THREAT ANALYSIS OF STEGANOGRAPHIC AND COVERT COMMUNICATION IN NUCLEAR I&C SYSTEMS

M. HILDEBRANDT
Otto-von-Guericke University
Magdeburg, Germany
Email: mario.hildebrandt@iti.cs.uni-magdeburg.de

R. ALTSCHAFFEL
Otto-von-Guericke University
Magdeburg, Germany

K. LAMSHÖFT
Otto-von-Guericke University
Magdeburg, Germany

M. LANGE
Magdeburg-Stendal University of Applied Sciences
Magdeburg, Germany

M. SZEMKUS
Magdeburg-Stendal University of Applied Sciences
Magdeburg, Germany

T. NEUBERT
Brandenburg University of Applied Sciences
Brandenburg, Germany

C. VIELHAUER
Brandenburg University of Applied Sciences
Brandenburg, Germany

Y. DING
Magdeburg-Stendal University of Applied Sciences
Magdeburg, Germany

J. DITTMANN
Otto-von-Guericke University
Magdeburg, Germany

Abstract

Steganographic and covert communication is increasingly used for hiding attacks. Often information hiding is used by attackers in advanced persistent threats in order to operate without being noticed. Attackers might use it to hide data exfiltration or control channels for persistent malware. The paper analyzes potential threats to operational technology by investigating different communication protocols towards their suitability as cover channels. Using a generalized network architecture model and the communication flows of nuclear power plants, the attack potential using information hiding is exemplary assessed. An example for a supply chain attack for command injection is given on the foundation of the OPC UA protocol and an off-the-shelf programmable logic controller. Subsequently, recommendations for a strategic and operational preparation for operators towards the prevention and detection of information hiding attacks are derived.

1. INTRODUCTION

Within the scope of industrial automation systems in general and nuclear instrumentation and control systems (I&C) in particular, information hiding is a new application field. I&C systems consist of:

(a) Sensors, measuring the physical environment of a process
(b) Computing units, which process the sensor data
(c) Actors controlled by the computing units, influencing the physical environment of a process

Those components are vital for the automation in modern power plants, the process industry as well as manufacturing companies. However, with the increasing deployment of such systems and their importance for the economy, more and more attacks are directed towards I&C systems.

The paper's objective is the analysis of existing attacks utilizing information hiding in the information technology (IT) domain, such as the Turla Lightneuron Backdoor [1] which could exfiltrate files, execute commands or download and execute files via an Exchange mail server, and their projection to the operational technology (OT, I&C) domain. In conjunction with that, attacker models are created, which might be used to extend the cyber design basis threat (cyber DBT). As a foundation the paper will summarize the general setup of I&C systems based on IEC 62443 [2] and zone concepts from IAEA Nuclear Security Series draft NST047 [3] in Section 2.

Based on this generalized architecture model, potential cover channels and the suitability of I&C specific communication for steganography (or in short stego) are identified and analyzed in Section 3. Such cover channels are formed by the usual communication within the I&C system including, but not limited to, Modbus/TCP, OPC UA, Syslog, etc. Using the generalized architecture model and the communication flows, the attack potential using information hiding is assessed exemplary for OPC UA communication in Section 4. Based on that, recommendations for a strategic and operational preparation for operators towards the prevention and detection of information hiding attacks are derived in Section 5. Subsequently, s summary and an outlook on potential future work is given in Section 6.

2. STATE-OF-THE-ART - NETWORK ARCHITECTURE, SERVICES AND HIDDEN COMMUNICATION CHANNELS

This section gives an overview regarding the network architecture of Nuclear Power Plants (NPP) and the hidden communication channels which might potentially be used by an attacker.

2.1. Network Architecture According to NST047

Computer systems in NPP environments carry with them requirements in term of computer security. In order to increase computer security, communication flows are restricted and various subsystem decoupled from each other. Such a Defensive Computer Security Architecture (DCSA) is described in NST47 [3] composing five different Security Levels and various Security Zones. The Security Levels are assigned based on the impact a given system could have on the underlying physical process within the NPP. Hence, the Security Levels are as follows:

- Security Level 1 (**SL1**): systems vital to the facility (e.g. physical emergency protection)
- Security Level 2 (**SL2**): operational control systems which require high security
- Security Level 3 (**SL3**): supervision systems not required for operations
- Security Level 4 (**SL4**): technical data management systems (e.g. used for maintenance)
- Security Level 5 (**SL5**): business systems

Each security level is divided from the adjacent security levels by routers, firewalls and/or data diodes. In general, information flow from a lower security level to a higher security level is allowed, while communication from higher to lower security levels is highly restricted. From SL2 to SL1 no communication is allowed, while SL3 to SL2 allows for necessary acknowledgements or control packets to pass. SL4 to SL3 allows for specific and limited activity, while SL5 to SL4 carries little restriction. SL5 might be connected to the internet.

In addition, systems on the same security level are grouped into various zones representing logical or physical groupings decoupled from other systems by using computer security measures (like firewalls).

2.2. Setup and Services of I&C systems based on IEC 62443

The IEC 62443 "Industrial communication networks - Network and system security" (see [2]) deals with the IT security of so-called "Industrial Automation and Control Systems" (ICS). It was developed as a standard for automation technology in the process industry. Developed and compatible with the international IT security standard ISO 27001/27002 (see [4]), IEC 62443 now covers all industries from discrete manufacturing to distributed supply systems. It deals with the topics of compliance (conformity with laws and regulations), technical requirements and implementation, as well as the protection of persons. This family of standards consists of 7 parts. Part 1-1 "Terminology, concepts and models" (see [2]) of IEC 62443 describes the life cycle of security requirements and measures in the interaction between component manufacturers, plant constructors and plant operators. The IEC62443-2-1 "Establishing an industrial automation and control system security program" (see [5]), on the other hand, describes a complete information security management system (ISMS). Since weak points and thus also attack vectors can be introduced by service providers, security requirements must be placed on their systems and processes. This is described in Part 2-4 of the 3rd part of this standard (see [6], which is a guide to IT security protection in industrial environments. Here, the IEC 62443-3 [7] concept comprises central components: security zones with the associated communication channels, security levels and risk analyses. IEC 62443-3-3 regulates the areas of authorization, authentication and identification of persons and systems from the operator's point of view and consists of 7 basic requirements (**FR** - Foundational Requirement, see TABLE 1).

TABLE 1 Foundational Requirements of IEC 62443 [7]

No.	Abbreviation	Title
1	IAC	Identification and Access Control
2	UC	User Control
3	SI	System Integrity
4	DC	Data Confidentiality
5	RDF	Restricted Data Flow
6	TRE	Timely Response to Events
7	RA	Resource Availability

For each of these requirements there are several system requirements (**SR** - System Requirement). Here the specific requirements for a system are described. In order to identify the correct system requirement for the operation, a priority risk analysis assigns it to a security level (**SL**). A distinction is made between 5 security levels. Those security levels are semantically different to the security levels in NST047. For better comprehension, we denote the use of IEC 62443 by adding a dash (SL-1, SL-2, SL-3, SL-4). In the context of the IEC 62443 SL-0 describes applications with no security requirements at all, whereas SL-4 describes the highest requirements in terms of security. Depending on the security level, there are different amounts and different sets of requirements. In addition, depending on the SL, no or several extensions of the requirements (**RE** - Requirement Enhancement) are added as exemplary shown in TABLE 2.

TABLE 2 Exemplary SR and RE based on IEC 62443 Security-Level

SR and RE	SL-1	SL-2	SL-3	SL-4
SR 2.8 Testable events and their recording	x	x	x	x
SR 2.8 RE1 Centrally managed system-wide event recording			x	x
SR 3.1 System shall protect integrity of transmitted information	x	x	x	x
SR 3.1 RE1 Use of cryptographic protective measures to maintain integrity			x	x

2.3. Steganography and Data Hiding

Hidden communication is often called secret communication or Prisoners' Problem in presence of a warden. A sender (often called Alice) and receiver (often called Bob) wants to hide a message and their communication activity in a cover channel or cover object, the warden is able to see and read the communication. There is a wide variety of literature and known approaches. An introduction to the main concepts and definitions can be found as open access in "Information Hiding" [8].

In the following, we mainly consider that, besides the confidentiality of the information hidden, the undetectability of the presence of an information or a communication channel from the point of the warden is the objective of steganography and information hiding.

To achieve both goals the hiding approaches need to fulfill the following requirements:

- use cover channels or cover objects which are plausible in the normal, realistic and expected behavior in the target IT System,
- protocol-compliant to ensure the information can be exchanged successfully over a cover channel or cover object between sender and receiver, and
- also warden-compliant to ensure that an observer listening to the communication channel or objects has no suspiciousness that a hidden channel exits at all.

In the literature (see [8]) two special cases of the warden are discussed: the passive Warden (observing only) with the possibility to read the communication and the goal to just detect the presence of hidden communication, and the active Warden (traffic alteration) with the means to actively change and manipulate the cover in order to disturb the communication or even expose and detect sender and receiver.

Approaches for steganography and information hiding can be divided into cover selection, cover synthesis and cover modification (see also [8] or [9]). Very simple cover modification approaches known as pure steganography use a plausible cover such as a network packet and modify/embed directly in the packet (e.g. header information) the plain-text secure and hidden message without any usage of a key.

The security of such pure approaches is very limited and if the warden learns about the embedding strategy the overall message can be detected and also directly read, modified or deleted. A first enhancement would be that both sender and receiver – share a secret to encrypt the message to in order to avoid disclosure of the message text. To establish a secure communication channel, secret key based steganography should be designed with the assumption, the algorithm is public and known (see also [8]: “Kerckhoffs’ Principle, which in cryptography states that the enemy should be assumed to know the system, by saying that the Warden is aware of Alice and Bob’s communication algorithms”).

If the secret message is embedded directly from the endpoint to another endpoint in a communication channel or object, the approach is called end to end steganography or active steganography/information hiding. If the message is embedded only in parts, such as one section of the communication, the approach is called passive steganography/information hiding, see in [10]. The warden might therefore have access to original data without a message and stego channels/objects with a message.

Another challenge for establishing the steganographic communication channel is the synchronization between the sender and the receiver (see e.g. [11]). Based on a known property of the messages, the beginning of a steganographic message must be indicated in order to be able to extract the embedded message. Such an indicator could be either a specific sequence, i.e. a preamble, specific times or conditions of the cover channel.

3. THE THREAT OF HIDDEN COMMUNICATION IN OPERATIONAL TECHNOLOGY

Unlike steganography in image data, where the sender is able to choose images that are particularly suitable for embedding additional information, steganography in operational technology (OT) networks needs to utilize the existing network traffic, ideally without influencing the controlled process in an unintended manner. In addition to that, the amount of available cover data is significantly lower in comparison to images. Overall, we can assume that there are three different objectives of steganographic channels in OT networks:

(a) Injection of Information,
(b) Exfiltration of Information,
(c) Command & Control, i.e. bi-directional communication.

All communication options could be performed between the original end points of the communication (active steganography) or by injecting or extracting information at any communication node between the sender and the recipient of a message (passive steganography). Hereby the overall goal remains the same as for any other steganographic or hidden channel, the detection of the communication channel should be avoided.

3.1. Hidden communication channels in IT networks

In the information technology (IT) domain information hiding techniques are increasingly used to minimize indicators of compromise. Here, several communication protocols, e.g. DNS requests or SMTP (email) traffic could be used as cover data. In addition to this various media data such as images or videos are common in IT networks. Such data requires a high bandwidth for its transfer and thus, allows for embedding larger amounts of data. In addition to that, an attacker might be at liberty of choosing especially suited cover data to hide the steganographic data in. Furthermore, the traffic in IT networks is usually more dynamic, i.e. less predictable, in comparison to OT traffic which simplifies the objective of hiding data. A recent example of a complex advanced persistent threat (APT) is the Turla Lightneuron Backdoor [1] which infects Exchange mail servers allowing for exfiltrating files, executing arbitrary commands or downloading and executing files. Similar to a SPAM-filter module, the backdoor registers itself as a transport agent within the Exchange server integrating a dynamic-link library (DLL) containing the malware's machine code. The backdoor utilizes a storage channel using the emails as cover data. Here, the hidden communication relies on specifically crafted PDF or JPG files attached to the emails. This is an example of active steganography which allows for establishing a command and control channel with the infected Exchange server.

An example of a data exfiltration using steganography is the malware UDPoS [12]. It is designed to steal credit card information in point-of-sale terminals. The gathered data is afterwards sent via crafted DNS requests. However, the data to be transferred is coded within the queried domain names. Thus, a detection of such a malware is possible by monitoring the DNS requests originating from the IT network.

3.2. Potential hidden communication in OT networks

Achieving the objective of avoiding a detection of the covert channel in OT networks is significantly more challenging, since the traffic and the controlled process is usually known well and any deviation from the known patterns might be an indicator of compromise. Besides the process control requirements of the IEC 62443 might lead to additional communication, e.g. Syslog in order to fulfill the requirement enhancement “Centrally managed system-wide event recording”.

In addition to that, not all possible cover data can be utilized depending on the access to low level information especially in active steganography. Based on the classification of potential channels from [13] we have distinguished between the high-level access, e.g. the programming of a PLC (Programmable Logic Controller - a computing unit common in OT networks), and low-level access which requires access to the PLC's firmware as depicted in Fig. 1.

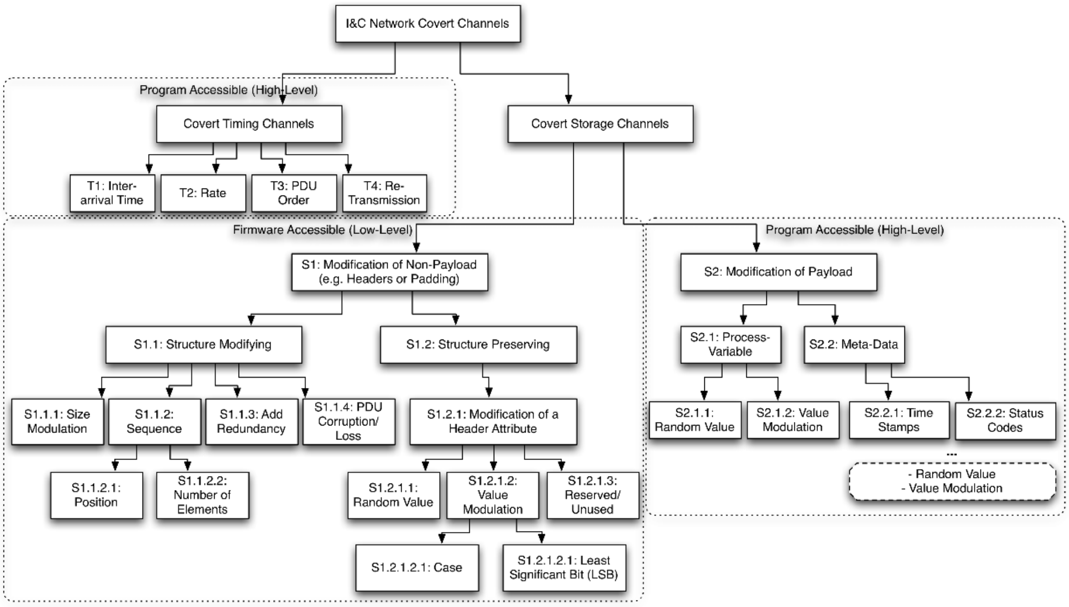
I&C Network Covert Channels
Program Accessible (High-Level)
Covert Timing Channels
T1: Inter-arrival Time
T2: Rate
T3: PDU Order
T4: Re-Transmission
Covert Storage Channels
Firmware Accessible (Low-Level)
S1: Modification of Non-Payload (e.g. Headers or Padding)
S1.1: Structure Modifying
S1.2: Structure Preserving
S1.1.1: Size Modulation
S1.1.2: Sequence
S1.1.3: Add Redundancy
S1.1.4: PDU Corruption/ Loss
S1.2.1: Modification of a Header Attribute
S1.1.2.1: Position
S1.1.2.2: Number of Elements
S1.2.1.1: Random Value
S1.2.1.2: Value Modulation
S1.2.1.3: Reserved/ Unused
S1.2.1.2.1: Case
S1.2.1.2.1: Least Significant Bit (LSB)
Program Accessible (High-Level)
S2: Modification of Payload
S2.1: Process-Variable
S2.2: Meta-Data
S2.1.1: Random Value
S2.1.2: Value Modulation
S2.2.1: Time Stamps
S2.2.2: Status Codes
...
- Random Value
- Value Modulation

Fig. 1. Covert Channels in I&C Networks Based on [12] and access requirements for active steganography, access, access requirements

From the programming of the PLC the payload (S2), including relevant meta-data and the timing (T1-T4) is accessible or modifiable. The modification of the headers usually requires low-level access to the TCP/IP stack. Thus, utilizing such data as cover data usually requires modifying the firmware of the PLC.

In I&C networks potential cover channels rely on the existing protocols for controlling the process, e.g. Modbus/TCP, OPC UA, S7comm as well as supporting services for logging or time synchronization such as Syslog, SNMP or NTP. From the perspective of hiding data, the advantage of those protocols is the regular traffic.

It can be expected that control messages or log messages occur in small, known intervals with small variations caused by the PLC's cycle times.

Timing channels alter this communication behavior in order to embed additional information. However, given the deterministic behavior of nuclear power plants, any deviation from the known traffic could be detected as an anomaly. For example, the alteration of the PDU order (T3) would be easy to implement on the programming level of the PLC, depending on the bit to embed, the order of queries to a process control server could be altered without influencing the controlled process. Likewise, an attacker with access to network infrastructure devices could easily alter the sequence of packets using a store-and-forward approach. Here, a detection approach could learn the order of the queries for the normal system behavior as a result any deviation from the previously observed order is considered as an indicator of compromise.

For storage channels on one hand the modified packets need to be protocol-compliant, on the other hand the process should not be influenced in order to avoid raising any suspicion. Within the scope of a supply-chain-attack both the firmware or the programming could contain additional code for embedding or extracting data in or from hidden communication channels.

Besides pure storage or timing channels a hybrid channel can be established by modifying a value for some arbitrary interval of time. Such a channel could be compared to simple wireless protocols with an on-off-keying. Within the scope of this paper we utilize the suffix 'h' in subscript for indicating a hybrid channel. For example, $S1.2.1.2.1_h$ would be a hybrid channel based on a modification of the case within a header attribute of the Ethernet frame.

3.3. First Attacker models for OT networks

Specific attacker models depend on the objective of the hidden communication channel. The injection of information (a) in I&C systems is usually directed towards the PLC controlling a process or a Human-Machine-Interface (HMI) displaying the status of a process. Since the attacker has no information about the current status of the controlled process, it is likely that the process should be manipulated in a pre-defined manner, e.g. in order to destroy components. As a requirement, the attacker needs to have a high knowledge level about the controlled process, the components and the process variables. The motivation for the attacker is likely to cause any kind of damage. For the exfiltration of information (b) only a limited knowledge level about the is necessary. However, the key information to be exfiltrated needs to be known to the attacker. The motivation for the data exfiltration usually aims towards a gain – either knowledge or economical – for the attacker. For example. in the context of NPPs an information about non-scheduled down-times could present insider information for trading shares and options of plant operators. The establishment of a command & control channel (c) combines the possibilities of the first two channels. In addition to that, an attacker with access to a command & control channel can gain more information about the plant and its processes allowing for more directed attacks. However, especially with the concept of zones and different security levels, establishing and maintaining such a channel is significantly more challenging.

4. EXEMPLARY SUPPLY-CHAIN ATTACK BASED ON OPC UA

For the exemplary assessment of the possibility of a supply chain attack based on steganography in OPC UA we use an off-the-shelf Siemens S7 1516-3 PN/DP PLC[1] running firmware version 2.6.1. For the experiments, we create a custom OPC UA server based on the open62541[2] library. In order to avoid modifying the firmware of the PLC the cover channels are limited to timing channels and storage channels within the payload (S2). To avoid influencing the process, we have selected the Meta-Data, specifically the time stamps of the values within the OPC UA read response.

The OPC UA client integrated into the PLC stores the results of the OPC UA responses in a data block. Within this block the values, the source timestamps and the status codes are stored. Thus, a steganographic channel could use any of those three types of information as cover data. However, the status code is usually zero indicating a valid reading (OPC UA Good). Thus, any deviation from this value should be considered as an anomaly, e.g.

[1] https://mall.industry.siemens.com/mall/en/WW/Catalog/Products/10204211

[2] https://open62541.org/

caused by a defect and hence triggering an investigation. A modification of the value reading could influence the process behavior but might be undetected if it is in the valid value range and not significantly different to the previous reading. Here, for example an embedding in the least significant bit might be an option for the stego channel. However, the capacity of this channel would be limited. The last option is the embedding in the source timestamp. The source timestamps in OPC UA are represented by a 64-bit value with a nanosecond resolution accompanying each value read from the server. This timestamp should indicate when a specific value has been updated the last time.

Given the cycle time of a PLC of at least 1 millisecond, the timestamp has enough space for embedding additional information without influencing the controlled process. However, in order to avoid discrepancies in the sequence of the value readings, it is crucial to ensure that the timestamps are increased in strictly monotonic manner.

In our exemplary evaluation the PLC is implementing a rather simple control process of a generator system. The PLC regularly reads the turbine speed, the grid frequency and a control signal for the generator breaker from an OPC UA server. If the turbine speed would result in electricity generated matching the grid frequency and the breaker command is set, the PLC will close the generator breaker by sending a digital output signal on the output 8.0. If the breaker command is not set or the turbine speed would result in a wrong frequency, the breaker will be opened by sending a digital output signal on the output 8.1. The PLC is running at a cycle time of 100 milliseconds.

Within the PLC we have implemented two steganographic communication channels in SCL, a storage channel based on S2.2.1 and a hybrid channel based on the timestamp S2.2.1$_h$.

For preparing the extraction of the injected information the PLC the data type LDT in the OPC UA client data block is converted into DTL which allows direct access to the data as summarized in TABLE 3.

TABLE 3 Data Fields in the DTL Data Type for Timestamps

Data Field	YEAR	MONTH	DAY	WEEKDAY	HOUR	MINUTE	SECOND	NANOSECOND
Data Type	UInt	USInt	USInt	USInt	USInt	USInt	USInt	UDInt

In the following subsections the two steganographic channels are described in detail. Both channels are designed to inject data to the PLC in order to trigger a specific behavior or modify the process.

4.1. Covert Storage Channel (CC$_S$) Based on the OPC UA Source Timestamp

Due to the cycle time we can expect that values would be updated in the PLC once in every tenth of a second. In order to embed information into the timestamp can utilize the remainder of the timestamp. The basic idea of our covert storage channel CC$_S$ is to embed enough information to set arbitrary digital outputs to a specific state. For that the function for indirect addressing 'POKE_BOOL' is used within the PLC's programming. This function allows us to access the output on the foundation of a numerical identifier without the need of a specific tag within the PLC. The embedded message is neither encrypted or obfuscated, thus in theory it could be detected if the OPC UA communication is not encrypted. The resulting value of the nanosecond part of the timestamp is summarized within TABLE 4.

TABLE 4 Structure of the nanosecond part of the timestamp as cover data S2.2.1 with an embedded steganographic message (Bold face denotes the digits of the steganographic message, italic dynamic digits)

Milliseconds			Microseconds			Nanoseconds		
1/10s	**4**	**2**	**M**	**N**	**O**	0	0	0

The tenth of a second is retained from the original timestamp followed by 42 as a header for the steganographic message for the synchronization. The actual nanoseconds remain constant at '000'. In theory, it would be possible to embed data in this part of the timestamp as well. However, such a manipulation would be rather obvious because it is a significant deviation from the normal behavior of the source timestamps. The three digits representing the microseconds within timestamp determine the parameters of the 'POKE_BOOL' function. The first digit M contains the byte offset within the range between 0 and 3. Within the PLC's function block the byte offset B is determined by calculating B = 8 + M. The bit offset represented by the second digit N. It is in the range between 0 and 8 and is directly passed to the function. The third digit O describes the target state of the output. In our example we use the digit two to indicate the target state 'true' whereas the all other options are considered as 'false'.

Our evaluation shows that the steganographic channel is not impacting the protocol compliance of the cover channel. However, due to the large amount of embedded data, the static part of the timestamp could be detected as an anomaly.

4.2. Covert Hybrid Channel (CC_h) Based on the OPC UA Source Timestamp

Our exemplary covert hybrid channel CC_h alters the second and the nanosecond areas of the source timestamp of an OPC UA variable within the read request by delaying updates of the timestamp. The embedded information relies on the cross sum of the two digits of the seconds. In order to send a signal, the cross sum is kept even or odd for an arbitrary interval. In order to ensure a monotonic increase of the timestamp, the nanosecond field is scaled. For example, for sending packets with an even cross sum the following two rules are applied on the OPC UA server:

(a) the two digits already have an even cross sum: divide nanosecond field by a factor of two

(b) the two digits have an odd cross sum: subtract 1 from the second digit, divide the nanosecond field by a factor of two and add half a second

On the PLC the time stamps have to be monitored and for each request the cross sum has to be determined. The PLC needs to store the timestamp when the last odd cross sum has been observed. If a certain interval of time has passed since the last observation of an odd cross sum, a specific action could be triggered.

In comparison to the storage channel this hybrid channel has a significantly lower capacity. Furthermore, it is not possible to influence the process in each cycle of the PLC.

5. STRATEGIC AND OPERATIONAL PREPARATION FOR THE PREVENTION AND DETECTION OF INFORMATION HIDING ATTACKS

This section discusses on how information hiding attacks can be detected. It starts with the necessary preparation and then goes on to the specific detection measures.

5.1. Strategic and Operational Preparation

Detecting information hiding attacks relies on the ability to identify anomalies in the communication between the various OT components. This implies two requirements. At first, the possibility to gather the communication within the OT network. It also requires an understanding of the normal communication flows. In a complex world of APTs, both measures need some special attention during the strategic preparation of the facility.

The data capturing must be performed in a way which is not identifiable by a potential attacker. Otherwise, a sophisticated attacker could identify the data capture and alter the communication in order to avoid detection until the capturing is discontinued. This basically requires the inclusion of interfaces which can be used for data capturing during the installation of the architecture. If following a DCSA (see section 2.1.), the network infrastructure elements at the borders between Security Levels and Security Zones can be used for this means by using port mirroring.

Getting a real understanding of the normal communication flows can be achieved by reviewing the specification of the attached OT systems. However, while the physical processes controlled by OT are usually well-defined, the protocols used are often proprietary and hence, constructing a correct model of a normal communication from the specification alone is often impossible. Another method is baselining, where a known-good of the communication is captured and used for creating a model of known-good communication. However, this requires the system to be in a known-good state. With supply chain attacks, this might not always be the case. The best method seems to be baselining and verification based on the specification of the components in questions and the protocols used.

5.2. Detection Approach for CC_S

We have designed a pattern recognition-based detection approach, for CC_S as introduced in section 4.1. The approach analyzes the OPC UA source timestamps in the recorded network traffic in the OT. Therefore, a classifier is trained with the data mining suite WEKA 3.8 [14]. We use its OneClassClassifier [15] with default parameters and a target-rejection-rate set to 0.001 to keep the false positives rate as low as possible. The OneClassClassifier is only trained with known-good OPC UA network data to define a so-called target class. When the target class is trained, the classifier has the ability to detect outliers of the target class.

To train the classifier of our pattern recognition-based approach, we have designed a 9-dimensional feature space which determines the percentage distribution of the third microsecond digit **O** (see TABLE 4) in the OPC UA timestamp. The distribution is determined of a network packet sequence of 25 packets. We use OPC UA source timestamps for the first two variables of every OPC UA response packet to analyze a packet sequence. Thus, the distribution of 50 digits is analyzed for one training sample. We assume, that OPC UA timestamps in unaltered network traffic follow a uniform distribution for the digit for O and that network traffic including steganographic messages should show anomalies here. Our detection approach determines if a packet sequence of 25 packets belongs to the trained target class or not.

For the exemplary evaluation of our detection approach, we create one training data set and two test data sets. The training data set consists of 489 authentic packet sequences (25 packets per sequence) and is used to train the OneClassClassifier based on the 9-dimensional feature space presented in the previous chapter. Therefore, we determine the distribution of the third microsecond digit O for every sequence of the training data and label the data as target. The test data includes two data sets: One data set includes only authentic packet sequences (93) and the other data set includes only packets with the integrated covert channel communication (208). Our trained classifier classifies all samples of the two test data sets correctly.

5.3. Detection Approach for CC_h

For the detection of CC_h from section 4.2, we use a pattern recognition-based detection approach, too. Therefore, we analyze time deltas of the OPC UA timestamp of a network packet to its two previous packets. Thus, for this detection approach we analyze a packet sequence of three packets We calculate the time deltas for the last recorded packet of a sequence to the two previous recorded packets. So, for every packet sequence we determine two delta times:

$$\delta\{x1, x2\}=\{\text{timestamp}[n] - \text{timestamp}[n-1], \text{timestamp}[n]-\text{timestemp}[n-2]\}$$

For the classification of test data, we have to define two target time intervals for the x1 and x2 to classify data with this approach. The time intervals are the features for our purpose-built custom classifier which specifies the range of x1 and x2 for authentic network traffic. When x1 or x2 of a sequence is outside this target time intervals, it is classified as outlier, if both are inside the range of the intervals, it is classified as target. The time intervals are defined with authentic network traffic by training data. To define the target time intervals, we simply calculate the minimum and maximum time delta of a training data set and add the standard deviation of the data set to the min and max boundaries of the intervals.

To evaluate our detection approach, we use 4080 packet sequences to train our classifier. Based on the training data, the classifier defines two target time intervals. If both time deltas (x1 and x2) of a packet sequence are inside the range of defined intervals the sample is classified as target, if not, the sample is classified as an outlier. For our exemplary evaluation we use two test data sets. Our first test data set has only authentic network traffic (781 packet sequences) and our second test data set includes authentic traffic and traffic with covert channel communication. All samples in both test data sets are classified correctly.

6. SUMMARY AND FUTURE WORK

This paper sheds light on the thread posed by hidden communication channels within OT networks. In APTs such channels can be created using various cover data in order to avoid being detected by minimizing indicators of compromise. In section 3 we present and analyze potential cover channels within the typical communication in OT networks and the requirements to create hidden channels on a PLC. We show two simple, exemplary channels based on the OPC UA source time stamps including possible detection approaches. The main purpose of those channels is to raise the awareness towards hidden communication channels in ICS/OT networks. In an ideal case an audit of all components could be performed at the source code and hardware specification level in order to avoid such threads caused by supply-chain attacks. However, even then passive steganography could be used to hide the data exfiltration e.g. by insiders.

In future work more complex methods for embedding the hidden data should be evaluated in order to develop new detection approaches for a broad variety of different steganographic attacks on OT networks.

ACKNOWLEDGEMENTS

The work in this paper has been funded by the German Federal Ministry for Economic Affairs and Energy (BMWi, Stealth-Szenarien, Grant No. 1501589A, 1501589B and 1501589C) within the scope of the German Reactor-Safety-Research-Program.

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

REFERENCES

[1] Faou, M., TURLA LIGHTNEURON One email away from remote code execution, ESET Research White papers, May 2019 (2019) https://www.welivesecurity.com/wp-content/uploads/2019/05/ESET-LightNeuron.pdf

[2] CAN/CSA-IEC/TS 62443-1-1:2017-10-01, Industrial communication networks - Network and system security - Part 1-1: Terminology, concepts and models (Adopted IEC technical specification 62443-1-1:2009, first edition, 2009-07) (2017)

[3] IAEA- INTERNATIONAL ATOMIC ENERGY AGENCY, Computer Security Techniques for Nuclear Facilities, Nuclear Security Series nº 47, NST047 DRAFT (2017)

[4] ISO 27001/27002

[5] CAN/CSA-IEC 62443-2-1:2017-10-01, Industrial communication networks - Network and system security - Part 2-1: Establishing an industrial automation and control system security program (Adopted IEC 62443-2-1:2010, first edition,) (2017)

[6] CAN/CSA-IEC 62443-2-4:2017-12-01, Security for industrial automation and control systems - Part 2-4: Security program requirements for IACS service providers (Adopted IEC 62443-2-4:2015, first edition, 2015-06) (2017)

[7] CAN/CSA-IEC 62443-3-3:2017-10-01, Industrial communication networks - Network and system security - Part 3-3: System security requirements and security levels (Adopted IEC 62443-3-3:2013, first edition) (2017)

[8] Ker, A.: Information Hiding (complete), (2016) http://www.cs.ox.ac.uk/andrew.ker/docs/informationhiding-lecture-notes-ht2016.pdf

[9] Chandramouli, R., Li, G., Memon, N. "Adaptive Steganography", In Security and Watermarking of Multimedia Contents IV (Proceedings of SPIE Vol. 4675) (2002) 69-78

[10] Dittmann, J., Hesse, D., Hillert, R. "Steganography and steganalysis in voice-over IP scenarios: operational aspects and first experiences with a new steganalysis tool set", In Security, Steganography, and Watermarking of Multimedia Contents VII (Proceedings of SPIE Vol. 5681) (2005) 607-618

[11] Grabski, S., Szczypiorski, K.: Steganography in OFDM Symbols of Fast IEEE 802.11n Networks, IEEE Security and Privacy Workshops, San Francisco (2013) 158-164

[12] Cylance Threat Research Team, Threat Spotlight: Inside UDPoS Malware, (2018) https://threatvector.cylance.com/en_us/home/threat-spotlight-inside-udpos-malware.html

[13] Wendzel, S., Zander, S., Fechner, B., Herdin, C., Pattern-Based Survey and Categorization of Network Covert Channel Techniques, CSUR 47 3 (2015)

[14] Hall, M, Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. H.: "The WEKA data mining software: An update", SIGKDD Explorations 11 1 (2009)

[15] 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