order to exchange manufacturing and resource data. For this purpose, the historian from L3 is fully or partially mirrored to a historian instance at the DMZ. The DMZ typically also hosts patch and antivirus servers as well as applications for secure file exchange and remote access to OT assets for third parties.

When correctly implemented, no communication connection should directly traverse the OT DMZ, all traffic should originate/terminate in the DMZ. For instance, historian data from L3 are first sent to the historian instance in the DMZ and then sent to L4 for consumption by business applications. Similarly, all communication flows from the IT network or Internet first land in the DMZ and are then routed further via newly established sessions.

Level 4: Enterprise network

Level 4 or the Enterprise level, is where IT systems, applications and functions exist tasked with the management of the entire industrial systems estate and business processes. This is the level that provides business direction and orchestrates operations with data flowing up from the shop floor (L1) and decisions flowing down from the boardroom. Examples include Enterprise Resource Planning (ERP) systems that drive long-term plant production schedules, material use, shipping, inventory levels and logistics processes, business-to-business and business-to-customer applications and services as well as various business applications such as emailing and voice communication, document management and data centers. This level also hosts various supporting functions such as advanced data analytics and associated applications, Research and Development (R&D), Security Operations Center (SOC) and others.

Level 5: Cloud applications

Level 5 or Cloud level does not officially exist in the Purdue or IEC 62443 reference architecture. However, we added it to illustrate two ongoing trends:

- Increasing desire of enterprises to reduce on-premise hardware footprint and reduce infrastructure maintenance cost. Various Software-as-a-Service (SaaS) applications such as Office 365 and modern endpoint security solutions are examples of cloud-hosted business applications. In this regard, Level 5 could be seen as an extension of the office network;
- Digitalization of industrial automation to achieve the next level of efficiency and operational excellence. This trend is also widely known as Industry 4.0 [174], where production and asset data are being directly fed into so-called data lakes and cloud-hosted applications for analysis and gathering insights.

In the past, collected data were stored in on-premise servers or data centers and consumed locally for local decision-making. Later data from various sites started to be aggregated at the enterprise level for enterprise-wide usage and decision making. With the wider adoption of cloud technologies and the evolvement of advanced machine learning techniques for big data analytics, data storage increasingly began to move to data lakes for ease of data sharing and simplified management of data access by the enterprise users, partners, third-party service providers and customers.

It would not be possible to cover all the benefits of industrial digitalization. Among the most mature and widely adopted examples of cloud-based applications is Predictive Maintenance [177]. This is a proactive maintenance strategy that utilizes condition monitoring applications to detect various deterioration signs, anomalies, and equipment performance issues to estimate when a piece of equipment might fail so that maintenance work can be performed ahead of a potential breakdown. For this purpose complex pieces of equipment are instrumented with additional sensors to gather additional asset data as shown in Figure 2.16. The instrumentation of assets with additional sensors is sometimes called asset digitization. Asset data is then used to create predictive models to detect signs of asset degradation.

Among the most recent advances in the utilization of asset data is Augmented Reality (AR) applications where asset maintenance personnel can visualize various asset status data directly in the field with the help of tablets either in the form of 2D and 3D asset

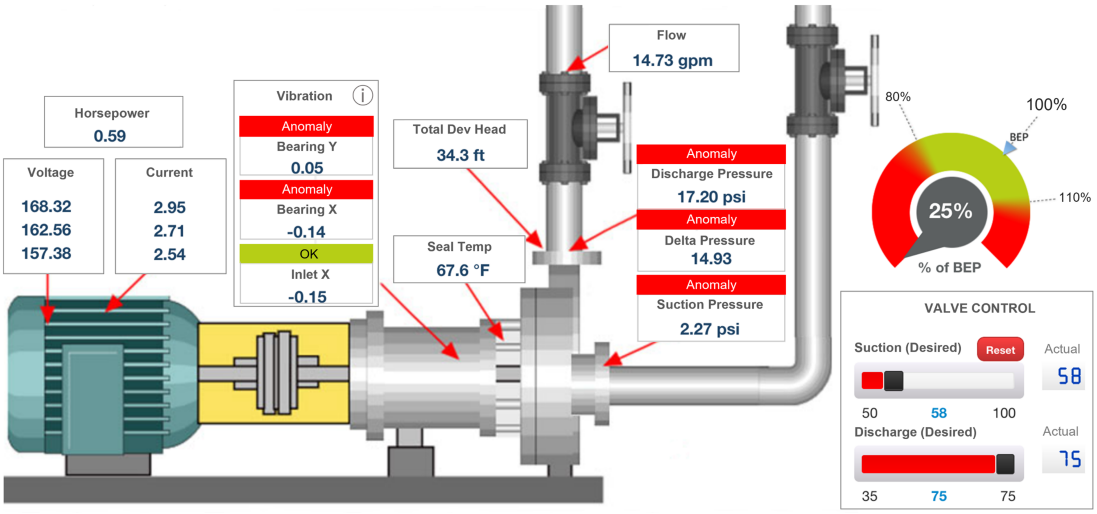


Figure 2.16: Pump monitoring solution by Flowserve [119]

models or by superimposing live photo or video and asset information. An overview of the visualization methods in industrial AR prototypes can be found in [81]. Among emerging trends are Cloud-Based Design (CBD) and Cloud-Based Manufacturing (CBM) which refer to a distributed collaborative cloud-based engineering design services and manufacturing model that exploits on-demand access to a shared collection of distributed manufacturing resources to form temporary, reconfigurable production lines in response to demand [234].

Enterprise Zone

Levels 4-5 collectively form a so-called *Enterprise Zone*. Note, that this zone contains both – systems and applications which belong to OT infrastructure and pure office/enterprise IT systems. Some infrastructure can be of double use with examples being SOC/NOC (Network Operations Center). Despite being located in the same trust zone, usage of shared infrastructure, e.g., hosting of the OT and IT applications on the same servers should be avoided. Furthermore, where possible, it is recommended to implement network segmentation between OT and IT systems. Similar considerations apply to cloud environments.

Transformation of the Purdue Model

Asset monitoring and plant optimization applications with data flows utilizing Internet-based communication and being consumed by cloud-based software platforms have received the name of Industrial Internet of Things (IIoT) [24]. With the added sensors for asset and process monitoring, the network bandwidth of the BPCS became insufficient to transfer large volumes of high-resolution telemetry data. As a consequence, it was decided to send IIoT data via dedicated communication channels which included edge processing devices and communication gateways. Naturally, the Purdue reference architecture does not provide any guidance neither about how such IIoT communication architecture should be implemented

nor how it should be integrated with the main OT infrastructure. To standardize IIoT implementation, NAMUR, an International User Association of Automation Technology in Process Industries, came up with the NAMUR Open Architecture (NOA) which provides guidance on the architectural design of IIoT data flows and intercommunication between IIoT and the core process control infrastructure. Figure 2.17 shows such a design for L0-L3/L4 for the reference architecture.

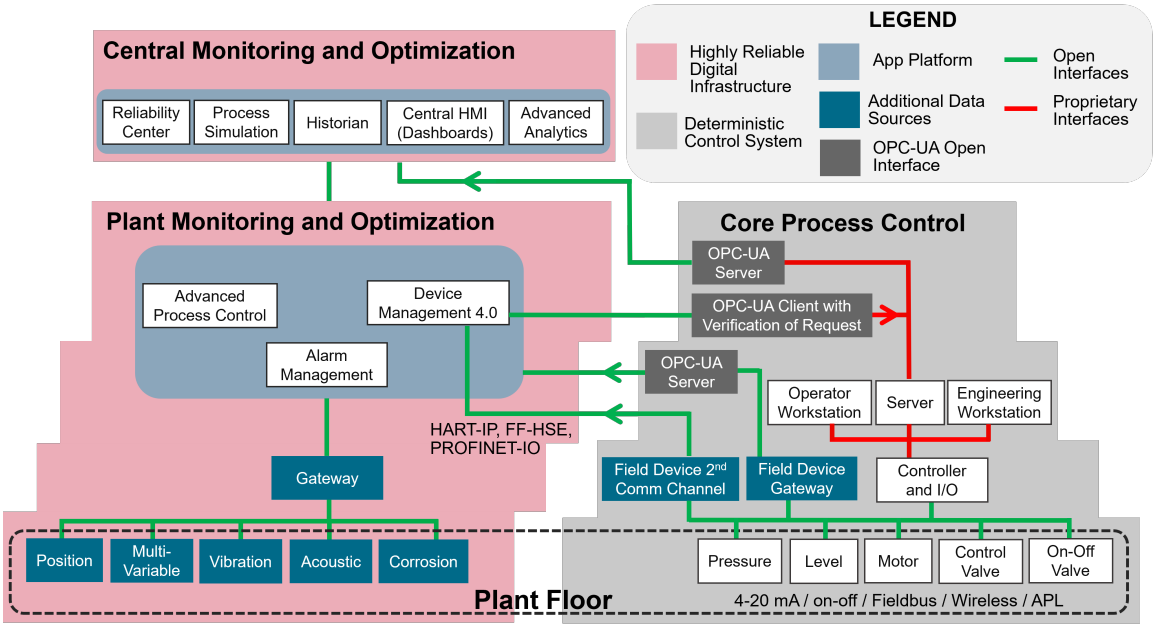


Figure 2.17: Open NAMUR reference architecture for IIoT applications [17]

In the future, the current Purdue model will continue to be challenged. With the industry's growing desire to reduce infrastructure deployment and maintenance costs, L3 and L2 computing resources are increasingly hosted in the cloud in the form of Infrastructure-as-a-Service (IaaS) [44]. This trend will inevitably compel the deployment of a DMZ infrastructure in the cloud as well, calling for rethinking security requirements and the architectural design of the Purdue reference model. In the future, with L3 and L2 applications being consumed as SaaS offerings, the Purdue architecture will transform even further.

Data processing in industrial automation

With the introduction of computer-aided manufacturing, *data* became the *most essential ingredient* of automation and processing this data into information became a substantial task of control systems.

The key to handling information was the establishment of a *transparent data flow* inside an automation system *with a strict subdivision of the data processing into a hierarchical model* which resulted in an architecture known as automation pyramid [48] and shown in Figure 2.18. Established by the ISA-95 work group, *this model became a foundation for the Purdue reference architecture* discussed above. Allocation of devices to automation

levels highlights the varying time constraints associated with data collection and response. The direction of the process data flow is bottom-up, while management and control data flows from top to bottom, with each layer adding latency, predominately related to data aggregation and analysis strategies.

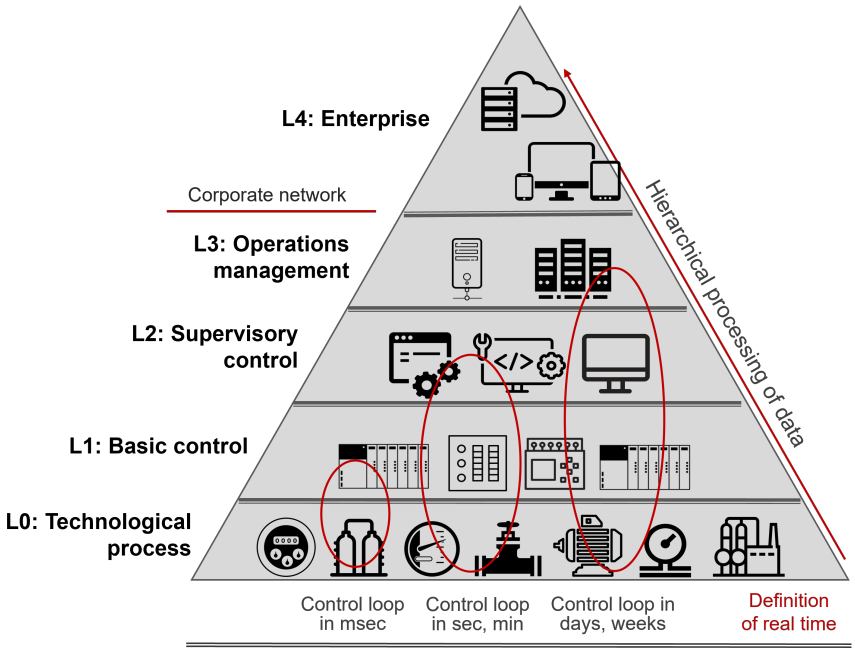


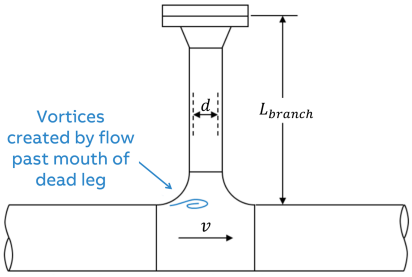
Figure 2.18: Automation pyramid according to ISA-95

At the level of regulatory/local control (L1) data has momentary significance and is used merely for real-time process corrections. Data exchange between field instrumentation and control equipment is happening in milliseconds or seconds intervals and data processing mostly addresses the quality of the sensed signals. While industrial controllers mostly operate on *raw data*, the systems in the upper layers of the automation pyramid operate on the *information* extracted from the data. Field data are contextualized for extracting useful observations by means of various data processing and analysis techniques such as filtering, conditioning, aggregation, transformation, correlation and others. Supervisory control is about short-term tactical control by process operators and medium-term strategic monitoring by process supervisors and process managers. Process information is used for displaying trends, alarm management and overall situational awareness with data significance lasting in sec-min for operators and min-hrs for supervisors. Information at the level of operations management is concerned with correcting plant performance drifts and strategic long-term planning.

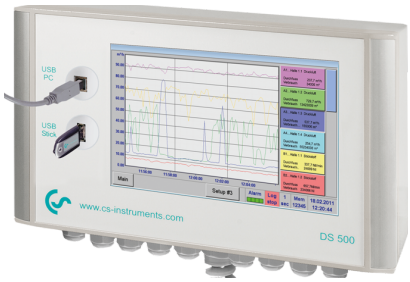
> Incorrect processing of process data may result in partial or complete loss of situational awareness.

In his public blogpost, Weiss [228] described a real-world use-case in which two *identical* plants reported to the plant vendor that one plant experienced a flow-induced vibration issue

and was operated at a reduced load, while the other plant did not have a vibration problem and was operated at full power. The latter plant ultimately reported significant damage to the nuclear fuel system caused by vibration. It was discovered during the investigation that both plants had flow-induced vibration issues (Figure 2.19a). The indicator of this physical problem was a resonant frequency detectable by the plants' in-core instrumentation recorders. However, one chart recorder (Figure 2.19b) reported the problem while the chart recorder in the other plant had the sensor signal filter altered to eliminate the higher frequency noise in the signal that was indicative of the flow-induced vibration. This resulted in the "filtered" plant operating in unsafe conditions, without the operators' awareness.



(a) Flow-induced vibration illustration [98]



(b) Example of chart recorder equipment [110]

Figure 2.19: Flow-induced vibration in piping systems and its detection

Such an erroneous change of sensor signal processing parameters resulted in unintentional data utility corruption. Data utility differs from data integrity in the sense that the data itself is not manipulated but the way of extracting information from the data is altered. The sensory unit and the sensor signal were authentic, change of the filtering parameter was done by the authorized personnel. However, the filtered content of the sensor signal no longer provided one of its intended functions, i.e., vibration monitoring.

> **Appendix A** provides an extended discussion on data security properties specific to cyber-physical systems.

In the past reliance of industrial equipment and software on proprietary protocols posed significant challenges to centralized acquisition of process data. To bridge the "communication gap", an industrial automation task force consisting of industrial vendors and end users developed an OPC standard for communication interoperability in 1996 [75]. OPC stands for Open Platform Communications or OLE (Object Linking and Embedding) for Process Control in the initial version of the protocol name. OPC serves as a bridge between both Windows-based software applications such as historians and various process control hardware as well as control equipment of different vendors. Figure 2.20 illustrates the OPC communication topology for data acquisition. To date, OPC has become one of the most widely used interfaces for accessing process data not only in OT but also in IIoT applications and allows building multi-vendor architectures in a "plug and play" fashion.

OPC and its latest version OPC UA (United Architecture) is not simply a protocol but a transport-agnostic industrial framework that is adaptable to different transport layers

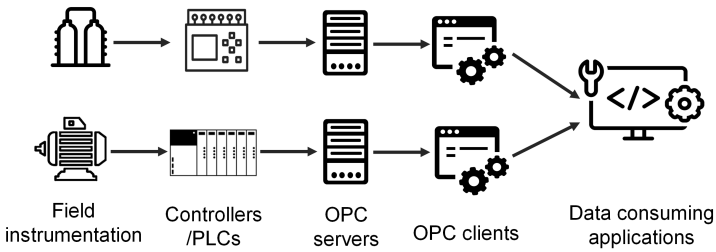


Figure 2.20: OPC communication topology

depending on the application-specific requirements and use cases. OPC UA makes use of a universal Quality-of-Service (QoS) concept, which includes real-time communication capabilities with guaranteed bandwidth and low latencies. The Field Level Communications (FLC) Initiative of the OPC Foundation developed OPC UA FX (Field eXchange) extension specification to enable uniform and consistent communication solution for vertical and horizontal integration, including field, edge and cloud as shown in Figure 2.21.

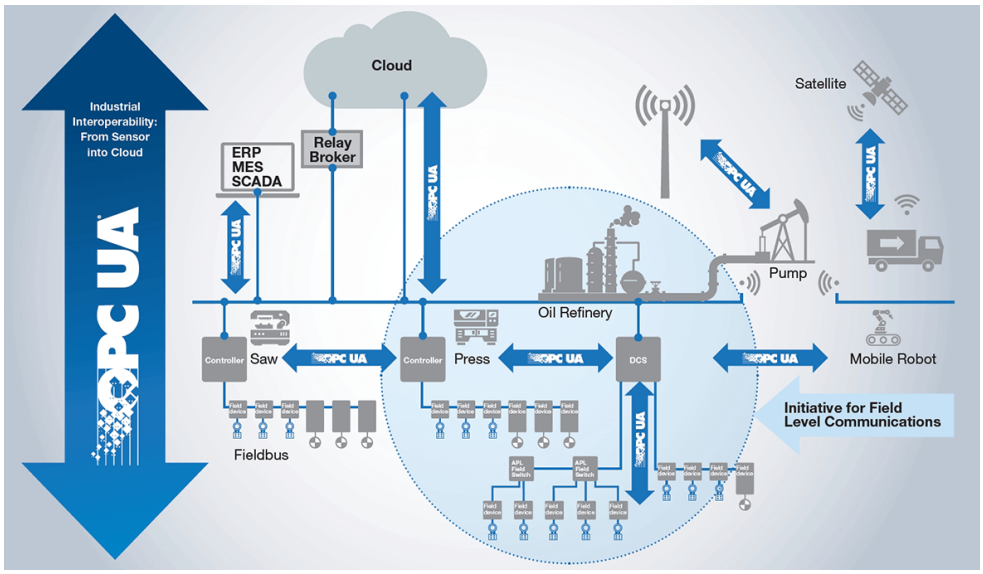


Figure 2.21: OPC UA framework with extensions for field exchange [145]

2.3 Cyber-Physical Attacks

ICS have traditionally been designed for dependability, durability, economic efficiency and safe use. Pervasive computerization and automation of control systems has enabled vertical and horizontal systems integration and introduced cyber interdependencies [183], which became a source of disturbances that are unusual and difficult to foresee. Subtle control loops, cascading failures and malware propagation were the price for increased

efficiency. Technology has improved efficiency but at the same time has become a source of concern. In the past few decades, plants have undergone tremendous modernization. What used to be a panel of relays became an embedded computer. What used to be a simple analog sensor is now an IP-enabled smart transmitter [158] with multiple wired and wireless communication modes, a large number of configuration possibilities, and even a web server so that maintenance staff can calibrate and setup the device without approaching it. While security engineers try to limit the attack surface, vendors keep introducing novel opportunities for remote exploitation of the physical processes and equipment.

Pervasive "cyberfication" became a source of concerns about plant vulnerability to both random cyber failures and security attacks. On one hand, embedded computers have enabled the governing of physical applications to achieve desired outcomes. On the other hand, physical systems can be instructed in the same way to perform actions that are not intended. As a result software code which does not inherently possess tangible force can potentially acquire destructive capacity through the ability to instruct physical systems to malfunction. Cyber attacks on physical systems are correspondingly called *cyber-physical attacks*. The implications of this class of cyber attacks, namely the ability to inflict physical damage, is the main difference between cyber-physical and conventional cyber attacks.

> What is not always understood is that breaking into a cyber-physical system and taking over its component(s) is not enough to achieve the desired physical impact. After all, breaking into a system is not the same as breaking a system.

Executing a successful cyber-physical attack requires a body of knowledge that is different from what is commonly employed in the information technology field. In particular, the attacker needs to understand signal processing [223], control principles [204], the physics of process behavior [148], the mechanics and failure conditions of the equipment [130, 205, 27], etc. Moreover, different types of CPS are subjected to fundamentally dissimilar failure modes the conditions of which first need to be discovered [130]. Figure 2.22 shows the distinction between the constituents of a cyber-physical attack.

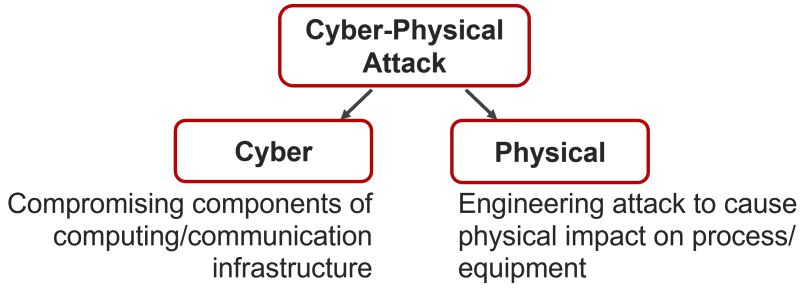


Figure 2.22: Multidisciplinary components of a cyber-physical attack

In the following sections, we describe some intrinsic aspects of cyber-physical exploitation as well as the current threat landscape and the state of threat actors abilities to execute cyber-physical attacks.

2.3.1 Timing Parameter

Cyber attacks in the IT domain do not generally depend on timing aspects. The few exceptions include attack categories such as race conditions, Time-of-Check to Time-of-Use vulnerabilities, or cross-site scripting attacks that rely on getting access to session cookies before they expire, where attackers need to make sure that their attacks occur within a tight window of time. In cyber-physical systems, however, timing plays a major role as the physical state of the system evolves continuously over time and some states might be more vulnerable to attacks than others. Timing also characterizes the sensitivity of a cyber-physical system, e.g., it may take minutes for a chemical reactor to burst [192], hours to heat a water tank or burn out a motor, and days to destroy centrifuges [129]. Understanding the timing parameters of the physical processes not only allows an attacker to construct a successful attack but also to maximize the desired impact (e.g., long-term damage to the system).

The dynamic evolution of process variables can be described with a simple model consisting of process gain, dead time, and time constant. The process gain describes how much the process will respond to a change in controller output, while dead time and time constant describe how quickly the process will respond (Figure 2.23). Precisely, dead time describes how long it takes before a process begins to respond to a change in controller output and the time constant describes how fast the process responds once it has begun changing. Controlling the processes with large time constants is a challenging task causing operator stress and fatigue [151]. The described timing parameters are not only important for the design of a control algorithm but also for the attacker to design an effective attack.

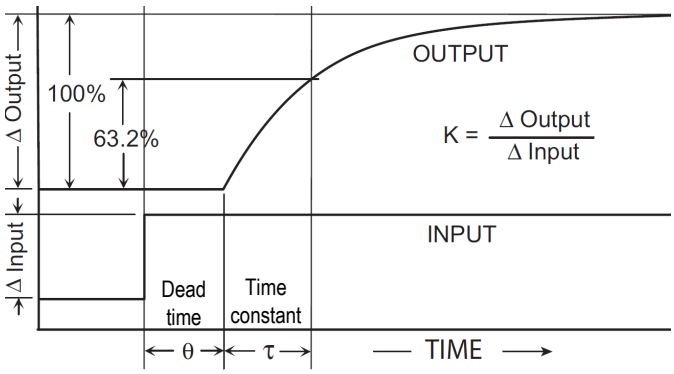


Figure 2.23: Time constants in process control, based on [204]

Attackers and researchers have shown numerous ways of compromising and controlling the digital systems involved in process control [193, 137, 41, 111, 5], thereby raising the possibility that catastrophic failures in the physical world can result from taking control of those digital systems. When an attack transitions from the control of a digital system to the control of a physical process, physics becomes the controlling factor instead of the digital rules encoded into the microcontroller.

> A single bit flip can engage the burner under a tank of water, but the water will still take hours to heat regardless of the state of the controller outputs.

> Changing the state of the outputs does not immediately put the process into a vulnerable state. An attacker needs to take into account the timing and state of the system and act when the process is in a vulnerable state.

The key questions are when the process will reach the vulnerable state and how long it will take after the start of the attack to achieve the attacker's goals. Since the exact answers to these questions may not be known to the attacker, there may be considerable uncertainty with regard to the timing of the attack. Feedback loops are capable of governing the process when it is operated within expected boundaries. However, as the process is pushed further away from normal conditions, the effectiveness of the control loops becomes less known. Correspondingly, timing parameters become more uncertain and the process may enter an uncontrollable state. Moreover, the field instrumentation is calibrated to measure the process within certain predefined limits and may be inadequate to describe the process when it operates outside its configured region [189].

Consider a piping infrastructure in a plant. Fluid dynamics play an important role in the exploitation of the piping systems. Pipes are typically designed to the maximum expected design pressure of the system, but a water hammer may amplify the system pressure by as much as six to ten, or more, times the original, intended design pressure and cause pipe rupture [135]. A classic water hammer or a hydraulic shock is caused when a fluid in motion is forced to stop or change direction suddenly, e.g., when a valve closes suddenly at an end of a pipeline system (Figure 2.24). The resulting pressure surge travels upstream at a sonic velocity causing pressure change in the pipe. Certain timing parameters such as speed and frequency of the valve closing, period and velocity of the wave propagation must be considered by an attacker while designing the damage scenario on the piping system.

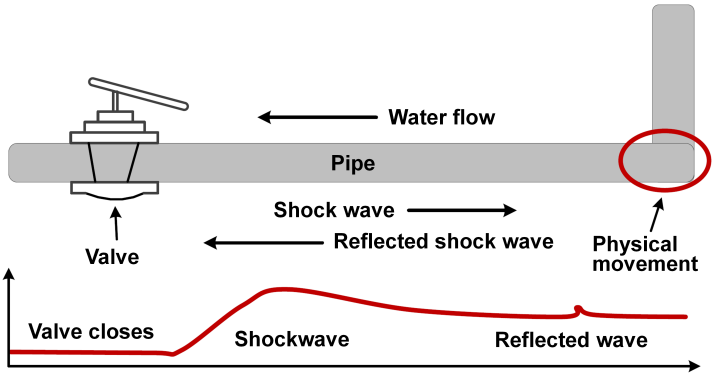


Figure 2.24: Visualization of water hammer occurrence mechanism

Every physical phenomenon in nature is best experienced in moderation and their extremes may yield disruptions and damages. Resonance and vibrations are common threads that run through almost every branch of physics and mechanics. Taken to an extreme, they occasionally cause a bridge to collapse, a helicopter to fall apart and other detrimental effects. Understanding the nature of their occurrence and the susceptibility of the physical and mechanical systems to their exposure is the foundation for many classes and types of cyber-physical attacks [130]. The equations to compute conditions for the occurrence of unwanted forces and determination of their unsafe levels are available in the

literature [205, 135]. However, the design and execution of such attacks in cyberspace would require strategic and timely manipulation of the process parameters.

To give an example of the above mentioned attack scenarios, we consider a catastrophic accident at Sayano-Shushenskaya Dam, Russia's Largest Hydroelectric Power Station in 2009 (75 killed) [96]. It was caused, among other reasons, by excessive vibrations which wore down the bolts that kept the turbine in place and by a consequent water hammer that ripped the turbine out of its seat. Additionally, another time-related parameter was involved in the accident. The working life defined by the manufacturer for the turbine was specified as 30 years. At the moment of the accident, the age of the turbine was 29 years and 10 months. Vibrations and shocks degrade equipment robustness over time. Therefore the exploitation regime of aged machinery must be selected with care in order not to overstress it. The fragility of the equipment at the end of its designed exploitation time can be abused by the attacker.

Depending on the attack objective and the attack progression, time typically plays a more critical/pressing role during the cyber-physical incident (the physical process is a target) compared to the cyber incident (information or IT infrastructure is a target).

> **IT domain.** On average, companies take about 207 days to identify and 70 days to contain a breach in 2022 according to IBM [102].
>
> **OT domain.** At 1:23 pm reactor cooling problem identified. At 1:33 pm the reactor burst and its contents exploded, killing 4 and injuring 38 people [192].

The difference in requirements to the *urgency* of attack reaction should be taken into account when designing security monitoring infrastructure, configuring detection parameters and establishing incident response processes. The speed of incident reaction in the OT domain is subject to continuous optimization. In recent years this resulted in the development of high performance HMI displays (Figure 2.25) aimed at improving the operator's cognition of process state and facilitating shorter incident resolutions.

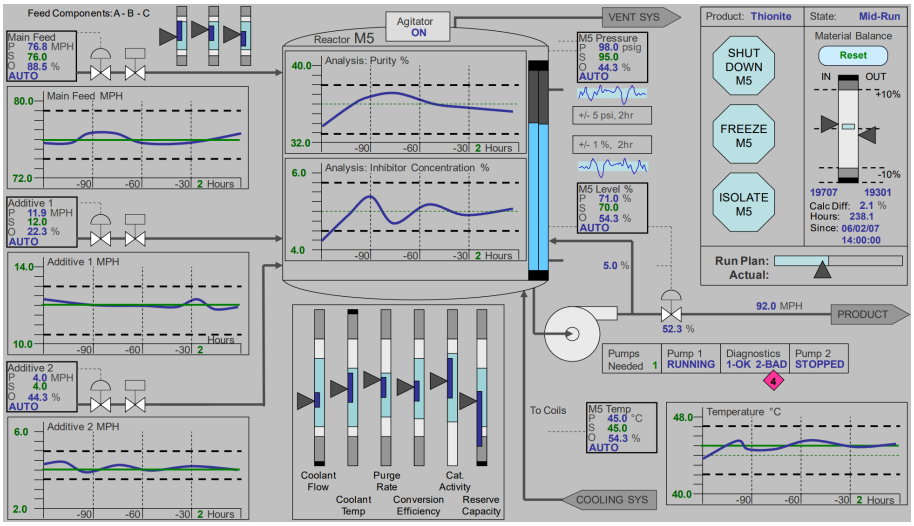


Figure 2.25: High performance HMI display [18]

Note, that the two fundamental design principles of modern HMIs remained the same as for previous generations:

> The operator interface allows operators to focus their mental resources on controlling the process, not on interacting with the underlying system platform;
>
> That means the HMI is consistent and easy to use in terms of making minimal demands on the console operators' mental and physical resources to understand and interact with the process control system [170].

Although more effective HMIs contributed to the overall improvement in process safety, even modern HMIs are not optimized for resolving process upsets caused by intentional/malicious actions. For instance, in a particular situation discussed in [188], the resolution of a tricky alarm required an operator to click and review 28 screens. As a result, the process was in an abnormal, running outside the operating limits state for 32 minutes, which resulted in increased maintenance costs from excessive stress on the equipment. Understanding of timing parameters involved in revolving abnormal process states can be taken advantage of by the attacker, e.g., by causing an unnatural alarm sequence or alarm flood to cause confusion and increase operator response time.

2.3.2 Safety vs. Security

In the physical world safety is a primary concern. In the context of CPS, safety systems have the critical function of detecting dangerous or hazardous conditions and taking actions to prevent catastrophic consequences for the users and the environment. The industrial control community has substantial experience in identifying and addressing potential hazards and operational problems in terms of plant design and human error, used to minimize the effects of atypical situations and to achieve a safe outcome from a situation that could have resulted in a major accident. Even before somebody is allowed to visit a plant, they usually must watch a safety training video. Many industrial companies have a large screen displaying the number of days elapsed since the last safety accident.

However, the evolution of safety systems is largely built on the ability to interconnect systems and automate notifications and alarms in the event of safety breaches. As a result, safety systems became vulnerable to cyber attacks. In the past, the relationship between safety and security was studied in the context of dependable computing (Figure 2.26). Compared to existing works on determining common approaches to safety and security that have their focus on IT or system design, e.g., [166, 138], we suggest also including the underlying physical processes into consideration.

Figure 2.26: Dependability and security attributes, based on [14]

Both cyber security and safety have distinct histories and have developed their own bodies of work. In both disciplines basic concepts have developed into a language that can be used to describe best practices. However, the current efforts to secure critical infrastructures have used the language of cyber security drawing little from the language of safety. Architectures are most often described in terms of security boundaries and not in terms of hazards. This cyber-oriented view of the world has been codified into standards and regulations governing process control.

One regulation illustrating this point is the NERC CIP standard [164]. Under this regulation, a control system is broken down into a set of "control centers". The communications between control centers and outside entities define an electronic security perimeter (ESP). Not all control centers are required to be defended. Simple tests are used to determine whether a particular control center is required to be defended to achieve compliance with the standard. However, most of the tests are cyber-oriented. The only safety-oriented test is that the control system should have the ability to shed 300 MW of load. All other hazards such as bringing a generator out of phase [238] or energizing a line during maintenance work are not considered by the standard. The NIST 800-53a standard has a similar flavor [109]. Its general hardening recommendations such as password lengths are applied broadly to the devices used in process control. The standard is meant to be applied to all industrial processes without any additional considerations for the specific product being produced or manufactured. There is no need for the implementer to understand the inherent hazards of the system. Hazards are simply seen as part of the nameless devices residing at the lowest level of the control system architecture, whereas cyber security controls are implemented as a barrier on top of these devices, frequently in the form of firewalls. The danger of this line of thinking is that all parts of the process are often grouped together into a single security compartment without any regard to how the parts of the process interact with each other and, specifically, how these interactions may impact safety of the process.

Consider a piping infrastructure at an industrial facility. Once two devices are inserted into a process, they can become related to each other by the physics of that process. The physical process then becomes a communication medium and may be used for delivering malicious payloads even if the devices are segregated electronically [119]. With that, the IEC 62443 *approach* to *defining zones* and conduits is *violated* when an attacker can use equipment in a lower security zone to deliver an attack payload to piece of equipment in a more secure zone via process physics, rather than in electronic form via a configured conduit as shown in Figure 2.27. Section 4.6 provides additional discussion on the topic.

Safety and security are sometimes described as two sides of the same coin [54]. If the attacker can compromise safety systems through cyberspace and prevent them from performing their intended protection function, a security incident may lead directly to a catastrophic event. Security and safety are interconnected but both have different missions and employ different vocabularies. In order to understand their *"separation of concerns"*, in the simplest way, the relationship between security and safety can be explained with the diagram shown in Figure 2.28. If the attacker manages to breach a security perimeter and gains access to the plant infrastructure, there are still several layers of safety protections in place (including mechanical ones) to prevent harm to the process, equipment, personnel and environment.

From a safety point of view, any device that can fail is an additional risk. A firewall might physically catch fire, or it might be misconfigured and suppress important messages.

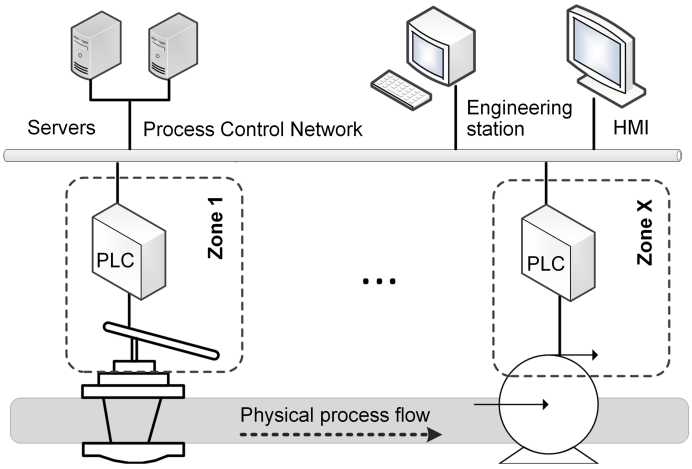


Figure 2.27: Illustration of network security zoning violation

As emergency cases are rare, uncommon messages will be sent especially often in cases when they are important. Message encryption makes it harder to monitor communication channels. Message authentication failures might refuse valid messages. Key management may introduce further potential points of failure.

Safety and security are related and require similar mindsets, e.g., investing resources for something not to happen, but there are also some fundamental differences in the way of thinking. In a safety system, failures are usually assumed to be accidental. Failure probabilities of individual components are thus by and large independent. If the probability of an engine failure in an airplane is 1 in 100.000 flights, the probability of two engines failing is 10^{-12}; this is then the failure probability for a plane that can land safely with one engine. In security, failures are not independent. An attacker attempting to manipulate one device is likely to also attack the fallback. An attacker's next step may *depend* on what has been observed so far. Events deemed so unlikely that they are simply not accounted for, such as a name containing valid SQL commands, may be triggered intentionally.

> Finally, security and safety have opposing update regimes. IT security employs *frequent updates* such as installing patches, upgrading firmware or adding new firewall rules to react to new *security advisories*. On the other hand, an engineering system once approved is expected to remain *safe if left untouched*.

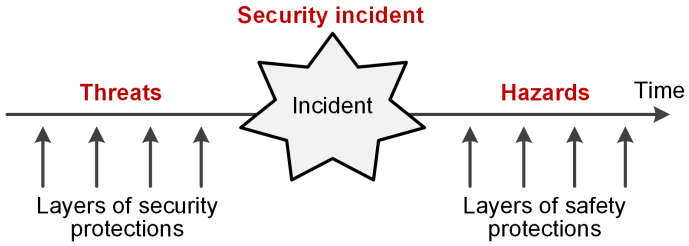


Figure 2.28: Relationships between security and safety

Any change in software or operational practices must be followed by an extensive safety revision. Failing to do so can result in casualties. For example, after an update of the SAP-based maintenance software at DuPont without a subsequent safety review, an alarm notifying on a due date for a hose change disappeared. As a result, a hose used to transfer phosgene from a cylinder to a process wore out and catastrophically failed spraying a worker in the face and resulting in his death [191]. With the increasing digitalization of the safety systems and their interconnectivity with other systems such as DCS and historians, there is a growing risk for these systems to be compromised and not being able to provide their protective function. This will be further discussed in the next section.

2.4 Industrial Control Systems Threat Landscape

In the IT domain, the ultimate objective of the attackers is often to obtain data (e.g., intellectual property or specific records). Misuses in the OT domain are much more concerning as the attackers' goal is to cause an impact in the physical world. The concerns are exacerbated by the fact that most control equipment, systems and protocols either lack fundamental security controls or do not have them properly configured. As compensating security measure, it is possible to segregate OT equipment from the high-risk networks and Internet via the OT DMZ. However, many industrial environments either do not have this network layer or have it implemented with very limited capability. As a result, the susceptibility of the OT zone to malware compromise or human-assisted intrusions remains high. Figure 2.29 shows the timeline of publicly-known attacks which had manifested in physical consequences. We added two targeted reconnaissance campaigns because their activities were directed at collecting information related to the execution of cyber-physical attacks.

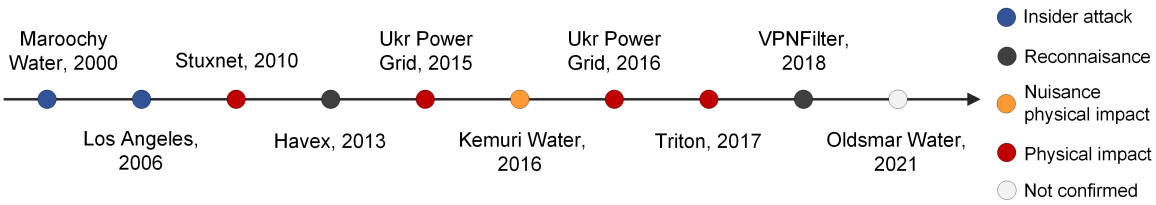


Figure 2.29: Historical time of cyber-physical attacks

In general, the concept of control systems misuse to achieve specific physical impact is not novel. Before becoming a target of external threat actors, control systems were subject to insider attacks. Among the most frequently discussed cases is an attack on the Maroochy Shire sewage treatment process in Queensland, Australia in 2000. A disgruntled contractor who had installed the radio-controlled waste management system used commercial radios to interact with control equipment and caused millions of liters of raw sewage to spill out into local parks and rivers, contaminating creek water and causing death to local marine life [203]. In 2006 two traffic engineers in Los Angeles, USA, remotely accessed the city's traffic control system and changed the duration of the red light at four busy intersections located near freeways and major destinations in Los Angeles as part of a labor union protest. They also inserted a code that prevented an easy fix of the hack. It took a few days to resolve the traffic gridlock caused [84]. As can be noticed from these two examples, insider

attacks can be especially damaging thanks to intimate knowledge of the systems possessed by the insider employees.

The epoch of external targeted attacks on control systems started with Stuxnet, malware that was designed to limit the pace of the Iranian uranium enrichment capabilities [73]. Among the distinctive features of Stuxnet is the amount of multi-national efforts invested into its development [240] which makes it a special case not only in the ICS space but also in the overall history of cyber security. The cyber-physical payload of Stuxnet was implemented in two distinctive variants [129]. The first payload was designed to over-pressurize centrifuges by keeping exhaust valves within the centrifuge cascades closed. This attack scenario turned out to be not sufficiently reliable. The second payload variant targeted the centrifuges' rotor speeds to over-speed centrifuge rotors and to take them through their critical (resonance) speeds/frequencies. While Stuxnet did not significantly set back the Iranian program due to its premature discovery, it opened Pandora's box of cyber-physical attacks and prompted other nation-state threat actors to ramp up their cyber activities in the ICS space.

Not long after the Stuxnet discovery, industrial organizations primarily in the USA and Europe were targeted with a malware campaign which hypothetically could have been a reconnaissance stage preceding a Stuxnet-like sabotage attack. A malware named Havex was designed to search for OPC servers and collect information about connected field instrumentation [103]. While no malicious code targeted at process manipulation was discovered, with the obtained process configuration data such capability could have been added to Havex with just a few lines of code. In another instance, an attack campaign targeted a wide range of networking devices in 2018, including components for detecting and logging Modbus TCP traffic [209]. It is not uncommon to relay industrial traffic over the Internet without protection [162]; the attackers used this knowledge to identify sources and destinations of such traffic together with process data payloads.

In the safety discipline, a *near miss* is an event not causing harm, but having the potential to cause injury or ill health [100]. A reported cyber-physical attack on an unnamed water utility in 2016, could have potentially caused significant harm to the population but resulted only in nuisances. Partly due to the attacker's "incompetence" and partly due to the safety measures in place [227, 140]. The threat actor exploited weaknesses in the Internet-facing system, obtained remote access to the control application and modified the ratios of chemical components entering the water treatment process. However, process manipulations were promptly identified and reverted to normal values. While there was no danger to public safety due to strict water quality controls before water leaves for public use, the attackers could cause damage to process equipment through chemical contamination. The amateur and naive way of conducting this attack suggests an opportunistic attacker who was not driven by a deterministic objective of achieving a specific outcome. A similar attack against another water utility in Florida, USA was reported in 2021 [70]. However, in 2023 it was revealed that the incident was caused by the own employee who also notified the supervisor as soon as the correct composition of chemicals was restored [176]. We, therefore, marked this attack with gray.

Although opportunistic attacks may potentially result in harmful outcomes, high-precision attacks are much more concerning. Besides Stuxnet, there were three other proven targeted cyber-physical attacks with physical consequences in past years, all happening in a short period between 2015-2017:

- Attacks on three power substations in Ukraine, 2015; malware family – BlackEnergy3 [66].

  On December 23rd, 2015 a major coordinated attack on electrical utilities in three different regions in Ukraine left hundreds of thousands of inhabitants without electricity for about six hours in total. The attackers compromised remote access to the operator console in the OT network and used HMI software to open circuit breakers. Simultaneously, in one of the regional utilities, the attackers updated Serial-to-Ethernet converters with malicious firmware to make them inoperable. This ensured that even if the operator workstations were recovered, it would not be possible to re-close the breakers in the affected substations remotely, requiring manual remediation. Only thanks to the timely restoration of the power supply the attack did not result in casualties or other significant consequences.

- Attack on a power substation in Ukraine, 2016; malware family – Industroyer [239].

  Just before midnight on December 17th, 2016, a $330kV$ substation north of the Ukrainian capital Kyiv went offline, affecting the population in a part of Kyiv and a surrounding area. Forensic investigation has discovered malicious software on several SCADA servers which was used to disconnect circuit breakers via scripted networking commands. In contrast to the attack in 2015, the final subversion activities were not executed manually by human threat actors but instead scripted into an autonomous malware framework [41]. As during the previous incident, the power supply was manually restored in a timely fashion, avoiding serious consequences. Another variant of this malware, *Industroyer2,* was discovered in Ukraine in 2022 [69]. However, the malware was detected before the threat actors were able to bring it into action. While Industroyer design has evolved since its initial deployment in 2015 [78], it remained unclear whether Industroyer2 could have been successfully executed in the victim environment.

- Attack on Safety Instrumented System in a refinery in Saudi Arabia, 2017; malware family – TRITON [111].

  While both attacks on the Ukrainian power grid could have potentially resulted in a significant impact on the civilian population via collateral consequences, the TRITON attack on the Triconix safety controller directly undermined the safety of human workers in the victim facility. TRITON is a passive software implant with *read*, *write* and *execute* capabilities, residing in the controller's memory and activated by a special network packet. While TRITON was discovered at the stage of its installation, once successfully installed the implant allows to remotely reprogram safety system with an attacker-defined payload. In the worst-case scenario, the TRITON capability could be extended to controlling I/O functions and denying safety shutdown functionality during hazardous events. Figure 2.30 shows the block diagram of the 3008 Triconex electronic board with main and I/O processors located on the same board and communicating with each other via shared memory. With this hardware design, passing commands to safety instrumentation could be implemented with minimal effort.

When mapping the attacks discussed above against a traditional layered representation of an ICS network, it becomes apparent that cyber-physical exploitation tactics are following an evolutionary path, clearly picturing the ongoing **"race-to-the-bottom"** trend between

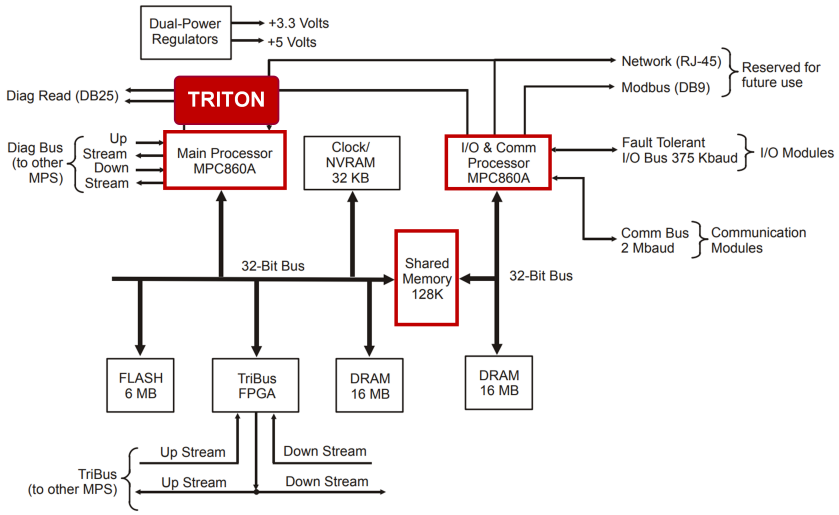


Figure 2.30: Architecture of a model 3008 main processor [45]

ICS threat actors and defenders. This trend refers to the tendency of the attackers to move their exploits one layer down as soon as security controls are introduced at some layer of the computer or network architecture. While OT asset owners begin to harden operator consoles and embrace ICS network monitoring solutions, the attackers are already moving their exploits into the control equipment as illustrated in Figure 2.31.

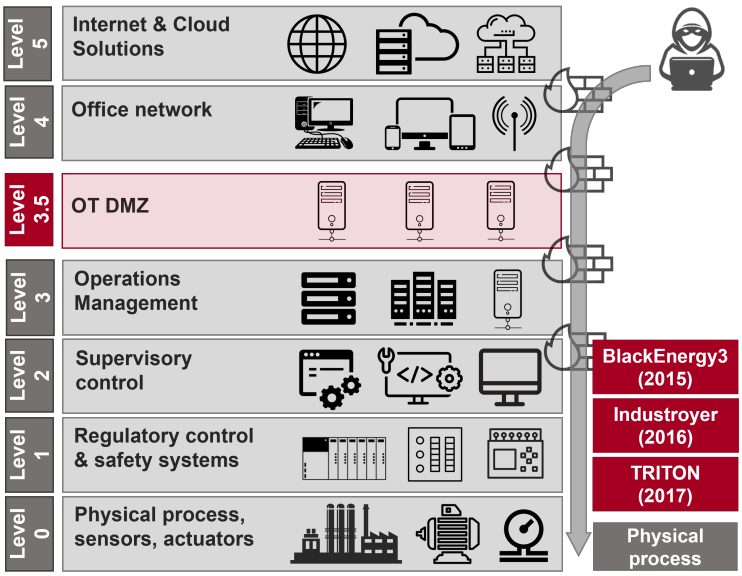


Figure 2.31: "Race-to-the-Bottom" exploitation trend in Industrial Control Systems

In the 2015 attack on the Ukrainian power grid, the attackers leveraged unauthorized access to the HMI via remote access software to open the breakers and de-energize the substation. With that, the attack was executed at the level of supervisory control by

compromising more familiar Windows-based IT systems. However, executing ICS attacks at this level is less reliable due to readily-available security controls for Windows systems. In the 2016 power grid incident the attackers launched their exploits at the industrial protocol layer. Before the 2015-2016 events very few industrial organizations invested in visibility solutions for their control networks. This is why Industroyer activities were not detected at the time of the attack. With increased awareness of security risks and OT network monitoring solutions readily available on the market, investing in offensive capabilities at the level of control networks has become less practical. In the 2017 TRITON attack, the attacker moved their exploit all the way into the control equipment at the regulatory control level. The reason for this strategy is the lack of exploit mitigation and detection capabilities in most of the embedded components deployed within the ICS and a lack of tools to support compromise assessment and forensic analysis of these systems. Therefore, in the foreseeable future embedded exploits are likely to remain undetected for an extended period. Also, the lack of international regulations in using cyber warfare against critical and vital infrastructures does not discourage the execution of cyber-physical attacks, further motivating threat actors to continue with the race.

It is not uncommon for sensing devices and even actuators to be accessible either physically or via exposed communication channels. Examples include wireless sensing devices located too close to the physical fence of an industrial facility or devices placed in accessible places. The attackers can take control over such end-point devices and use their connectivity to a wider network to compromise systems located in the upper layers of the Purdue reference architecture. This control-systems specific exploitation scenario is sometimes called *"hacking upstream"*. In one instance the attackers were able to exfiltrate a large volume of data from a casino by compromising an Internet-connected thermostat in the casino's aquarium [153]. In another example, the researcher demonstrated how the attacker can trigger software vulnerabilities in an ERP system such as SAP in a multistage attack which starts with compromising HART transmitters [21]. Industrial facilities in remote locations which rely on radio transmission lines are especially susceptible to such attack scenarios [144].

Every industrial environment is unique. Creating customized payloads is laborious and not scalable. It is natural for the attackers to search for (1) design payloads that require little proprietary knowledge about the target environment (e.g., configuration parameters, analysis of process data), (2) *reusable* cyber-physical attack scenarios. The second variant of the Stuxnet payload is an early example of a reusable damage scenario. Its payload targets a common weakness of all rotating equipment: namely, equipment overstress through operation at harmful frequencies. With minor modification, this payload can be re-used against other types of rotating equipment such as motors or turbines.

The Industroyer malware is among the more recent examples of *reusable cyber-physical payloads*. Its attack logic is relatively simple: find all or selected controllable signals on RTUs and OPC servers and change their status to off. The attackers used publicly available protocol specification documentation to discover the needed points in the control equipment instead of relying on the substation's configuration information, which makes the Industroyer attack scenario applicable to any substation worldwide. To maximize their success chances, the attackers implemented attack logic in four different modules for four widely used industrial protocols in the electrical sector as shown in Figure 2.32. A launching module was pre-programmed to execute all four payloads simultaneously at the predefined

time. This success maximization strategy can be compared to a heap spray attack on computer systems. The Industroyer attack framework is also extensible in that it can be easily enhanced to support additional protocols.

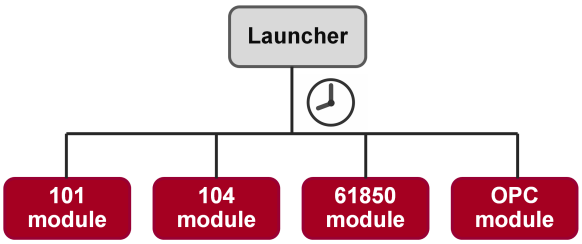


Figure 2.32: Industroyer payload components

Previously, limited access to industrial equipment, software and proprietary engineering documentation were thought to be the main barrier to entering the ICS exploitation field. The situation has changed significantly over the past decade. If previously one would need to be a legitimate buyer to buy certain types of industrial equipment such as safety systems, nowadays such equipment and related software can be freely purchased at e-commerce platforms such as eBay [106] or Ali Baba [90]. Also, restricted and proprietary documentation such as various engineering documentation, equipment manuals and development guides, specification sheets, trainings, purchasing documentation and others are frequently uploaded to public file repositories such as Scribd [108]. Industrial software executables, firmware for industrial equipment, control logic, and configuration files are often uploaded to Virus Total [220] and GitHub [107]. It is also not uncommon when engineers and subcontractors expose sensitive documentation via improperly configured file servers [161].

To make use of the vulnerabilities discovered in the ICS equipment, they need to be weaponized. There is a growing effort in the security community to produce exploits for publicly disclosed ICS vulnerabilities for both research and commercial purposes. This includes stand alone exploits, e.g., as Metasploit modules [178] and commercial frameworks such as SCADA+ Pack [82], which includes both public and 0-days vulnerabilities in one exploit pack. To close the gap in lacking tools for ICS exploitation, researchers and security practitioners developed repositories with ICS attack instruments, e.g., [199, 58, 211, 167]. Additionally, a significant effort was dedicated to developing password bruteforces and creating libraries of the default/commonly used credentials for industrial equipment [217, 194, 59, 57]. As anything publicly available, such repositories of attack tools lower the bar for threat actors with an interest in ICS exploitation. It is worth mentioning that it took only a few months until a detailed technical analysis of Industroyer [41] became publicly available and just a couple of weeks for the TRITON attack framework to be uploaded to the GitHub repository after incident disclosure [64]. Although pulling off successful cyber-physical attacks still requires a significant amount of specialized skills and expertise, the accessibility of tools and exploit modules is making it easier for threat actors with limited resources to bridge their knowledge gap.

Although ransomware became a prominent threat to ICS environments in the past few years [171], we do not consider this threat in the white paper because

its impact on the physical process is either collateral, caused by poor or missing segmentation between OT and IT networks [26], or tailored to increase the chance of receiving the ransom by stopping some critical monitoring and controls services [29]. We limit the discussion scope to *targeted cyber-physical threats* directed at causing a *pre-defined impact* in the *physical world.*

Note, that we did not include cyber-physical attacks claimed on social media Telegram and Twitter in 2022 [216] as their authenticity could not be validated. We also did not include an attack on the KA-Sat communication satellite in 2022 which resulted in the loss of communication with almost 6000 wind turbines in Central Europe as it did not affect photovoltaic systems [232]. Last but not least, we did not include the physical destruction of critical infrastructure in Ukraine with kinetic weapons as at this time there is no public evidence that cyber attacks could have been contributing factors to the attack outcome. Also, a combination of cyber and kinetic warfare falls into the realm of *hybrid warfware* that is not in the white paper scope. On a similar note, we omit analysis of the so-called "Vulkan files" [11, 224] due to a requirement for a longer assessment between known cyber-physical attacks and leaked documentation.

2.5 Conclusion

With the introduction of computer-aided manufacturing, significant effort was put into the standardization of industrial architectures and communication flows resulting in concepts like the automation pyramid and the Purdue reference architecture. With the growing business need to exchange data with enterprise systems and users as well as provide remote access to complex pieces of machinery, the industry has focused on finding best practices for securing automation systems from potential external threats and avoiding any disruptions in control systems functioning. However, being a conservative industry, the adaptation of security practices in automation has been much slower than the desire of the business for the better economic performance of the plants, resulting in increasing discrepancies between technological advancements such as the usage of cloud technologies and efforts to identify and mitigate associated security risks.

The Industrial Control Systems threat landscape has changed dramatically over the past few years. New threats have emerged, exacerbating global concerns brought about by Stuxnet, the first known cyber-physical attack. Although modernization and wider connectivity introduced opportunities for remote exploitation of control systems, attempts to disrupt a process without clearly understanding the consequences of the attack actions on the physical process are likely to result in a minor nuisance instead of an actual disruption. In order to achieve the desired cyber-physical effect, the attacker requires a body of knowledge that is different from what is commonly employed in the IT field. For instance, it takes a specially-crafted attack to bring the process into a needed vulnerable state or the attacker needs to wait for a process to be in a particular state before executing her attacks. In the pre-Stuxnet era, ICS exploitation was considered a boutique skill mostly possessed by state-sponsored threat actors. However, the previous status quo has been challenged over the past decade thanks to wider accessibility of OT equipment, open-source attack tools, and information, allowing less resourceful threat actors to participate in the ICS domain in support of economic, political and military interests.

3. Vinyl Acetate Monomer Process

Vinyl acetate (VAC) monomer is a commodity chemical and is an essential building block used in a wide variety of industrial and consumer products. VAC is a key ingredient in resins, intermediates used in paints, adhesives, coatings, textiles packaging, automotive plastic fuel tanks and many other final products. Detailed information about the product, including regulatory, health, environmental, and physical hazard information can be, for instance, found on the web page of Dow Chemical Company [213]. In 1997 Luyben and Tyreus published details of the industrial process for the production of Vinyl Acetate to make it available for public/open-source research. This model involves real non-ideal chemical components, a realistically large process flow sheet and consists of several standard unit operations that are typical of many chemical plants with the recycle stream and energy integration. We chose this model for its realism, size (large-scale) and complexity (nonlinear dynamic model). In addition, the advantage of the VAC model is that it consists of realistic components. Therefore, one can easily modify and enhance the process model with further simulation code, e.g., change catalyst decay conditions or add dynamics of the pumps. There is a large body of literature on the specifics of the individual unit operations and chemical reactions. The availability of literature on the VAC process specifics allows to design *high-precision targeted* attacks.

The model can be simulated in Matlab (see Section C). While Matlab is not an open-source software [155], it is one of the most popular software environments for engineers and scientists, and it is more affordable than other commercial dynamic process simulators. Note, that we adopted vinyl acetate (VAC) abbreviation instead of more commonly used vinyl acetate monomer (VAM) because of the original naming of this process in the introductory paper [38].

3.1 Plant Model Description

In the VAC plant model, there are ten basic unit operations, which include a vaporizer, a catalytic plug flow reactor, a feed-effluent heat exchanger (FEHE), a separator, a gas

compressor, an absorber, a carbon dioxide CO_2 removal system, a gas removal system, a tank for the liquid recycle stream, and an azeotropic distillation column with a decanter.

The route for vinyl acetate manufacturing used in the process model is the same as employed in today's manufacturing and involves seven chemical components. Ethylene C_2H_4, oxygen O_2, and acetic acid HAc are provided as both fresh and recycled feeds and are converted into the vinyl acetate with water H_2O and carbon dioxide CO_2 as byproducts. The fresh C_2H_4 stream contains an inert component C_2H_6. The following reactions take place in the reactor:

Main reaction: $C_2H_4 + CH_3COOH + \frac{1}{2}O_2 \rightarrow CH_2{=}CHOCOCH_3 + H_2O$,
Side reaction: $C_2H_4 + 3O_2 \rightarrow 2CO_2 + 2H_2O$.

The reactor contains tubes packed with a catalyst. Both reactions are highly exothermic and require tight control of the reactor cooling. The side reaction of ethylene combustion to CO_2 is highly undesirable as it lowers the conversion and complicates the removal of the reaction heat. Details of the ethylene combustion kinetics in the synthesis of vinyl acetate are presented in [97].

The reactor effluent is sent to the separator, where gas and liquid are separated. The vapor from the separator goes to the compressor and the liquid stream becomes a part of the feed to the distillation column. The gas from the compressor is recycled back to the reactor through the absorber and the CO_2 removal system. The liquid products, VAC and water, are withdrawn from the decanter. Figure 3.1 illustrates the process flowsheet with corresponding control structure.

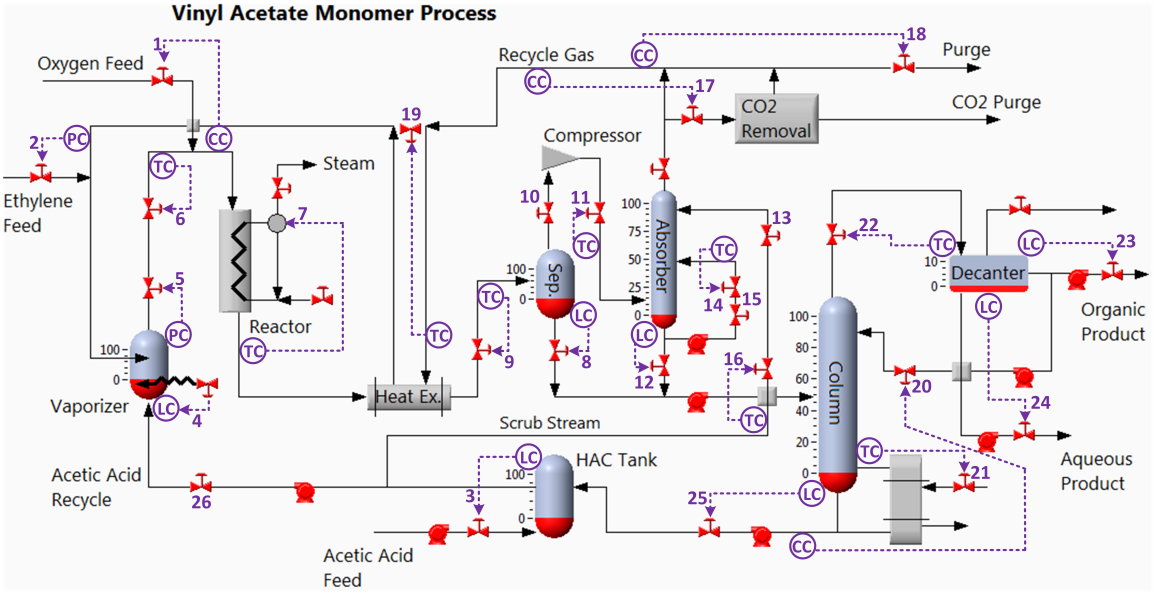


Figure 3.1: Vinyl Acetate Monomer plantwide process control structure

Detailed process description, including the reaction rate expressions, steady state process data and the major aspects of each unit operation are covered in [147] and Chapter 10 in [61]. To protect the proprietary information of any specific VAC production facility, the

kinetic data, process flowsheet information, equipment data and modeling formulation in the published process came from sources in the open literature, e.g., in [147] and references therein.

Safety constraints

Two key safety constraints exist in the process. Exceeding either of the safety limits will shut down the process via interlocks:

- O_2 concentration must not exceed 8 $mol\%$ anywhere in the gas recycle loop to remain outside the explosivity envelope of ethylene. More on the limits of oxygen concentration in gas mixtures can be found in [241].
- The pressure in the gas recycle loop and distillation column must not exceed 965 kPa (140 psi) because of the mechanical construction limit of the vessels.

Operating constraints

The process constraints must be maintained to ensure efficient production without interruptions for maintenance. They are specified as the following upper and lower bounds for key process variables:

- The peak reactor temperature along the length of the tube must remain below 200°C to prevent mechanical damage to the catalyst, which would require shutdown and catalyst exchange.
- Liquid levels in the vaporizer, separator, absorber base, distillation column base, and decanter must operate within the limits of $10-90\%$.
- Reactor inlet temperature and the hot side exit temperature form the heat exchanger must remain above 130°C to avoid condensation of liquid.
- Organic phase in the decanter must contain less than 600 $mol/million$ of acetic acid to prevent product contamination.
- The VAC composition in the bottoms stream must remain below 100 $mol/million$ to prevent polymerization and fouling in the reboiler and vaporizer.

The VAC plant model is not accompanied with the operating objective costs function for process control optimization. Instead, the economic objective is formulated as balancing trade-offs in maximizing vinyl acetate production and recovery of essential chemicals from the gas recycle loop to prevent yield losses while minimizing the carbon dioxide production and energy consumption.

3.1.1 Control Model

Luyben and Tyreus did not suggest any process control scheme. Instead, they proposed a set of control requirements and challenged the research community to come up with their own control approaches. The examples of developed control strategies include [148, 61, 197, 168]. The majority of the control design implementations were kept proprietary due to high modeling costs.

The process model includes 246 states, 26 manipulated variables (valves) denoted as XMV{1-26}, and 43 process measurements (sensors) denoted as XMEAS {1-43}. Readers are referred to [38] for a complete description of the process model formulation, assumptions,

and implementation. The process model used in the white paper utilizes a control scheme proposed in [148]. Figure 3.1 depicts the location of the control loops. The numbering of the control loops follows the numbering in Appendix 2 in [38]. Some manipulated variables are fixed, their control loops are not shown. Appendix B provides a listing of all control loops and measurements with descriptions. In [148] Luyben et al. specify a number of complexities and interesting dynamic effects of the VAC process. The process includes control loops with small and large time constants and exhibit fast and slow dynamics, depending on the change in the operating conditions. However, in general, the process quickly becomes unstable if pushed outside of the steady state conditions.

The original process model was executed as C-routines in Matlab, providing no user interface to process parameters and process control model. We enhanced the model with a Simulink control model, graphical user interface, and simulation results visualization to streamline model usage for cyber security research. Appendix C provides further details.

3.2 Attack modeling

The adversary's goal is to cause tangible impact on the process, either on its safety or on its economy. At the level of the regulatory or supervisory process control the attacker can either tamper with the sensor readings or modify the manipulated values issued by the controller via a *data integrity attack* (Man-in-the-Middle (MITM), packet injection, replay, etc.) or deny communication between process and controller via a *DoS attack* on the communication network or equipment (flood, starvation, packet drop, packet delay, etc.).

Let $S_i(t)$ be a PV or MV sample at time t, where $0 \leq t \leq T$, and T the duration of the simulation; time is *discrete*. The attack interval T_a is arbitrary and is limited by the simulation run time T. In our setting, we simulate the manipulated signal S'_i during a data integrity attack as follows:

$$S'_i(t) = \begin{cases} S_i(t), & \text{for } t \notin T_a \\ S^a_i(t), & \text{for } t \in T_a, \end{cases}$$

where $S^a_i(t)$ is the modified PV or MV measurement (attack value).

During a DoS attack on a sensor signal new sensor measurements do not reach the controller. During a DoS attack on a controller signal new manipulated variables do not reach an actuator. Translated into the real world scenario, the controller's input register assigned to storing the measurements of a particular sensor will not be overwritten by a fresh value during the next control cycle run as would happen in a normal case. If $X_i(t)$ is a measurement of sensor i, $Y_j(t)$ is a manipulated variable for actuator j and the attack starts at time t_a, we have:

$$X^a_i(t) = X_i(t_a - 1) \quad \text{and} \quad Y^a_j(t) = Y_j(t_a - 1),$$

where X^a_t is the stale data reading (the last process value received by the controller before the DoS attack) and Y^a_t is stale manipulated variable (the last MV received from the controller before the DoS attack).

In the context of Process Control Systems DoS attacks are similar to integrity attacks. The only difference is in *how* the attack value is brought about: by choosing a DoS approach the attacker has to attack at a *specific time* (e.g., when a valve is all the way open or closed).

To a certain degree, in the context of industrial control systems, DoS attacks are similar to data integrity attacks with the main difference being that the adversary cannot directly set the "attack value". Instead, the adversary can decide on the timing of an attack, such as its starting time t_a, to select an attack value of interest.

The advantage of the DoS attacks is that they can be used to manipulate the process even if control traffic is authenticated and the integrity of data packets is protected. This is because during DoS attacks, malicious actors simply delay or drop data packets instead of modifying them. Considering the impact of the DoS attacks in the process domain, they are also called *Stale Data* attacks, a term which we coined in [126].

3.2.1 Stale Data Attack

Stale data can be caused by both an intentional event (cyber attack) and an unintentional event (cyber incident). Thus, on 10 January 2019, 21:02 CET, the Continental Europe Power System which stretches across 26 countries registered for nine seconds the largest absolute frequency deviation since 2006. Among the main causes of the incident was a failure of a communication line, which resulted in stale data being used for various calculations [214]. Figure 3.2 show the difference between real import/export (yellow curve) and wrong measurement (green line) values caused an "unreal additional power flow" from one control area (Germany) to another (Austria). Since measured values are only transmitted if there is a change from the previous value, the fact of missing measurement updates was not immediately obvious.

Upon incident investigation, it became apparent that the current monitoring tools and alarm systems are not adequate for detecting a kind of error like stale data. Therefore, all relevant stakeholders were recommended to check the existing definition and implementation of fail-safe measurement and telecommunication standards, for all interconnector values used by Load Frequency Controller (LFC) across Continental Europe (CE). Also, it was decided to define and implement control system functionality standards to detect "frozen/stale" LFC values across CE.

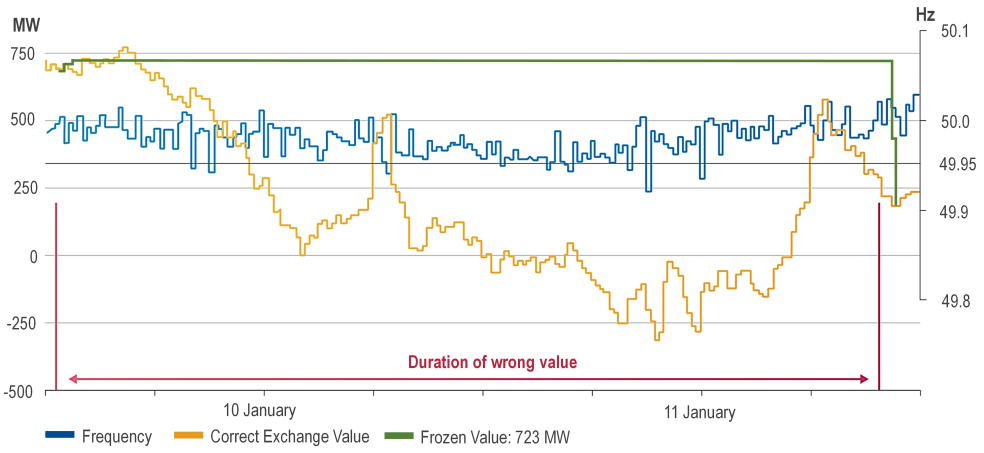


Figure 3.2: Frequency deviation due to "frozen" measurement [214]

A stale data attack can be an effective approach in cases when the scaling or units of data are unknown, e.g., as shown in Figure 3.3. After accessing the target environment, an attacker neither knows the sensitivity of the process to the random manipulations nor the maximum possible process variable range which keeps the process stable. This means that the attacker might be uncertain about which attack value to choose for a data integrity attack. Instead, the attacker may implement an algorithm that would allow her to launch a DoS attack at the time of the observing the highest or lowest process value within a certain time frame [126]. Choosing peak values allows for maximizing the attack impact and expediting the time to achieve said impact.

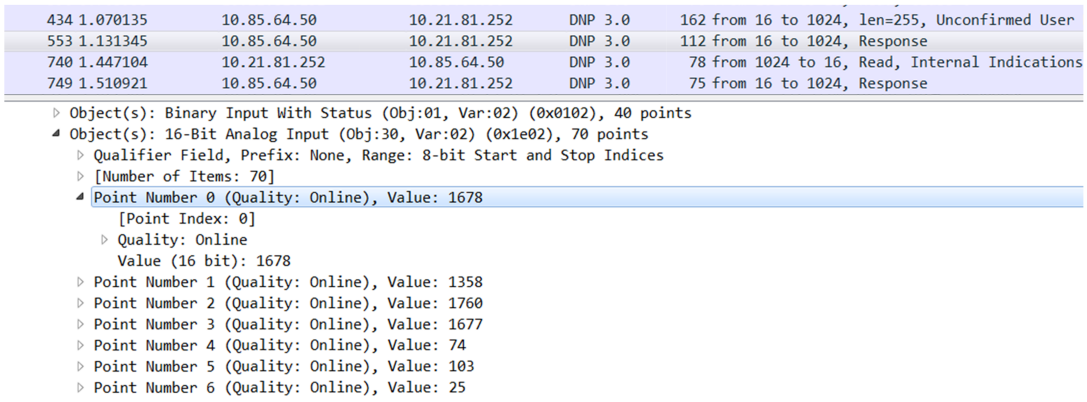


Figure 3.3: Data packet with process variables represented in unknown units

The issue of stale data may become especially prominent with the increasing usage of wireless communications such as 5th generation mobile radio (5G) in plants as shown in Figure 3.4. Even though 5G technology significantly improves important parameters for industrial networks such as latency and jitter compared to previous generations of cellular networks, the attacker may still strategically jam communication links in a temporary fashion or increase packet delivery latency.

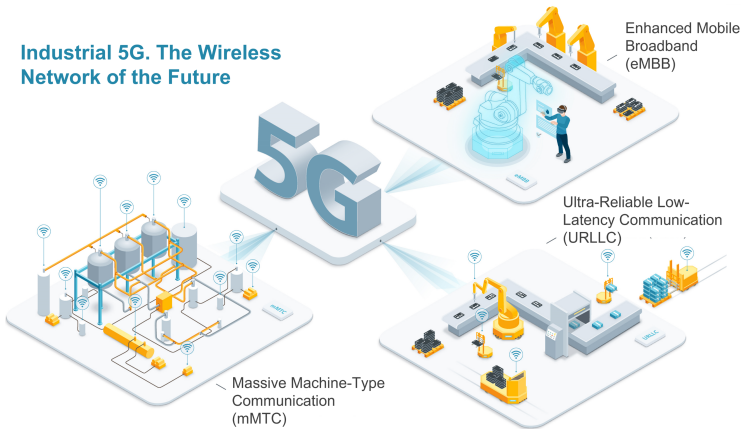


Figure 3.4: Usage of wireless 5G communication in an industrial plant [201]

It is worth noting that DoS can rarely be achieved instantaneously. Typically it takes some to achieve a long enough packet delay or ensure no successful packet retransmission. A large body of literature is dedicated to understanding the best strategies for causing DoS attacks on various communication protocols [143, 142, 79, 141, 40]. A significant advantage of DoS attacks is that complete parsers for proprietary protocols are not required to launch an attack. As a side note, in practical Stale Data attack implementations with the objective of selecting a specific stale attack value, it might be required to optimize the delay between the time the desired process value is observed and the time needed to achieve DoS effect.

4. Cyber-Physical Attack Lifecycle

4.1 Introduction

Threat intelligence and adversarial studies constitute an important part of cyber security defense strategies. Understanding attack strategies and tools allows defenders to proactively implement effective security controls and monitoring tactics. The attacks on selected organizations or systems are not executed at random. An attacker goes through a sequence of steps to successfully infiltrate a network and, e.g., exfiltrate data from it. An "attack lifecycle" or "kill chain" [101] is a common method to describe the process of conducting cyber attacks. One of the popular attack models is the attack lifecycle proposed by Mandiant [150]. Figure 4.1 shows the major attack stages, from initial reconnaissance where the attacker conducts research on a target to the final attack mission. Note that at each stage the adversaries have specific goals and will use appropriate methods to accomplish their goals. If one of the attack vectors and/or methods fails and the attack progress is impeded, the adversary would iterate the stage until success or circle back to a previous stage.

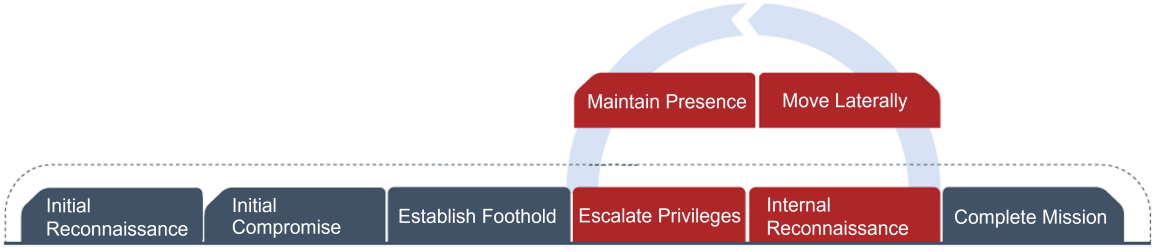


Figure 4.1: Mandiant Attack Lifecycle [150]

The attack steps leading to the final stage of the attack lifecycle "Complete Mission" are required to obtain persistent control over the victim's infrastructure. This sequence of steps is generic and applicable to most organizations and targets. In the final stage, the attackers

act upon their motivation and take the required action to achieve their planned mission. The attack mission can be simple or complex and may consist of additional steps. In information security, the attacker's goal may vary from stealing financial assets or exfiltrating sensitive information like intellectual property to extorting the victim organization or destroying networking equipment.

In the context of cyber-physical systems, the attacker aims at the interaction with the physical process to achieve a desired outcome in the physical world. This may include denial of control loop communication, degradation of equipment performance, or causing process upsets. When weaponizing a buffer overflow in a software application, the shellcode is constructed to make the target system perform actions desired by the attacker. Similarly, cyber-physical payloads must contain a set of instructions that manipulate the process, and the choice of instructions depends on the specific impact the attacker wants to have on the process.

Industrial environments are complex and their comprehension requires *multidisciplinary* knowledge. In the context of this white paper we consider the control infrastructure and physical parts of the attack mission and propose a **cyber-physical attack lifecycle** model. Similar to the "traditional" IT attack lifecycle, the proposed framework describes the attack steps necessary to design and execute a cyber-physical attack. We examine the hurdles an attacker might face when trying to manipulate physical processes, using a realistic simulation model of a *Vinyl Acetate plant* as a case study. We demonstrate a complete attack, from start to end, directed at *prolonged/persistent economic damage* to a production site while avoiding the attribution of production loss to a cyber event. Such an attack scenario could be used, e.g., by a manufacturer aiming at putting competitors out of business or as a pressing argument in an extortion attack.

4.2 Classes of Cyber-Physical Attacks

Modern industrial plants face multiple challenges such as the delivery of products at consistent quality and low cost, management of plant dynamics altered by material recycling and energy integration, strict adherence to environmental and safety regulations, and having a sufficient degree of flexibility to handle fluctuations such as production rate changes in response to changing market demand and feed quality. Maintaining these complex requirements is the responsibility of an efficient and reliable process control infrastructure. Due to a high level of interdependencies and efficiency requirements, any malicious intervention may result in significant consequences to plant performance and compliance. This knowledge can be exploited by the attacker.

The process of designing a cyber-physical attack scenario starts with an attack objective or a precisely desired impact. In general, the effects of cyber-physical attacks can be classified into three groups. Admittedly, the classes outlined are interrelated as damage of one kind may lead to another kind of damage. For example, production can be disrupted through the breakage of equipment. Runaway reactions can cause serious safety accidents and equipment breakage. However, the distinction between attack objectives allows for a better judgment of how different damage types can be maximized and provides guidance for more accurate risk assessment.

4.2.1 Equipment damage

This class of attacks aims at physical damage of equipment or infrastructure (e.g., pipes or valves). Classes of physical damage can be found in [130]. In general, equipment damage can be achieved under two conditions.

Overstress of equipment. Every piece of equipment wears out or breaks at the end of its expected lifecycle. Prolonged overstress of equipment can accelerate this process. An example is wear-off attacks on valves due to unstable process control. This type of attack was implemented in the second version of the Stuxnet worm [129].

Violation of safety limits. Violation of the equipment's safe operating/design typically results in rapid deterioration of the equipment state and its subsequent damage. This scenario was implemented by engineers at Idaho National Labs to remotely destroy a power generator [238]. This attack idea was also realized in the first version of Stuxnet [129]. Equipment safety limits are public information and are readily available in books, professional articles and equipment manuals. For instance, safety limits of piping infrastructures and related equipment can be found in [135].

4.2.2 Production damage

Instead of breaking equipment, an attacker can target the production process to degrade the product or make production more expensive. Attacks on production can be divided into three groups.

Product quality and production rate. The attacker may reduce the value of the product or make it unusable. Every product has its specification and its market price varies with the quality grade, car fuel being a familiar example. In certain instances, the price of a product may rise exponentially with product purity. Table 4.1 presents relative prices for paracetamol. As can be seen, not achieving the desired product quality can result in significant revenue losses. The similarly detrimental economic impact can be caused by reducing the production rate.

Purity	Price, Euro/kg
98%	1.0
99%	5.0
100%	8205.0

Table 4.1: Relative paracetamol prices [202]

Operating costs. After the process is configured and tuned, the operator's primary task is to keep the process as close as possible to the economically optimal operating conditions. Every plant has an objective cost function consisting of several components which impact the operating costs. It may be loss of raw materials in the purge, premature deactivation of the catalyst, or increased energy usage. Increased operating costs reduce revenue and plant profitability.

Maintenance efforts. The attacker can impact a production process by increasing the maintenance workload. Maintenance includes troubleshooting process disturbances and equipment malfunctions. For example, the rapid operation of a flow valve causes a damaging cavitation process. Cavitation eventually wears off the valve and leads to leaks, requiring

valve replacement. Additionally, the bubbling of liquid results in turbulent liquid flows and substantially complicates process control, requiring the involvement of external specialists.

4.2.3 Compliance violation

Industrial sectors tend to be strongly regulated to ensure employee safety and protect the environment. Non-compliance can result in fines and bad publicity (unlike economic attacks whose effect can be kept internal to a company).

Safety. Most damaging would be attacks on occupational and environmental safety as they may result in lethal accidents and serious environmental damage. This type of attack may also yield unpredictable collateral damage, resulting in an extended damage scope.

Environmental pollution. A less serious environmental impact is described as pollution. An attack targeted at exceeding regulatory pollution limits such as concentration and/or volume of gaseous emissions or insufficiently treated wastewater discharge will induce fines, and recurrent offenses can even lead to plant shutdown. A negative impact on reputation may be a further consequence.

Contractual agreements. Some industries may be required to adhere to legally binding production schedules, vaccine production being one example. Reactions to outbreaks of a disease often lead to political and public pressure to supply needed pharmaceuticals in a timely fashion. Missing delivery schedules may cause contractual sanctions and negative publicity.

Classes of cyber-physical attacks, their categories and predominant impact are summarized in Table 4.2.

Equipment damage	**Production Damage**	**Compliance violation**
Equipment overstress Safety limits violation	Product quality Production rate Operating costs Maintenance efforts	Safety Pollution Contractual agreements
Downtime	Financial	Publicity

Table 4.2: Classes of cyber-physical attacks and their major impacts

4.3 Cyber-Physical Attack Lifecycle

The stages of a SCADA attack with a goal to achieve physical impact were first presented in 2006 by Larsen as shown in Figure 4.2. However, Larsen did not provide descriptions of attacker activities related to each stage or supplied any other additional information. In the absence of supporting information, we decided to follow the proposed sequence of attack stages and used our judgment to identify necessary attacker activities at each of the attack stages.

Through experimental work and empirical observations, we were able to arrive at a description of the attack stages and enhanced the initial model to better describe essential attack steps. Specifically, we added **Prevent Response** and **Obtain Feedback** steps as independent attack sub-stages due to their significance in planning and executing cyber-physical attacks. Figure 4.3 illustrates the proposed cyber-physical attack lifecycle. Note,

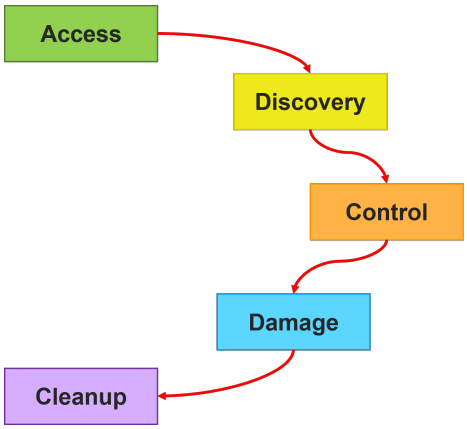


Figure 4.2: Stages of SCADA attacks [130]

that perfect knowledge about the target is never achieved and the attacker may need to circle back to previous stages or recursively repeat her exercises at the

> The main purpose of the white paper is to introduce the cyber-physical attack lifecycle and illustrate its applicability in the example of a targeted attack on a chemical plant with the objective of causing a prolonged negative economic impact. A developed set of attack payloads can be delivered as a series of remotely executed commands/scripts or programmed into an autonomous attack payload.

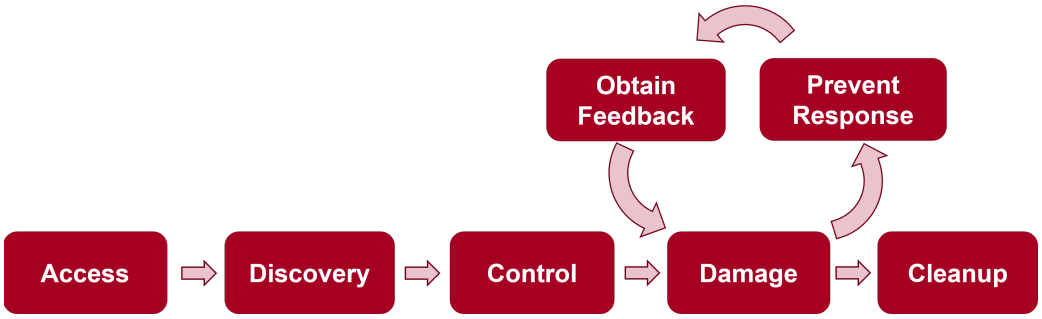


Figure 4.3: Proposed cyber-physical attack lifecycle

4.3.1 Access

Access is the stage closest resembling traditional intrusion attacks in the IT domain. In general, the attacker requires code executing capability in the victim's industrial control network to manipulate the process and thus has to find a way in. Even though this stage is largely the same as intruding into any other network, the attacker may utilize knowledge about the ICS ecosystem to find alternative ways of obtaining access.

Industrial control systems are increasingly getting connected to corporate networks and to external organizations like regulatory bodies, vendors and third-party providers, making them reachable from the Internet. Additionally, in some cases, ICS equipment is directly connected to the Internet for ease of remote access. Any of these data flows can be a potential way into a control network or provide direct access to control equipment. Among the latest known examples of successful execution of the Access stage is the Triton attack where a suspected state-sponsored threat actor obtained remote access to a Safety Instrumented System in a petrochemical facility [111]. Figure 4.4 illustrates the lateral movement of the attacker in the victim environment. The attackers obtained remote access credentials of an internal employee and took advantage of the Remote Desktop Protocol (RDP) access to the control system network from the Demilitarized Zone. Further, they leveraged architectural weaknesses to gain access to the SIS engineering station and attempted to implant a Remote Access Trojan (RAT) into the memory of an SIS controller.

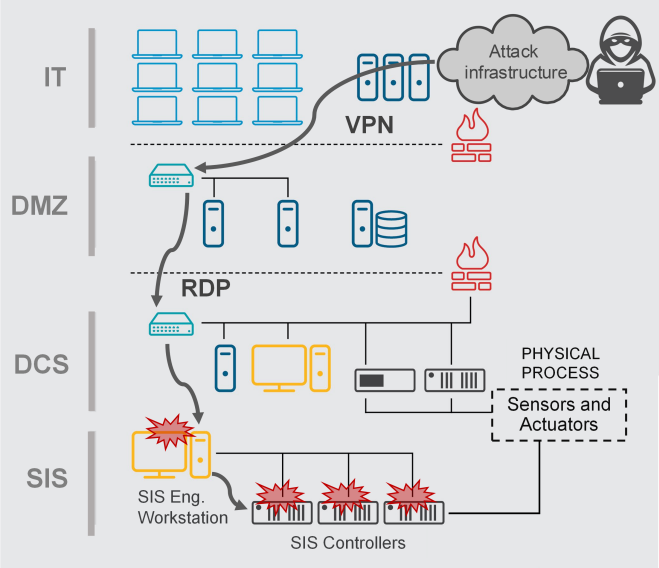


Figure 4.4: Attack path to safety systems in Triton attack [111]

Although the threat actor behind the Triton attack has shown that compromising individual industrial facilities is possible given sufficient time and effort, the execution of such operations does not scale easily. It took the attack team 12 months to obtain access to the assets needed in the target facility. In March 2018, the U.S. Cybersecurity and Infrastructure Security Agency (CISA) published an advisory about threat actors conducting multi-stage intrusion campaigns which used so-called “staging targets” to obtain access to the intended targets [104]. The staging targets are typically trusted third-party suppliers and partners with relatively insecure networks that are selected because of their pre-existing relationships with the intended targets. The threat actors have been using these third-party networks to host attack tools and eventually pivot into the infrastructures of their intended critical infrastructure targets. The advisory does not elaborate on whether the attackers were interested in specific industrial organizations or were acting opportunistically with the objective of obtaining access to as many industrial organizations as possible. One downside

of such intrusion campaigns is that initially unintended parties can become opportunistic targets and collateral victims simply because of being customers or partners of "staging targets".

Supply chain attacks in the industrial domain can be traced to at least 2012 with examples being the complete compromise of SCADA software vendor Telvent [116] and the distribution of malware that was digitally signed by the whitelisting software vendor Bit9 [115]. In one of the most successful supply chain campaigns, Havex, the attackers trojanized legitimate software update installers on at least three ICS vendor web sites [103, 71]. Once installed, Havex attempted to call back to a Command and Control server (C2). When successful, malware would send back information gathered about the victim environment and download further attack instruments (see Section 2.4). While Havex appeared to only conduct reconnaissance activities, its level of access to the control networks would have allowed an attacker to interact with the physical process and execute cyber-physical attacks of high precision.

Due to their size, complexity and interdisciplinary nature, many industrial environments rely on a large number of integrators, engineering contractors and subcontractors throughout the entire lifecycle of their operational sites. These organizations hold a vast amount of engineering documentation such as plant designs and network diagrams, calculations and project files which are indispensable when designing a targeted cyber-physical attack making these organizations valuable attack targets. In the past threat actors have already conducted large-scale operations to obtain access to such organizations [35].

Despite growing awareness of cyber security threats and risks to the process control infrastructures, thousands of ICS assets are being exposed to the Internet without proper security measures [160, 10] helping attackers to gain access to the control devices and applications. Researchers routinely conduct industry- and country-focused as well as worldwide scans to quantify the online exposure of ICS devices and identify exploitable and weakly protected systems [235, 196, 33, 162]. Among frequently identified issues are poor configurations, usage of insecure protocols, missing security patches and insufficient level of access control (e.g., absent or default passwords). The attackers could use public device search engines, e.g. [152, 32] or develop their own scripts to search and identify ICS devices of interest. For instance, threat actors may search for PLCs with specific protocol support [136], exposed control panels or HMIs [99] or complete systems such as wind turbines [215]. Once the devices of interest are discovered, threat actors can further use open-source intelligence (OSINT) methods to geolocate devices and identify host organizations.

4.3.2 Discovery

This stage refers to discovering information about the physical process of interest from the documentation and through traditional network-based reconnaissance activities. In the context of a production site, the attacker must reconstruct the layout and configuration of the plant and how it carries out its functions, find process design weaknesses or flaws which could be taken advantage of and identify the exploitability of the related ICS components. Below we describe the types of activities at the Discovery stage.

Industrial facilities are usually uniquely customized and even operational sites within the same industrial organization can be very dissimilar. This does not only relate to the sector-specific aspects and architecture of the installation but also to the proprietary configuration and operational methods. Therefore offensive capabilities in the process control domain

exist in relation to a specific target, which must be scoped by the attacker upfront. Without detailed knowledge about the victim environment, it is unlikely that an attacker can achieve more than a nuisance. For instance, simply trying to destroy a process by overheating a tank will probably only result in exercising the emergency shutdown logic and the pressure relief valves.

There are several data sources that describe the production process. The attacker may first study general information on the chemistry, kinetics, and thermodynamics of the physical processes of interest. This can be done by consulting the open literature as well as proprietary information of process design companies. Regulatory filings describe the inner workings of safety or environment-related subsystems. Engineering diagrams may be stored in cloud-based change management systems. Operator consoles offer a human-readable overview of process design and operating constraints [104].

Espionage and OSINT. Most data related to industrial processes are proprietary in nature and are therefore not easily available to a wider audience. This has been realized by attackers almost two decades ago manifesting in non-stop massive espionage campaigns against ICS-related research institutions, industrial organizations and suppliers. Early examples are [9, 157, 42, 179] and lately [104, 184, 34]. The attackers appeared to be especially interested in design documents, formulas, manufacturing processes and research materials, engineering diagrams, HMI screens and other related information.

At the same time, with the wider usage of the Internet, more and more information became available on the web, presenting new challenges for critical infrastructures. Open-source intelligence, which refers to legal information gathering from public sources, became a popular method of intelligence collection. Konstantinou et al. [114] showed how public information can be utilized in order to identify critical regions of the smart grid and which specific electrical utilities need to be targeted to achieve large-scale power outages. Further, the authors provided examples of public sources which may reveal the technologies (e.g., communication protocols, models of controllers, etc.) employed in the target utilities. While it remains uncertain how a specific substation was selected to carry out the Industroyer attack on the Ukrainian power grid, there is a possibility that the affected substation was targeted because of its detailed network diagram, an HMI screen with the electrical connections and switchgear locations were publicly available on the web page of its contractor [67]. Additionally, several official videos [225, 51] showed the exact brands and models of the control and protective equipment used in the substation together with an overview of the field equipment and the supporting OT infrastructure.

Virus Total [220], an online public service for analyzing files for malicious content, became a popular resource among organizations and individuals to evaluate suspicious files. However, in some cases, service users upload sensitive and even confidential files without understanding that other users can search for and download uploaded files. In 2018 the security company FireEye conducted an analysis of the OT-related files uploaded to Virus Total and discovered a vast amount of highly-sensitive engineering, contractual and purchasing documentation [237]. In a similar effort, security company Otorio discovered thousands of project files (control logic) uploaded to Virus Total [62]. Besides actual process automation logic, project files may contain network configuration, screen definitions, hardware and software configurations, and therefore could be of great assistance in developing targeted attack payloads. To confirm their hypothesis, the researchers created tools for automated parsing and analysis of project files and illustrated the process of deducing

tailored damage payloads.

Information gathered through espionage and OSINT efforts can be both target-specific and broad to support a general learning curve. These activities can be carried out as a preliminary step that precedes targeting as well as happening in parallel to or after the Access stage.

Reconnaissance. Similar to intrusion into IT networks, the attacker needs to discover the target plant from the networking perspective, enumerate important file shares, HMIs, engineering workstations and control equipment. While many of the reconnaissance objectives can be achieved with traditional IT reconnaissance means, discovery and enumeration of the ICS assets require specialized tools which collect relevant asset information such as shown in Figure 4.5. In the past years researchers published several public repositories with ICS reconnaissance tools for popular brands of control equipment [58, 211, 199]. Also, threat actors were observed to design such capabilities for their attack campaigns with an OPC scanner in the Havex attack and asset discovery modules in the Industroyer and Triton exploitation frameworks being a few examples.

```
Order Code           Module Type Name    Firmware Version    Module Name    Serial Number    Rack/Slot
-------------------  ------------------  ------------------  -------------  ---------------  -----------
6ES7 412-2EK06-0AB0  CPU 412-2 PN/DP     V 6.0.3                            SVPF126xxxx      0/3
```

Figure 4.5: Results from the S7comm asset discovery module from the ICSSPLOIT toolkit [58]

Process Comprehension. Plants can be highly proprietary. Even if restricted to the exact same chemical process, different engineers would make different choices when selecting vendors for pumps and valves. That, in turn, would influence the sizing and placement of various pipes and holding tanks. That in turn would change the way the plant is controlled, the design of the control loops and the programming of the controllers.

There exists little documented knowledge with regard to the collection of information about a target process, its analysis, and inclusion within a targeted attack. **"Process comprehension"** is a term coined in [86] to describe the understanding of system characteristics and components responsible for the safe delivery of a service (e.g., treatment of water). This includes all relevant physical and computational attributes. This is the most difficult and time-consuming activity in the Discovery stage. The inability to obtain vital contextual information can become a major hurdle for attackers and result in the inability to compile a final attack payload.

Among the most essential plant design documents are Piping and Instrumentation Diagrams (P&ID), Flow Diagrams, One-Line Diagrams which often contain information on safety interlocks, Cause & Effect Diagrams, Cable Schedule Diagrams, Project Interconnection Diagrams and others. Instrument I/O Lists contain the listing of instruments with vital information such as the type of instruments and their location in the process, range of set points, instrument tags and loop numbers, service descriptions, etc.

In the past such documentation was stored in form of standalone files on data servers, in change management systems and in the engineers' personal computers. The attacker would need to compromise and search multiple hosts to collect the necessary information. In response to the growing complexity of control and field instrumentation, demand for

	Instrument Datasheet			PRESSURE TRANSMITTER						
1	Tag No.			01-PT-510		Manufacturer: Yokogawa				
2	Loop Service			Reactor01-R-510		ModelNo: EJA110A				
3	P&ID No:	Line Number		01-220-004	01-P007-80-B1					
4	Area Classification			Zone 1, GrIIC, T3						
5	Ingress Protection			IP 57						
6	PROCESS CONDITIONS									
7	Fluid	State		HC	Vapour	Process Design Conditions				
8	Pressure	Normal	Max	1450 Kpag	1650 Kpag	Design Pressure	Min/Max	-	-	
9	Temperature	Normal	Max	100 C	149 C	Design Temperature	Min/Max	-	-	
10	TRANSMITTER									
11	Instrument Range	LRV / URV / Units		-0.5	14	MPa	Output Signal Type	4-20 mA		
12	Callibration Range	LRV / URV / Units		0	1700	KPag	Protocol/Version	HART		
13	Accuracy			+/- 0.75% of span			Burnout	Downscale		
14	Elevation	Supression		-		-	Installation Style	Horizontal Impulse		
15	LP Proc. Conn.	HP Proc. Conn.		1/4" NPT-F(Vent to Atm)		1/4" NPT-F	Mounting	Via Manifold		
16	Conduit Connection	Power Supply		2xM20 Female		Nominal 24VDC IS	Other	See Note 6		
17	Housing	Paint		Low Copper Cast Alu		Epoxy Resin Coating	Tag Plate	SS304 Permanent		
18										
19	ELEMENT									
20	Element Type	Element Material		DP Capsule		SUS316L	Temperature Limits	Min/Max	-40 C	-
21	Measurement (Gauge / Abs / Vac etc)			Gauge			Pressure Limits	Min/Max	-	-
22	Body Material	Body Rating		SCS14A		16 Mpa				
23	Bolts	Seals		SUS630		Teflon Coated SUS316				
25	Fill Fluid			Silicone Oil						

Figure 4.6: Instrumentation and Control System engineering software [13]

operational efficiency and digitalization trends, ICS vendors introduced dedicated applications for centralized management of process engineering efforts and documentation storage. Figure 4.6 shows a user interface for a pressure transmitter in an Instrumentation and Control Systems Engineering application from Aveva. The consolidated data include instrument service (Reactor control loop), tag number in the control logic, process conditions in which it is intended to be operated, instrument calibration and range, vendor and model, associated equipment and instrument tag on the P&ID diagram and even a user name of the maintenance personnel who has permissions to modify the instrument configuration. An example described in [86] illustrates non-trivial challenges faced by the attacker in achieving the level of process comprehension needed for achieving tailored physical impact by means of the network- and host-based MITM attacks. Access to engineering applications that concentrate large volumes of data on the process infrastructure would significantly reduce attacker efforts in locating and collating the information required for process comprehension.

4.3.3 Control

Some information about the process can be studied statically, but other information can only be determined empirically via dynamic testing. No engineering diagram is detailed enough to accurately predict the travel time of a shock wave down a pipe to the accuracy needed to, e.g., set up a resonance wave to cause a water hammer effect (see Figure 2.24). Hence such data must be extracted from the live process. The control stage studies what actuators of interest do and the side effects of their manipulation.

In dynamic systems such as cyber-physical systems, the values of process variables change with time according to the laws of physics. The transitioning of a process from

one state to another is in most cases not instantaneous and adheres to the well-known fact that "things take time". In the control stage, the attacker tries to discover the dynamic behavior of the process which can be described in the form of a simple differential equation $dy/dt = f(y, u)$, where u is an independent variable and y is the dependent variable which are related by cause-and-effect relationships.

Most physical processes are inherently non-linear systems system in which the changes in the output are not proportional to the change of the inputs. Depending on the conditions, changes in process variables over time may be unpredictable and counterintuitive to a non-expert observer as can be seen in Figure 4.7. This is why making a process misbehave in a predictable way can be a non-trivial task.

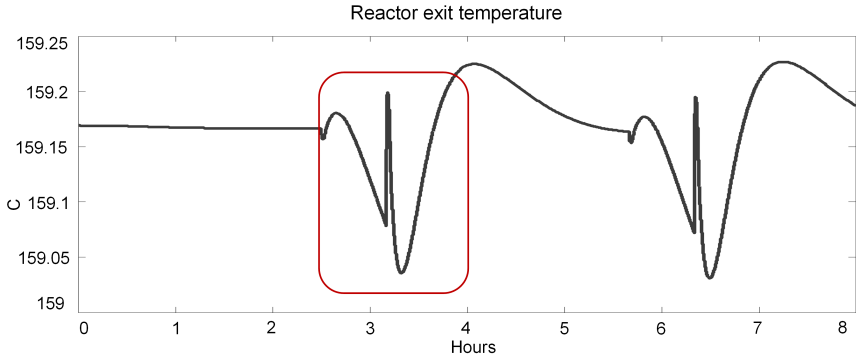


Figure 4.7: Nonlinear response of a control loop

> In general, even with a comprehensive knowledge of the exact plant configuration, it is not always possible to correctly determine whether a certain process response is related to a fundamental property of the plant or is a result of a specific attack parameter choice.

An attacker facing an unknown process is challenged by a lack of knowledge about the tuning and responses of control loops. Ultimately, the attacker may apply trial-and-error process probing, where errors should not trigger alarms as they may attract the operator's attention. To date, there are no well-established solutions to the problem of how to probe a given control loop with unknown responses without raising alarms. A rare exception is a work by Winnicki et al. [233] which offers a general approach to stealthy control loop testing. Note that simply knowing operational and safety thresholds is not sufficient as the non-linear responses of control loops can trigger alarms when not expected. Therefore, understanding basic factors that influence the dynamic effects of process behavior as shown in Figure 4.8 is fundamental to process control analysis. Obtaining such information is part of the process comprehension process at the Discovery stage.

Many safety accidents have previously happened because of the late identification of critical process degradation and/or wrong operator response. Human operators react to alarms. One of the key elements of alarm activation configuration is the time needed for the operator to respond to an alarm (Figure 4.9). Once the alarm is acknowledged by the operator, the clock is started. The operator has a predetermined time window to understand

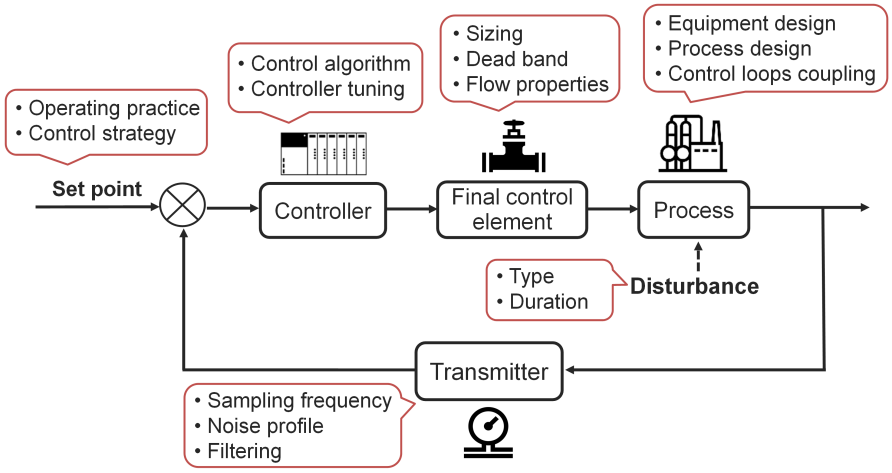


Figure 4.8: Parameters which influence control loop response

the alarm and select the course of action. Malicious process manipulation may invalidate the assumptions about the time needed for the operator to correctly identify corrective actions. Discovery of alarm activation and operator response timing parameters is part of the control stage. In those situation when the damage happens before operator may complete corrective action, the attacker does not need to spoof process.

Attack objectives may require the attacker to push a process beyond its normal operating conditions. Testing the performance of the control loop in abnormal conditions is essential for attack success. Every controller is configured, tuned and tested to perform within certain operational conditions and range of input signals, and may not be effective in different operating conditions. For instance, the dynamic behavior of a process in a temperature

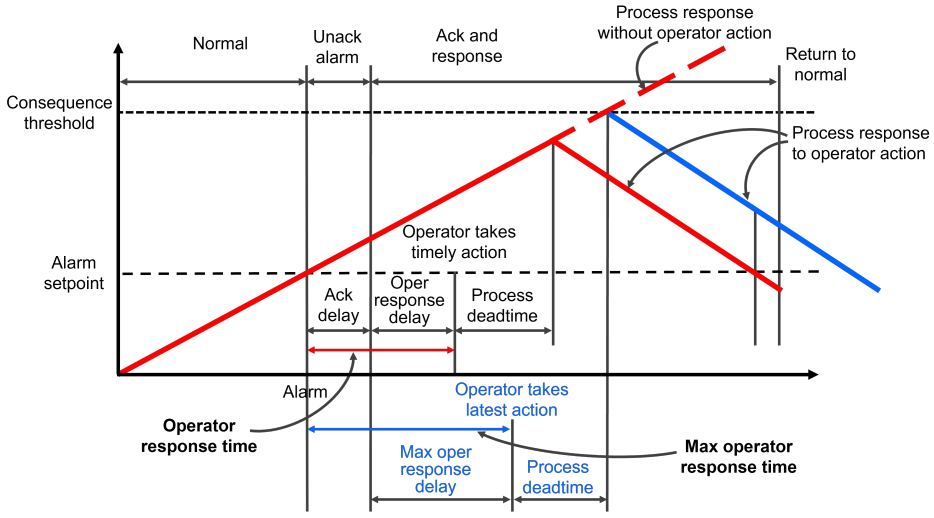


Figure 4.9: Calculation of the response time for the operator, based on [165]

range from 140 to 150°C may be different from when it is further heated to 160°C. If the control algorithm had never been tested for its performance in atypical temperature ranges, its control commands may result in erratic and even out-of-control process behavior.

Control loop coupling is another challenge that adds uncertainty to interpreting the process response. For two control loops to operate successfully in tandem each loop must "know" what the other is doing. Otherwise trying to achieve their respective objectives the two control loops may act against the interest of the other. This phenomenon is known as loop interaction. Some control loop interactions occur naturally as a result of their physical and chemical makeup. Some loop interactions may arise as a consequence of process design. Typical examples are heat integration and recycle streams which create the potential for disturbance propagation and the alteration of dynamic system behavior.

In continuous processes even distant parts of a process may be connected together in physical relationships. Increasing the temperature in a reactor may distort the chemical reaction. Improperly reacted chemicals will propagate downstream in form of imbalanced flow composition and cause irregularities in the refining section of a plant. The attacker thus needs to develop knowledge of all side effects of process manipulation as well as understand the extent of process upset propagation. In essence, the attacker needs to "reverse engineer" the physics of the plant and determine a model of its dynamic behavior. Given the size and complexity of industrial plants, manual acquisition and systematization of the dynamic process behavior can be prohibitively arduous. A semi-automated method for identification, cataloging and visualization of process dynamics by capturing sensor readings on-the-fly is proposed in [233]. Figure 4.10 shows a proposed visualization of the plant-wide response to a crafted impulse in the example of the Tennessee Eastman process. The purpose-built algorithm extracts behavioral patterns from sensor measurements with the help of lightweight data approximation procedures paired with estimation methods.

> The ability to control the actuators does not guarantee the ability to achieve control over the process. Adjusting one part of the process for malicious purposes may aggravate the process state in other parts of the plant.

For instance, it may be possible to turn off a pump, but as a side effect the pressure would build up quickly in an upstream pipe rendering the process inoperable. Additionally, an action taken in one logical layer of the cyber-physical system model might be prevented by a system design in another layer. Instructing a pump to run while the flow is off with a network packet in the cyber layer may be prevented by a hardware interlock in the physical layer. Note that such physical protections are not always specified on the engineering diagrams. The control stage also involves the study of timings. For instance, if the damage occurs in seconds, a safety shutdown minutes later will not stop the attacker.

The control stage is completed when the attacker identifies at least a minimum set of controls that can be applied in the Damage stage to obtain attack objectives. Note that successful control over the process will not necessarily result in the attacker's ability to cause the desired damage. This will be further discussed in the next section.

4.3.4 Damage

Once the attacker understands the process and how to control it, she needs to decide on the specifics of achieving her attack goals. It often requires the input of subject matter

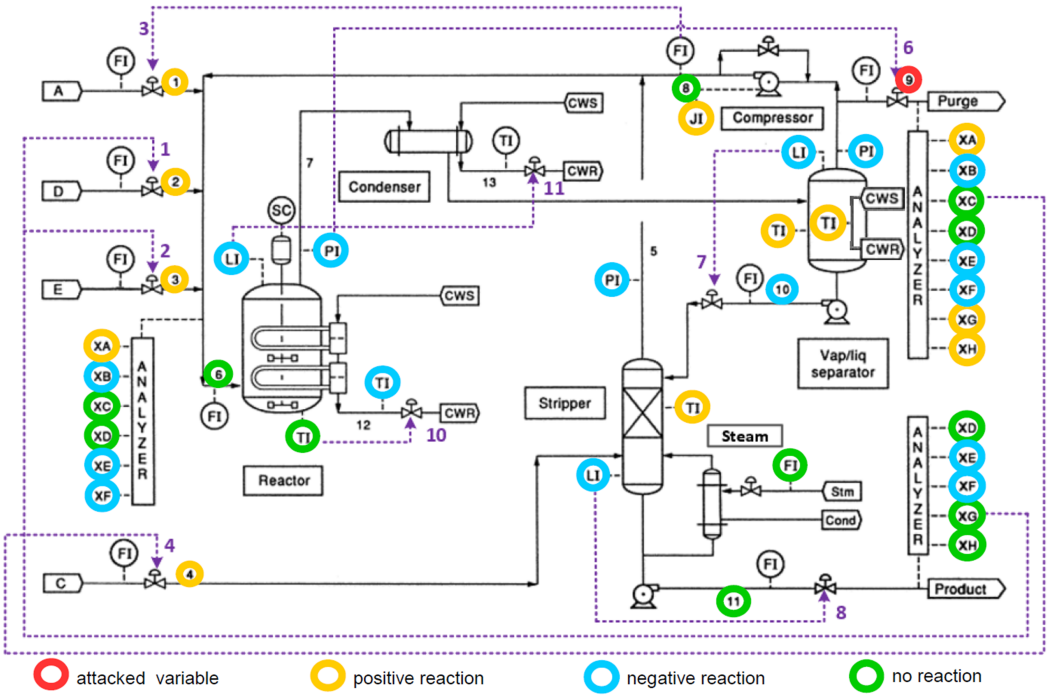


Figure 4.10: Visualization of the plant-wide process response to a manipulation of the valve [233]

experts to assess the full range of possibilities. However, physical damage scenarios may not come easily even to a process engineer's mind. Engineers asked to attack a process tend to come up with what is known as "salty cookie" scenarios. A group of engineers asked to attack a cookie factory all hit on variations of putting too much salt in the cookies so they would become unpalatable. However, in a real-world situation, the actual damage to the food factory was caused by the simultaneous presence of too much product in the pipes and a failed emergency flushing system. As a result, the pipes got clogged and had to be physically replaced after both water pressure and chemical means of clearing them failed.

Accident data can be a good starting point when brainstorming ideas for causing damage. If a particular type of process has gone wrong in the past by accident, it stands to reason that an attacker may be able to make the process fail in a similar way by design. There is a number of freely available such as [190], commercial or membership-based such as [37] incident databases with detailed investigation information.

Some of the damage scenarios are straightforward such as operating rotating equipment at its natural frequency at which a mechanical resonance and equipment vibrations may occur. Vibrations significantly reduce the expected equipment life span and lead to equipment breakage. The skip frequency parameters in the Variable Frequency Drives are used to set up a frequency band through which the drive output may pass, but never continuously operate in it to protect equipment from harmful operating conditions. Many VFDs allow a remote user to both read and modify the skip bands without authentication (Figure 4.11). The attacker strategy for causing vibration damage is quite simple: read the skip band

information from the VFD, modify the skip band value to a new range, and set the output frequency of the drive to the center of the previously-observed skip band, i.e., to the dangerous natural vibration frequency. Many VFDs allow this complete set of changes to be made while the equipment is still running.

Code	Name	Logic address	Access	Type
TDC1	IDC injection time	16#28A2 = 10402	R/W	UINT(Unsigned16)
JOG	Jog assignment	16#2B66 = 11110	R/WS	WORD (Enumeration)
JGF	Jog frequency	16#2B67 = 11111	R/W	UINT(Unsigned16)
PS2	2 preset speeds	16#2C89 = 11401	R/WS	WORD (Enumeration)
PS4	4 preset speeds	16#2C8A = 11402	R/WS	WORD (Enumeration)
PS8	8 preset speeds	16#2C8B = 11403	R/WS	WORD (Enumeration)
SP2	Preset speed 2	16#2C92 = 11410	R/W	UINT(Unsigned16)
SP3	Preset speed 3	16#2C93 = 11411	R/W	UINT(Unsigned16)
SP4	Preset speed 4	16#2C94 = 11412	R/W	UINT(Unsigned16)
SP5	Preset speed 5	16#2C95 = 11413	R/W	UINT(Unsigned16)
JPF	Skip frequency	16#2C25 = 11301	R/W	UINT(Unsigned16)
PIF	PID: PI function feedback assignment	16#2E7D = 11901	R/WS	WORD (Enumeration)
TLS	Time limited speed (LSP)	16#2DB5 = 11701	R/W	UINT(Unsigned16)
FBS	PID Feedback scale factor	16#2E7F = 11903	R/W	UINT(Unsigned16)

Figure 4.11: Tunable parameters of a VFD [20]

It is important to note that cyber-physical exploitation is not necessarily achieved by delivering bits of malicious exploit over electronic media. Consider an electrical analog motor. Is it not "hackable" in the traditional way as it does not have a programmable element in it. However, engineers use Digital Maintenance and Test Equipment (M&TE) for motor configuration, calibration and diagnostics. M&TE includes firmware, which is potentially exploitable. Thus, it may be possible to place malicious code on the M&TE used when conducting a motor test. This code could be created so as to report a bad motor as being good and/or a good motor as being bad [218]. As a result, a facility may incur replacement costs before the end of the useful life of a motor, or downtime costs from an unhealthy motor failing during operation.

Not all potential damage scenarios and their successful realization will result in an effective attack. Consider a water treatment plant [154]. The process begins with the inflow of raw water to be treated through an inlet pipe. Firstly, the water is treated with chemical disinfectants such as chlorine, sodium hypochlorite and others to eliminate pathogens such as bacteria and viruses. In the next step, the water is filtered with ultrafiltration (UF) membrane filters to eliminate unwanted inclusions. Filtered water is de-chlorinated using ultraviolet (UV) lamps and then fed into a Reverse Osmosis (RO) system for final filtering. The filtered water from the RO unit is supplied to consumers and the reject water is stored in the UF backwash tank. A backwash process cleans the membranes in UF at preset times or conditions.

We assume that the attacker's goal is to damage the UF filter which would result in an extended water supply interruption and monetary loss due to filter replacement. Filter membranes are sensitive to several hazards such as grease or overpressure. This information can be found in the UF filter operating or service manuals, e.g., [92, 91]. While increasing the number of grease particles in the water via a remote cyber attack can be difficult or impossible to achieve, creating overpressure conditions appears to be a more realistic

scenario. The attacker can turn to the engineering documentation of a water plant to work out an attack that causes overpressure. For instance, it can be discovered from a P&ID diagram such as found in [181] and highlighted in Figure 4.12 that there are two water inflows in the filter section: treated water flow and backwash flow. It is logically understood that these are mutually exclusive flows that should never be allowed to flow at the same time. However, the attacker may find a way to overcome the restraining interlock and make two flows run through the filter simultaneously, causing high pressure inside of the filter.

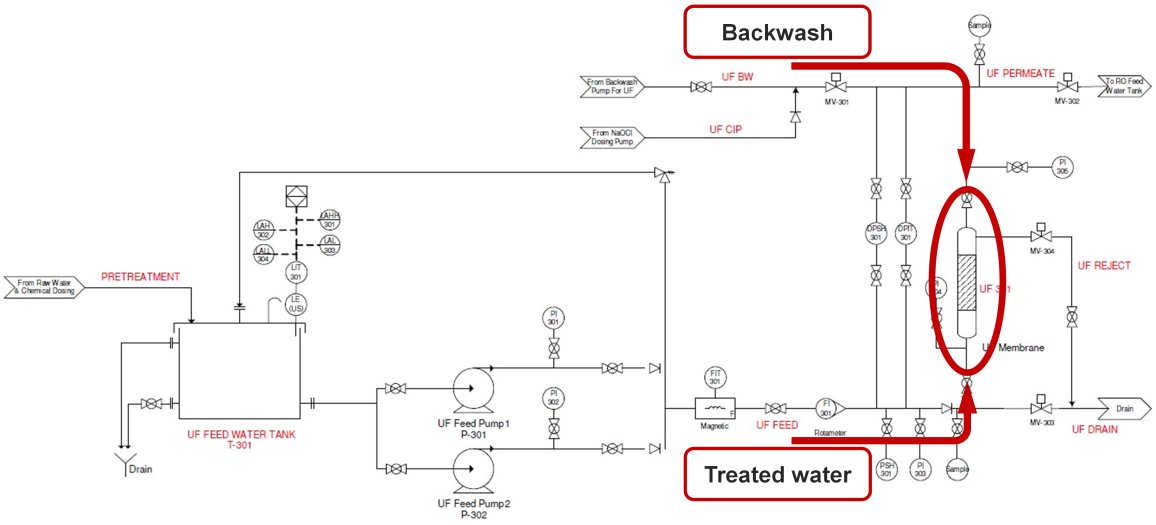


Figure 4.12: P&ID of the ultrafiltration section of a water treatment plant [181]

To practically implement this potential attack scenario the attacker needs to analyze the PLC code. There are three conditions that can trigger the backwash process, each guided by a state machine in a PLC: a preset timer (every 30 minutes), UF filter differential pressure (DP) $\geq 40\ kPa$ or plant shutdown. In collaboration with the iTRUST research center [180] we implemented an attack based on a preset timer and enabled the concurrent flow of both streams through the filter. The result of our experiment is shown in Figure 4.13.

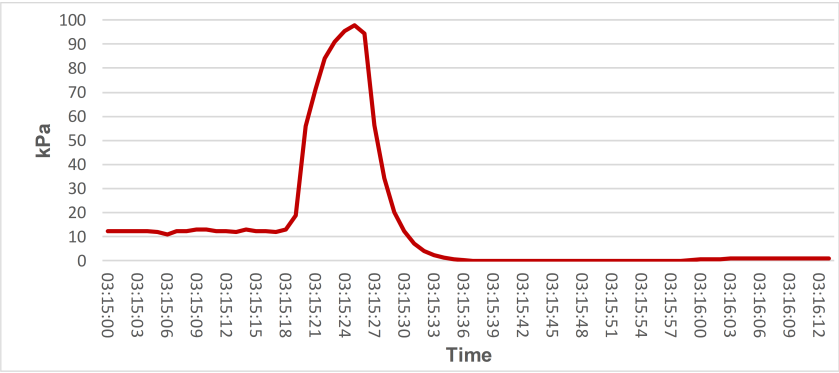


Figure 4.13: Differential pressure on a membrane in UF water filter

The normal average UF filter DP is $\approx 12-13$ kPa. The maximum differential pressure we were able to achieve was ≈ 98 kPa (1 bar). According to the filter documentation, this pressure is not enough to damage filter membranes. This example illustrates that the successful implementation of a potentially viable attack will not necessarily cause the damage intended. In such a situation the attacker has two options. The attacker may return to the Control stage and examine whether the flow properties (e.g., velocity) of both or any of the two flows can be modified to achieve the desired pressure in the filter. Alternatively, the attacker may return to the Discovery stage and work out new potential scenarios to disrupt the water treatment process.

An attacker is not restricted to a single scenario and different attacks are not mutually exclusive. It may be possible to attack one part of the process and then later attack another part based on the success of the first. An attacker may never have full knowledge of the process and the environment. It may also be impossible for the attacker to fully test her code before deploying it. Chaining together multiple attacks into a single payload maximizes the chance that one of them will have the desired effect.

> Note that while we adhered to the usage of "damage" throughout the chapter, it would be useful to distinguish between **"damage"** and **"degradation"** attack strategies. Degradation is easier to achieve and therefore such attack scenarios pose a higher risk.

Prevent Response

The main task of the control system ecosystem is to keep the process state as close to its set points as possible. The attacker will inevitably move the process away from its optimum, "provoking" compensating reactions.

> During the attack, the control system and the attacker compete for control over the process. Therefore, the attacker may require to obstruct the control system from "performing its duties".

In this attack stage it is important to remember that the control loop is not a single control loop but rather a hierarchical structure of control loops as shown in Figure 4.14. The control ecosystem includes the control logic, process operators and all other applications which contribute to control decisions, e.g., advanced process control or equipment monitoring applications. The attacker may need to hide attack artifacts across all control loops to be successful. For instance, it may be sufficient to keep the process operator's attention away to prevent a timely response. However, process optimization applications may automatically act on suboptimal operational process data and, e.g., calculate new process set points or adjustments to relevant valve positions to balance out process inefficiency.

Approaches to preventing a control (eco)system response include but are not limited to techniques such as "record-and-playback" [129], spoofing sensor signal inside smart sensors [125], poisoning of Historians with pre-calculated data, various methods of alarm suppression such as alarm hiding, relaxation, shelving and others [230], avoidance of alarm activation and disruption of feedback and control communication channels. An attacker may have to use different strategies across different hierarchical control loops.

Process state and alarm information propagate not only vertically but also horizontally, affecting process variables both upstream and downstream. To identify the *extent* and

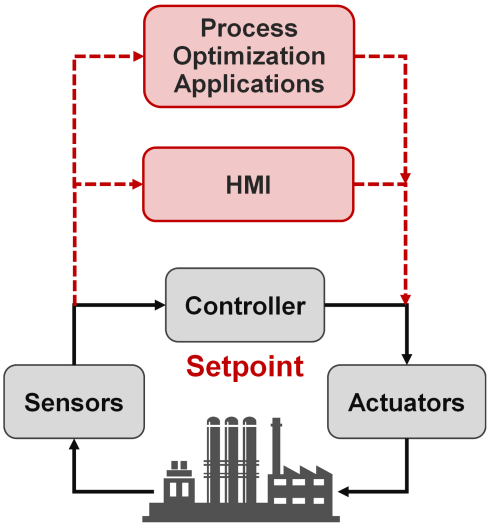


Figure 4.14: Hierarchical architecture of control loops

duration of process upset propagation, the attacker needs to see the process status and alarms for the related process unit(s) or even the entire plant. Good places to access such information could be historian applications. They are also often mirrored in near real-time to the corporate network so that they can be queried by various regulatory processes.

> Note that in the experiment with the UF filter described above, the maximum differential pressure was achieved in 8 sec. If the high pressure damages membranes instantaneously, such an attack may not even need concealment measures as the operators would have very little time to acknowledge, react and prevent the accident. However, if damaging the filter takes some time, the operators may intervene and the attack may fail. Therefore, *timing parameters* of the physical process (Figure 2.23), operator response (Figure 4.9) and data flow aggregation (Figure 2.18) need to be taken into consideration when designing a cyber-physical attack.

Obtain Feedback

There may be several competing attack designs with varying degrees of execution reliability. Therefore, the attacker may want to develop some kind of measure or metric to choose between them.

> To monitor and manage attack execution as well as measure attack success (or failure), the attacker needs to establish a feedback loop to observe the process under attack. This need is directly correlated to the core principle of process control that the process needs to be observable to be controllable (manageable).

This principle is also similar to the utilization of a Command & Control communication channel in attacks on IT/enterprise infrastructures. Depending on the attack scenario and scope of access obtained, the attacker may need to obtain a feedback loop between the

placed exploit and the process itself and/or coordinate attack activities between several implants. In this case process physics will be utilized as a communication channel. For instance, observation of state A in component B needs to trigger payloads X, Y, Z. This could be seen as a C2 mechanism for embedded implants based on process state detection. Figure 4.15 demonstrates the process change detection using CUSUM algorithm.

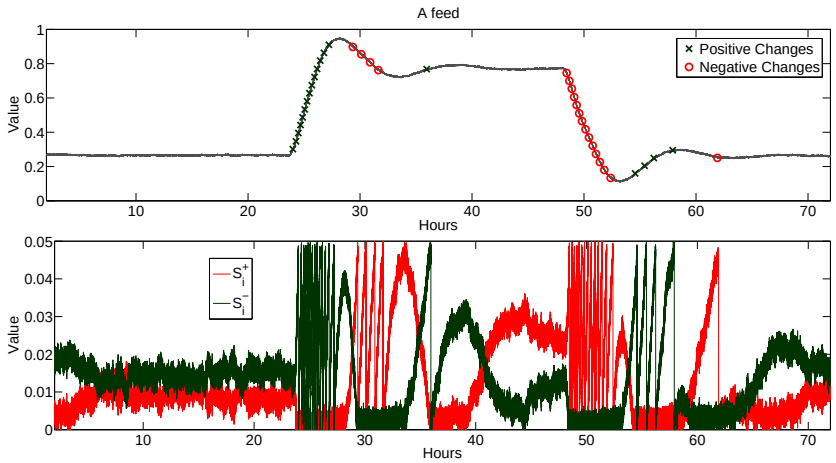


Figure 4.15: Detection of the plant state change

In the event of remote access to the target facility, the attackers may use familiar C2 capabilities to relay feedback information into their infrastructure on the Internet and to monitor and adjust attack execution in near real-time. Previously, threat actors showed the capability to obtain persistent remote access to SCADA servers [41] and safety systems [111]. A possibility to establish a covert C2 channel utilizing unused PLC memory in a distributed ICS/SCADA architectures has been shown in the research literature [117]. Figure 4.16 shows an architecture of a multi-stage C2-based feedback loop.

Figure 4.16: Architecture of C2 and feedback channels in cyber-physical attack lifecycle

The target plant may not have been designed in an "attacker-friendly" way and may not measure the values needed to monitor attack performance. As a result, the attacker may face the challenge of lacking sensors or other measuring instruments needed to monitor her attack. Additionally, process information may spread across disparate subsystems forcing the attacker to invade a greater number of devices. Generally speaking, there are two types of measurements available to the attacker:

- *"Technician" measurement* is a qualitative measure, e.g., whether some process value is decreasing or increasing. A so-called "sensor proxy" can be used to measure an aspect of the process where direct measurement is not possible. For instance, pressure (proxy measurement) in some cases can be used to infer temperature (measurement of interest). Similarly, the temperature in some cases can be used to estimate flow.
- *"Engineering" measurement* is a quantitative measure, e.g., by how much some process value is decreasing or increasing. When exploiting a process, attack scenarios requiring an engineering answer are harder to implement as they often rely on data unique to the plant and its current operating mode.

In the case when observation of the process state is not achievable via available measurements, the attacker may use alternative techniques which include but are not limited to the usage of proxy sensors (see Section 4.5.4), detection of process state as shown in Figure 4.15, reading the status of the state machine in the PLC code (mentioned above), building a process model, estimations and engineering calculations (see Section 4.5.4).

Figure 4.17 illustrates the execution of an implosion attack on a metal barrel. There is no "roundness" sensor to directly measure the success or failure of the attack, requiring finding alternative ways to monitor what is happening with the barrel. In this case a pressure sensor is used to detect the engagement of a vacuum breaker which opens an air vent and allows air into the system. Figure 4.17 shows how pressure starts to raise upon vacuum breaker engagement, signaling the unsuccessful outcome of the attack.

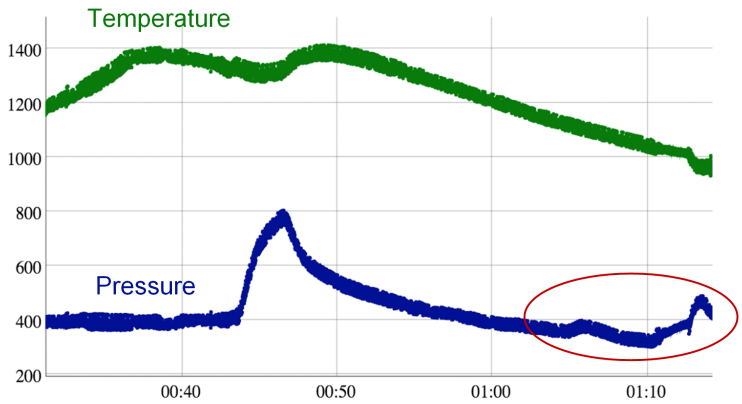


Figure 4.17: Indicator of an unsuccessful implosion attack [131]

4.3.5 Cleanup

In traditional IT attacks the goal is often to remain undetected. In most process control scenarios, this is not an option. If a piece of equipment is damaged or if a plant suddenly becomes less profitable, a team of experts will be dispatched to investigate. A cyber-physical attack changes things in the real world that cannot be obscured by, e.g., erasing or modifying log files.The cleanup phase is about creating a *forensic footprint* for investigators by manipulating the process, historical data and logs such that the investigation draws wrong conclusions. The goal is to get the attack blamed on operator error or equipment failure instead of a cyber event.

An example of a cleanup phase would be to show the operator a process out-of-control, making her take appropriate actions. When investigators ask the operator if she manually manipulated the process when it malfunctioned, she would answer in the affirmative. If malicious plant upsets are regularly timed so that they fall in a particular employee's shift, one might end up investigating the employee rather than the process.

The human operator is one of the key elements in monitoring industrial processes. Having a human operator in the control loop (Figure 4.18) turns the process control system from a pure cyber-physical system into a *socio-technical system* (STS). To take advantage of the operator's "vulnerabilities" the attacker needs to understand the specifics of the operator's job and act according to identified weaknesses in the operator's attention, judgment process or standard procedures he has to follow (see also Section 2.3.1 and 4.3.3).

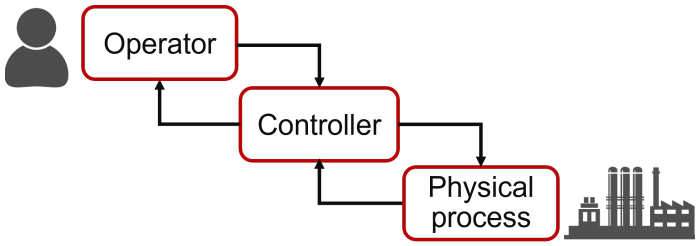


Figure 4.18: Socio-technical system

The investigator may be persuaded that a change is environmental in nature if attack patterns are timed to rainy or hot days. Replaced equipment is another good candidate for assigning blame. Most chemical plants are harsh environments and components fail regularly. If an attacker either intentionally makes a controller inoperable or waits for a component to fail naturally before beginning with her cyber-physical attack, the investigator may believe that process issues result from the suboptimal tuning of a new component instead of a cyber attack. The list of potential events that can be used as a decoy for an attack includes but is not limited to:

- Change in operating conditions: change of set point, change of raw material supplier, new equipment, etc.,
- Maintenance work – scheduled or unscheduled,
- Specific events: change in weather conditions, a particular operator on duty, etc.

An attacker may also hinder an immediate root-cause analysis of a malicious process upset by causing a so-called alarm flood [185]. An alarm flood is an event of too many alarms being raised in a short period of time, making recognition of the causal alarms difficult or impossible. Similarly, the attacker may intentionally bring about more critical alarms requiring immediate operator attention and with that taking their focus away from the ongoing attack.

4.4 Targeted Attack for Production Damage

We consider an attack scenario of a negative impact on plant economics over an extended period ("economic damage"). Figure 4.19 illustrates various cost elements in relationship to process efficiency. As can be seen, even insignificant deviations from the optimal operating conditions may impact plant profitability. Therefore even low-precision, non-strategic attacks can have an unwanted effect on plant operations. This is the advantage of economic attacks over high-precision physical damage attacks. Additionally, due to large factors which influence local process imbalances and overall plant profitability, it is much harder to identify root-cause of economic attacks.

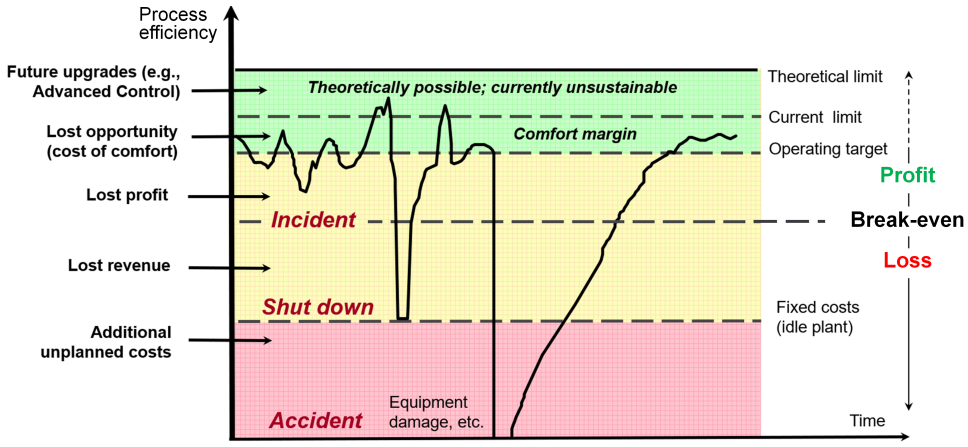


Figure 4.19: Relationship between process performance and plant profit [127]

> At the outset of attack planning, the attacker needs to consider key factors which influence attack design decisions.

One way to influence *plant profitability* is to make a process *inefficient* by inducing process disturbances and/or provoking control loops instabilities to increase maintenance efforts. However, creating loop instabilities can be a risky option as in the case of unskillful manipulations, the process can become completely uncontrollable and result in plant shutdown. Plant shutdown is risky for the attacker as it may instigate an investigation. For this reason, an attacker would want to remain in control of the process and to adjust disruptive actions over time.

To *persist* with her malicious objective, the attacker would want to attract as little attention from process operators as possible, e.g., by suppressing alarm activation on the operators' screens. However, alarm suppression adds complexity (and cost) to attack design. Out of this consideration we decided to manipulate the process without triggering alarms to save efforts on alarm suppression or process measurements spoofing. Another condition of persistence is avoiding attribution of the prolonged process misbehavior to malicious causes. This can be achieved by, e.g., timing the attacks to specific events in order to misdirect operators' suspicion as discussed in the previous section.

4.4.1 Preliminary Analysis of VAC Process

To illustrate the applicability the of cyber-physical attack lifecycle, we use the Vinyl Acetate Monomer process introduced in Chapter 3. Its plant model is sufficiently complex and includes known realistic components. This allows us to investigate the attack execution process close to the real-world conditions. Economic damage to a plant can be achieved in several ways. To narrow down the most effective attack scenario the attacker can first take a bird's-eye view on the plant structure. The VAC plant can be roughly divided into three parts: reaction, refinement and finished product output as shown in Figure 4.20.

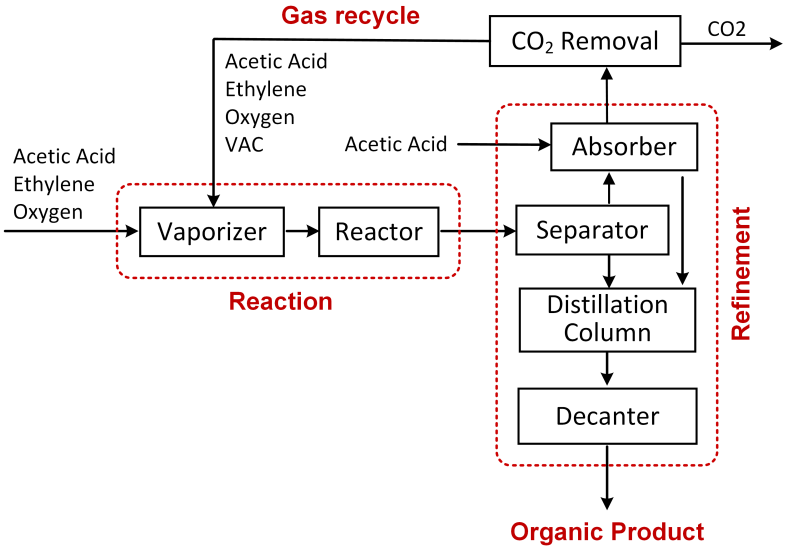


Figure 4.20: Simplified scheme of the VAC plant

Distilled products represent the most valuable commodities leaving a refinery or chemical plant. Therefore maximum economic damage could be achieved by destroying the pipe that carries the organic product into the storage vessels. This attack is certainly effective, but due to clear physical evidence of its consequences it will also be noticed and fixed quickly. For a prolonged damaging effect the attacker needs less conspicuous scenarios.The *refinement section* is responsible for VAC distillation to ensure a final product meeting rigid specifications. The refinement process consists of multiple units so the attacker has many opportunities to tamper with the process, but the operator also has many opportunities to notice changes and take compensating actions. Moreover, the operator may send an impure product for an extra cycle of refinement. In contrast, upsetting the *reaction process* in the reaction unit reliably yields reduced production of a useful product. This is why we focus attack efforts on the reactor unit.

Most factors influencing the reaction process fall under the following three categories:

- Catalyst deactivation caused by high temperatures and other factors,
- Reduction of reaction rate due to wrong ratio and/or preconditioning of the chemical components,
- Reduction of the primary reaction due to wrong material and energy balance.

While the catalyst deactivation attack vector could potentially offer the most reliable

reduction of production rate and the largest potential for a negative economic impact, we were not able to realize this scenario in our simulations. Further details are given in Section 4.6. Therefore we were left with two scenarios for reducing the reaction rate. Both scenarios naturally occur as a result of manipulating actuators in the refinement session and bringing the process out of the optimal production conditions.

4.5 Achieving Attack Objectives

After deciding on a general approach to achieving the attack goal, the attacker proceeds with designing a detailed attack plan. Following the stages described in Section 4.3, the attacker develops a set of attack instances which can be scripted into individual payloads to form an overall attack framework/tooolkit or included in a single final payload.

4.5.1 Access

When working with a model, it is easy to look at a process environment as static and isolated. When considering an attack on a VAC process somewhere in the world, one must remember that this is a real facility with real-world needs. As any other production facility, a VAC plant would rely on dedicated teams of subcontracting workers, engineers, consultants vendor- and third-party service providers to support plant operation, maintenance and troubleshooting all year around. Additionally, a large number of data needs to be exchanged between corporate and third parties networks for predictions, equipment maintenance and hardware, software and security updates. An attacker can take advantage of any of the above mentioned interactions to obtain access to the plant, specifically to the reactor unit.

4.5.2 Discovery

The chemical approach to VAC manufacturing is not a trade secret. There is a wealth of information on the process itself and how it is typically implemented in a factory. We assume that the attacker has already obtained knowledge about the process such as presented in Section 3.1. Since we are working with the process model which does not include networking components, the discovery activities are directed at the process itself. Specifically, we concentrate on discovering measurements and actuators available to the attacker at the reactor unit as well as on a high-level understanding of the inner workings and configuration of the target plant.

Six sensors, XMEAS{1-6} and one flow composition analyzer are available to the attacker at the reactor unit, vaporizer{P;L;T}, heater exit{T}, reactor exit{T;F}, and molar concentrations of the seven chemical elements in the reactor feed stream from the analyzer, XMEAS{37-43}. Specifically, the O_2 concentration is used to monitor the hazardous conditions related to the explosivity of ethylene in the presence of oxygen.

There are seven degrees of freedom XMV{1-7} available for control in the reactor unit, three reactants fresh feeds {O_2;C_2H_4;HAc}, two valves to control vaporize{heater; vapor exit}, reactor preheater valve, and steam drum valve to control the reactor temperature. The location of the valves in the reactor unit is shown in Figure 4.21. All valves except XMV(3) control effects within the reactor unit itself. However, part of the acetic acid inflow controlled by XMV(3) is also sent into the vaporizer (reactor unit) and into the absorber (refinement section). Also, acetic acid comes from a supply tank. It means any attack on this feed will be buffered by the acetic acid holdup in the tank. Additionally, we discovered

a disparity between process documentation and its software implementation. Specifically, inflow of ethylene XMV(2) is used to control the pressure in the absorber and not the pressure in the gas recycle loop. Thus any manipulation of this control loop will have effects on both the reactor and refinement sections. For the attacker this poses an additional requirement to closely monitor process state in the refinement section when manipulating XMV(2).

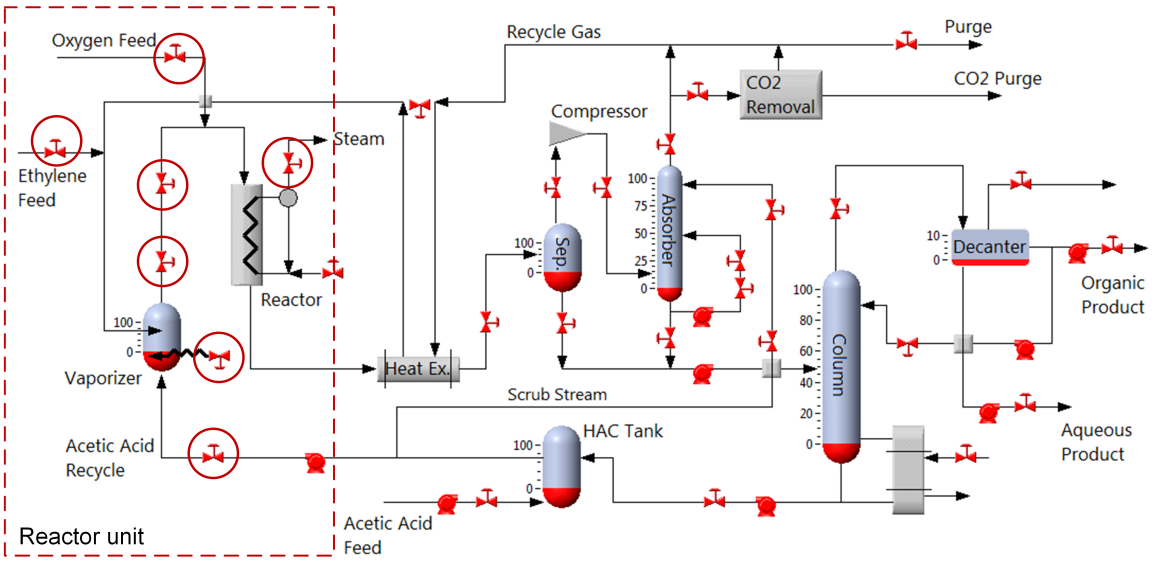


Figure 4.21: Location of control valves in the reactor unit of the VAC plant

A sensor measuring a process variable important for safety or operating constraints will have an alarm or interlock set at certain operating ranges. A quick search of chemical engineering journals turns up information on such constraints and specifics of the VAC process (see Section 3.1). Exact numeric limits can be discovered in operator screens, controller logic, one-line diagrams, etc. The basic plumbing of the process can be understood from its flow diagram (Figure 3.1). In a real world scenario, this information would have to be gathered from configuration files and other sources as described in Section 4.3.2.

4.5.3 Control

At the control stage the attacker explores dynamic process behavior by triggering system responses. Typically, digital controllers are designed based on process models and with very few exceptions the designs begin with the specification of some desired closed loop properties. Collecting such information in advance at the Discovery stage may improve attacker understanding of the observed process reactions and accuracy of system identification.

Figure 4.22 illustrates VAC process response to an attack on the heater exchange and the propagation of the resulting disturbance across the plant. Even though the overall plant-wide process response can be considered as consistent, the responses of individual control loops are very dissimilar. Some of the simulation results of the VAC plant to setpoint changes are described in [148]. It was specifically mentioned that some dynamic behaviors of this process are not intuitively obvious. As discussed in Section 4.3.3, this is due to the

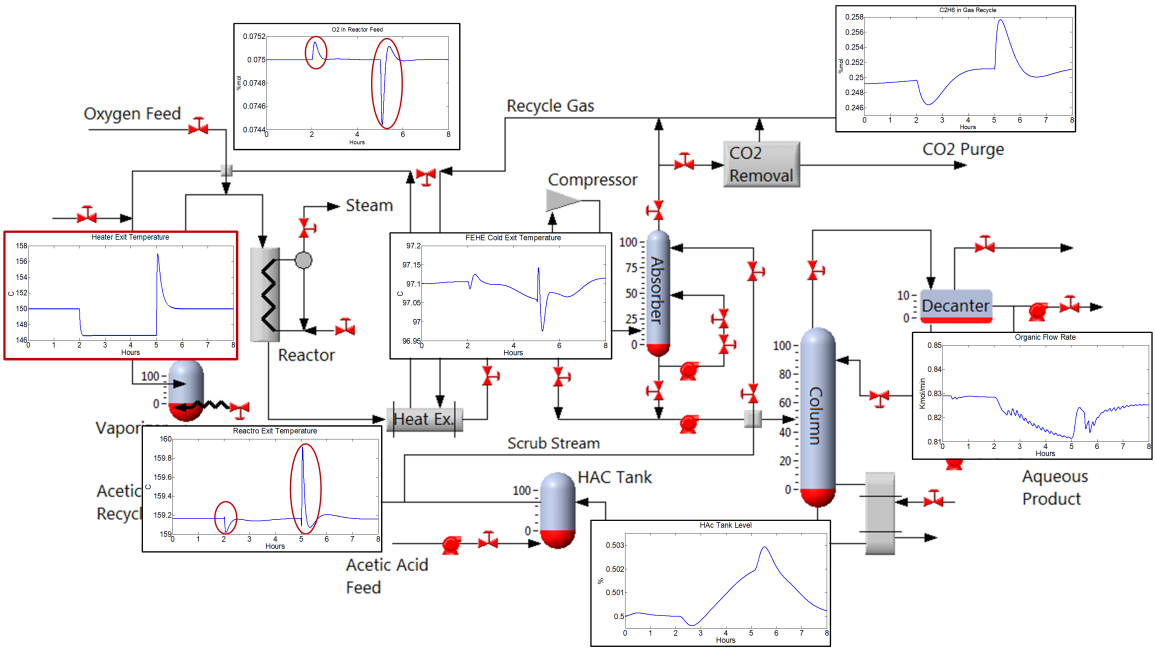
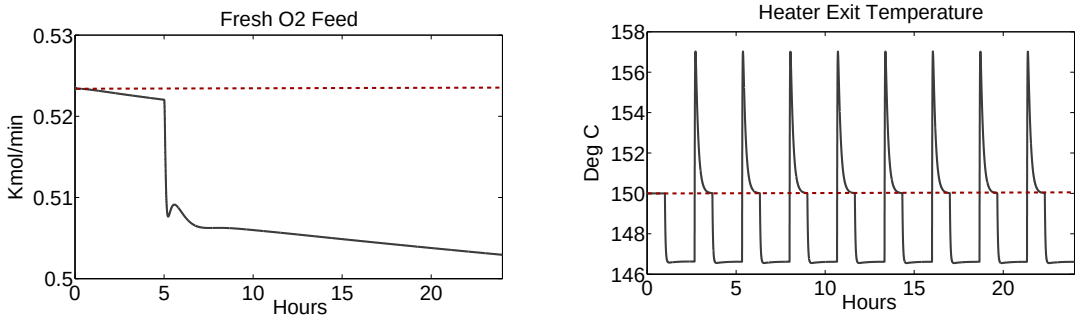


Figure 4.22: Illustration of process response across the entire plant in VAC process

unique hardware plant setup, configuration of each control loop and control loop coupling.

While the attacker may apply many different attack patterns, in this work we consider the two following attacks:

- **Steady state attacks (SSA)** – step-like attacks which bring the process into the new state and leave it there (Figure 4.23a). This attack is similar to a setpoint change.
- **Periodic attacks (PA)** – recurring attacks interleaved with natural process recovery phases (Figure 4.23b). This attack is similar to causing a short disturbance.



(a) Steady state attack

(b) Periodic attack

Figure 4.23: Types of attacks on process behavior (red line denotes steady state value)

A large part of the control phase is concerned with mapping out the dependencies between

each actuator and all of the downstream measurements. Mapping of the dependencies can be done through detailed modeling of the process (resource and time consuming) or observed on the live process. One possible approach is to make small changes to the process and note process response as it propagates through the various sensors. Whilst working with a model one can safely apply any input to the process and observe process response, including alarm generation and propagation. When working with the actual process, the attacker would need to make minor changes and then extrapolate how large a change was necessary to cause an unwanted result while avoiding alarm activation.

We applied the following strategy to discovering dynamic process behavior. We first identified the steady state MV values. We than increased or decreased MV by approximately 1% for 30 *sec* and observed the process response measured by sensors. Depending on the response we increased the magnitude and duration of the manipulation and monitored the process variables for reaching operational or safety constraints.

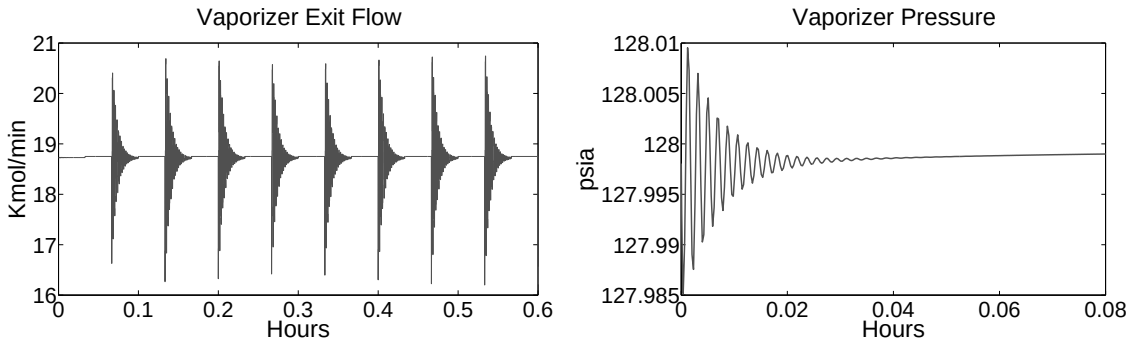
Steady state attacks. Not all actuators are suitable for carrying out steady state attacks. Manipulation of XMV{4;5} moves the process to its operational or safety constraints within a short time (from minutes to few hours) even if adjusted only slightly.

Periodic attacks. This attack scenario can be seen as pulse-width modulation of a steady state attack in which pulse amplitude represents *attack value* (position of the valve), pulse width stands for *attack duration* and inter-pulse distance is *process recovery time*. Examining the sensitivity of control loops to periodic attacks turned out to be laborious and challenging due to the large number of attack parameters.

One of the challenges we faced was management of the recovery (inter-attack) phase. Initially we directly set the position of the valves to their steady state values or lower. However, this strategy was unsuccessful as the majority of control loops continued drifting towards their operational and safety constraints. This was due to the fixed MV values which removed the ability of the valve to adjust dynamically to the process state. Therefore we decided to leave administration of the process recovery phase to the controllers.

Another challenge was the control loop *ringing* while manipulating valve XMV(5), which regulates the vaporizer exit flow. Sometimes, digital controllers produce a control signal that oscillates with decreasing amplitude around the final equilibrium level. This phenomenon is known as "ringing" and is caused by negative real controller poles. Ringing is an unwanted effect as it increases the wear and tear of valve components and can cause process instability within a multi-loop environment. Manipulation of XMV(5) in the negative direction (closing the valve) causes ringing of high amplitude as can be observed in Figure 4.24a. Figure 4.24b shows that the process could in general absorb this disturbance as the controlled variable oscillation is negligible, with a ratio of 1:150 compared to manipulated variable. However, the oscillation of the vaporizer exit flow variable did not allow us to achieve the desired reduction in the flow.

Despite an unwanted ringing effect in the XMV(5) control loop, we still pursued this attack as it reduces the inflow of reactants into the reactor and would subsequently result in decreased production of vinyl acetate. To overcome control loop ringing challenge we decided to take advantage of the *negative compensation reaction* in the *process recovery phase*. Specifically, we slightly increased the flow for $1-2$ minutes and let the process recover for 2 minutes or longer. In the recovery phase the controller decreased the flow to bring the controlled variable to its set point. In this way we were able to achieve a



(a) Ringing of the manipulated variable (b) Ringing of the controlled variable

Figure 4.24: Outcome of the vaporizer exit flow valve manipulation

reduced flow (when averaged over time). In total, there are three control loops in the reactor unit which can become unstable under certain attack parameters. Those are XMV{2;4;5}.

Outcome of the control stage

We have exhaustively tested all control loops for their sensitivity against a large number of attack settings combinations {attack value; attack duration; recovery time}. Although we could establish a good mental model of process behavior, we needed to find a systematic way to categorize controlled loops. We chose two parameters:

- **Sensitivity to magnitude of manipulation (MM)** – how much can we change the process in response to control loop manipulation;
- **Required recovery time (RT)** – if the process can recover in a time equal to the attack time or shorter we consider such such control loops of low sensitivity.

Table 4.3 gives the results of our analysis. Sensitive control loops are riskier to operate and therefore less suitable for reliable control from the attacker perspective.

Sensitivity	**MM**	**RT**
High	XMV{1;5;7}	XMV{4;7}
Medium	XMV{2;4;6}	XMV(5)
Low	XMV(3)	XMV{1;2;3;6}

Table 4.3: Sensitivity of control loops

We finalized our testing with the analysis of alarm activation withing the plant in response to SSA and PA attacks as shown in Table 4.4. Additionally, for each manipulated variable in the reactor section we noted upper limits of all possible attack parameters which would allow to manipulate the process without triggering an alarm (stored in a large excel spreadsheet).

Alarm	SSA	PA
Gas loop O_2	XMV(1)	XMV(1)
Reactor feed T	XMV(6)	XMV(6)
Reactor T	XMV(7)	XMV(7)
FEHE effluent	XMV(7)	XMV(7)
Gas loop P	XMV{2;3;6}	XMV{2;3;6}
HAc in decanter	XMV{2;3;7}	XMV(3)

Table 4.4: Activation of alarms

4.5.4 Damage

In the previous stage we evaluated the potential to control the process. In the damage phase the attacker uses identified controls to achieve the desired damage scenario. This stage is similar to the "what occurs if" approach in a HAZOP analysis ("what happens if this valve is closed").

As mentioned above, we opted for attack strategies that do not activate alarms. This means that the overall process state would remain within the "normal" operating envelope, however, the production economics or equipment operating conditions would deteriorate. Therefore, the results obtained in *Control* stage will be used in the *Prevent Response* sub-stage to avoid detection and prevent immediate corrective actions.

For reliability, the attacker must choose one or more attack scenarios to deploy in the final payload. These could be arbitrarily chosen based on gut feeling. However, given the amount of effort it takes to mount a real attack, an attacker may use a metric to measure potential of success for each attack scenario. For an economic attack a meaningful metric would be the amount of *monetary loss* to the victim. In order to determine monetary loss one needs to measure how much of vinyl acetate is produced in the reactor.

Obtaining Feedback

To measure the exact amount of a specific chemical in a production stream two measurements are needed: the total flow (FT) and the molar concentrations of the total flow composition. What would immediately catch the attacker's eye is the absence of an analyzer in the reactor exit stream (Figure 4.25). Chemical composition analysis systems are expensive; their installation must be justified by important considerations such as safety or significant product quality improvement. This is a significant disadvantage for the attacker as she will not be able to directly obtain measurements of the molar concentration of the produced vinyl acetate in the reactor outflow. Therefore the attacker cannot immediately obtain engineering measurements to monitor and measure attack execution.

The rate of the reaction can be qualitatively determined from the reactor exit temperature (TT). This would provide an attacker with a technician measurement. A decrease in temperature signals that less reaction is happening in the reactor, so less product is being produced (Figure 4.26a). This measurement can be sufficient to determine whether a specific attack has an effect on the reaction rate, but it does not allow to quantify the effect of an attack and select the most effective one. Looking at the process flowsheet, the only location where the attacker would be able to determine the exact amount of VAC produced is the decanter exit. However, this number would be available to the attacker only after

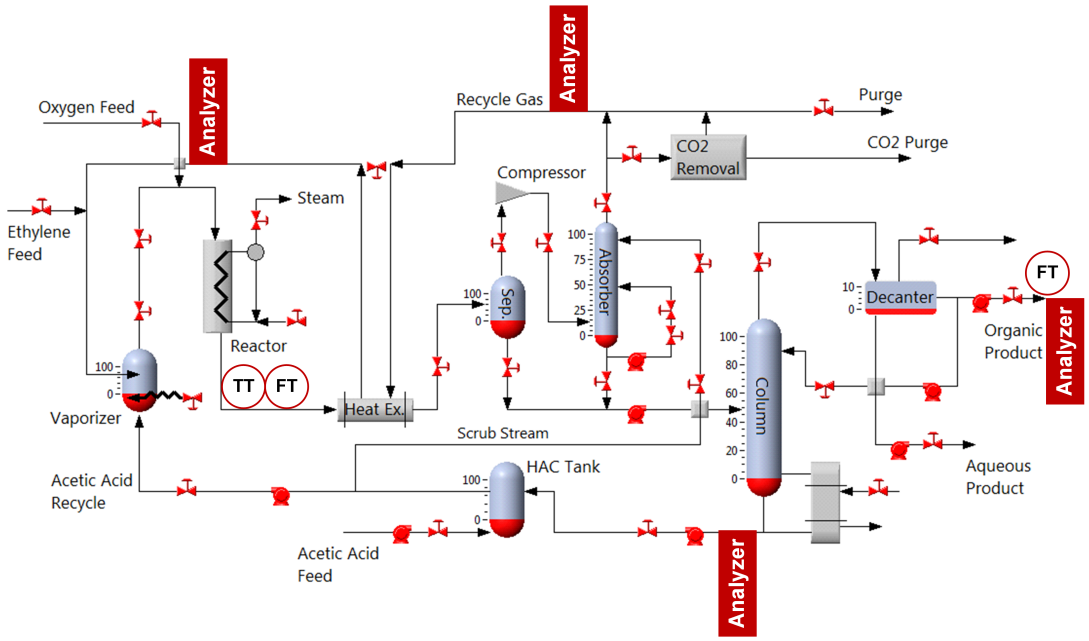
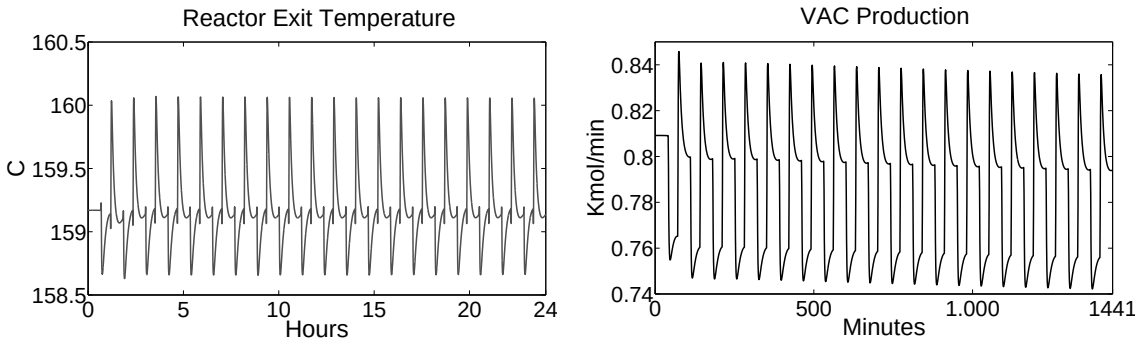


Figure 4.25: Location of chemical analyzers in VAC plant

hours, at the end of the refinement phase. This may not be a satisfactory option. Moreover, this would require the attacker to exploit additional network segments and devices. In our analysis we have not found a way to solve this challenge.



(a) Reactor exit temperature as proxy measure of reaction rate

(b) Direct measurement of VAC molar flow in reactor exit

Figure 4.26: Comparison of indirect and direct VAC production measurements

In the real world such an obstacle would force the attacker to look into controller code or search for process models in the test plant. In our case, we decided to look into the implementation code of the process model. In particular we were interested in the state variables in the reactor unit, used in the internal computations of the process model. We

were able to locate "concentration" variables of seven chemical components in ten sections of the reactor as shown in Figure 4.27.

```
//This is the main file for the reactor
//The plug-flow reactor is modeled as a 10-section distributed model
/*NR: section number + 1 */
/*calculate derivatives*/
for (n=1;n<NR;n++)
{
    /*Use component balance to get concentration change: */
    /*dC/dt=-delta(C*v)/deltaZ+sum(vij*rj)              */
    /*Use single backward                               */
    C_O2_t[n-1]=(-(C_O2[n]*v[n]-C_O2[n-1]*v[n-1])/dz Coeff1[0]*r_all[n][0]+Coeff2[0]*r_all[n][1])/cata_porosity;
    C_CO2_t[n-1]=(-(C_CO2[n]*v[n]-C_CO2[n-1]*v[n-1])/dz + Coeff1[1]*r_all[n][0]+Coeff2[1]*r_all[n][1])/cata_porosity;
    C_C2H4_t[n-1]=(-(C_C2H4[n]*v[n]-C_C2H4[n-1]*v[n-1])/dz + Coeff1[2]*r_all[n][0]+Coeff2[2]*r_all[n][1])/cata_porosity;
    C_VAc_t[n-1]=(-(C_VAc[n]*v[n]-C_VAc[n-1]*v[n-1])/dz + Coeff1[4]*r_all[n][0]+Coeff2[4]*r_all[n][1])/cata_porosity;
    C_H2O_t[n-1]=(-(C_H2O[n]*v[n]-C_H2O[n-1]*v[n-1])/dz + Coeff1[5]*r_all[n][0]+Coeff2[5]*r_all[n][1])/cata_porosity;
    C_HAc_t[n-1]=(-(C_HAc[n]*v[n]-C_HAc[n-1]*v[n-1])/dz + Coeff1[6]*r_all[n][0]+Coeff2[6]*r_all[n][1])/cata_porosity;
    Q_rct[n]= UA*(Tg[n]-Shell_T); /*kcal/min m^3*/
    Tg_t[n-1]=1/(cata_porosity*CCP[n] + cata_heatcapacity *cata_bulk_density)*(-FCP[n]*(Tg[n]-Tg[n-1])/dz -
                                                                    - r_all[n][0]*E_r1-r_all[n][1]*E_r2-Q_rct[n]);
};
```

Figure 4.27: Reactor code to calculate the concentration change of chemical components

Through extensive analysis of the obtained data and the working principle of the plug flow reactor employed in the VAC plant we could determine that the concentration of chemicals in the tenth section would be the same or about the same as in the reactor outflow. However, the obtained numerical values were of an unknown unit measure and very small with examples being 0,00073; 0,00016; 0,0007. The sum of all concentration values did not sum up to one/hundred or to the total flow. As the result we could not determine the utility of the obtained values.

After further investigations and consultation with the literature, we concluded that molar concentrations of chemical components can be computed according to the derived by us formula:

$$MOL_{comp}(t) = \frac{CONC_{comp}(t)}{\sum_{cmp} CONC_{comp}(t)},$$

where $CONC_{comp}(t)$ are the concentrations of the individual chemical components and cmp is the index of a current component.

We verified the numbers obtained with those provided in Table 5 in [147]. Since the total reactor exit flow is directly measured in the plant, we could compute the amount of vinyl acetate produced as:

$$Ouflow_{comp}(t) = MOL_{comp}(t)[\%mol] \times F_{reac}[Kmol/min],$$

where F_{reac} is measured reactor outflow, XMEAS(6).

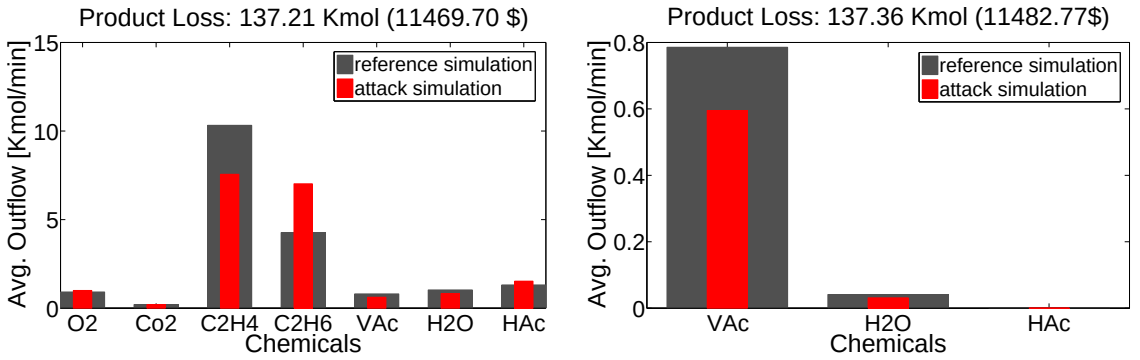
Figure 4.26b shows the production of vinyl acetate. The rate of VAC production indeed coincides with the reactor outflow temperature profile. Knowing the molar production of

VAC we could finally quantify production loss in dollar equivalents as:

$$Cost = VAC_{out}[Kmol] \times 86.09[g/mol] \times 0.971[\$/kg],$$

where $86.09[g/mol]$ is the VAC molar weight and $0.971[\$/kg]$ is the VAC price as given in [147].

To verify the numbers obtained we compared the amount of VAC produced in the reactor over a time period (Figure 4.28a) with the amount of VAC leaving the factory as final organic product (Figure 4.28b); the numbers matched.



(a) VAC production in reactor

(b) Final product in decanter

Figure 4.28: Vinyl acetate production, 24 hours

Evaluation of Attack Scenarios

With a suitable metric for evaluating the impact of the attack we could finally start deciding on the most effective attack design and their implementation. We have established the reference value of the steady-state production to determine loss (or gain) as a result of the attack.The categorization of the control loops based on their economic damage potential is given in Table 4.5. Note, that certain attacks have caused an increase of the vinyl acetate produced. This does not necessarily mean that the overall financial gain of the plant as such increased as production may have yielded increased operating costs and thus may resulted in revenue loss. However, we did not pursue this investigation. Therefore we did not consider attacks causing product gain as "successful". Lastly, attacks on XMV(1), oxygen feed, only have little attack potential, but also that this control loop easily becomes unstable. In addition, this control loop must be manipulated with great care as it quickly reaches its safety limit. Among all XMVs we marked XMV(1) as of least use.

Outcome of the damage stage

The damage phase is concluded with a portfolio of attacks which can be deployed at any opportune time. By scheduling attack value, attack duration and process recovery time

Production loss	SSA	PA
High, $\geq 10.000\$$	XMV(2)	XMV{4;6}
Medium, $5.000\$ - 10.000\$$	XMV{6;7}	XMV{5;7}
Low, $2.000\$ - 5.000\$$	–	XMV(3)
Negligible, $\leq 2.000\$$	XMV{1;3}	XMV{1;2}

Table 4.5: Categorization of control loops based on damage potential

we can control the amount of economic damage we would like to bring about. Note that our analysis has only determined how much money will be lost. Ideally economic damage numbers should be multiplied by the chance of success so that a risky high-damage scenario can be compared with a low-risk low-damage scenarios. Precise risk metrics may never be available, but in general attacks that require manipulating more components are considered a higher risk. Attacks that require an engineering answer (high-precision attack scenario) are more complex to design and execute and therefore are riskier than attacks only requiring a technician answer. Additionally, attacks that must hit a particular measure value or fail are riskier than attacks that simply get more effective the closer they are to the optimal value. Finally, the cost of attack implementation should be considered and compared against attack "gain".

4.5.5 Cleanup

In our attack scenario we were aiming for attacks that do not move the process towards operational limits or unsafe conditions. Since the attack execution does not trigger alarms, the operator may not notice immediately that the process has drifted from its economically optimal operating state. However, process operators may get concerned after noticing a persistent decrease in VAC production and may try to fix the problem. There can be numerous reasons for a process upset and operators are used to them. In the following we discuss how to influence the operators' beliefs about what is happening with the process.

An industrial process, just like software, has to be debugged when it malfunctions. If the attacker changes her attacks based on the debugging efforts of the maintenance engineers, future attacks may be attributed to the efforts of the engineer rather than a cyber attacker. Figure 4.29 illustrates outcome of attacks on four different control loops which cause symmetric fluctuations of different amplitudes in reactor exit temperature. These attacks can chained together and rotated in response to debugging efforts such as sensor calibration or reactor troubleshooting.

If a reactor is suspected to malfunction, a group of experts will be invited to investigate. It is not possible to see directly into the reactor. The investigators will apply specific metrics allowing them to evaluate the chemical processes in the reactor and –hopefully– determine potential reasons for the deterioration of reactor efficiency. They will then schedule maintenance work in accordance with identified issues. The attacker can execute different attacks which have the same effect on specific chemical processes in the reactor making engineers believe that their maintenance efforts are not bringing the expected results.

Typical examples of such metrics would be selectivity and conversion rate. Selectivity is a metric to control catalyst activity. Catalyst selectivity determines the fraction of the ethylene consumed that makes the desired VAC product or, in other words, how much (in

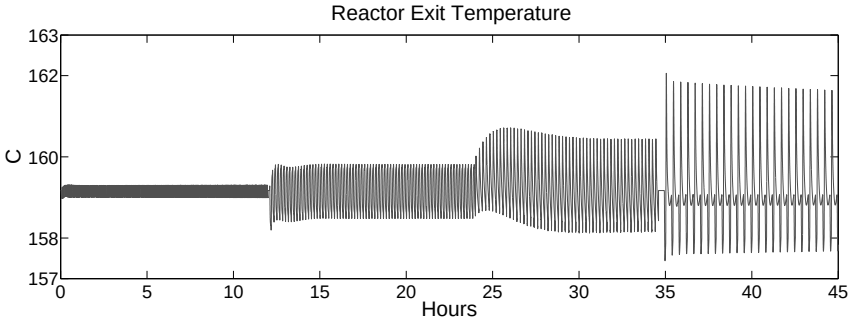


Figure 4.29: Increasing variation of reactor exit temperature caused by different attacks

percent) of the primary reaction has been induced by the catalyst:

$$SEL(t) = \frac{VAC_{out}(t)}{VAC_{out}(t) + 0.5 \times CO2_{out}(t)} \times 100,$$

where $VAC_{out}(t)$ and $CO2_{out}(t)$ are molar flows of the respective chemical components in the reactor outflow.

Conversion determines the fraction of the chemicals consumed (converted into product and byproducts) during the reaction. This metric is informative in several ways. For instance, there are certain safety limits and best-practice conversion rates, which should not be exceeded. Thus, the reduced conversion rate of acetic acid and increased conversion of ethylene suggest an increase in the amount of the undesired secondary reaction. Conversion is computed as:

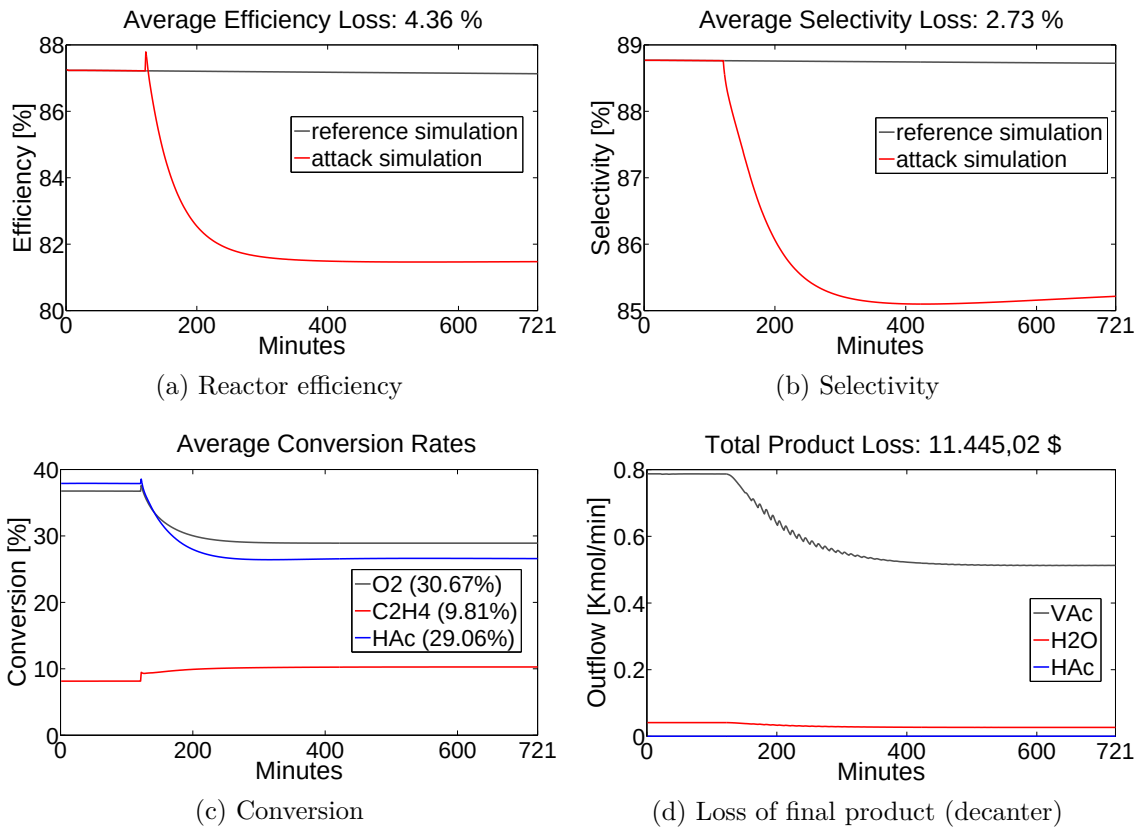
$$CONV_{comp}(t) = \frac{COMP_{in}(t) - COMP_{out}(t)}{COMP_{in}(t)} \times 100,$$

where $COMP_{in}(t)$ and $COMP_{out}(t)$ are the molar masses of the chemical components in the reactor in- and outflows.

In addition, we introduced a metric to measure **reactor efficiency** (EFF). It computes how much molar mass of acetic acid has reacted, and compares this value to the amount of reacted ethylene. Since the reaction ratio of ethylene and acetic acid in the primary reaction is $1:1$, the amount of reacted acetic acid is equal to the amount of correctly reacted ethylene. Relating this value to the amount of total reacted ethylene indicates the percentage of the primary reaction. Efficiency allows similar conclusions as catalyst selectivity, however, it is calculated based on the converted reagents rather than on the produced products:

$$EFF(t) = \frac{HAC_{in}(t) - HAC_{out}(t)}{C2H4_{in}(t) - C2H4_{out}(t)} \times 100.$$

Figure 4.30 illustrates the processes in the reactor during the attack on XMV(2). We decrease ethylene feed at time $t = 120$ minutes.



(a) Reactor efficiency

(b) Selectivity

(c) Conversion

(d) Loss of final product (decanter)

Figure 4.30: Analysis of the physical processes in the reactor

One can see how the attack affects the ratio between primary and secondary (ethylene combustion) reaction: the percentage of the primary reaction drops from 87% to under 82% and the amount of the secondary reaction increases by 4.32% on average (Figure 4.30a). Selectivity also drops to a lower level (Figure 4.30b). Since selectivity is calculated based on the ethylene consumed in both primary and secondary reactions, we can conclude that an increase of the secondary reaction has a stronger effect on the reagents consumed than it has on the products produced. In other words, the consumption of reagents is the more expressive metric in this case.

Figure 4.30c plots the conversion rates for the main reagents in the reactor. Ideally, the conversion rate of acetic acid is $\approx 37\%$, around 2% higher than the oxygen conversion. However, due to the attack, the conversion rate of acetic acid drops beneath oxygen conversion, indicating that the reaction kinetics have changed. This is because the newly induced secondary reaction also consumes oxygen (even more than the primary reaction). At the same time ethylene consumption has increased. Therefore we can conclude that the amount of the primary reaction has decreased (less acetic acid is converted), and the amount of ethylene combustion has increased. The result of the attack on XMV(2) is a significant reduction in the production of the final product (Figure 4.30d).

4.6 Discussion

Our initial attacker model restricts the attacker from triggering operational constraints. In reality, the attacker may suppress alarms while supplying the operator with good process values, e.g., using sensor signal spoofing techniques as proposed in [125].

> We intentionally chose not to activate alarms to understand attacker challenges when discovering dynamic process behavior in a stealthy manner. This turned out to be a laborious manual process with currently no established methods in the public literature to automate or support the testing and cataloging of dynamic process responses in cyber-physical systems.

We examined whether it could be beneficial for the attacker to violate operational constraints to cause more damage. Whereas we could almost double the loss with steady state attacks, the increase of damage in periodic attacks was modest. In certain cases violation of operational alarms eventually moved the process into an unsafe state, triggering process shutdown and attack interruption. It is, therefore, important for the attacker to monitor the plant state at all times to avoid unintentional plant shutdown. Not only does such an event slow down testing and/or execution of attack scenarios, it may also prompt an investigation. Thus an investigation of plant shutdown in the Triton attack resulted in the discovery of attack tools and triggered a full-fledged incident response [111].

> The attacker is not almighty and some damage scenario might be unachievable for various reasons.

For instance, initially, we attempted to achieve economic damage via catalyst deactivation. A chemical catalyst is a substance that increases the rate of a reaction without itself being consumed. It is frequently made of precious metals and therefore tends to be expensive. Catalyst productivity reduces over time requiring its replacement. The typical catalyst lifetime of VAC catalyst is $1-2$ years [61] after which plant profitability decreases. Preferred operation conditions are temperatures around $150-160°C$, higher temperatures within reactor tubes promote catalyst activity decay. Hot spots above $200°C$ lead to permanent catalyst deactivation requiring a costly production shutdown and catalyst replacement [147]. We were not able to realize this damage scenario as we were not able to overheat the reactor for long enough to damage the catalyst.

To conceal the ongoing attack, the attacker may implement a routine to suppress the digital alarm in the heat exchanger control loop to hide information about the deteriorated state of the physical process from the controller and from the human operator. However, alarm suppression does not change the physical state of the process. The "unhealthy" process state propagates into a neighboring plant section causing a low liquid level in an absorber vessel. Even though the absorber alarm is also suppressed to prevent compensating actions from the operator, the degraded state of the process keeps propagating downstream, eventually reaching an unsafe limit in the distillation column causing a safety shutdown of the plant as shown in Figure 4.3.4.

> This example also illustrates why the control stage precedes the damage stage – designing a specific damage scenario prior to testing control loops for controllability conditions such as operational/safety alarm activation may result in attack failure during execution.

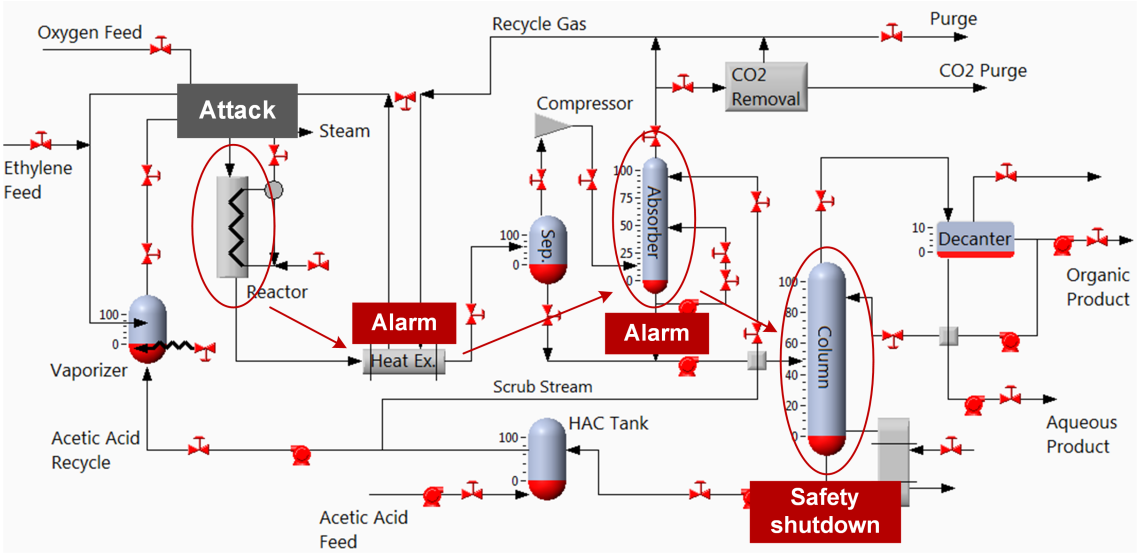


Figure 4.31: Propagation of alarms and deteriorated process state

Importantly, even though the flow of the digital alarms over the communication link is interrupted, the plant units/subsystems continues to communicate with each other via the physics of the process. *This is a natural information flow in physical processes that cannot be prevented.* However, it is critical to be aware of such physical information flows as they may be used by attackers for coordinating attack activities. For example, by implementing routines for detecting specific process states, the attacker can effectively coordinate attack activities between multiple malicious implants throughout the plant without sending messages via digital communication links as discussed in Section 4.3.4. Further discussion on physical process being a conduit and an information flow could be found in [119, 124].

> It is worth mentioning, because dynamic process behavior cannot be deduced from the plant documentation or control logic and can only be measured on a live process, the Control stage provides an opportunity for the defenders to detect anomalous short-term process fluctuations or unusual operations on equipment. Additionally, implementing defense-in-depth security controls and reducing the threat of unauthorized remote access to live process may significantly limit attacker abilities to develop viable attack scenarios.

Due to the high effort and extensive knowledge required overall to engineer and implement tailored damage attacks, it is more likely for threat actors to search for so-called **"easy button"** attack options. These are simple attack scenarios with easily identifiable damage conditions (e.g., operating rotating equipment at skip frequency) and reliable attack outcome (e.g., equipment damage due to vibrations). The Industroyer attack can be considered an "easy button" because it simply turns off all discovered control signals [41] resulting in the predictable outcome of substation shutdown. Control logic inherently contains information about exceptions and unsafe conditions making it a prime candidate for discovering "easy button" attack ideas.

As briefly discussed in Chapter 3 building and maintaining realistic testing environments is a significant financial investment. It has been suggested in [88, 172] that both attacks on the Ukrainian power grid were live drill operations by a state-sponsored threat actor to test cyber-physical sabotage capabilities.

> The chance exists that in the future threat actors will continue using real-world ICS environments for capability development and testing purposes. Understandingly, industrial plants with weak digital perimeters are more prone to becoming such testing environments.

With the growing awareness of cyber-security threats, industrial organizations start deploying conservative remote access solutions and apply zero trust principles to OT architectures requiring attackers to search for alternative pathways into the operational environments. Project files are considered to be trusted files coming from trusted sources like subcontractors or service providers. Such files are typically directly loaded to the control network and executed with high privileges. Gurkin [93] has shown that it is possible to embed complex exploits such as callback trojan or execution of OPC methods into the project file code itself which will be invisibly launched upon project startup. The emergence of such innovative methods for delivering attack code calls for the development of additional security controls such as ICS-aware sandbox applications.

4.7 Related Work

Attack Lifecycle

The overall concept of describing targeted cyber attacks in form of an attack lifecycle is relatively new. It was first introduced by researchers from Lockheed Martin Corporation in 2011 under the Kill Chain trademark name [101]. In 2015 SANS Institute adapted the Lockheed Martin Cyber Kill Chain model to the ICS needs and proposed The Industrial Control System Cyber Kill Chain [12]. The model consists of two phases. Stage 1: Cyber Intrusion Preparation and Execution and Stage 2: ICS Attack Development and Execution. While this model comprehensively covers cyber activities of a targeted cyber-physical attack, the physical part of the attack is represented as a single stage called *Attack Development and Tuning*. In contrast, the cyber-physical attack lifecycle proposed in the white paper focuses on the physical part of the attack and details the steps needed for achieving an objective of the above mentioned stage.

In the same year, Hahn et al. [95] proposed their variation of Lockheed Martin's Kill Chain model applied to the cyber-physical systems. Specifically, they added *Perturb Control* and *Physical Objective* stages to describe attack activities in the physical and control layers of the cyber-physical system. However, this work focuses on the formal descriptions of attack activities from the control theory standpoint whereas our proposed attack lifecycle is focused on the practical execution of an attack in real-world production environments.

In 2016 Larsen [131] described a process of developing a cyber-physical exploit chain with the help of Timing and State Diagrams (TSD). The sequence of the exploit building activities can be directly mapped to the stages of the cyber-physical attack lifecycle as shown in Figure 4.32. Larsen's approach is complimentary to the cyber-physical attack lifecycle as it allows to compare different attack design decisions in terms of development effort and execution reliability.

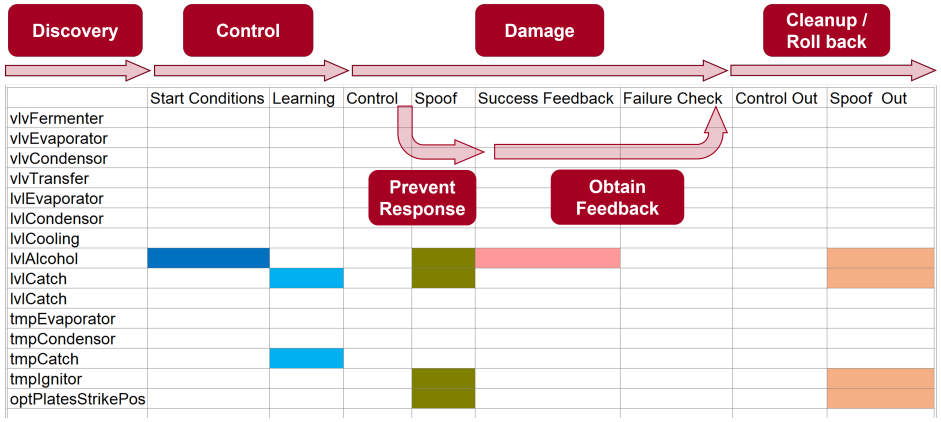


Figure 4.32: Relationship between TSD and cyber-physical attack lifecycle

Finally, in 2019 MITRE Corporation published the ICS ATT&CK knowledge base of the tactics and techniques that adversaries can use when attacking an industrial control systems [53]. While ICS ATT&CK provides a useful summary of known ICS attack techniques, the focus of the framework is on the execution and detection of the attacks in the cyber layer. In contrast, the framework proposed in the white paper puts greater focus on which attack tactics and techniques need to be implemented in the control and physical layers. The MITRE ICS ATT&CK was later expanded by Menendez [159] to better fit the attack lifecycle in the context of the electrical sector. However, the focus of the framework remained on the cyber aspects of ICS attacks.

To verify the applicability of the cyber-physical attack lifecycle to discrete cyber-physical systems, we tested attack stage on the example of a traffic light system [121]. As an application of moderate complexity, traffic light systems have a relatively small and constant process state space which can be fully tested by the attacker in the Control stage. Figure 4.33 shows an excerpt from a control logic where traffic light (TL) states are predefined by a fixed logic based on Boolean operators AND and OR.

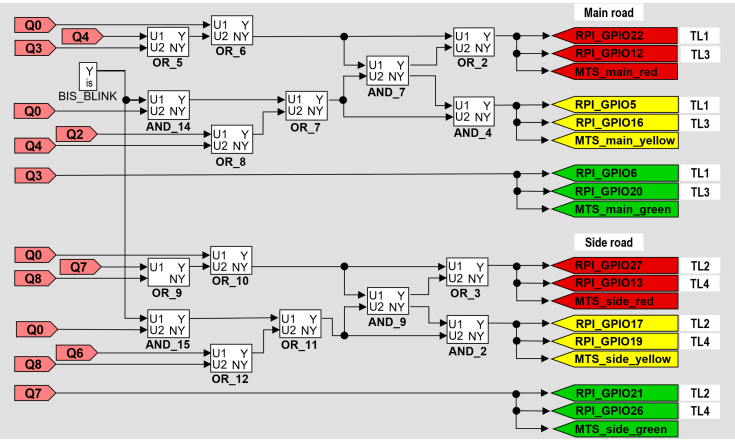


Figure 4.33: Snippet of a control logic for a traffic light system

Once useful controls are identified and tested, the attacker can develop a portfolio of ready-to-use attack instances and apply them in the Damage stage at an opportune time. Exploits can be combined into a final payload depending on desired attack outcome. Examples of damage scenarios include large-scale gridlocks to cause economic damage or short-term traffic congestions to slow down the arrival of vital services such as police or ambulance.

Cyber-Physical Attacks

ICS and cyber-physical security research achieved a significant level of maturity in exploiting devices and designing attack payloads in the past years. While the cyber-physical attack lifecycle guides through the attacker's logical sequential steps and associated activities when targeting physical process or equipment, it only provides an indication of attack types and exploits needed to be utilized by the attacker at each attack stage. Below we provide references to works which illustrate some of the attack Tactics, Techniques, and Procedures (TTPs) involved in cyber-physical exploitation in the ICS environments.

The lateral movement is an indispensable post-infiltration tactic where attackers navigate through networks and systems seeking for information/data, high-value hosts and services to expand their foothold and achieve persistence in the targeted network and, eventually, its end goal. While it was previously assumed that lateral movement via proprietary OT protocols (including backplanes) is prohibitively hard/challenging, Wetzels shown that it is an achievable task for a motivated threat actor on a practical cyber-physical attack scenario [229]. Figure 4.34 shows sequential steps of the OT lateral movement path described in the technical report. Additionally, Wetzels proposed an initial taxonomy of OT lateral movement techniques with illustrative examples.

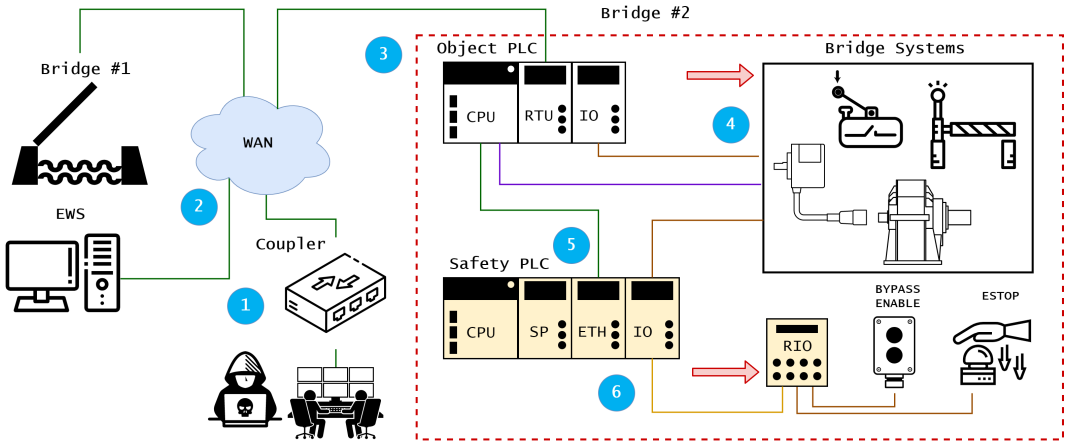


Figure 4.34: Illustration of attack path in the movable bridge control infrastructure [229]

One of the time-consuming attack design activities is *process comprehension* (see Section 4.3.2) or developing a comprehensive understanding of physical, control configuration, and computation characteristics of the targeted process as described in [80, 86, 5, 4], including finding exploitable attacks scenarios in control logic [207, 31, 85, 236]. Several works covered topics of dealing with compiled control logic [72, 198] and automation of project file

analysis [149, 169]. Discovered engineering attack scenarios then need to be implemented in the form of malicious payloads to manipulate control logic and possibly firmware to achieve the attack mission. Security weaknesses of the popular controllers' runtime environments can be found in [2, 112, 3, 187]. Research on manipulating controllers and field instrumentation internals is discussed in [132, 15, 5, 80, 105, 222].

Some of the involved attack TTPs are potentially detectable with tools available on the market. Translated into MITRE ICS ATT & CK [53] language, examples include Persistence (TA0110) [73, 111, 206, 28, 74], Privilege escalation (TA0111) [111], Evasion (TA0103) [73, 187], Command and Control (TA0101) [119, 85], Exfiltration (TA0010) [1, 85] as well as components of the OT payloads with a major focus on Impairing Process Control (TA0106) and Inhibiting Response Function (TA0107) [73, 28, 5, 230, 80].

The above research examples are by no means exhaustive but are meant to showcase the *diversity* of *non-trivial* attack activities involved in cyber-physical exploitation and the significant effort required to be invested even prior to developing final damage/degradation payloads.

4.8 Conclusions

Cyber-physical systems security concerns specifically attacks that cause physical impact. To achieve such an impact, the attacker has to find ways of manipulating the physical processes in the target environment. Cyber attacks on control networks may allow the attacker to obtain access to sensor measurements and manipulate instructions sent to actuators. However, to appreciate the effect of such manipulations the attacker has to understand the physical part of her target. Similar to attacks executed in the IT domain, a cyber-physical attack goes through several attack stages before achieving the desired damage. Along the cyber-physical attack lifecycle, some activities may resemble well-known IT exploitation methods and some will require expertise on the physical part of the cyber-physical system, expertise not commonly found in the IT security community. Depending on the attack scenario, state of the process and encountered challenges, the attacker may skip certain stages or may need to circle back to previous stages. While numerous techniques exist to exploit ICS systems (assets, protocols, applications), the cyber-physical attack lifecycle describes how an attacker would identify assets of their interest and arrive at the engineering design of their attack. Specifically, it provides greater insights into the detours an attacker may have to take to reach her goal, the types of data sources to consult to achieve the required level of process comprehension and which vital reconnaissance tasks need to be executed on the live process. We illustrated the utility of the proposed framework in the example of a simulated model of the Vinyl Acetate Monomer plant and demonstrated non-trivial hurdles to overcome at each attack stage. It took almost *one and half years* to finish this research undertaking from the initial familiarization with the VAC process to obtaining final results. Understandably, since such a work was conducted for the first time, some time was consumed by trial-and-error activities.

From the attacker's perspective understanding the desired state of a physical process that achieves certain attack goals, and knowing how to reach that state are two distinct problems. The process is not designed for the attacker. The attacker may be impeded by automatic safety interlocks and/or may not have access to observations that allow her to monitor the effect of her actions. Due to the significant effort required to execute sophisticated

high-precision attacks such capabilities are likely to remain in the hands of state-sponsored threat actors with sufficient resources and reserved for critical targets. In contrast “easy button” attacks are more likely to be executed on a larger scale.

Successful execution of cyber-physical attacks relies on access to specific information sources and assets. Understanding attacker needs assists defenders with identifying “crown jewels” assets such as historians, specialized engineering applications, OPC servers, engineering workstations, cloud-based (I)IoT backends and strategizing defense and detection efforts.

5. Conclusion

Cyber-physical systems are engineered systems that are built from and depend upon the seamless integration of computational and physical components. While advances in computing technologies gave industry new opportunities and functionalities for interacting with the physical world, they have also changed the attack surface and introduced new forms of attacks. In contrast to IT security which can trace its roots back several decades, the discipline of CPS security is comparatively young, and the research community has not yet achieved the same clarity about relevant threat scenarios and best security practices. While compromising computing infrastructure and communication channels is a powerful facilitator for launching attacks aimed at disrupting physical processes, the damage from an attack will be limited if the attacker is unable to manipulate the control system in a way needed to achieve the desired physical outcome.

The attacker is not bound to adhere to assumptions made by the defender about possible adversarial behavior. In fact, it is an effective attack strategy to search for assumptions that can be violated, including previously unconsidered attack scenarios. We have shown that the security concepts from the IT domain are not sufficient to describe security demands on cyber-physical systems nor are the defenses effective against all attacks which are part of the cyber-physical kill chain. We, therefore, suggest treating the cyber part of a cyber-physical system as a control system instead and focusing on the interfaces between physical space and cyberspace. For the attacker, the control system is a weird machine [25] commandeered by unconventional inputs used in ways not intended by the designer. To constrain the attacker, systems ultimately have to be designed and operated in a way so that as little as possible can be done with them beyond their intended use. This may include replacing critical digital components with immutable alternatives. Among the main goals of the white paper was to show that defense efforts should extend to the design of process control applications and process physics itself as visualized in Figure 5.1.

We conclude with the *encouragement* to further research attacker strategies when targeting physical processes and engineered control systems as well as to extend the defense strategies/considerations beyond IT security approaches from the digital space. This, for

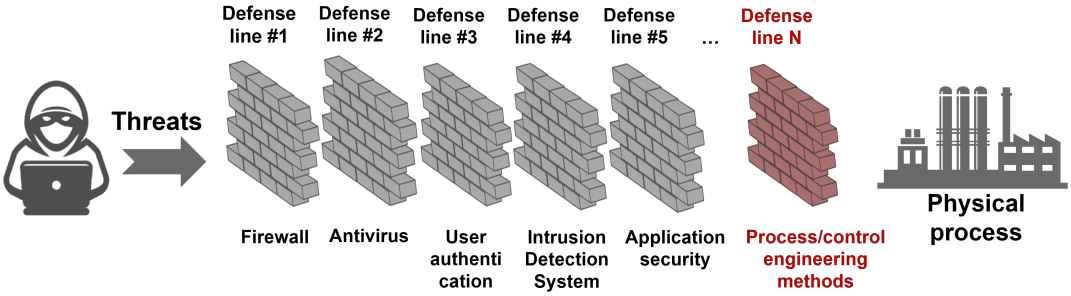


Figure 5.1: Visualization of defense-in-depth concept in cyber-physical systems

instance, includes redesigning cyber-physical systems to make them more robust to the compromised physical environment (both in immediate proximity and at a distance). CPS should be further enhanced with an enlarged forensic footprint to enable improved situational awareness in the presence of attacks. Once physical and/or IT barriers are compromised, adversaries are potentially able to take control of the physical process. While the latter is a non-trivial task and requires expert knowledge, it would be negligent to assume that attackers are unable or not incentivized enough to obtain and apply the required knowledge. Technical means should be complemented with the appropriate international regulations where the use of cyber-physical attacks against civilian critical infrastructures should be prohibited to preclude potential humanitarian crises.

References

[1] D .Atch and G. Lashenko. *Exfiltrating Reconnaissance Data from Air-Gapped IC-S/SCADA Networks*. Black Hat EU, https://www.blackhat.com/docs/eu-17/materials/eu-17-Atch-Exfiltrating-Reconnaissance-Data-From-Air-Gapped-Ics-Scada-Networks.pdf. Retrieved: Jan, 2023. 2017.

[2] A. Nochvay (Kaspersky ICS CERT). *Security research: CODESYS Runtime, a PLC control framework*. Research Report, https://ics-cert.kaspersky.com/media/KICS-CERT-Codesys-En.pdf. Retrieved: Jan, 2023. 2019.

[3] A. Nochvay, A. Zinenko, E. Goncharov (Kaspersky ICS CERT). *ISaPWN – Research on the Security of ISaGRAF Runtime*. Research Report, https://ics-cert.kaspersky.com/media/Kaspersky-ICS-CERT-ISaPWN-Research-on-the-security-of-ISaGRAF-Runtime-En.pdf. Retrieved: Jan, 2023. 2022.

[4] A. Palanca, L. Cremona and R. Gordon. *UWB Real Time Locating Systems: How Secure Radio Communications May Fail in Practice*. Black Hat USA, https://i.blackhat.com/USA-22/Wednesday/US-22-Gordon-UWB-Real-Time-Locating-Systems.pdf. Retrieved: Jan, 2023. 2022.

[5] Ali Abbasi et al. "Stealth Low-Level Manipulation of Programmable Logic Controllers I/O by Pin Control Exploitation". In: *International Conference on Critical Information Infrastructures Security*. Springer. 2016, pp. 1–12.

[6] Simumatik Academy. *OPCUA integration with SIEMENS TIA Portal + PLCSIM Advanced*. https://academy.simumatik.com/wp-content/uploads/static-html-to-wp/data/41d1c51267e75b39c055ec32675ee246/ThirdParty/PLC/siemens_TIA/. Retrieved: Nov, 2019.

[7] Raimo Ahola. "The future of calibration is integration". In: *InTech* 61.2 (2014), p. 44.

[8] Emad Ali. "Understanding the Operation of Industrial MSF Plants Part I: Stability and Steady-State Analysis". In: *Desalination* 143.1 (2002), pp. 53–72.

[9] Dmitri Alperovitch. *Revealed: Operation Shady RAT*. Tech. rep. McAfee, 2011.

[10] Oxana Andreeva et al. *Industrial Control Systems and Their Online Availability*. Tech. rep. Kaspersky Lab, 2016.

[11] Nikolai Antoniadis et al. *The Vulkan Files: A Look Inside Putin's Secret Plans for Cyber-Warfare*. https://www.spiegel.de/international/world/the-vulkan-files-a-look-inside-putin-s-secret-plans-for-cyber-warfare-a-4324e76f-cb20-4312-96c8-1101c5655236. Retrieved: April, 2023. 2023.

[12] Michael J. Assante and Robert M. Lee. *The Industrial Control System Cyber Kill Chain*. Tech. rep. SANS Institute, 2015.

[13] Aveva. *Aveva Instrumentation Software*. https://www.aveva.com/en/products/electrical-instrumentation. Retrieved: Nov, 2019.

[14] A. Avizienis et al. "Basic Concepts and Taxonomy of Dependable and Secure Computing". In: *IEEE Transactions on Dependable and Secure Computing* 1 (2004).

[15] Z. Basnight et al. "Firmware modification attacks on programmable logic controllers". In: *International Journal of Critical Infrastructure Protection* 6.2 (2013).

[16] R.D. Bell and K.J. Åström. *Dynamic Models for Boiler-Turbine aAlternator Units: Data Logs and Parameter Estimation for a 160 MW Unit*. 1987.

[17] Jonas Berge. *Implementing the NAMUR Open Architecture (NOA)*. https://www.linkedin.com/pulse/implementing-namur-open-architecture-noa-jonas-berge. Retrieved: Feb, 2020. 2019.

[18] Bill Hollifield, PAS. *A High Performance HMI: Better Graphics for Operations Effectiveness*. White paper, https://https://isawaterwastewater.com/wp-content/uploads/2012/07/WWAC2012-invited_BillHollified_HighPerformanceHMIs_paper.pdf. Retrieved: May, 2020. 2012.

[19] Irena Bojanova and Jeffrey Voas. "Securing the Internet of Anything (IoA)". In: *IEEE Computer Society* (Nov. 2015).

[20] A. Bolshev et al. "A Rising Tide: Design Exploits in Industrial Control Systems". In: *10th USENIX Workshop on Offensive Technologies (WOOT 16)*. Austin, TX: USENIX Association, 2016.

[21] Alexander Bolshev. *ICSCorsair: How I will PWN your ERP through 4-20 mA Current Loop*. Black Hat USA, https://www.blackhat.com/docs/us-14/materials/us-14-Bolshev-ICSCorsair-How-I-Will-PWN-Your-ERP-Through-4-20mA-Current-Loop.pdf. Retrieved: Nov, 2019. 2014.

[22] Alexander Bolshev and Marina Krotofil. *Never Trust your Inputs: Causing 'Catastrophic PHysical Consequences' from the Sensor (or How to Fool ADC)*. Black Hat Asia, https://www.blackhat.com/docs/asia-16/materials/asia-16-Bolshev-Never-Trust-Your-Inputs-Causing-Catastrophic-Physical-Consequences-From-The-Sensor.pdf. Retrieved: Nov, 2019. 2016.

[23] Niklas Borselius. "Mobile Agent Security". In: *Electronics & Communication Engineering Journal* 14.5 (2002), pp. 211–118.

[24] Hugh Boyes et al. “The industrial internet of things (IIoT): An analysis framework”. In: *Computers in Industry* 101 (2018), pp. 1–12.

[25] Sergey Bratus et al. “Exploit programming: From buffer overflows to weird machines and theory of computation”. In: *USENIX: ;LOGIN:* (2011).

[26] Bill Briggs. *Hackers hit Norsk Hydro with ransomware. The company responded with transparency*. `https://news.microsoft.com/transform/hackers-hit-norsk-hydro-ransomware-company-responded-transparency`. Retrieved: Nov, 2019. 2019.

[27] J. T. Broch. *Mechanical Vibration and Shock Measurements*. Bruel & Kjer, 1984.

[28] N. Brubaker et al. *Integrated but separate: Advances in integrated and safety control*. Mandiant blog, `https://www.mandiant.com/resources/blog/incontroller-state-sponsored-ics-tool`. Retrieved: Jan, 2023. 2022.

[29] Nathan Brubaker et al. *Financially Motivated Actors Are Expanding Access Into OT: Analysis of Kill Lists That Include OT Processes Used With Seven Malware Families*. `https://www.fireeye.com/blog/threat-research/2020/07/financially-motivated-actors-are-expanding-access-into-ot.html`. Retrieved: Nov, 2020. 2020.

[30] Richard Candell, Dhananjay Anand, and Keith Stouffer. “A Cybersecurity Testbed for Industrial Control Systems”. In: *Proceedings of the 2014 Process Control and Safety Symposium*. 2014.

[31] J. H. Castellanos et al. “AttkFinder: Discovering Attack Vectors in PLC Programs Using Information Flow Analysis”. In: *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses*. RAID ’21. 2021.

[32] Censys. *Censys search engine*. `https://censys.io`. Retrieved: Dec, 2019.

[33] J.M. Ceron et al. *Online Discoverability and Vulnerabilities of ICS/SCADA Devices in the Netherlands*. Tech. rep. Universiteit Twente, 2019.

[34] Kaspersky ICS CERT. *MontysThree: Industrial Espionage with Steganography and a Russian Accent on Both Sides*. `https://ics-cert.kaspersky.com/reports/2020/10/08/montysthree-industrial-espionage-with-steganography-and-a-russian-accent-on-both-sides/`. Retrieved: Dec, 2020. 2020.

[35] Kaspersy Lab ICS CERT. *Spear phishing attack hits industrial companies*. `https://ics-cert.kaspersky.com/alerts/2016/12/16/spear-phishing-attack-hits-industrial-companies`. Retrieved: Feb, 2017. 2016.

[36] *Chapter header image source*. Online: `http://www.cls-soudage.fr/wp-content/uploads/2016/02/img-industrie.png`. Retrieved: July, 2017.

[37] American Institute of Chemical Engineers. *Process Safety Incident Database*. `http://www3.aiche.org/PSID/Home.aspx`. Retrieved: November, 2019.

[38] Rong Chen et al. “A Nonlinear Dynamic Model of a Vinyl Acetate Process”. In: *Industrial & Engineering Chemistry Research* 42.20 (2003), pp. 4478–4487.

[39] Rong Chen et al. *Code repository: A Nonlinear Dynamic Model of a Vinyl Acetate Process*. `http://www.isr.umd.edu/~mcavoy/VAC%20Material`. Retrieved: June, 2013. 2003.

[40] Peng Cheng, Heng Zhang, and Jiming Chen. *Cyber Security for Industrial Control Systems: From the Viewpoint of Close-Loop*. CRC Press, 2016.

[41] Anton Cherepanov. *WIN32/INDUSTROYER: A new threat for industrial control systems*. Tech. rep. ESET, 2017.

[42] Eric Chien and Gavin O'Gorman. *The Nitro Attacks: Stealing Secrets from the Chemical Industry*. Tech. rep. Symantec, 2011.

[43] D. D. Clark and D. R. Wilson. "A Comparison of Commercial and Military Computer Security Policies". In: *IEEE Symposium on Security and Privacy*. 1987, pp. 184–184.

[44] A.W. Colombo et al. *Industrial Cloud-Based Cyber-Physical Systems*. Springer, 2014.

[45] United Stated Nuclear Regulatory Commission. *Planning and Installation Guide for Tricon v9–v10 Systems*. `https://www.nrc.gov/docs/ML0932/ML093290420.pdf`. Retrieved: Dec, 2017. 2009.

[46] IEC/SC 65B committee. *IEC 61132-3:2013 - Programmable Controllers - Part 3: Programming Languages*. Standard. International Electrotechnical Commission, 2013.

[47] ISA 18.2 committee. *IEC 62682:2015 - Management of Alarms Systems for the Process Industries*. Standard. International Electrotechnical Commission, 2015.

[48] ISA-95 Committee. *Enterprise-Control System Integration*. `https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95`. Retrieved: Nov, 2019.

[49] ISA-99 Committee. *IEC 62443 - International Series of Cybersecurity Standards for Operational Technology in Automation and Control Systems*. Standard. International Electrotechnical Commission, 2010.

[50] ISA-99 Committee. *Industrial Automation and Control Systems Security*. `https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa99`. retrieved: Nov, 2019.

[51] Kyiv Energy Construction Company. *Reconstruction of the substation "Severnaya"*. `https://www.youtube.com/watch?v=VLwgprHrqMI`. Retrieved: Feb, 2020. 2014.

[52] J.P. Contreras et al. "Vinyl Acetate from Ethylene, Acetic Acid and Oxygen Industrial Plant Simulation". In: *Proceedings of the Computing and Systems Technology Division, American Institute of Chemical Engineers (AIChE) Annual Meeting*. AIChE. 2009, pp. 249–259.

[53] MITRE Corporation. *MITRE ATTCK Matrix for ICS*. `https://attack.mitre.org/matrices/ics/`. Retrieved: Nov, 2019.

[54] John Cusimano and Eric Byres. "Safety and Security: Two Sides of the Same Coin". In: *ControlGlobal* (2010).

[55] U.S. Cybersecurity and Infrastructure Security Agency. *MEMS Accelerometer Hardware Design Flaws*. `https://www.cisa.gov/uscert/ics/alerts/ICS-ALERT-17-073-01A`. Retrieved: Jan, 2022. 2017.

[56] D. V. Reising and P. Bullemer. *An Introduction to the ASM Guidelines Effective Operator Display Design*. Online, https://www.asmconsortium.net/Documents/2009%20ASM%20Displays%20GL%20Webinar%20v014.pdf. Retrieved: June, 2014. 2009.

[57] dark-lbp. *Credentials wordlists*. https://github.com/dark-lbp/isf/tree/master/icssploit/wordlists. Retrieved: Feb, 2020. 2017.

[58] dark-lbp. *Industrial Control System Exploitation Framework*. https://github.com/dark-lbp/isf. Retrieved: Nov, 2019. 2017.

[59] dark-lbp. *S7 Bruteforce Modules*. https://github.com/dark-lbp/isf/tree/master/icssploit/modules/creds. Retrieved: Feb, 2020. 2017.

[60] Food Department for Environment and Rural Affairs. *Waste water treatment in the United Kingdom – 2012. Implementation of the European Union Urban Waste Water Treatment Directive – 91/271/EEC*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/69592/pb13811-waste-water-2012.pdf. Retrieved: Nov, 2019. 2012.

[61] Alexandre C. Dimian and Costin Sorin Bildean. *Chemical Process Design: Computer-Aided Case Studies*. WILEY-VCH Verlag GmbH & Col, 2008.

[62] Matan Dobrushin and Yoav Flint Rosenfeld. *Project RunAway, 2019*. CS3sthlm, https://cs3sthlm.se/program/presentations/matan-yoav. Retrieved: Nov, 2019. 2019.

[63] J. J. Downs and E. F. Vogel. "A plant-wide industrial process control problem". In: *Computers & Chemical Engineering* 17.3 (1993), pp. 245–255.

[64] Marcin Dudek. *TRITON original samples*. https://github.com/MDudek-ICS/TRISIS-TRITON-HATMAN/tree/master/original_samples. Retrieved: Feb, 2020. 2017.

[65] Gregory Duranso. *Development of High-Performance HMI Philosophy*. https://realpars.com/hmi-philosophy. Retrieved: Sept, 2022. 2021.

[66] E-ISAC. *Analysis of the Cyber Attack on the Ukrainian Power Grid*. https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf. Retrieved: Nov, 2019. 2016.

[67] LLC "NTK ENPASELECTRO". *Completed projects: Control systems*. https://enpaselectro.com/ru/tehnologii/sistemy-upravleniya.html. Retrieved: Feb, 2020.

[68] Adriana Erickson. *RTU/PLC Communication*. https://control.com/forums/threads/rtu-plc-communication.2951. Retrieved: Nov, 2019. 2001.

[69] ESET Research. *Industroyer2: Industroyer reloaded*. https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/. Retrieved: Apr, 2022. 2022.

[70] Jack Evans. *Someone tried to poison Oldsmar's water supply during hack, sheriff says*. https://www.tampabay.com/news/pinellas/2021/02/08/someone-tried-to-poison-oldsmars-water-supply-during-hack-sheriff-says. Retrieved: March, 2021. 2021.

[71] F-Secure. *Havex Hunts for ICS/SCADA Systems*. https://www.f-secure.com/weblog/archives/00002718.html. Retrieved: July, 2014. 2014.

[72] F. 'FX' Lindner (Recurity Labs). *Building Custom Disassemblers: Instruction Set Reverse Engineering*. 27C3, https://data.proidea.org.pl/confidence/9edycja/materialy/prezentacje/FX.pdf. Retrieved: Jan, 2023. 2010.

[73] Nicolas Falliere, Liam O Murchu, and Eric Chien. *W32.Stuxnet Dossier*. Symantec, https://www.wired.com/images_blogs/threatlevel/2010/11/w32_stuxnet_dossier.pdf. Retrieved: Feb, 2020. 2011.

[74] Forescout. *Cybersecurity in Building Automation Systems (BAS)*. Resources, https://www.forescout.com/resources/bas-research-report-the-current-state-of-smart-building-cybersecurity-2/. Retrieved: Jan, 2023. 2020.

[75] OPC Foundation. *What is OPC?* https://opcfoundation.org/about/what-is-opc. Retrieved: Feb, 2020.

[76] *Front image source*. Online: https://www.nae.edu/File.aspx?id=14929. Retrieved: July, 2015.

[77] Kevin Fu and Wenyuan Xu. "Risks of Trusting the Physics of Sensors". In: *Communications of ACM* 61.2 (Jan. 2018), pp. 20–23.

[78] G. Tsaraias and I. Speziale. *Industroyer vs. Industroyer2: Evolution of the IEC 104 Component*. https://www.nozominetworks.com/downloads/US/Nozomi-Networks-WP-Industroyer2.pdf. Retrieved: Apr, 2022. 2022.

[79] Jianlei Gao et al. "Research about DoS Attack against ICPS". In: *Sensors* 19 (Mar. 2019), p. 1542.

[80] L. Garcia et al. "Hey, My Malware Knows Physics! Attacking PLCs with Physical Model Aware Rootkit". In: *Network and Distributed System Security Symposium*. 2017.

[81] Michele Gattullo et al. "What, How, and Why are Visual Assets used in Industrial Augmented Reality? A Systematic Review and Classification in Maintenance, Assembly, and Training (from 1997 to 2019)". In: *IEEE Transactions on Visualization and Computer Graphics* (2020), pp. 1–15.

[82] GLEG. *SCADA+ Pack - 0day exploits*. http://gleg.net/agora_scada.shtml. Retrieved: Nov, 2019.

[83] Dieter Gollmann. "Veracity, Plausibility, and Reputation". In: *Information Security Theory and Practice. Security, Privacy and Trust in Computing Systems and Ambient Intelligent Ecosystems*. Vol. 7322. LNCS. 2012, pp. 20–28.

[84] Shelby Grad. *Engineers who hacked into L.A. traffic signal computer, jamming streets, sentenced*. https://latimesblogs.latimes.com/lanow/2009/12/engineers-who-hacked-in-la-traffic-signal-computers-jamming-traffic-sentenced.html. Retrieved: Nov, 2019. 2009.

[85] B. Green et al. "PCaaD: Towards Automated Determination and Exploitation of Industrial Systems". In: *Comput. Secur.* 110 (2021).

[86] Benjamin Green, Marina Krotofil, and Ali Abbasi. "On the Significance of Process Comprehension for Conducting Targeted ICS Attacks". In: *Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy (CPS-SPC)*. CPS '17. ACM, 2017, pp. 57–67.

[87] Benjamin Green, Marina Krotofil, and David Hutchison. "Achieving ICS Resilience and Security through Granular Data Flow Management". In: *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. CPS-SPC '16. 2016, pp. 93–101.

[88] Andrew Greenberg. *How an Entire Nation Became Russia's Test Lab for Cyberwar*. `https://web.archive.org/web/20170624161851/https://www.wired.com/story/russian-hackers-attack-ukraine`. Retrieved: Nov, 2019. 2017.

[89] Alarm Systems User Group. *EEMUA 191. Alarm systems - A guide to design, management and procurement*. Standard. EEMUA, 1999.

[90] Alibaba Group. *Alibaba e-commerce platform*. `https://www.ebay.com`. Retrieved: Feb, 2020.

[91] Gruenbeck. "Operating instructions: Ultrafiltration system GENO-Ultrafil 450/900 with GENO-matic". In: (2017).

[92] GRUNDFOS. "AQpure UF, Modular water treatment system. Service instructions". In: (2010).

[93] Yuriy Gurkin. *SCADA Projects - Hackers' Point of View*. ZeroNights, `https://2018.zeronights.ru/wp-content/uploads/materials/21-SCADA-projects-from-the-point-of-view-of-hackers.pdf`. Retrieved: Dec, 2018. 2018.

[94] H. Tanaka. *Optimize Alarm Management*. Online, `https://www.chemicalprocessing.com/automation/control-systems/article/11309974/optimize-alarm-management`. Retrieved: May, 2020. 2018.

[95] Adam Hahn et al. "A Multi-Layered and Kill-Chain based Security Analysis Framework for Cyber-Physical Systems". In: *International Journal of Critical Infrastructure Protection* 11 (2015), pp. 39–50. ISSN: 1874-5482.

[96] F. Hamill. "Sayano Shushenskaya accident – presenting a possible direct cause". In: *Plant Power and Dam Construction: online magazine* (2010). Retrieved: December, 2013. URL: `http://www.waterpowermagazine.com/features/featuresayano-shushenskaya-accident-presenting-a-possible-direct-cause`.

[97] Y.-F. Han et al. "Kinetics of ethylene combustion in the synthesis of vinyl acetate over a Pd/SiO_2 catalyst". In: *Journal of Catalysis* 224.1 (2004), pp. 60–68.

[98] Chris Harper. "AIV and FIV in Pipelines, Plants, and Facilities". In: vol. Volume 1: Pipelines and Facilities Integrity. International Pipeline Conference. Sept. 2016.

[99] Stephen Hilt et al. *Exposed and Vulnerable Critical Infrastructure: Water and Energy Industries*. Tech. rep. TrendMicro, 2018.

[100] HSE. *Accidents and investigations*. `https://www.hse.gov.uk/toolbox/managing/accidents.htm`. Retrieved: Feb, 2020.

[101] Eric M. Hutchins, Michael J. Cloppert, and Rohan M. Amin. *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*. Tech. rep. Lockheed Martin Corporation, 2011.

[102] IBM. *Cost of a data breach 2022*. Online report: https://www.ibm.com/reports/data-breach. Retrieved: May, 2023.

[103] ICS-CERT. *ICS Focused Malware*. https://us-cert.gov/ics/alerts/ICS-ALERT-14-176-02A. Retrieved: Feb, 2015. 2014.

[104] ICS-CERT. *Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure Sectors*. https://www.us-cert.gov/ncas/alerts/TA18-074A. Retrieved: March, 2018. 2018.

[105] U.S. ICS-CERT. *HATMAN: Safety System Targeted Malware*. https://www.us-cert.gov/sites/default/files/documents/MAR-17-352-01%20HatMan%20-%20Safety%20System%20Targeted%20Malware%20%28Update%20A%29_S508C.PDF. Retrieved: Dec, 2017. 2017.

[106] eBay Inc. *eBay e-commerce platform*. https://www.ebay.com. Retrieved: Feb, 2020.

[107] GitHub Inc. *GitHub: Internet hosting for software development*. https://github.com. Retrieved: Feb, 2020.

[108] Scribd Inc. *Scribd digital library*. www.scribd.com. Retrieved: Nov, 2019.

[109] Joint Task Force Transformation Initiative. "NIST SP 800-53A Rev.1. Guide for Assessing the Security Controls in Federal Information Systems and Organizations". In: *National Institute of Standards and Technology* (2010).

[110] CS INSTRUMENTS. *Intelligent chart recorder DS 500 for compressed air and gases*. https://www.cs-instruments.com/products/d/chart-recorder/ds-500-intelligent-chart-recorder. Retrieved: Nov, 2019.

[111] Blake Johnson et al. *Attackers Deploy New ICS Attack Framework "TRITON" and Cause Operational Disruption to Critical Infrastructure*. https://www.fireeye.com/blog/threat-research/2017/12/attackers-deploy-new-ics-attack-framework-triton.html. Retrieved: Nov, 2019. 2017.

[112] A. Keliris and M. Maniatakos. "ICSREF: A Framework for Automated Reverse Engineering of Industrial Control Systems Binaries". In: *Network and Distributed System Security Symposium*. 2019.

[113] Sinclair Koelemij. *Interfaced or integrated?* https://otcybersecurity.blog/2020/05/10/interfaced-or-integrated-2/. Retrieved: June, 2020. 2020.

[114] C. Konstantinou, M. Sazos, and M. Maniatakos. "Attacking the smart grid using public information". In: *2016 17th Latin-American Test Symposium (LATS)*. 2016, pp. 105–110.

[115] B. Krebs. *Security Firm Bit9 Hacked, Used to Spread Malware*. http://krebsonsecurity.com/2013/02/security-firm-bit9-hacked-used-to-spread-malware/. Retrieved: Dec, 2013. 2013.

[116] Brian Krebs. *Chinese Hackers Blamed for Intrusion at Energy Industry Giant Telvent*. https://krebsonsecurity.com/2012/09/chinese-hackers-blamed-for-intrusion-at-energy-industry-giant-telvent/. Retrieved: Sept, 2012. 2012.

[117] M. Krotofil and R. Derbyshire. *Greetings from the '90s: Exploiting the Design of Industrial Controllers in Modern Settings*. Black Hat EU, https://i.blackhat.com/EU-21/Wednesday/EU-21-Krotofil-Greetings-from-the-90s-Exploiting-the-Design-of-Industrial-Controllers-in-Modern-Settings.pdf. Retrieved: Dec, 2021. 2021.

[118] M. Krotofil and J. Larsen. *What You Always Wanted and Now Can: Hacking Chemical Processes*. HITB, https://archive.conference.hitb.org/hitbsecconf2015ams/wp-content/uploads/2015/02/D2T1-Marina-Krotofil-and-Jason-Larsen-Hacking-Chemical-Processes.pdf. Retrieved: Dec, 2021. 2015.

[119] Marina Krotofil. *EVIL BUBBLES or How to Deliver Attack Payload via the Physics of the Process*. Black Hat USA, https://www.blackhat.com/docs/us-17/wednesday/us-17-Krotofil-Evil-Bubbles-Or-How-To-Deliver-Attack-Payload-Via-The-Physics-Of-The-Process.pdf. Retrieved: Nov, 2019. 2017.

[120] Marina Krotofil. *What I Learned about ICS Security from Cyber-Physical Hacking*. DHS ICSJWG, https://www.slideshare.net/MarinaKrotofil/dhs-icsjwg-2015v3. Retrieved: Nov, 2019. 2015.

[121] Marina Krotofil and A. D. *Hack Like a Movie Star: Step-by-step guide to crafting SCADA payloads for physical attacks with catastrophic consequences*. ZeroNights, https://repo.zenk-security.com/Conferences/ZeroNights/12-Krotofil.pdf. Retrieved: Feb, 2020.

[122] Marina Krotofil and Alexander Isakov. *Damn Vulnerable Chemical Process - Tennessee Eastman Process*. http://github.com/satejnik/DVCP-TE. Retrieved: Jan, 2015. 2015.

[123] Marina Krotofil and Alexander Isakov. *Damn Vulnerable Chemical Process - Vinyl Acetat Monomer*. http://github.com/satejnik/DVCP-VAM. Retrieved: Jan, 2015. 2015.

[124] Marina Krotofil, Klaus Kursawe, and Dieter Gollmann. "Securing Industrial Control Systems". In: *Security and Privacy Trends in the Industrial Internet of Things*. Springer International Publishing, 2019, pp. 3–27.

[125] Marina Krotofil, Jason Larsen, and Dieter Gollmann. "The Process Matters: Ensuring Data Veracity in Cyber-physical Systems". In: *Proceedings AsiaCCS'15*. 2015, pp. 81–94.

[126] Marina Krotofil et al. "CPS: Driving Cyber-physical Systems to Unsafe Operating Conditions by Timing DoS Attacks on Sensor Signals". In: *Proceedings of the 30th Annual Computer Security Applications Conference*. ACSAC '14. New Orleans, Louisiana, USA: ACM, 2014, pp. 146–155. ISBN: 978-1-4503-3005-3.

[127] Jason Laberge. *Introduction to ASM*. https://www.asmconsortium.net/resources/presentations/Pages/default.aspx. Retrieved: Nov, 2019. 2008.

[128] Canary Labs. *A Guide to the Best Data Historian Software: A Review of the Canary Historian Versus Rockwell FactoryTalk or OSIsoft Pi*. https://blog.canarylabs.com/2016/06/27/a-guide-to-the-best-data-historian-software-a-review-of-the-canary-historian-versus-rockwell-factorytalk-and-osisoft-pi. Retrieved: Nov, 2019. 2016.

[129] Ralf Langner. *To kill a centrifuge*. Tech. rep. Langner Communications, 2013.

[130] Jason Larsen. *Breakage*. Black Hat Federal, `http://www.orkspace.net/secdocs/Conferences/BlackHat/Federal/2008/SCADA%20Security.pdf`. Retrieved: May, 2013. 2007.

[131] Jason Larsen. *Hacking Critical Infrastructure Like You're Not a N00b*. RSA, `https://www.slideshare.net/cisoplatform7/hacking-critical-infrastructure-like-youre-not-a-n00b`. Retrieved: Nov, 2019. 2016.

[132] Jason Larsen. *Miniaturization*. Black Hat USA, `https://www.blackhat.com/docs/us-14/materials/us-14-Larsen-Miniturization-WP.pdf`. Retrieved: Feb, 2015. 2014.

[133] Truls Larsson and Sigurd Skogestad. "Plantwide control – A review and a new design procedure". In: *Modeling, Identification and Cotnrol* 21.4 (2000), pp. 209–240.

[134] Edward A. Lee. "The Past, Present and Future of Cyber-Physical Systems: A Focus on Models". In: *Sensors* 15.3 (2015), pp. 4837–4869.

[135] Robert A. Leishear. *Fluid Mechanics, Water Hammer, Dynamic Stresses, and Piping Design*. ASME, 2013.

[136] Eireann Leverett and Reid Wightman. "Vulnerability inheritance programmable logic controllers". In: *Proceedings of the Second International Symposium on Research in Grey-Hat Hacking*. 2013.

[137] Éreann Leverett and Ried Wightman. "Vulnerability Inheritance Programmable Logic Controllers". In: *GreHack'13* (2013).

[138] N. G. Leveson. *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2012.

[139] Timothy Levine, Hee Park, and Steven McCornack. "Accuracy in Detecting Truths and Lies: Documenting the "Veracity Effect"". In: *Communication Monographs - COMMUN MONOGR* 66 (June 1999), pp. 125–144.

[140] John Leyden. *Hacker jailed for revenge sewage attacks*. `https://www.theregister.com/2016/03/24/water_utility_hacked`. Retrieved: Nov, 2019. 2016.

[141] Y. Li et al. "Jamming attack on Cyber-Physical Systems: A game-theoretic approach". In: *2013 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*. May 2013, pp. 252–257.

[142] Z. Lu, W. Wang, and C. Wang. "From Jammer to Gambler: Modeling and Detection of Jamming Attacks Against Time-Critical Traffic". In: *2011 Proceedings IEEE INFOCOM*. Apr. 2011, pp. 1871–1879.

[143] Z. Lu et al. "Review and Evaluation of Security Threats on the Communication Networks in the Smart Grid". In: *2010 - MILCOM 2010 MILITARY COMMUNICATIONS CONFERENCE*. Oct. 2010, pp. 1830–1835.

[144] Lucas Apa and Carlos Mario Penagos Hollman. *Compromising Industrial Facilities from 40 Miles Away*. Tech. rep. IOActive, 2013.

[145] Peter Lutz. *OPC UA – from automation pyramid to information network*. https://iebmedia.com/technology/iiot/opc-ua-from-automation-pyramid-to-information-network/. retrieved: May, 20122.

[146] Michael L. Luyben, Bjorn D. Tyreus, and William L. Luyben. "Plantwide control design procedure". In: *AIChE Journal* 43.12 (1997), pp. 3161–3174.

[147] Michael L. Luyben and Björn D. Tyréus. "An industrial design/control study for the vinyl acetate monomer process". In: *Computers & Chemical Engineering* 22.7–8 (1998), pp. 867–877.

[148] William L. Luyben, Bjorn D. Tyreus, and Michael L. Luyben. *PlantwideProcess Control*. McGraw-Hill, 1998.

[149] M. Dobrushin and F. Rosenfeld. *OTORIO "Project RunAway": How Manufacturers Unknowingly Leak Classified Project Files*. Resources, https://www.youtube.com/watch?v=sj87AXoa4SI&t=629s. Retrieved: Jan, 2023. 2019.

[150] FireEye Mandiant. *Mandiant Attack Lifecycle*. https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/ds-threatspace.pdf. Retrieved: Nov, 2021. 2021.

[151] Sam Mannan. *Lees' Loss Prevention in the Process Industries: Hazard Identification, Assessment and Control*. Vol. 1. Butterworth Heinemann, 2005.

[152] John C. Matherly. *SHODAN: Industrial Control Systems*. https://www.shodan.io/explore/category/industrial-control-systems. Retrieved: Dec, 2019.

[153] Lee Mathews. *Criminals Hacked A Fish Tank To Steal Data From A Casino*. https://www.forbes.com/sites/leemathews/2017/07/27/criminals-hacked-a-fish-tank-to-steal-data-from-a-casino/#656917b732b9. Retrieved: Nov, 2019. 2017.

[154] A. P. Mathur and N. O. Tippenhauer. "SWaT: a water treatment testbed for research and training on ICS security". In: *2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*. 2016, pp. 31–36.

[155] MathWorks. *Matlab*. https://www.mathworks.com/products/matlab.html. Retrieved: Nov, 2013.

[156] MathWorks. *Simulink*. https://www.mathworks.com/products/simulink.html. Retrieved: Nov, 2013.

[157] McAfee Foundstone Professional Services and McAfee Labs. *Global Energy Cyberattacks: "Night Dragon"*. Tech. rep. McAfee, 2011.

[158] Craig McIntyre. "Using Smart Instrumentation". In: *Plant Engineering: online magazine* (2011). http://www.controleng.com/single-article/using-smart-instrumentation/a0ec350155bb86c8f65377ba66e59df8.html.

[159] A. F. Menendez. *CAFFEINE: Attack Framework for Electric Sector*. https://cybercamp.es/sites/default/files/contenidos/videos/adjuntos/cybercamp19_afm_caffeine.pdf. Retrieved: Nov, 2019. 2019.

[160] Ariana Mirian et al. "An Internet-wide view of ICS devices". In: *2016 14th Annual Conference on Privacy, Security and Trust (PST)*. 2016, pp. 96–103.

[161] Sheppard Mullin. *Power Company Slammed With Hefty 2.7M Fine After Data Breach*. https://www.eyeonprivacy.com/2018/03/power-company-regulator. Retrieved: Feb, 2020. 2018.

[162] M. Nawrocki, T. C. Schmidt, and M. Waehlisch. "Uncovering Vulnerable Industrial Control Systems from the Internet Core". In: *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*. 2020, pp. 1–9.

[163] Lee Neitzel and Bob Huba. "Top ten differences between ICS and IT cybersecurity". In: *InTech* 61.3 (2014), pp. 12–18.

[164] NERC. *Critical Infrastructure Protection Standards*. http://www.nerc.com/pa/Stand/Pages/CIPStandards.aspx. Retrieved: Nov, 2014.

[165] Ian Nimmo and Stephen Maddox. *A lasting plan for managing alarms*. https://www.controlglobal.com/articles/2015/a-lasting-plan-for-managing-alarms. Retrieved: Nov, 2019. 2015.

[166] T. Novak and A. Gerstinger. "Safety- and Security-Critical Services in Building Automation and Control Systems". In: *IEEE Transactions on Industrial Electronics* 57.11 (2010), pp. 3614–3621.

[167] NullArray. *s7scan*. https://github.com/NullArray/AutoSploit. Retrieved: Nov, 2019. 2018.

[168] D. G. Olsen, W. Y. Svrcek, and B.R. Young. "Plantwide Control Study of a Vinyl Acetate Monomer Process Design". In: *Chemical Engineering Communication* 192.10 (2005), pp. 1243–1257.

[169] OTORIO Research. *Project Files are the Blueprints of the Industrial Processes*. Blog post, https://www.otorio.com/blog/manufacturers-unknowingly-leak-classified-project-files/. Retrieved: Jan, 2023. 2019.

[170] P. Bullemer. *Creating an ASM-compliant HMI goes deeper than screen color selection*. Online, https://www.controleng.com/articles/creating-an-asm-compliant-hmi-goes-deeper-than-screen-color-selection/. Retrieved: June, 2014. 2014.

[171] Danny Palmer. *Ransomware attacks are now targeting industrial control systems*. https://www.zdnet.com/article/ransomware-attacks-are-now-targeting-industrial-control-systems. Retrieved: Nov, 2019. 2019.

[172] Donghui Park, Julia Summers, and Michael Walstrom. *Cyberattack on Critical Infrastructure: Russia and the Ukrainian Power Grid Attacks*. https://jsis.washington.edu/news/cyberattack-critical-infrastructure-russia-ukrainian-power-grid-attacks. Retrieved: Nov, 2019. 2017.

[173] Donn B. Parker. "Toward a New Framework for Information Security". In: *The Computer Security Handbook*. Ed. by Seymour Bosworth and M. E. Kabay. 4th ed. John Wiley & Sons, 2002. Chap. 5.

[174] Ulyana A. Pozdnyakova et al. "Genesis of the Revolutionary Transition to Industry 4.0 in the 21st Century and Overview of Previous Industrial Revolutions". In: *Industry 4.0: Industrial Revolution of the 21st Century*. 2019, pp. 11–19.

[175] T. Zimmerman R. Candell and K. Stouffer. *Industrial Control System Cybersecurity Performance Testbed*. Tech. rep. NIST, 2015.

[176] Bob Radvanovsky. *Florida city water cyber incident allegedly caused by employee error*. http://scadamag.infracritical.com/index.php/2023/03/29/florida-city-water-cyber-incident-allegedly-caused-by-employee-error/. Retrieved: April, 2023. 2023.

[177] Yongyi Ran et al. "A Survey of Predictive Maintenance: Systems, Purposes and Approaches". In: *arXiv* 1912.07383 (2019).

[178] Rapid7. *Vulnerability Exploit Database*. https://www.rapid7.com/db/?type=metasploit. Retrieved: Feb, 2020.

[179] ESET Research. *ACAD/Medre.A: 10000's of AutoCAD Designs Leaked in Suspected Industrial Espionage*. Tech. rep. ESET, 2012.

[180] iTRUST Research Centre. *Garden of Testbeds: Booking Requests*. https://itrust.sutd.edu.sg/itrust-labs_overview. Retrieved: Nov, 2019.

[181] iTRUST Research Centre. *Secure Water Treatment (SWaT) Testbed*. http://locnuoctienthanh.net/document/water/ReverseOsmosisPID.pdf. Retrieved: Feb, 2020. 2018.

[182] N. L. Ricker. *Tennessee Eastman Challenge Archive*. http://depts.washington.edu/control/LARRY/TE/download.html. Retrieved: May, 2013.

[183] S.M. Rinaldi, J.P. Peerenboom, and T.K. Kelly. "Identifying, understanding, and analyzing critical infrastructure interdependencies". In: *Control Systems, IEEE* 21.6 (2001), pp. 11–25.

[184] Control Risks. *Iranian Cyber Espionage Group Targets Suppliers of Industrial Control Systems*. https://www.controlrisks.com/our-thinking/insights/iranian-cyber-espionage-group-targets-suppliers-of-industrial-control-systems. Retrieved: Dec, 2020. 2020.

[185] Douglas H. Rothenberg. *Alarm Management for Process Control: A Best-Practice Guide for Design, Implementation, and Use of Industrial Alarm Systems*. Momentum Press, 2009.

[186] Adepu S., Kandasamy N.K., and Mathur A. "EPIC: An Electric Power Testbed for Research and Training in Cyber Physical Systems Security". In: *Security of Industrial Control Systems and Cyber-Physical Systems (CyberICPS 2018)*. Vol. 11387. LNCS. 2018, pp. 20–28.

[187] S. Brizinov (Claroty). *The Race to Native Code Execution in PLCs*. Blog, https://claroty.com/team82/research/the-race-to-native-code-execution-in-plcs-using-rce-to-uncover-siemens-simatic-s7-1200-1500-hardcoded-cryptographic-keys. Retrieved: Jan, 2023. 2022.

[188] S. Maddox. *Alarms and Operator Intervention – A Flawed Safety Layer*. Online, https://https://isawaterwastewater.com/wp-content/uploads/2019/10/isa-wwid_2018-webinar2_Alarms-a-Flawed-Safety-Layer-1.pdf. Retrieved: May, 2020. 2019.

[189] U.S. Chemical Safety and Hazard Investigation Board. *BP America Refinery Explosion: Final Investigation Report*. https://www.csb.gov/bp-america-refinery-explosion. Retrieved: May, 2013. 2007.

[190] U.S. Chemical Safety and Hazard Investigation Board. *Completed Investigations*. https://www.csb.gov/investigations/completed-investigations. Retrieved: November, 2019.

[191] U.S. Chemical Safety and Hazard Investigation Board. *DuPont Corporation Toxic Chemical Releases: Investigation Report*. https://www.csb.gov/dupont-corporation-toxic-chemical-releases. Retrieved: May, 2013. 2011.

[192] U.S. Chemical Safety and Hazard Investigation Board. *T2 Laboratories Inc. Reactive Chemical Explosion: Final Investigation Report*. https://www.csb.gov/t2-laboratories-inc-reactive-chemical-explosion. Retrieved: May, 2013. 2009.

[193] Ruben Santamarta. *Here be Backdors: A Journey into the Secrets of Industrial Firmware*. Black Hat USA, https://media.blackhat.com/bh-us-12/Briefings/Santamarta/BH_US_12_Santamarta_Backdoors_Slides.pdf. Retrieved: December, 2012. 2012.

[194] SCADAStrangeLove. *SCADAPASS*. https://github.com/scadastrangelove/SCADAPASS/blob/master/scadapass.csv. Retrieved: Feb, 2020. 2016.

[195] Thorsten Schneider. *Damn Vulnerable Linux*. https://www.linux.com/tutorials/damn-vulnerable-linux. Retrieved: Nov, 2019. 2010.

[196] Kudelski Security. *6 Months of ICS Scanning*. https://research.kudelskisecurity.com/2017/10/24/6-months-of-ics-scanning. Retrieved: Feb, 2018. 2017.

[197] Hiroya Seki et al. "Plantwide Control System Design of the Benchmark Vinyl Acetate Monomer Production Plant". In: *Computers & Chemical Engineering* 34.8 (2010).

[198] S. Senthivel et al. "Denial of Engineering Operations Attacks in Industrial Control Systems". In: *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*. 2018.

[199] Kaspersky Lab Security Services. *s7scan*. https://github.com/klsecservices/s7scan. Retrieved: Nov, 2019. 2018.

[200] Yasser Shoukry et al. "Non-Invasive Spoofing Attacks for Anti-Lock Braking System". In: *Cryptographic Hardware and Embedded Systems*. CHES 2013. Springer, 2013, pp. 55–72.

[201] Siemens. *Industrial 5G – The Wireless Network of the Future*. https://press.siemens.com/global/en/feature/industrial-5g-wireless-network-future. Retrieved: Nov, 2019. 2019.

[202] Sigma-Aldrich. *Paracetamol*. https://www.sigmaaldrich.com/DE/de. Retrieved: Jan, 2015.

[203] Tony Smith. *Hacker jailed for revenge sewage attacks*. https://www.theregister.com/2001/10/31/hacker_jailed_for_revenge_sewage. Retrieved: Nov, 2019. 2001.

[204] Jacques F. Smuts. *Process Control for Practitioners*. OptiControls Inc, 2011.

[205] Anthony Sofronas. *Case Histories in Vibration Analysis and Metal Fatigue for the Practicing Engineer*. John Wiley & Sons, 2012.

[206] R. Spenneberg, m Brueggemann, and H. Schwartke. *PLC-Blaster: A Worm Living Solely in the PLC*. Black Hat Asia, https://www.blackhat.com/docs/asia-16/materials/asia-16-Spenneberg-PLC-Blaster-A-Worm-Living-Solely-In-The-PLC-wp.pdf. Retrieved: Jan, 202023. 2016.

[207] P. Sun, L. Garcia, and S. Zonouz. "Tell Me More Than Just Assembly! Reversing Cyber-Physical Execution Semantics of Embedded IoT Controller Software Binaries". In: *49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 2019, pp. 349–361.

[208] O. Taha and M. Rashid Khan. "Advanced process control for clean fuel production: smart plant of the future". In: *Advances in Clean Hydrocarbon Fuel Processing*. Woodhead Publishing Series in Energy. Woodhead Publishing, 2011, pp. 496–508.

[209] Talos. *New VPNFilter malware targets at least 500K networking devices worldwide, 2018*. https://blog.talosintelligence.com/2018/05/VPNFilter.html. Retrieved: Nov, 2019. 2018.

[210] DVWA team. *Damn Vulnerable Web Application (DVWA)*. https://github.com/ethicalhack3r/DVWA. Retrieved: Nov, 2019. 2010.

[211] ICSMASTER Security Team. *Industrial Security Exploitation Framework*. https://github.com/w3h/isf. Retrieved: Nov, 2019. 2017.

[212] National Telecommunications and Information Administration. *Software Bill of Materials*. https://ntia.gov/SBOM. Retrieved: June, 2022. 2020.

[213] The Dow Chemical Company. *Product Safety Assessment: Vinyl Acetate*. http://www.dow.com/productsafety/finder/vinyl.htm. Retrieved: June, 2014.

[214] The European Network of Transmission System Operators for Electricity. *Continental Europe Significant Frequency Deviations - January 2019*. Tech. rep. ENTSO-E, 2019.

[215] Thomas Brewster. *Hundreds Of Wind Turbines And Solar Systems Wide Open To Easy Exploits*. https://www.forbes.com/sites/thomasbrewster/2015/06/12/hacking-wind-solar-systems-is-easy. Retrieved: Feb, 2015. 2015.

[216] Joe Tidy. *Predatory Sparrow: Who are the hackers who say they started a fire in Iran?* https://www.bbc.com/news/technology-62072480. Retrieved: July, 2022. 2022.

[217] Aleksandr Timorin. *SCADA tools*. https://github.com/atimorin/scada-tools. Retrieved: Feb, 2020. 2014.

[218] Michael Toecker. *Didgital Maintenance and Test Equipment and Impact on Control Systems Security*. https://www.slideshare.net/MichaelToecker/maintenance-and-test-equipment-cyber-security. Retrieved: March, 2016. 2015.

[219] Control Tools. *Split Range Control Application using PLC Ladder Logic*. https://instrumentationtools.com/split-range-control-plc-ladder-logic/. retrieved: May, 2018.

[220] Virus Total. *Virus Total online service*. www.virustotal.com. Retrieved: Nov, 2019.

[221] Yazhou Tu et al. “Trick or Heat? Manipulating Critical Temperature-Based Control Systems Using Rectification Attacks”. In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. CCS ’19. Association for Computing Machinery, 2019, pp. 2301–2315.

[222] D. Tychalas, H. Benkraouda, and M. Maniatakos. “ICSFuzz: Manipulating I/Os and Repurposing Binary Code to Enable Instrumented Fuzzing in ICS Control Applications”. In: *30th USENIX Security Symposium*. 2021.

[223] George Tzagkarakis et al. “Signal and Data Processing Techniques for Industrial Cyber-Physical Systems”. In: Oct. 2015, pp. 181–226.

[224] J. Bartz und U. Stoll. *Die Vulkan Files Die geheimen Waffen russischer Cyberkrieger*. `https://www.zdf.de/politik/frontal/doku-vulkan-files-cyberangriff-russland-ukraine-krieg-leak-daten-100.html`. Retrieved: April, 2023. 2023.

[225] NPC Ukrenergo. *Press tour on substation 330 kV "Severnaya"*. `https://www.youtube.com/watch?v=AUoiKZBqIo0`. Retrieved: Feb, 2020. 2015.

[226] Arturo Urquizo. *PID controller overview*. `https://commons.wikimedia.org/wiki/File:PID_en.svg`. Retrieved: Nov, 2019. 2011.

[227] Verizon. *Data breach digest. Scenarios from the field.* `https://maritimecyprus.com/wp-content/uploads/2016/03/verizon_data-breach-digest_en-1.pdf`. Retrieved: Dec, 2017. 2016.

[228] Joe Weiss. *Marina Krotofil’s presentation on how to hack a chemical plant and it’s implication to actual issues at a nuclear plant*. `https://www.controlglobal.com/blogs/unfettered/marina-krotofils-presentation-on-how-to-hack-a-chemical-plant-and-its-implication-to-actual-issues-at-a-nuclear-plant`. Retrieved: Nov, 2019. 2015.

[229] J. Wetzels. *OT Deep Lateral Movement: When Perimeter is Not a Perimeter*. Technical report: `https://www.ibm.com/reports/data-breach`. Retrieved: May, 2023.

[230] J. Wetzels and M. Krotofil. *A Diet of Poisoned Fruit: Designing Implants and OT Payloads for ICS Embedded Devices*. TROOPERS, `https://troopers.de/downloads/troopers19/TROOPERS19_NGI_IoT_diet_poisoned_fruit.pdf`. Retrieved: March, 20120. 2019.

[231] Theodore J. Williams. “The Purdue Enterprise Reference Architecture”. In: *Computers in Industry* 24.2–3 (Sept. 1994), pp. 141–158. ISSN: 0166-3615.

[232] Marian Willuhn. *Hackerangriff auf Satelliten legt Steuerung von elf Gigawatt Windkraftanlagen lahm*. `https://www.pv-magazine.de/2022/03/01/hackerangriff-auf-satelliten-legt-steuerung-von-elf-gigawatt-windkraftanlagen-lahm/`. Retrieved: March, 2022. 2022.

[233] Alexander Winnicki, Marina Krotofil, and Dieter Gollmann. “Cyber-Physical System Discovery: Reverse Engineering Physical Processes”. In: *Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security*. CPSS ’17. 2017.

[234] Dazhong Wu et al. “Cloud-Based Design and Manufacturing: A New Paradigm in Digital Manufacturing and Design Innovation”. In: *Computer-Aided Design* 59 (), pp. 1–14.

[235] Xuan Feng et al. “Characterizing industrial control system devices on the Internet”. In: *2016 IEEE 24th International Conference on Network Protocols (ICNP)*. 2016.

[236] Z. Yang et al. “Reverse Engineering Physical Semantics of PLC Program Variables Using Control Invariants”. In: *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. SenSys ’22. 2023.

[237] Daniel Zafra. *Fantastic Information and Where to Find It: A Guidebook to Open-Source OT Reconnaissance*. `https://www.virusbulletin.com/uploads/pdf/magazine/2019/VB2019-Kapellmann-Zafra.pdf`. Retrieved: Dec, 2019. 2019.

[238] M. Zeller. “Myth or reality - Does the Aurora vulnerability pose a risk to my generator?” In: *Protective Relay Engineers, 2011 64th Annual Conference for*. 2011.

[239] Kim Zetter. *Analysis of the Cyber Attack on the Ukrainian Power Grid*. `https://www.vice.com/en_us/article/bmvkn4/ukrainian-power-station-hacking-december-2016-report`. Retrieved: Nov, 2019. 2017.

[240] Kim Zetter. *Revealed: How a secret Dutch mole aided the U.S.-Israeli Stuxnet cyber-attack on Iran*. `https://news.yahoo.com/revealed-how-a-secret-dutch-mole-aided-the-us-israeli-stuxnet-cyber-attack-on-iran-160026018.html`. Retrieved: Feb, 2020. 2019.

[241] Isaac A. Zlochower and Gregory M. Green. “The limiting oxygen concentration and flammability limits of gases and gas mixtures”. In: *Journal of Loss Prevention in the Process Industries* 22.4 (2009), pp. 499–505.

A. Data Security in Cyber-Physical Systems

With the progressive digitization of industrial control systems, process data collection, processing, storing and presentation became an essential task of the control infrastructures with process data becoming one of the "assets" to manage. Control system applications supervise all of the functions and operations in a plant by taking in real-time data from the field instrumentation and using the trends extracted from the process data to make decisions about each process unit performance. When speaking about software applications, they are seen as applications that run on top of an untrusted IT infrastructure. Input validation is performed to ensure that only properly formed data is entering the application workflow, preventing the unexpected outcome of code execution. Physical processes similarly rely on the process data being appropriate and not harmful to the physical application.

Security properties should fit the application's needs. For a long time it was thought that if data security policies were guided by the Confidentiality, Integrity, Availability triad, the data was "secure". With the change in the volume of data being produced and the way how and where data was stored and used, the CIA model began to be seen as limiting. In 2002 Donn Parker extended the CIA model to better describe data security requirements as shown in Figure A.1 [173].

Parker's six attributes of information security are non-overlapping in that they refer to unique aspects of information with *utility* or *data usefulness* being the only property that is not necessarily binary in nature. Not only utility can have degrees of usefulness, it is also highly contextual and can only be evaluated in relation to a specific application.

In process automation, field devices such as sensors are considered to be monolithic units that are fully *trusted* [163] and the data originated by the devices are correspondingly considered *trustworthy*. In many cases the controller and operator can observe a physical process only through process measurements, therefore their accuracy is critical. As discussed in Chapter 2, acquired sensor signals pass through a variety of functions that process them in a variety of ways such as amplification, scaling, conversion, filtering, aggregation and normalization to name a few. Furthermore, data sources are combined through computation formulas prior to being consumed by additional control logic and applications. In essence,

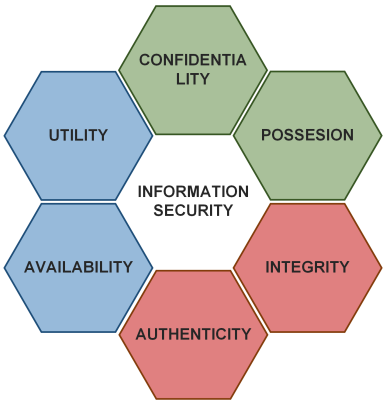


Figure A.1: The Parkerian hexad

data processing is conducted to provide usable, i.e., actionable information, based on the requirements defined by data consuming circuits/devices/applications at each stage in a data chain. Any error in data acquisition or processing along data routes harbors the potential to degrade and even lose visibility of the process state. This is why the understanding of data sources and their pathways is essential to the assessment of the undesirable impact on process operations, caused by errors or intentional manipulation of data streams. In process control, data condition/content is addressed in the context of *"data quality"* or *"data reliability"*. Telemetry engineers are frequently responsible for this task.

Data utility refers to the data accuracy and usefulness as a result of its transformation within the computing infrastructure. The utility does not address how trustworthy or accurate the source of the data is, and whether the information extracted from such data is correctly representing the reality. Process data are originated in the physical process and measured by the sensors. It would be prudent to establish methods able to determine if the data received truthfully capture the attributes of the physical world relevant to the application under control. This leads to a new security requirement called *veracity*, or trustworthiness of data. The relationship between data veracity and data utility is shown in Figure A.2.

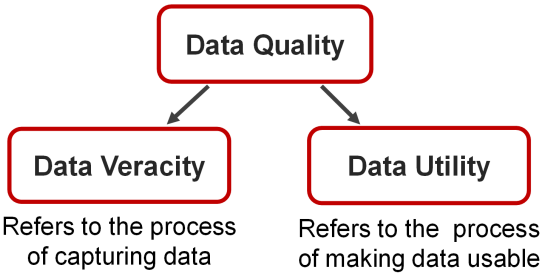


Figure A.2: Data quality in process control systems

A.1 Data Utility

Utility stands for data usefulness. Reduced data utility may result in the loss of vital or critical information resulting in significant consequences. The utility of process data can be impaired due to different causes. Consider a use case of obligatory regulatory monitoring of certain process parameters for conformity to certain operational requirements. For instance, in the UK regulators such as the Department for Environment, Food and Rural Affairs (DEFRA) require data that identifies wastewater treatment facilities' conformity to predetermined operational restrictions such as quality and quantity of final effluent [60]. Data must be collected from the appropriate sources within the operational process, conforming to any defined time constraints. While the regulatory instructions impose specific operational requirements to the process data, the same data may also be used for performance analysis, and by decision makers at the board level, to determine the financial viability of operational sites, and for identification of required investments to streamline processes. Furthermore, the same data may also be applied to the remote alarm monitoring of unmanned operational processes. In some cases, this remote alarm management may form part of an organizational safety reporting process. It is evident that without complete knowledge of all stakeholders of the individual data sources and understanding of their needs on data content, there is a high risk of losing or misrepresenting critical process information due to erroneous configuration of data processing points.

Figure A.3 presents the flow of data from a single sensor to the remote access connectivity residing within the corporate level of the OT reference architecture. When an instrumentation engineer recalibrates or replaces a sensor, she will likely know to liaise with a PLC configuration engineer to account for the new properties of sensor signals within the PLC logic, therefore maintaining accurate control of the physical process. This fulfills the primary requirement of sensors, that of providing accurate input to operational decision making applications such as control logic. However, when re-configuration requirements arise in relation to neighboring devices such as in-line monitors or data historians if their configuration is left unadjusted, regulatory, alarm monitoring, safety, and performance analytics data may become compromised. In the given example, awareness of sensor data consumption beyond the PLC could fall outside the scope of an instrumentation engineer's role. As a result, while the data processing formulas in the PLC's control logic were modified to maintain the accuracy of calculations, the other data processing points across the data path were not, leading to the corruption of calculations and stored values in the historian server.

When a situation like this occurs, from the initial suspicions around data quality, historian support engineers would drill down into complex mathematical calculations to identify the root cause of spurious data. It is not infrequent that the calculations are derived from up to 30+ operational tags (signal inputs), therefore the investigation process could prove to be time-consuming. Furthermore, if data is processed at multiple points in the system, the visibility into data processing points may end quickly, requiring interaction with Level 1 control engineers to better understand any changes further downstream, across suspected data flows. It is evident that preserving data utility from the source and along its path is critical for ensuring business continuity and reducing troubleshooting efforts.

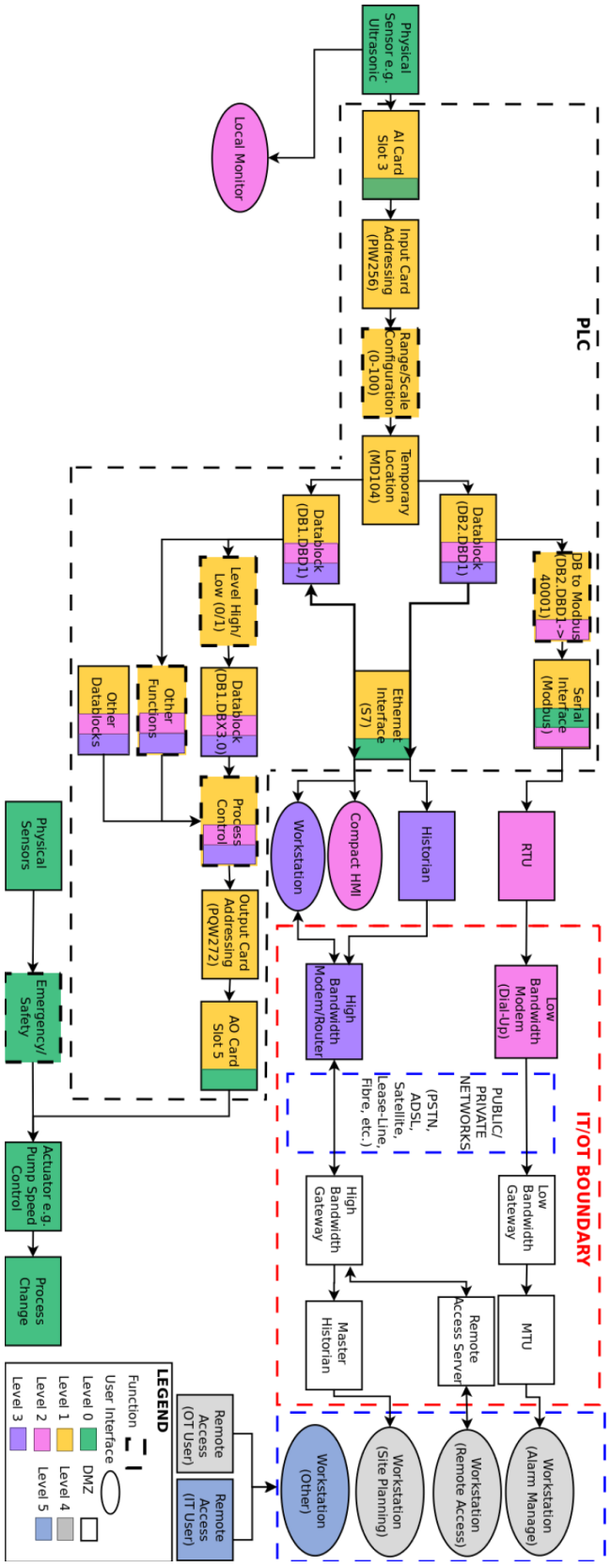


Figure A.3: Data flow of a single sensor through the levels of ICS reference architecture [87]

Data Utility Impact

In a public example [68] it was discussed how a control engineer in a pumping station responsible for PLC programming used the 0 to 4095 analog-to-digital converter counts range to convert the $4-20$ mA analog signal from the measuring instrument whereas the engineer responsible for the telemetry used the 819 to 4095 range on the RTU. The 819 counts offset was introduced to detect signal underrange as shown in Figure A.4. As a result, when the PLC received zero pressure (4 mA), it converted it to 0 counts whereas the RTU program output 819 counts for zero pressure. The SCADA read process data from the RTU whereas the HMI at the site received data from the PLC. Consequently, the mismatch in data scaling leads to inconsistencies in data readings across different applications. It is apparent that such inconsistencies in converting measured data into digital counts and subsequently into engineering units may result in undesired consequences.

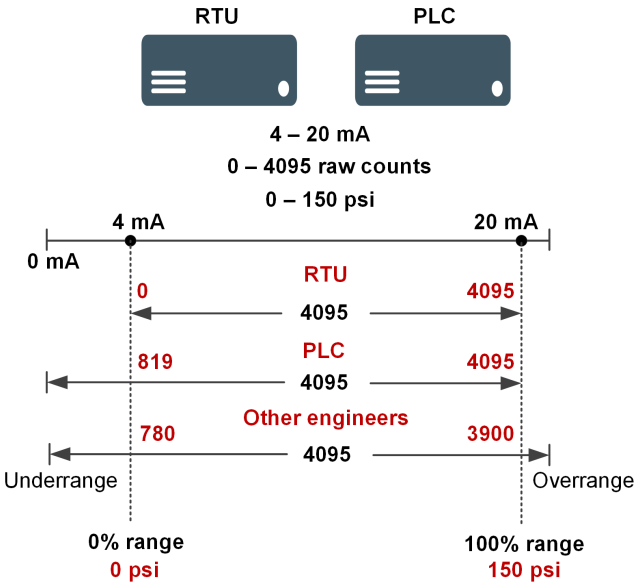


Figure A.4: Inconsistency of control equipment configuration

In process automation, field devices such as sensors are considered to be monolithic units that are fully trusted [163] and the data originated by the devices are correspondingly considered trustworthy. To provide useful measurement units, an analog signal must be scaled first. The accuracy of the scaling is validated through *calibration*. Improperly calibrated measuring instruments can be a source of safety accidents. An example is an explosion at BP Texas City Refinery (15 people killed, 180 injured) [189]. The root cause of that tragedy was critical alarms and control instrumentation providing false indications. Due to calibration errors the splitter tower level indicator showed that the tower level was declining when it was actually overfilling with flammable liquid hydrocarbons. As a result, the operator kept filling the tower. A chain of further events eventually led to an explosion. Hence, manipulation of instrument calibration is a potentially dangerous attack vector.

Calibration systems are increasingly transforming from stand-alone systems or work processes to software-based integrated solutions [7]. Although the automation of the calibration processes and tight integration into the documentation and maintenance management

systems was driven by the desire to reduce human error, the integration also introduced an easy-to-exploit and dangerous attack vector. Quoting a software-induced incident revealed by safety systems vendor HIMA [120]:

> Due to an unknown bug in the engineering software, all scaling of the SIS AI (analog input) got altered from 0 to 100% automatically. The altered value got loaded and activated automatically based on in unknown bug at the same system.

In the past, sensors used to be purely analog and the measured analog signal was converted into a digital signal by the controller. These days it is more common for a microcontroller to be embedded into the sensor itself [158]. The presence of a digital stage in the sensor allows attack scenarios where a sensor behaves maliciously but always passes calibration tests.

Figure A.5 shows how the change of sensor sampling frequency (blue signal) would allow an attacker to conceal process oscillations (red signal) and impair forensic investigation.

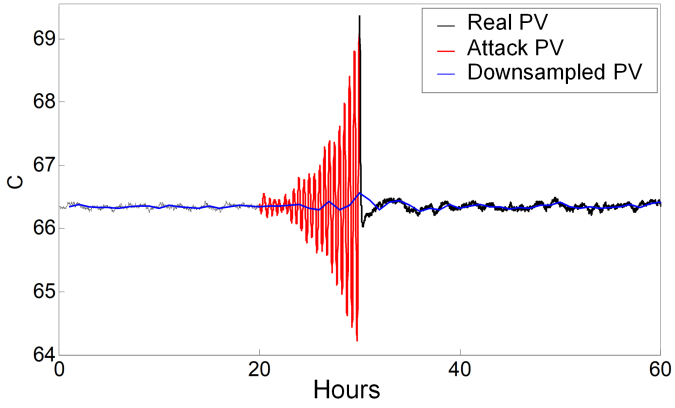


Figure A.5: Concealment of attack trace (oscillation) by changing signal sampling frequency

Currently, there are no methods to identify whether some information is potentially missing in sensor signals. This is a research question that deserves attention, especially when forensically sound data are needed for legally-bounded decisions.

A.2 Data Veracity

The word veracity derives from the Latin adjective vērāx – "truthful", which in turn comes from the earlier vērus – "true". Initially, the property of veracity gained its popularity in the area of criminology – the ability to detect whether a witness is veracious or not [139]. The term veracity is used both in relation to actors such as witnesses and their statements. Borselius extended veracity to the property of mobile agents namely that they will not knowingly communicate false information [23].

We refer to statements about aspects relevant in a given application domain as *assertions*. Assertions may be true or false. In this white paper we use the definition of veracity proposed in [83].

Veracity: *The property that an assertion truthfully reflects the aspect it makes a statement about.*

Restating, data veracity is the degree of accuracy or truthfulness of a data set. Veracity is not a property guaranteed by any of the familiar IT infrastructure security services. Veracity refers to aspects outside the IT infrastructure: the adversary is not an entity launching an attack in the infrastructure but an entity making false assertions. The data are already false when passed to the infrastructure. "Trusted" sensor data just because they have been digitally signed do not help when the sensor readings are incorrect or have been manipulated before being submitted to the communication protocol stack. Veracity is a new security property relevant at the application level.

There is an old engineering aphorism that says *"one must measure a process in order to control it"*, which is still as current today as it was in the past. A corollary to this saying is that trustworthy process measurement information is required so that users feel comfortable enough to risk making tighter control improvements [158]. For data to be regarded as trustworthy, particularly to support the decision-making process, it is vital that data accurately reflect the reality. However, data veracity is a complex property that deals with uncertainties in data brought about by a diversity of causes with biases, noise, abnormalities and missing values being a few examples. Enforcing veracity can be challenging. Quoting the words of Prof. Vijay Varadharajan in [19]:

> Every physical and virtual device in the IoT infrastructure generating huge quantities of data presents immediate and direct consequences. Just because data are accessible doesn't mean it's trustworthy or reliable for making decisions, or even ethical to access and use it.

To secure an application, one thus may have to go beyond securing the infrastructure and in addition provide mechanisms for detecting false data. Bellow we examine examples of how data veracity can be impacted in process control applications and then briefly outline approaches to detecting non-veracious sensor signals.

Violation of Data Veracity

In many cases the operator can only observe the process through sensors and must have faith that the sensors describe the true underlying physical reality. This faith can be used by an attacker even when error handling is in place and used to mathematically solve the error in the system. For example, in the power grid complementary disturbances caused in adjacent nodes are canceled out as a residual error in a state estimator.

Violation at the Process Level

The lowest level of manipulation possible is the actual physics of the process itself. The most intuitive attack scenario is a manipulation of the environment around a sensor. This attack vector is especially applicable to sensors that can be "approached" by the attacker such as weather and ambient sensors or proximity and object detection sensors. Increasingly, data veracity is becoming of concern in the automotive sector. Modern cars are equipped with a large number of sensors some of which are exposed to the outer surroundings without physical protection. Shoukry et al. [200] showed that by modifying the physical environment around analog sensors such as Antilock Braking System (ABS) it is possible to

alter wheel speed measurements and potentially induce life threatening situations. ABS relies on magnetic–based wheel speed sensors which are exposed to an external attacker from underneath the body of a vehicle. By placing a thin electromagnetic actuator near the ABS wheel speed sensors, an attacker can inject magnetic fields to both cancel the true measured signal and inject a malicious signal, thus spoofing the measured wheel speeds.

Industrial processes are typically located within the protective boundaries of a physical fence or similar and therefore are not easily approachable by attackers. However, it is possible to achieve an adversarial impact on data veracity by creating special conditions in the physical process itself via remote attacks. In continuous processes, even if two sensors are segregated electronically, the process physics may connect them creating a "data flow". When the logic of a field device operates on a particular datum, that datum may actually be an aggregate of other data even if that value was directly measured from the process. In such case, the *"unseen"* process data may have a negative impact on the control loop performance. Consider a pressure measurement. Depending on the process, changes in pressure may be the result of temperature, flow, volume, or reaction speed. A simultaneous change in temperature and volume may be incorrectly interpreted as a change in reaction speed. Incorrect interpretation of unmeasured quantities (here: reaction speed) is a frequent concern in process control.

Consider another example of a process unit that produces ammonia with pressure in a vessel being maintained with an inflow pump. Maintaining the right ammonia pressure is critical to the financial health of the plant, therefore its control loop is closely monitored. A second pump is responsible for the outflow of ammonia from the vessel but is not considered critical to the plant economy. It may be possible to set up a standing wave between the two pumps that has a direct impact on the ability of the first pump to perform its function. In that case, the unseen state of the second pump is critical to the functionality of the first pump even though there are no electronic data flows between these two control loops. A variation of this attack scenario has been implemented on a realistic hardware setup to degrade the performance of an analog pump. By modifying the position of the inlet valve, an attacker was able to cause a cavitation process which impacted the working of the pump up to full stoppage of the flow [119].

Although data veracity attacks at the layer of the physical process are hard to detect and diagnose, they are also much harder to execute from the design and attack reliability standpoints. During such attacks both the operator and the adversary operate in conditions of limited visibility of the underlying process physics. Depending on the attack scenario, prior testing on a mockup system may be required to determine the exact attack parameters.

Violation at the Sensor Level

Input devices such as sensing devices are inherently susceptible to attacks with unvalidated, unwanted inputs. Attackers can exploit the physics of materials to fool sensors into becoming unintentional receivers of malicious signals to manipulate sensor output or induce intentional errors. This class of attack is called a transduction attack [77]. This is one of the most frequent attack scenarios covered in the CPS security research literature. For instance, Tu et al. [221] implemented an attack on temperature sensors that exploits an unintended rectification effect in analog amplifiers that can be induced by injecting electromagnetic interference (EMI) at a certain wavelength. The attack scenario bypassed conventional noise filtering and generated a controllable DC (direct current) voltage offset at the analog-

to-digital converter (ADC) input. Due to the altered offset, the resistance of a resistance temperature detector (RTD) was converted into a wrong albeit "valid" value within the operating range. As stated by the researchers, "from meters away or an adjacent room, an attacker could trick the internal control system of an infant incubator to heat or cool to unsafe temperatures."

In another example, Bolshev et al. [20, 22] exploited the ADC's input signal sampling process in industrial equipment. Specifically, the authors discussed a scenario of injecting a specially crafted analog signal which would be converted into different digital values by the serially connected equipment on an analog line. In the architecture depicted in Figure A.6, a PLC responsible for the regulatory control sends an analog control command to the actuator (e.g., a motor) whereas a so-called "safety PLC" performs checks of that command for unsafe conditions. It was shown that it was possible to generate an analog signal that would be converted into different digital values by different equipment. For instance, the actuator converted the analog signal into 1.5 V (ON command) and the safety PLC output 0 V (OFF command on the HMI). If such an attack is successfully executed on the plant floor, the operator would lose awareness of the true state of the process and could make wrong and potentially harmful control decisions. Similarly, the safety protection would not engage due to a wrong notion of process state. The perceived threat of the transduction attacks to the industrial-grade sensing equipment has grown such that in one instance U.S. ICS-CERT warned manufacturers of the related hardware design flaws in a security advisory [55].

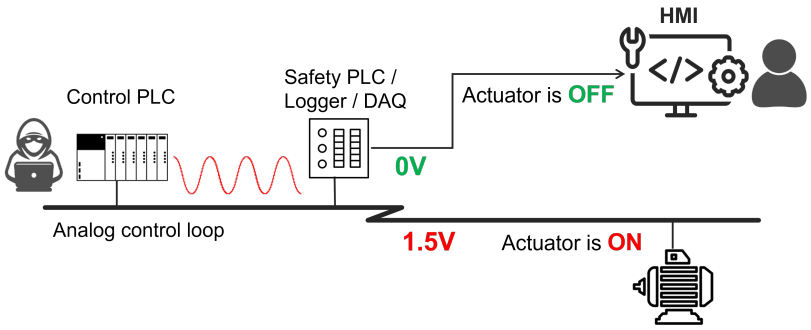
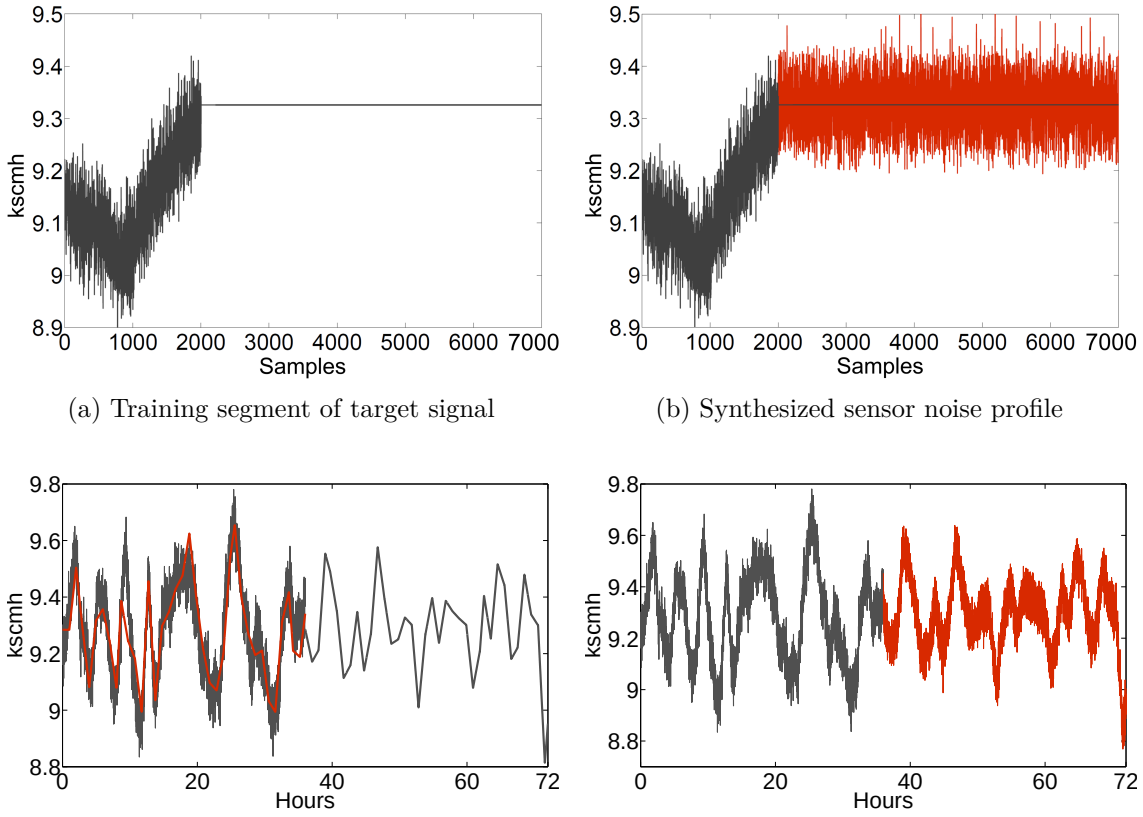


Figure A.6: Threat scenario for analog control loop [20]

While the implementation of the above scenarios requires close proximity or physical access to the target environment for attack execution, the evolution from simple analog and digital sensors to sophisticated "smart" transmitters makes it possible to execute an attack on data veracity directly on the sensing device, and with that to undermine the assumed trust in field equipment. Modern smart sensors are equipped with microprocessors and IP-based communication capabilities making them attractive targets for remote and supply chain attacks. We described an approach to violating data veracity directly on the smart sensor through its compromise in [125]. Specifically, we proposed to forge sensor signal directly on the microprocessor of the transmitter, before sensor data are fed to the network protocol stack. The step-wise approach to such sensor signal spoofing with custom *Runs* and *Triangle Approximation* algorithms is illustrated in Figure A.7.



(a) Training segment of target signal

(b) Synthesized sensor noise profile

(c) Approximation of signal with line segments

(d) Spoofed sensor signal with added noise

Figure A.7: Implementation steps of the sensor signal spoofing

Ensuring Data Veracity

The definition of veracity does not differentiate between assertions that are false on purpose and assertions that are false by accident. This distinction does not matter for the application; the application would react to a false assertion in the same way in both cases. The distinction matters in the design of countermeasures. For instance, redundancy can be used for accidental violations but not for malicious violations. In this regard, Borselius [23] commented that veracity cannot be effectively enforced in mobile agent systems as the required redundancy for such a system is likely to make the system useless.

For security purposes, data veracity can be achieved in a few ways. First, if the environment warrants a sufficient degree of physical protection tamper-resistant "trustworthy" sensors can be deployed, i.e. reliance on physical security. Second, hardware and/or software-based validation of the input signals may help to reject malicious signals or filter out harmful signal components (e.g., high frequencies). Third, if the environment and input signals cannot be easily controlled, countermeasures can take the form of plausibility and consistency checks on the received sensor inputs. Redundancy and consistency checks such as majority voting have been used for detecting accidental sensor failures. Since the defenses against

intentional attacks cannot be built on the basis of statistical independence, countermeasures may take the form of plausibility and/or consistency checks. In this case, models of the physical space under observation are used to judge to which extent individual sensor readings are consistent with the overall state of the system derived from all the readings. It is possible to further model the relationship between different aspects of the physical process (e.g., temperature and pressure) in order to detect spurious sensor readings or flag implausible readings as suspicious. Changes in the plant configuration are not required to implement such countermeasures which makes them practical.

As mentioned in Section 4.6, the attacker may require to probe the target process to determine control loop responses. Simultaneously, the attacker may apply attack hiding/concealment techniques such as sensor signal spoofing to avoid detection as shown in Figure A.7. In one of our works, we implemented the detection of such spoofed sensor signals using correlation entropy in a cluster of related signals [125, 118].

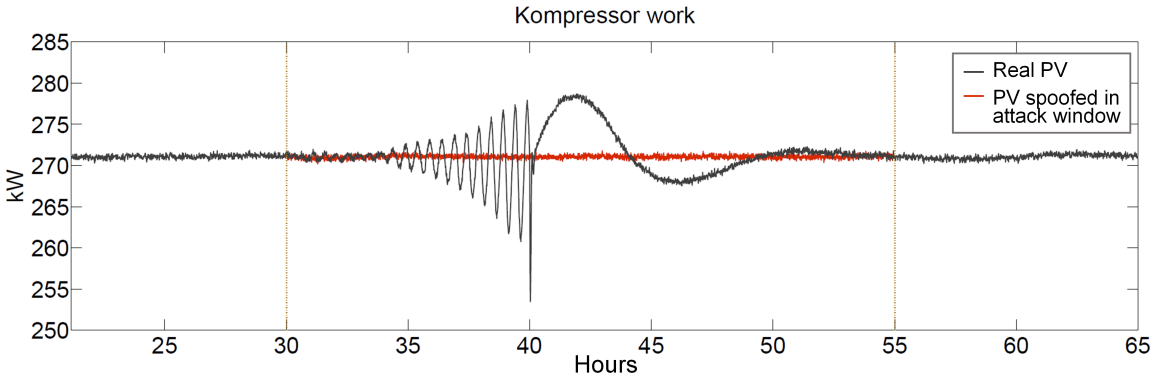


Figure A.8: Spoofed and real signal, steady state

A.3 Conclusions

The famous phrase *Don't trust your inputs* from "Writing Secure Code" book by Howard & LeBlanc is relevant for cyber-physical systems as much as for software applications. Data quality plays a paramount role in the correct, efficient and safe functionality of a cyber-physical system. On one hand, it is critical to correctly configure all *data processing points* along a data flow to preserve *data utility* and the ability to extract correct informational insights from the process data. On the other hand, ensuring the correct or truthful acquisition of process data or *data veracity* is the most fundamental security requirement for process data. In contrast to data utility, veracity security requirement is not typically studied in information security as data content and context are not considered in the threat model and as a result, security approaches to guarantee data veracity are not part of the standard IT security toolkit.

Further research is required into approaches for violating both data veracity utility, and their impact on physical applications. Better awareness of relevant threats and security requirements is essential to formulating pertinent security policies and guiding the implementation of effective security controls in an organization [43]. Additionally, at the moment there

are no open-source or commercial tools to assist with the discovery and visualization of the attack surface for both data veracity and utility. The development of approaches to capturing the inventory of data flows and detailing data process points is another research question to address. We call this concept the *Data Bill of Materials* (DBOM), echoing the concept of the Software Bill of Materials (SBOM) in software applications [212]. Application-specific DBOMs could be captured already at the application design phase as shown in Figure A.9. If later made available to application integrators, DBOM-related diagrams could facilitate a better decision-making process when selecting appropriate security controls.

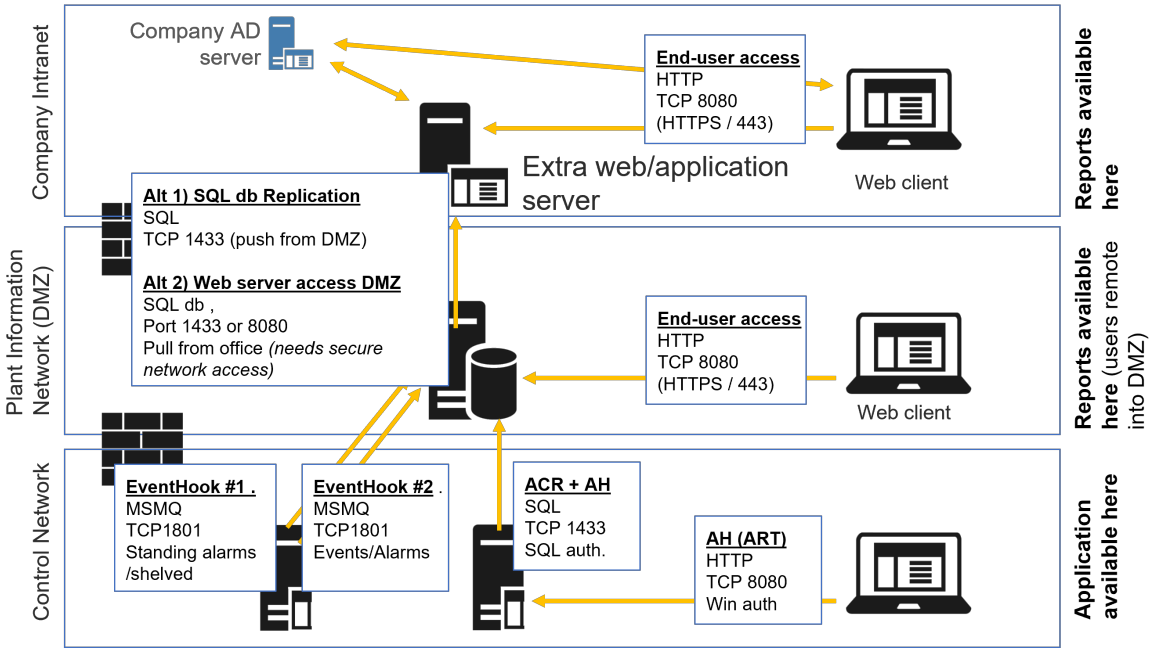


Figure A.9: Visualization of data flows during application design (the source is intentionally withhold)

B. Vinyl Acetate Plant: Listing of Variables

When implementing the VAC control model in the Simulink we applied the same naming to manipulated variables and process measurements as used in the TE process for compatibility, XMV and XMEAS correspondingly. The numbering of both XMVs and XMEAs follows the numbering of variables in Appendices 1-3 in [38].

B.1 Manipulated Variables

The below table provides a list of manipulated variables (valves) together with the corresponding controlled variables and PID controller (ctrl.) types. The table combines information from Appendix 1 and Appendix 2 in [38]. The numbering of XMVs corresponds to the numbering of control loops in Figure 3.1. Some valve positions are fixed at specific values to generate process behavior as defined in [146].

XMV	Description	Controlled Variable	Manipulated Variable	Ctrl.	Unit
XMV (1)	Fresh O2 Feed	%O2 in the Reactor Inlet	O2 fresh feed sp	PI	$Kmol/min$
XMV (2)	Fresh C2H4 Feed	Gas Recycle Stream Pressure	C2H4 fresh feed valve	PI	$Kmol/min$
XMV (3)	Fresh HAc Feed	HAc Tank Level	HAc fresh feed valve	P	$Kmol/min$
XMV (4)	Vaporizer Steam Duty	Vaporizer Level	Vaporizer Heater Valve	PI	$Kcal/min$
XMV (5)	Vaporizer Vapor Exit	Vaporizer Pressure	Vaporizer Vapor Exit Valve	PI	$Kmol/min$
XMV (6)	Vaporizer Heater Duty	Heater Exit Temperature	Reactor Preheater Valve	PI	$Kmol/min$
XMV (7)	Reactor Shell Temp.	Reactor Exit Temperature	Steam Drum Pressure sp	PI	$^{\circ}C$
XMV (8)	Separator Liquid Exit	Separator Level	Separator Liquid Exit Valve	P	$Kmol/min$
XMV (9)	Separator Jacket Temp.	Separator Temperature	Separator Coolant Valve	P	$^{\circ}C$
XMV (10)	Separator Vapor Exit	Separator Vapor Flowrate	Separator Vapor Exit Valve	Fixed	$Kmol/min$
XMV (11)	Compressor Heater Duty	Compressor Exit Temp.	Compressor Heater Valve	PI	$Kcal/min$
XMV (12)	Absorber Liquid Exit	Absorber Level	Absorber Liquid Exit Valve	PI	$Kmol/min$
XMV (13)	Absorber Circulation	Absorber Scrub Flowrate	HAc Tank Exit Valve 2	Fixed	$Kcal/min$
XMV (14)	Circulation Cooler Duty	Circulation Stream Temp.	Absorber Scrub Heater Valve	PI	$Kmol/min$
XMV (15)	Absorber Scrub Flow	Absorber Circulation Flowrate	Absorber Circulation Valve	Fixed	$Kmol/min$
XMV (16)	Scrub Cooler Duty	Scrub Stream Temperature	Circulation Cooler Valve	PI	$Kcal/min$
XMV (17)	CO2 Removal Inlet	%CO2 in the Gas Recycle	CO2 Purge Flowrate sp	P	$Kmol/min$
XMV (18)	Purge	%C2H6 in the Gas Recycle	Purge Flowrate sp	P	$Kmol/min$

XMV (19)	FEHE Bypass Ratio	FEHE Hot Exit Temp.	Bypass Valve	PI	$[0-1]$
XMV (20)	Column Reflux	%H2O in the Column Bottom	Column Reflux Flowrate sp	PI	$Kmol/min$
XMV (21)	Column Reboiler Duty	5th tray Temperature	Reboiler Steam Valve	PI	$Kcal/min$
XMV (22)	Column Condenser Duty	Decanter Temperature	Column Condenser Duty	PI	$Kcal/min$
XMV (23)	Column Organic Exit	Decanter Organic Level	Organic Product Flowrate	P	$Kmol/min$
XMV (24)	Column Aqueous Exit	Decanter Aqueous Level	Aqueous Product Flowrate	P	$Kmol/min$
XMV (25)	Column Bottom Exit	Column Bottom Level	Column Bottom Flowrate	P	$Kmol/min$
XMV (26)	Vaporizer Liquid Inlet	Liquid Recycle Flowrate	HAc Tank Exit Valve 1	Fixed	$Kmol/min$

Table B.1: Categorization of control loops based on damage potential

B.2 Process Measurements

The bellow table provides a listing of the process measurements (sensors) with their correspondent measuring units as given in Appendix 3 in [38].

XMEA	Description	Unit
XMEA (1)	Vaporizer Pressure	$Psia$
XMEA (2)	Vaporizer Level	
XMEA (3)	Vaporizer Temperature	°C
XMEA (4)	Heater Exit Temperature	°C
XMEA (5)	Reactor Exit Temperature	°C
XMEA (6)	Reactor Exit Flowrate	$Kmol/min$
XMEA (7)	FEHE Cold Exit Temperature	°C
XMEA (8)	FEHE Hot Exit Temperature	°C
XMEA (9)	Separator Level	
XMEA (10)	Separator Temperature	S°C
XMEA (11)	Compressor Exit Temperature	°C
XMEA (12)	Absorber Pressure	$Psia$
XMEA (13)	Absorber Level	
XMEA (14)	Circulation Cooler Exit Temperature	°C
XMEA (15)	Scrub Cooler Exit Temperature	°C
XMEA (16)	Gas Recycle Flowrate	$Kmol/min$
XMEA (17)	Organic Product Flowrate	$Kmol/min$
XMEA (18)	Decanter Level (Organic)	
XMEA (19)	Decanter Level (Aqueous)	
XMEA (20)	Decanter Temperature	°C
XMEA (21)	Column Bottom Level	
XMEA (22)	5th Tray Temperature	°C
XMEA (23)	HAc Tank Level	
XMEA (24)	Organic Product Composition (VAc)	$\%kmol$
XMEA (25)	Organic Product Composition (H2O)	$\%kmol$
XMEA (26)	Organic Product Composition (HAc)	$\%kmol$
XMEA (27)	Column Bottom Composition (VAc)	$\%kmol$
XMEA (28)	Column Bottom Composition (H2O)	$\%kmol$
XMEA (29)	Column Bottom Composition (HAc)	$\%kmol$
XMEA (30)	Gas Recycle Composition (O2)	$\%kmol$
XMEA (31)	Gas Recycle Composition (CO2)	$\%kmol$
XMEA (32)	Gas Recycle Composition (C2H4)	$\%kmol$
XMEA (33)	Gas Recycle Composition (C2H6)	$\%kmol$
XMEA (34)	Gas Recycle Composition (VAc)	$\%kmol$
XMEA (35)	Gas Recycle Composition (H2O)	$\%kmol$
XMEA (36)	Gas Recycle Composition (HAc)	$\%kmol$
XMEA (37)	Reactor Feed Composition (O2)	$\%kmol$

XMEA (38)	Reactor Feed Composition (CO2)	$\%kmol$
XMEA (39)	Reactor Feed Composition (C2H4)	$\%kmol$
XMEA (40)	Reactor Feed Composition (C2H6)	$\%kmol$
XMEA (41)	Reactor Feed Composition (VAc)	$\%kmol$
XMEA (42)	Reactor Feed Composition (H2O)	$\%kmol$
XMEA (43)	Reactor Feed Composition (HAc)	$\%kmol$

C. Vinyl Acetate Plant: Model Enhancement

This section provides details on how the initial model code base of the model was enhanced with a Simulink control model, Graphical User Interface (GUI), and simulation results visualization to streamline model usage for cyber security research.

The original authors of the Vinyl Acetate plant built a rigorous nonlinear dynamic model of the process to verify the feasibility of simulating the plant. Details on the assumptions and details of the modeling are described in Chapter 5 in [147]. Initially, the simulation model was implemented in TMODS, DuPont's in-house dynamic simulation environment, which was not available for public use. In several academic works on vinyl acetate, process models were implemented in specialized commercial simulation tools such as HYSYS [52], Visual Modeler [197], Aspen Plus [197] and others. To make the process model available for a wider range of users McAvoy et al. have developed a simulation model of VAC for Matlab [39] which we adopted for our experimental framework. Both the steady state and dynamic behavior of the Matlab model were designed to be close to the TMODS model.

Originally, process equations had been modeled in Matlab and then translated into C-routines. The C codes were compiled into "MEX functions" and could be called within the Matlab environment. A separate m-file was responsible for the control of the VAC process (scheduling of the C-routines) with a developed multiloop Single-Input, Single-Output (SISO) controller architecture. Additional four Matlab routines were developed for plotting the results of the simulations. No simulation data were output to the workspace for further analysis. Also, the initial model did not provide any interface to the code of the VAC process mode. Therefore, any manipulations of the model inputs (e.g., setpoint or controller update) had to be carried out directly in the C code, requiring its re-compilation.

Considering the number of variables in the complex VAC process and the inconvenience of manipulating the process within the source code, we have developed a *Simulink model* of the process. Initially, we developed a user interface and attack codes without building a Simulink model. Several months of experimentation have revealed the inconveniences and limitations of this approach. **Simulink** [156] provides an interactive, graphical environment for modeling, simulating, and analyzing dynamic systems at any level of detail. Simulink

models are compact enough to be understood with moderate effort. Additionally, the interactive nature of Simulink allows easy experimenting by changing the model and its parameters, and immediately observing what happens. Thus, modeling attacks on a selected process component can be done easily by adding a function block with several lines of code. Such functionality suits well the *"what if"* nature of cyber-physical exploitation. Specifically, we implemented data integrity (Figure C.1) and Denial of Service (DoS) attacks on the process variables (sensor signals) and the manipulated variables (controller output to actuators) in the form of MITM attacks. Since process models include controllers, the simulator allows the implementation of attacks on controllers (e.g., false control logic upload) directly in the Simulink model as well.

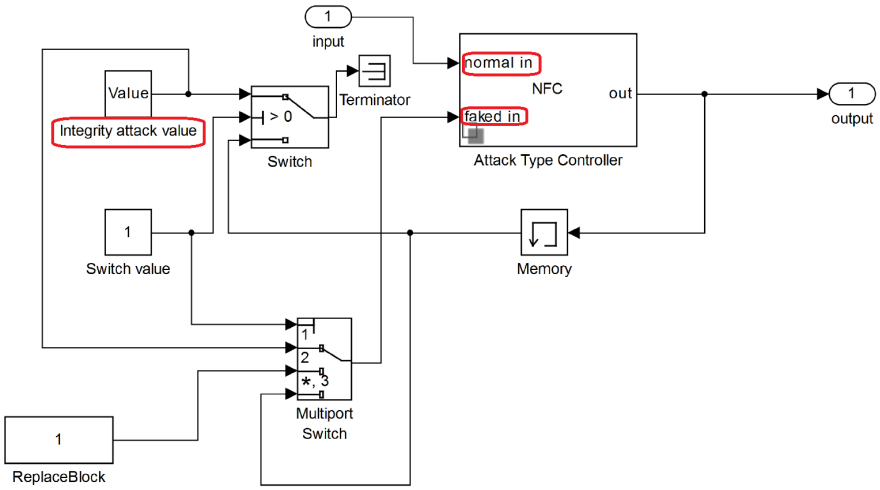


Figure C.1: Functional implementation of integrity attack on sensor signal in Simulink

We designed the Simulink Model for the VAC process similar to the one used in the Tennessee Eastman (TE) process, one of the most popular and widely-used public process model [63, 182] to preserve user experience. We also named manipulated variables as XMVs and controlled variables as XMEAS to maintain consistent notation between the TE and VAC models. We instrumented the VAC Simulink model with a user interface for a convenient update of simulation parameters and setting up desired attack parameters on individual elements of the control infrastructure as shown in Figure C.2. For instance, for an integrity attack on sensor signals, the user can choose the attack value, attack time and duration (predefined or random) as well as the frequency of the attacks (single or periodic). We also implemented an option of feeding a predefined sensor values stream from a file. The user may run several attacks in parallel or sequentially (chain several attacks). We also implemented an output of the simulated data to the workspace for storage and analysis and enabled their automatic visualization (Figure C.3). Besides, we fixed several implementation mistakes in the process code and made several improvements to its control model to make the process more stable.

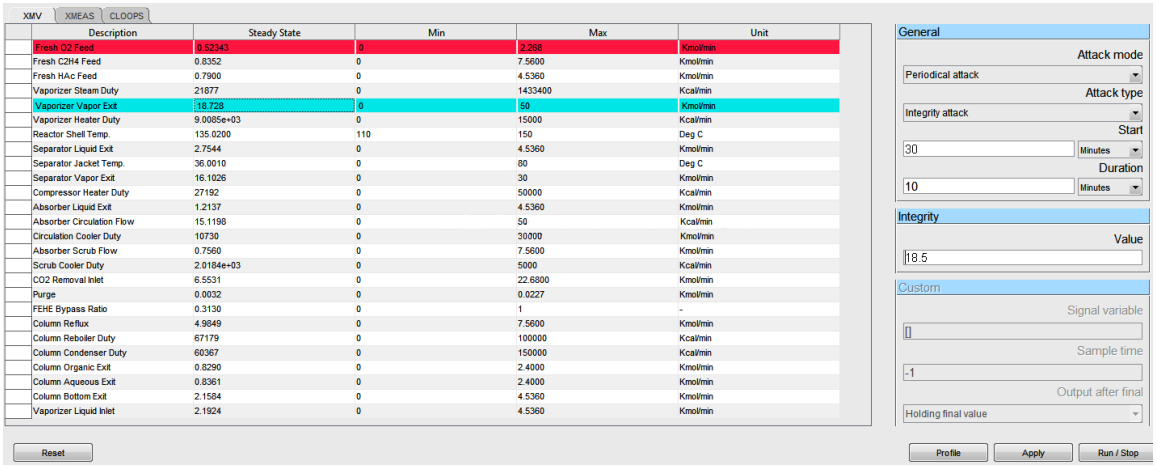


Figure C.2: User interface for setting attack parameters (implemented in Matlab environment)

C.1 Damn Vulnerable Chemical Process Framework

One of the challenges of cyber-physical security research is the lack of realistic large-scale testbeds for studying complex cyber attacks and their effects on physical processes. Well-known realistic testbeds are Secure Water Treatment (SWaT) [154] and Electric Power and Intelligent Control (EPIC) [186], hosted by the iTRUST research center at the Singapore University of Design and Technology (SUTD). These testbeds were conceived, designed and built collaboratively by SUTD faculty members, international consultants, and engineers from public utilities in Singapore. Each testbed mimics a smaller in scale realistic industrial process and is built from popular models and brands of industrial equipment. The high fidelity of these experimental environments is certainly very beneficial to multidisciplinary researchers when conducting cyber-physical security experimentation. However, building and

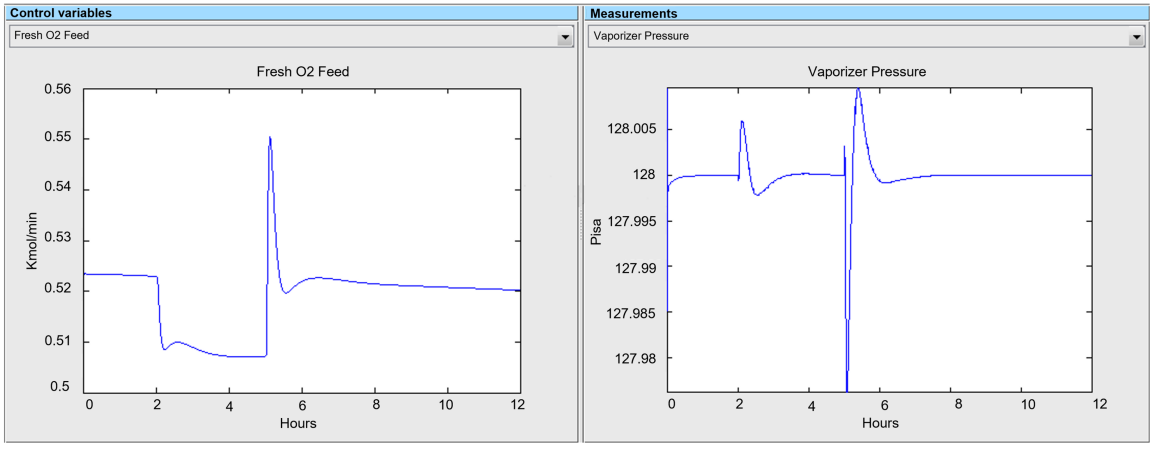


Figure C.3: Visualization of simulation results

maintaining such testbeds requires specialized personnel and significant financial investments. Thus, at the time of writing, external users were charged an hourly rate of $900 ($400 for internal students) for testbed usage [180]. Provided that security research requires extensive experimentation to evaluate exploitation scenarios' reliability and discover edge-cases, research on a realistic testbed can prove economically prohibitive. Going further, additional repair expenses are inevitable if the testbed is damaged as a result of a successful experiment aimed at physical disruption. Additionally, unsafe experiments are not permitted by safety regulations. This means that testing on the realistic testbeds is limited to a set of well-calculated experimental use cases, which do not jeopardize the safety and physical integrity of the equipment. Clearly, this restriction imposes further significant limitations when conducting research on realistic testbeds. Out of these considerations, a more affordable and safe way of conducting cyber-physical security experimentation is using simulated process models.

> In the process engineering domain, developing models of physical processes is an integral part of a cost-effective R&D process.

Process simulation is a model-based representation of chemical and other processes in software, intending to optimize designs and operating procedures of plants and individual units. Also, models allow for close collaboration between chemical, control and process engineers. While the chemical engineers model complex thermodynamic interactions within the plant, control and process engineers devise suitable control algorithms and plantwide control schemes and optimize equipment design. The ability of the control system to support plant chemistry is then tested in closed-loop simulations.

Developing process models is time-consuming and costly, therefore the majority of them are kept proprietary. Additionally, running complex process simulations requires highly-specialized software, which is also often kept proprietary or available with expensive licenses. For these reasons, the number of open-source models of realistic industrial processes is limited (e.g., [63, 16, 38, 8]). However, such public models allow researchers to focus on industry-relevant problems and to allow the comparison of research results.

> To support open-source research, we combined two public models, Tennessee Eastman and Vinyl Acetate into an open-source experimental framework which we called **Damn Vulnerable Chemical Process (DVCP)** [122, 123].

The naming of the framework follows the naming of two well-known "damn vulnerable" frameworks from the IT security domain, namely Damn Vulnerable Linux (DVL) [195] and Damn Vulnerable Web Application (DVWA) [210]. These frameworks are intentionally vulnerable Linux distribution and web applications, created to study a variety of security topics and legally practice exploitation techniques in a controlled environment.

Since the control strategies for both plant models were designed without considering any potential malicious interaction with the plant, it makes them a suitable test case for researching what it takes to convert a cyber attack into a successful cyber-physical attack. The frameworks can be utilized for Red-Team (offensive) type of research questions such as the discovery of novel attack vectors, designing individual attack instances and complex coordinated attacks. Consequently the "defenders" can study the resilience of processes to cyber attacks, develop and evaluate risk assessment methods, robust control algorithms,

attack detecting techniques, etc. The framework can be used as a standalone experimental environment or as a physical layer in a distributed industrial control system testbed, e.g. as implemented National Institute for Standard and Technology (NIST) [175, 30]. The main objective of the testbed design was "to emulate real-world industrial systems as closely as possible without replicating an entire plant or assembly system". Figure C.4 shows a high-level architecture of the testbed with the Tennessee Eastman process. The testbed was conceived to serve as a guide on how to implement security safeguards effectively without negatively affecting process performance and to measure the performance of industrial control systems while undergoing a cyber-attack, with resiliency being a central focus of the experimentation.

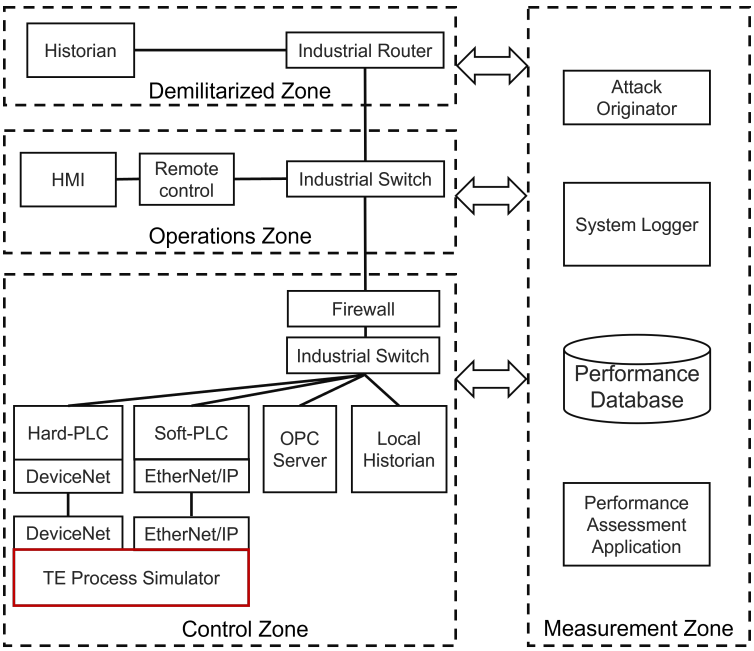


Figure C.4: Distributed industrial control system testbed with TE process [175]

C.2 Conclusions

Software-based testbeds are frequently seen as less credible for scientific research. This opinion is predominately caused by the perceived reduced complexity of the simulated environments and assumed limited implementation of real-world behaviors. While this assessment might be valid for certain research areas or specific research questions, in the process- and control engineering fields, simulation-based testing environments have become the primary type of experimental testbeds. With the help of specialized software, engineers can model and simulate dynamic behaviors of physical processes and equipment with the required degree of fidelity. This allows for conducting cost-effective and safe "trial and error" type of experimental activities. Working with the combination of control equipment/protocols and the physical process can be arduous. Once a cyber vulnerability is understood it is easier to *abstract* its potential impact and to apply it directly to the

process model. Owing to an elaborate emulation of the physical and control layers of a cyber-physical system, the framework allows to abstract from the research factors specific to the cyber layer and concentrate on the aspects which are intrinsic to control- and process engineering.

The past few years were marked by the active development of various types of experimental, testing and visualization platforms to support the growing needs of the ICS security industry in research, education and testing. The emerging trends include $3D$ visualization and gamification, cyber ranges and digital twins. It is expected that these technologies will further stimulate progress in the cyber-physical security discipline. In our research, we use a realistic model/simulation of the Vinyl Acetate Monomer plant with its underlining control infrastructure. Since the plant design and control strategy was designed "without security in mind", it makes it suitable for researching what it takes to convert a cyber attack into a successful cyber-physical attack.