Resilience Against and Detection of Information Hiding in Nuclear Instrumentation and Control Systems within the Scope of NSS 17-T

Kevin Lamshöft, Mario Hildebrandt, Robert Altschaffel
Otto-von-Guericke University Magdeburg
Universitätsplatz 2, 39104 Magdeburg, Germany
firstname.lastname@ovgu.de

Oliver Keil, Ivo Hempel, Jana Dittmann
Otto-von-Guericke University Magdeburg
Universitätsplatz 2 39104 Magdeburg, Germany
firstname.lastname@ovgu.de

Tom Neubert, Claus Vielhauer
Technische Hochschule Brandenburg
Magdeburger Straße 50, 14770 Brandenburg an der Havel, Germany
firstname.lastname@th-brandenburg.de

ABSTRACT
The graded approach of the IAEA NSS 17-T, in conjunction with highly restricted and deterministic traffic in a computer network, increases the importance of Information Hiding (IH) technologies for attackers. Thus, it is necessary to provide detection mechanisms and resilience against IH in the architectural designs. We reflect hidden communication channels discovered in both common and industrial network protocols within the scope of the recently updated NSS 17-T. We discuss the potential deployment of detection techniques as well as potential attack vectors such as hidden supply-chain attacks, insider threats and conventional attacks covered by the full depth of the graded approach of NSS 17-T.

INTRODUCTION
Information hiding (IH) is an important technique usable by advanced persistent threats (APT) for staying undetected over a long duration. While in information technology (IT) the mere presence of different types of communication might be sufficient in order to conceal the compromise of the systems, the graded approach of NSS 17-T [4] in conjunction with highly restricted and deterministic traffic increases the importance of IH technologies for attackers even further. Thus, the general awareness provided by an architectural design with resiliency against Information Hiding and potential detection mechanisms are a necessity. Additionally, such design decisions likely assist the mitigation of (sensitive) digital asset (SDA) vulnerabilities. In this paper, we reflect on discovered hidden communication channels in industrial protocols such as Modbus/TCP and OPC UA, as well as commonly used supporting protocols such as Syslog and NTP within the scope of the newly updated NSS 17-T in order to support the risk informed approach against potential sabotage and unauthorized access to sensitive nuclear information. Based on the graded approach (including computer security levels and computer security zones), the impact of physical access control and decoupling mechanisms for data flows are evaluated in order to assess their suitability for preventing IH-assisted cyber-attacks.

General Structure of IT and OT in the Nuclear Domain
Operational Technology (OT) differs from Information Technology (IT) in function, structure and employed components. OT is used to supervise and control the process of energy generation within the nuclear domain. The general structure of OT is described in [1] as relying on 1) the use of sensors to gather information about the physical processes and/or the environment, 2) on computing units to generate control signals based on these sensors and/or user input, 3) and on actuators implementing these control signals and therefore altering the physical processes and/or the environment. The user interaction is handled by Human-Machine-Interfaces (HMI) with the most prominent being situated

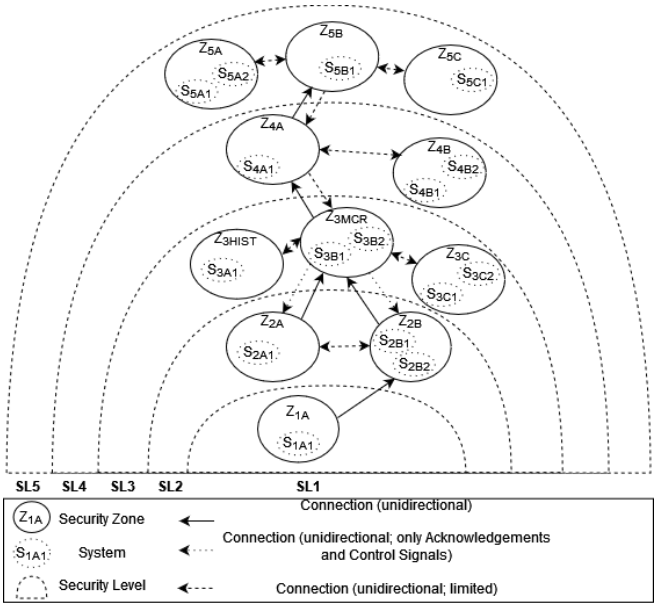

Figure 1: Communication Flow within a DCSA. Example based on IAEA NSS 17-T [4]

within the Main Control Room (MCR). Within the OT domain, the computing units are referred to as Programmable Logic Controllers (PLCs). Differences between IT and OT and their impact on security are explored in [2]. Due to the specifics of components (low memory capacity and computing power), the OT components have to rely on constant communication between sensors, computing units and actuators in order to perform the given task. Hence, this communication is of central importance. Since the OT controls a critical physical process which could be severely impacting if the actuators act in any unwanted manner, any attack on the (cyber-)security of the system carries threat to safety (as explored in [3]).

The IAEA NSS 17-T & DCSA

Since communication between the various components of OT in the nuclear domain is highly critical and requires security from manipulation (integrity), security architectures are used during the planning of information flows. One such model is the Defensive Computer Security Architecture (DCSA) proposed in the NSS 17-T (Rev. 1) [4]. This approach is based on the older approach of Graded Security as described in [5] and [6]. This model includes the separation of components based on the impact of a potential compromise of the function. The more critical the function of the component is, the higher are the security requirements. Within the specific security levels, security zones form trusted areas for internal communication. The communication flow between security levels and security zones are restricted. The more critical a function, the more restrictions are in place to limit communication toward the specific component. Fig. 1 presents an example based on [4] including 5 Security Levels and various zones. Security Level 1 (S_1) covers the physical safety systems with an unidirectional connection to other security levels (only used to send a signal that the physical safety system was triggered and associated data). Security Level 2 (S_2) covers most control functions. In this case, these are separated into two zones (Z_{2a} and Z_{2b}) with trust between the specific systems within each zone (S_{2b1} and S_{2B2}). Security Level 3 (S_3) includes supervision functions (e.g. MCR – Z_{3MCR}). Security Level 4 (S_4) and 5 (S_5) contain classical IT systems performing functions required for process analysis or internal maintenance scheduling (S_4) or business functions (S_5).

INFORMATION HIDING IN NUCLEAR INSTRUMENTATION AND CONTROL SYSTEMS – A THREAT?

Communication protocols found in I&C can generally be divided in protocols used for automation and protocols of infrastructure services. Common examples for automation protocols are open protocols like Modbus (TCP), DNP3, OPC (UA) and proprietary protocols like Siemens' S7Comm and Allen-Bradley's PCCC. Infrastructure service protocols are network protocols that viable for the upkeep and establishment of communication in Ethernet networks, for example time synchronization protocols like

NTP and PTP and address resolution protocols like DNS and ARP. However, such communication can be misused by adversaries to secretly communicate and infiltrate or exfiltrate information using methods of network steganography. The core idea here is to conceal hidden communication within already existing network traffic. Such communication channels are referred to as covert channels. In the past, several protocols found in I&C were analysed in regards to their potential as a carrier for such hidden communication. For example, in 2020 we identified 14 covert channels in Modbus/TCP that could be used information-hiding based attacks [2]. In the same year we also described a supply-chain based attack scenario using a covert channel in OPC UA, leveraging timestamps as cover [3]. This covert channel enables an adversary to arbitrarily control I/Os from compromised PLCs. Time synchronization protocols like NTP and PTP have recently been shown to be vulnerable to covert channels as well, providing adversaries a mean to infiltrate information even from outside a facility using radio and satellite transmissions [4, 5]. Another work in 2021 puts the focus on the transmission of sensor values and how these could be used to exfiltrate sensible information from high security networks to lower security levels using process historians [6].

NSS 17-T Threat Scenarios and Information Hiding

In the following we take a look at the threat scenarios provided in the appendix of the NSS 17-T and reflect on how techniques of information hiding might aggravate such threats. In the second provided scenario called “Exploitation of the transitive trust between reporting servers on the perimeter network and internal SDAs“ the communication path between SDASs of high security levels towards servers on lower security levels is exploited to gain access. This scenario can be easily extended using techniques of IH to improve its stealthiness. Options here include to use text steganography or network covert channels in the protocols used for communication with these reporting servers, for example Syslog which also had recently been analysed [7]. In Scenario IV adversaries obtain sensitive information about nuclear plant operations directly from inappropriately decommissioned equipment. Using methods of IH this scenario can be further extended as well. For example, a malware can deliberately enforce a replacement of hardware by damaging hard drives through massive read/write operations. Additionally, using methods of filesystem steganography the malware could hide sensible information within these hard drives before decommissioning. Other covert channels, like writing to unused memory spaces of mainboards can further increase robustness against decommissioning guidelines which would result in loss of (hidden) information. Finally, Threat Scenario V “Strategic Social Engineering on the Facility Security Officer” even considers covert channel without directly mentioning it. By using methods of social engineering the security officer is tricked to open email attachments containing malware. This malware then covertly exfiltrates sensible information and files from the security officer’s computer towards the adversary. This covert transmission can be achieved by using network covert channels, for example using DNS and HTTPS as carrier. In summary, these threat scenarios provided by the IAEA indicate that covert channels increase the stealthiness of already known attack paths. Therefore, it is important to achieve better resilience against such threats.

TOWARDS BETTER RESILIENCE AGAINST INFORMATION HIDING BASED THREATS

Within the constrained and deterministic network communication of nuclear I&C systems implementing the graded approach of NSS 17-T utilizing the existing communication as unsuspicious cover data for concealing malicious communication can be a key for a successful attack. A key lesson from hidden communication in IT networks is that the ability to create and send arbitrary requests is sufficient to transmit and receive arbitrary data (see e.g. [9] for an overview or [8] for an arbitrarily selected actual malware). In order to increase the resilience against information hiding based threats the possibility of communication across multiple security levels needs to be limited to transfer well-known and well-defined information instead of complex protocols. In particular protocol conversions should be used in conjunction with no or very limited metadata. Additionally, the resolution of values should be strictly limited to the required accuracy for the specific process controls. However, if it is necessary to provide a certain level of assurance regarding the communicated information or confidentiality of credentials some cryptographic methods might be required as well. Since this would increase the complexity of the protocols as well as the entropy of the data an increased possibility of information hiding is inevitable. In order to mitigate such a threat without sacrificing the general security requirements it is necessary to be able to decrypt the data at least at the boundary of a zone and/or security level in

order to allow for the utilization of detection mechanisms. In addition to that, the decrypted authentication and authorization data should follow a simple protocol - as a result undiscovered hidden communication would be limited to intra-zone communication. Furthermore, any embedding method that alters the cryptographic data is likely to be detected if the decryption on zone/level-boundaries result in garbled - i.e. non-protocol compliant - data. In order to achieve resilience, the packet should be only forwarded to the adjacent zone/security level if nothing suspicious is detected. However, if the communication is time critical, the latency of the detection process must be taken into account for the design of the nuclear I&C system.
Another countermeasure is a further increase of the deterministic behaviour of the nuclear I&C systems. In particular cycle times and packet intervals should be as constant as possible. As a result any significant deviation from the known pattern can be considered as a potential anomaly.

Detection of Information Hiding in I&C Networks
In order to detect IH in I&C networks usually machine learning (ML) driven approaches based on supervised statistical pattern recognition are used to elaborate suitable forensic detection and defense mechanisms. Especially in forensics, it is desirable to elaborate approaches that deliver an explainable and comprehensible prediction for a sample to justify and understand the decision of a detector to ensure an appropriate reaction to the detection of potentially malicious data. Thus, ML-based detection approaches for IH mostly use comprehensible handcrafted feature spaces that are extracted from I&C network data with lean classification algorithms (classifiers) to deliver explainable prediction results on test samples.
For example, in [16] a so-called One-Class-Classifier [17] is trained with known-good OPC UA network data with a handcrafted feature extractor to define a “target” class. This enables the possibility for a ML-based approach to detect outliers or anomalies based on the trained target class in the network traffic with no need of malicious data for training. However, it is also in common use to train detection approaches with two classes to detect IH in I&C networks. On the one hand, this leads to higher detection accuracies on malicious data, because obviously the machine learning approach “knows” in this case how the malicious data (i.e. the anomaly) should look like, but on the other hand malicious data has to be generated before (which could be time consuming and challenging). Another trend is to use convolutional neural networks (CNNs) for the classification of (malicious) data in I&C networks [12, 15] because they revolutionized feature based learning with the ability to extract self-learned features from data with no need of elaboration a handcrafted feature space. CNNs usually have a high level of complexity including a lot of simple processing units with a high degree of interconnectedness between the individual units. On the one hand, this leads to the ability of adaptive learning and error tolerance skills with convincing detection results but on the other hand it leads to a lack of explainability because the delivered prediction from the CNN is less comprehensible.

CONCLUSION
Information Hiding and Covert Channels are a viable threat to computer security and highly suspected to be used in advanced targeted attacks to cover its tracks. This is especially true for IT and OT networks were active measures are taken against common threats. This cat-and-mouse game leads to an increase in adversaries’ stealth abilities and methods to compromise and breach networks. In this paper, we describe further countermeasures to mitigate such threats of Information-hiding based attacks, making use of the computer security techniques described in the IAEA NSS 17-T and its’ graded approach to security as well as the DCSA and extend these by further methods and architectural changes to gain resilience. Moreover, we show how machine-learning based techniques and convolutional neural networks as part of artificial intelligence can be a method of threat hunting in this advanced threat scenarios.

ACKNOWLEDGEMENTS
The presented work is funded by the German Federal Ministry of Economic Affairs and Energy (BMWi) within the projects STEALTH (Grant No.: 1501589A) and SMARTEST2 (Grant No.: 1501600B) in the framework of the German reactor safety research program.

REFERENCES

[1] R. Altschaffel, M. Hildebrandt, S. Kiltz, and J. Dittmann. Digital Forensics in Industrial Control Systems. In Proceedings of 38th International Conference of Computer Safety, Reliability, and Security (Safecomp 2019), pages 128–136. Springer Nature Switzerland, 2019.
[2] R. Alschaffel, 2020. Computer forensics in cyber-physical systems: applying existing forensic knowledge and procedures from classical IT to automation and automotive. Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik. DOI: http://dx.doi.org/10.25673/35364
[3] M. St John-Green, 2020. If it is not secure, it is not safe.
[4] Computer Security Techniques for Nuclear Facilities, IAEA Nuclear Security Series No. 17-T (Rev. 1), https://www.iaea.org/publications/14729/computer-security-techniques-for-nuclear-facilities, 2021
[5] https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1527_web.pdf
[6] https://www-ns.iaea.org/downloads/security/security-series-drafts/tech-guidance/nst047.pdf
[7] M. Hildebrandt, K. Lamshöft, J. Dittmann, T. Neubert, and C. Vielhauer. 2020. Information Hiding in Industrial Control Systems: An OPC UA based Supply Chain Attack and its Detection. In Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '20). Association for Computing Machinery, New York, NY, USA, 115–120. DOI: https://doi.org/10.1145/3369412.3395068
[8] K. Lamshöft, J. Dittmann, Assessment of Hidden Channel Attacks: Targetting Modbus/TCP, IFAC-PapersOnLine, Volume 53, Issue 2, 2020, Pages 11100-11107, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2020.12.258.
[9] J. Hielscher, K. Lamshöft, C. Krätzer, & J. Dittmann, (2021, August). A Systematic Analysis of Covert Channels in the Network Time Protocol. In The 16th International Conference on Availability, Reliability and Security (pp. 1-11).
[10] K. Lamshöft. "The Threat of Covert Channels in Network Time Synchronisation Protocols." Journal of Cyber Security and Mobility (2022): 165-204.
[11] K.Lamshöft, T. Neubert, C. Krätzer, C. Vielhauer, and J. Dittmann. 2021. Information Hiding in Cyber Physical Systems: Challenges for Embedding, Retrieval and Detection using Sensor Data of the SWAT Dataset. In Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '21). Association for Computing Machinery, New York, NY, USA, 113–124. DOI:https://doi.org/10.1145/3437880.3460413
[12] K. Lamshöft, T. Neubert, J. Hielscher, C. Vielhauer, J. Dittmann, Knock, knock, log: Threat analysis, detection & mitigation of covert channels in syslog using port scans as cover, Forensic Science International: Digital Investigation, Volume 40, Supplement, 2022, 301335, ISSN 2666-2817, https://doi.org/10.1016/j.fsidi.2022.301335.
[13] Cybersecurity & Infrastructure Security Agency, "Alert (AA20-302A) Ransomware Activity Targeting the Healthcare and Public Health Sector", 2020, [Online] https://us-cert.cisa.gov/ncas/alerts/aa20-302a
[14] W. Mazurczyk, S. Wendzel, and K. Cabaj. 2018. Towards Deriving Insights into Data Hiding Methods Using Pattern-Based Approach. In Proceedings of the 13th International Conference on Availability, Reliability and Security (Hamburg, Germany) (ARES 2018). Association for Computing Machinery, New York, NY, USA, Article 10, 10 pages. https://doi.org/10.1145/3230833.3233261
[15] T. Neubert, C. Krätzer, and C Vielhauer. 2021. Artificial Steganographic Network Data Generation Concept and Evaluation of Detection Approaches to secure Industrial Control Systems against Steganographic Attacks. In The 16th International Conference on Availability, Reliability and Security (ARES 2021), August 17–20, 2021, Vienna, Austria. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3465481.3470073

[16] Hildebrandt, M, Altschaffel, R., Lamshoeft, K., Lange, M., Szemkus, M., Neubert, T., Vielhauer, C., Ding, Y., and Dittmann, J. (2020). Threat Analysis Of Steganographic and Covert Communi- cation in Nuclear I&C Systems. In Proceed- ings of IAEA ICONS 2020: International Conference on Nuclear Security: Sustaining and Strengthening Efforts, 10-14 February 2020, Vienna, Austria, https://event.do/iaea/a/#/events/3301/f/29007

[17] Hempstalk, K., Frank, E., Witten. I. H.: One-Class Classification by Combining Density and Class Probability Estimation. In: Proceedings of the 12th European Conference on Principles and Practice of Knowledge Discovery in Databases and 19th European Conference on Machine Learning, ECMLPKDD2008, Berlin (2008) 505-519