Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P6265-cd

Possible measures for the safety of process plants in case of cyber attacks

Gabriele Baldissone

DISAT, Politecnico di Torino, Italy. E-mail: gabriele.baldissone@polito.it

Davide Fissore

DISAT, Politecnico di Torino, Italy E-mail: davide.fissore@polito.it

Salvina Murè

DISAT, Politecnico di Torino, Italy E-mail: salvina.mure@polito.it

Micaela Demichela

DISAT, Politecnico di Torino, Italy E-mail: micaela.demichela@polito.it

With the digitalization of industrial plants, cybersecurity is becoming an increasingly relevant problem. The most commonly used method for managing this problem is to use computer protection systems (e.g. firewalls). The purpose of these programs is to make impossible or at least difficult to intrude and alter computer systems.
In the process industry, intrusion into information systems can lead to malicious alteration of plant parameters, leading to significant risks to the safety of people and property.
For these reasons, this paper presents two additional barriers that can be adopted to prevent intrusions into computer systems from causing accidents.
The first possible barrier that can be proposed is to use a digital twin. To compare the measured variables with the data obtained from the digital twin in order to identify the deviation between the two values indicating the presence of a problem.
The second potential barrier can be in the ability of control room operators to recognize deviations in process variables following the intrusion, as well as to take corrective measures, including manual ones.

Keywords: Cyber security, Cyber Physics Systems, Swiss Cheese model.

1. Introduction

Nowadays many industrial plants are connected to the network in order to elaborate the process data in real time, improving efficiency and control of the plant (Xianghui, et al. 2013). But at the same time, this connection open new opportunity of cyber intrusion to the plant, creating new threats (Ani, He e Tiwari 2016).
In the past, cyber-attacks have caused the injurie of people (Leyden 2008) or serious production problems (Langner 2011). These incidents demonstrate that in certain industrial sectors cyber-attacks can cause damage in the real world.
For these reasons, over the years, much work has been done to improve the cyber security of the industrial plants, acting in various fields such as: security architectures (Falco, et al. 2002, Jones e Horowitz 2012), policies (Bertolotti, et al. 2013), vulnerability scanning (Stamp, et al. 2003, Coffey, et al. 2018), authentication (Chakravarthy, Hauser e Bakken 2010, Genge, Haller e Duka 2018), access control (Yalcinkaya, Maffei e Onori 2017). data encryption (Xu, et al. 2017) and intrusion detection (Dolgikh, et al. 2011). All these systems strengthen the cyber part of industrial plants, but once overcome, interventions can still make to avoid damage in the physical world.
In this way, between the cyber attack and the physical consequences, different layers of protection can be interposed, some IT, others process-related, human or passive systems.
This paper describes how the Swiss Cheese model can be adapted to the cyber security of Cyber Physics Systems, focusing especially on process-related and human measures.

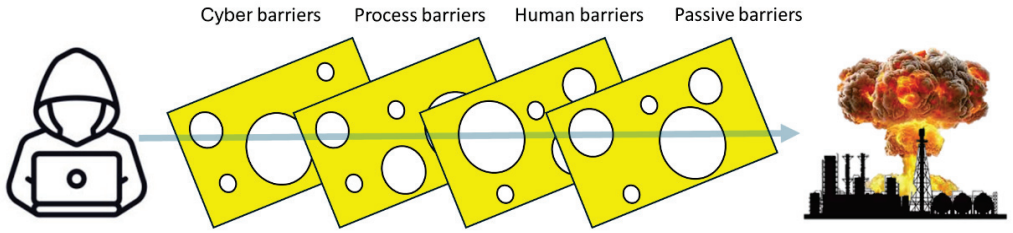

Fig 1 Swiss Cheese model in the cyber security

2. Swiss Cheese model in cyber security

In the last century, to describe incidents, Reason (Reason 2000) introduced the Swiss Cheese model. In this model is assumed that between the causes and the consequences there are some layers, but these layers have holes. The incident occurs when the holes of the various layers are aligned. In this paper are proposed to apply this model to cyber security. In which, as can be seen in Fig 1, between the cyber-attack and the real incident there are different layers of protection. Each layer can be divided into sub layers (for example Fig 2). In each layer there are hole and the consequences of the cyber-attack occur when these are aligned.
In this scheme the following layers can be found:

- Cyber barriers: in this layer can find all those IT measures that try to make impossible or difficult for malicious people to access cyber data and systems. This layer only concerns the IT part.
- Process barriers: in this type of layer are insert those measures that can only be developed with in-depth knowledge of the process. These layers help to understand if a cyber attack has modified the state of the process and that it can lead to consequences in the real world. This aspect will also be discussed in a later paragraph.
- Human barriers: Even highly automated processes require a certain degree of supervision and control by human operators. In this case, the presence of the operator makes possible for him to notice the consequences of the cyber attack on the system, intervening manually, regaining control of the system and placing it in safe conditions. This aspect will also be discussed in a later paragraph.
- Passive barriers: these include mechanical safety systems that can prevent major accidents, but the activation of these systems still results in consequences for the system. These systems include, for example, PSVs and Rupture Discs. By their nature, these systems cannot be compromised by cyber attacks.

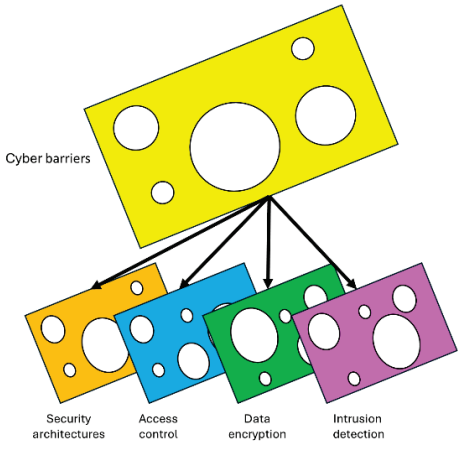

Fig 2 Cyber barriers measures

2.1. *Process barriers*

Process barriers are measures that, thanks to the knowledge of the process, can limit the consequences or allow the identification of a possible cyber-attack. In particular, we try to develop methodologies for identifying cyber-attacks which interfere with the plant and which can potentially create an accident in the real world. This type of barriers try to identify the inconsistencies between the measures taken in the field and the process parameters modified by a cyber-attack.
As part of the SERICS project, the Safer group of the Politecnico di Torino aims to develop and analyze these types of barriers. This type of barrier can be classified into different categories (Fig 3):

- A possible solution may consist in analyzing some key variables of the plant operation and in particular evaluating their congruence. In this way, this approach involves selecting a series of key variables and monitoring them over time and evaluating whether these variables are consistent.
- Another possible solution consists in analyzing the trend of the key variables and comparing the trends of the different variables. This approach is already used for fault identification (Maurya, Rengaswamy e Venkatasubramanian 2007). In this case, rather than checking the values of the key variables, the trends of the variables are compared. This comparison takes into account the congruence of the trends (that the sign of the slope of the trend is correct), and that the magnitude of the slope is congruent. In this way, it is possible to observe whether the cyber-attack interferes with the normal behavior of the plant.
- Another possible methodology consist in the use of the Digital Twin, in order to compare the Digital Twin forecasts with the variables measured in the field. In order to identify any significant deviations between the two data caused by a cyber attack. Logically this system is a more complex methodology but it should be able to identify the part of the plant involved in the cyber attack.
- Use of data reconciliation, deepening of the approach proposed by Reibelt et al. (2021) . This methodology is based on data reconciliation.

Within the SERICS project we aim to deepen these methodologies in the case of cyber attacks on industrial plants. Also identifying their operating limits by investigating the ability of these methodologies to distinguish between failures and cyber attacks.

2.1.1. *Developing of the Process barriers*

The development of an application system based on the Process Barrier concept requires an in-depth understanding of the industrial facility under analysis. Depending on the complexity of the facility and the potential risks to people in the event of a cyber attack, a more or less sophisticated approach can be adopted.

For the most complex scenarios, a Digital Twin can be used to compare its simulated results with real-world measured data, detecting any discrepancies that may indicate a cyber attack.

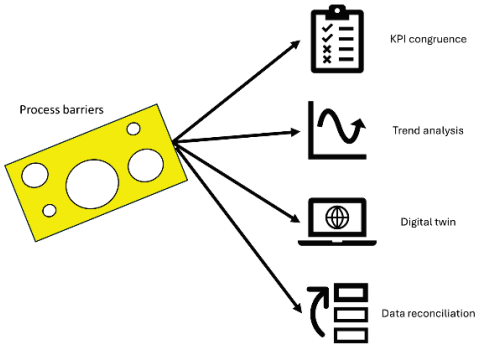

Fig 3 Process barriers measures

- To be effective, the Digital Twin must remain independent from the facility's operational systems to minimize the risk of being compromised by a cyber attack. This requires:
- Selecting a minimal but essential set of data necessary for the Digital Twin to model the system's state.
- Using an independent data acquisition system and sensors, reducing the likelihood of cyber threats compromising the measurements.
- If the Digital Twin also integrates data from the facility's control system, these values should be considered less reliable due to the potential risk of manipulation.

Once the Digital Twin processes the data, its results are compared with real-time measurements from the facility. The system analyzes for:

- Significant deviations between expected and actual values.
- Irregular trends in the data over time.

If discrepancies are detected, the control room operator receives an alert and must investigate the cause of the inconsistency, collaborating with field operators if necessary. This process allows for the early detection of cyber attacks that may have altered key operational parameters.

As part of the SERICS project, we aim to apply this approach to a case study, testing its ability to detect

various types of cyber attacks and evaluating its effectiveness in enhancing industrial cybersecurity.

2.2. Human barriers

Even the most automated systems require human supervision, in case of cyber-attacks the operator can notice the problems in the IT system. If the operator notices the cyber-attack, the operators can intervene both on the IT systems and by operating manually and in the worst case by operating a manual shutdown.
For these reasons within the SERICS project we aim to investigate how the operator of the control room can notice the cyber-attack and how he can intervene.
In this area we propose to work in the following areas:

- The first field of research to investigate concerns whether the operator is able to notice the alterations caused by a cyber-attack, investigating the situational awareness of the operator in these cases. Focusing on the ability of the operator to identify early process deviations following cyber-attacks, identifying any discrepancies between the various variables. This activity is conducted with the approach proposed by Winifred Amazu et al (2024), carrying out tests on the operators.
- Another area to investigate is the study of the procedures to adopt in case the operator becomes aware of a cyber-attack. The study of these procedures can be important because they are complex and involve multiple figures and the exchange of information between them. In fact, in the event of a cyber-attack, the level of compromise of the system cannot be known, so readings of variables in the field and manual actions are necessary, making the management of the system complex.
- The last area to analyze is the study of the emergency procedures to adopt in the event of a cyber-attack, with particular attention to any manual shutdown procedures in safety and without damage to the system, that procedures are very complex and required high coordination between different operators.

2.2.1.Testinf of the human barriers

As part of the SERICS project, we aim to assess the effectiveness of the human barrier against cyber attacks and develop training strategies to help operators detect and respond to threats at an early stage.
To achieve this, we plan to adopt an approach similar to the one presented by Amazu et al (2024), which involves testing control room operators using a simulator. This simulator replicates the operation of a control room and consists of two main components:

- The user interfaces that facilitate interaction between the control room and the operator (Fig 4).
- A dynamic system model that simulates the behavior of the facility in response to the operator's commands.

Fig 4 Example of the control test system

The goal is to study how effectively operators can recognize and respond to cyber attacks. To this end, we will simulate different attack scenarios, including:

- Sensor Data Manipulation: A cyber attack alters sensor readings, leading the automatic control system to create dangerous process deviations. The operator must detect the deviation and manually restore normal conditions.
- System Parameter Tampering: A cyber attack modifies control system parameters and data, causing system instability. The operator must regain control by identifying the issue and bypassing it manually.
- Combined Attack on Process and Manual Controls: In addition to creating a process deviation, this attack disables part of the manual control system. The operator must recognize the issue and manage the situation

by adjusting unaffected parameters until the maintenance team restores full functionality.

The three test scenarios will vary in difficulty, allowing us to evaluate operators' preparedness at different levels.

Based on the results of these tests, we aim to design a comprehensive training program to equip operators with the necessary skills to detect and respond to cyber attacks. This program will incorporate hands-on training using the simulator to provide realistic experience in managing cyber threats.

A second area of analysis focuses on more severe cases where a cyber attack results in partial or complete loss of visibility and control from the control room. In such situations, field operators play a crucial role in executing a safe manual shutdown procedure.

To evaluate the effectiveness of this procedure, we plan to conduct a dynamic simulation of the system, including interactions between operators. This analysis will help assess the efficiency of manual shutdown operations and identify potential areas for improvement.

3. Conclusions

In recent years, industrial plants will be increasingly connected to the network, increasing their efficiency. But at the same time opening up new cyber threats.

For this reason, we propose to apply the Swiss Cheese model to cyber security by identifying the following families of barriers: Cyber barriers, Process barriers, Human barriers and Passive barriers.

In particular, within the SERICS project, we aim to delve into the process barriers, examining the applicable methodologies, trying to identify their effectiveness and limits.

Also within the SERICS project, we aim to investigate the ability of control room operators to understand early on that they are under a cyber attack and the development of effective procedures to adopt.

Acknowledgement

This work was partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU

References

Amazu C.W., J. Mietkiewicz, A.N: Abbas, H. Briwa, G. Baldissone, M. Demichela, D. Fissore, A.L. Madsen, M.C. Leva (2024) Experiment data: Human-in-the-loop decision support in process control rooms. *Data in Brief, 53,* 110170.

Ani, U. D., H. He and A. Tiwari (2016) Review of cybersecurity issues in industrial critical infrastructure: manufacturing in perspective. *Journal of Cyber Security Technology, 1.* 32-74.

Bertolotti, I. C., L. Durante , T. Hu and A. Valenzano (2013). A Model for the Analysis of Security Policies in Industrial Networks. *1st International Symposium for ICS & SCADA Cyber Security Research 2013 (ICS-CSR 2013),* 66-77.

Chakravarthy, R., C. H. and D. E. Bakken (2010):. Long-lived authentication protocols for process control systems. *International Journal of Critical Infrastructure Protection, 3(3-4)*, 174-181.

Coffey, K., R., Smith, L. Maglaras, H. Janicke (2018) Vulnerability Analysis of Network Scanning on SCADA *Systems. Security and Communication Networks, 1.*

Dolgikh, A., T. Nykodym, V. Skormin, J. Antonakos, and M. Baimukhamedov (2011). Colored Petri nets as the enabling technology in intrusion detection systems. *2011 - MILCOM 2011 Military Communications Conference,* 1297-1301.

Falco, J., F. M. Proctor, K. A. Stouffer and A. J. Wavering (2002). *IT security for industrial control systems.* U.S. DEPARTMENT OF COMMERCE.

Genge, B., P. Haller, A.V. Duka, (2019) Engineering security-aware control applications for data authentication in smart industrial cyber–physical systems. *Future Generation Computer Systems, 91,* 206-222.

Jones, Rick A., B. Horowitz (2012) A System-Aware Cyber Security architecture. *Systems Engineering, 12(2),* 225-240.

Langner, R (2011). Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security and Privacy, 9(3),* 49-51.

Leyden, J. (2008). Polish teen derails tram after hacking train network. *The Register, 11*

Maurya, M. R,, R. Rengaswamy and V. Venkatasubramanian (2007). Fault diagnosis using dynamic trend analysis: A review and recent developments. *Engineering Applications of Artificial Intelligence, 20,* 133-146.

Reason, J. (2000). Human error: models and management. *British medical journal, 320,* 768-770.

Reibelt, K., H. B. Keller, V. Hagenmeyer and J. Matthes (2021).Dynamic Model Based Detection of Cyberattacks in Industrial Facilities. *PROCEEDINGS OF THE 31 European Symposium on Computer Aided Process Engineering,* 1339-1344.

Stamp, J., J. Dillinger, W. Young and J. DePoy (2003). *Common vulnerabilities in critical infrastructure control systems*. Sandia National Laboratories.

Winifred Amazu, C.; J. Mietkiewicz, A. N Abbas,. H. Briwa, A. A. Perez, G. Baldissone, M. Demichela, D. Fissore, A. L. Madsen and M. C. Leva (2024) Experiment data: Human-in-the-loop decision support in process control rooms. *Data in Brief, 53,* 110170.

Xianghui, C., C. Peng, C. Jiming and S. Youxian (2013). An online optimization approach for control and communication codesign in networked cyber-physical systems. *IEEE Transactions on Industrial Informatics, 9(1),* 439-450.

Xu, T., Y. Yang, T. Li, J. Ju and Q. Wang (2017). Review on cyber vulnerabilities of communication protocols in industrial control systems. *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2),* 1-6.

Yalcinkaya, E., A. Maffei and M. Onori (2017). Application of Attribute Based Access Control Model for Industrial Control Systems. *International Journal of Computer Network and Information Security, 9(2),* 12-21.