How to Find out What's Going on in Encrypted Smart Meter Networks – without Decrypting Anything

Oliver Eigner
University of Applied Sciences, St. Pölten, Department of Computer Science & Security
oliver.eigner@fhstp.ac.at

Hubert Schölnast
University of Applied Sciences, St. Pölten, Department of Computer Science & Security
hubert.schoelnast@fhstp.ac.at

Paul Tavolato
University of Vienna, Research Group Security and Privacy
paul.tavolato@univie.ac.at

ABSTRACT

Smart meter networks are part of the critical infrastructure and therefore central to IT security consideration. Besides various forms of access control a permanent monitoring of the network traffic is of utmost importance to the detection of malicious activities taking place. Such monitoring must happen in real time and should possibly be implementable everywhere in the network. These requirements do not allow for the decryption of the network traffic. The paper describes a method by which network packets can be assigned to use cases common in smart meter infrastructures without the need for decryption. It is based solely on metadata and reliably can establish the relationship between a network packet and a use case. The information calculated with this method can be used to detect packets that are not pertaining to any of the allowed use cases and hence are highly suspicious. Moreover, the execution of use cases not initiated by the central server become evident, too, and should raise corresponding alerts. The method was implemented as a proof-of-concept and tested in the real-world environment of a medium-sized city.

CCS CONCEPTS

• **Security and privacy** → Intrusion/anomaly detection and malware mitigation; Intrusion detection systems.

KEYWORDS

Smart Meter Networks, Anomaly Detection, Security Monitoring, Machine Learning

ACM Reference Format:
Oliver Eigner, Hubert Schölnast, and Paul Tavolato. 2024. How to Find out What's Going on in Encrypted Smart Meter Networks – without Decrypting Anything. In *The 19th International Conference on Availability, Reliability and Security (ARES 2024), July 30–August 02, 2024, Vienna, Austria.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3664476.3670925

1 INTRODUCTION

The progressively increasing automation and digitization of distribution networks for electrical energy opens up a large new range of possibilities for cyber-attacks. To guarantee the resilience of such critical infrastructure the implementation of effective protection mechanisms is essential. Security measures generally used today are limited to access control methods. State of the art security, however, demands "security in depth": additional measures must be implemented that become effective if an attacker has succeeded in overcoming the access barriers. Among others this can be realized by monitoring the network traffic with the goal of detecting anomalies in the system's behavior.

Intrusion detection and anomaly detection systems for smart meter networks do have quite different requirements from general intrusion detection systems: they must cope with specialized industrial protocols like Power Line Communications (PLC), scalability, availability in environments with restricted resources and others; see [1]. On the other hand, smart grid infrastructures are highly dedicated systems, which means that there is only reduced variability during normal operation. Hence anomaly detection systems based on a strict definition of regular operations are especially apt for securing such systems.

The first step to implement an anomaly detection system in automation networks of energy distribution systems is the development of a model of normal system behavior. The monitoring process uses this model and compares it with ongoing network traffic to detect anomalies. It's an advantage of smart meter networks that only a restricted set of use cases (such as meter readings, turning meters on or off and others) is executed on the network. If it turns out that each use case has a distinguished traffic profile, network packets could be assigned to use cases. And packets not belonging to a valid use case would indicate an anomaly. Moreover, use cases that were not initiated by the server are suspicious, too. In this paper we discuss a classification algorithm that assigns packets to use cases.

Network traffic in smart meter systems, however, nowadays usually comes along encrypted. And due to the fact that anomaly detection must occur in real-time, a workable detection system should work with the encrypted network packets only, as a decryption process would take too much time. Moreover, a detection algorithm that is distributed over the network is confined to using the encrypted packets only (decryption should be done solely by a safeguarded central server to prevent attackers from using it). A model of normal network behavior must therefore rely on metadata only.

The approach in this paper is characterized by the following aspects, the combination of which is to our knowledge not used so far:

1. It defines normal behavior by assigning packets to use cases.

ARES 2024, July 30–August 02, 2024, Vienna, Austria
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1718-5/24/07
https://doi.org/10.1145/3664476.3670925

2. It does not work with statistical values only, but tries to base its detection on semantically relevant data by relating network packets to use cases.
3. It works with encrypted packets and bases the classification of packets to use cases on metadata only.

Section 2 discusses related work, section 3 gives an overview of smart meter network traffic and section 4 describes the classification method. Section 5 shows the method in a real-world example (the project includes a proof-of-concept implementation of the monitoring and anomaly detection system). Conclusions and an outlook to future work completes the paper.

2 RELATED WORK

General discussions about security in distribution networks of electrical energy can be found among others in [2–7]. The concept of anomaly detection in general is described fully in [8], showing the multifaceted requirements of such systems; for example how to deal with future changes in normal behavior.

Most of the papers on anomaly detection in smart meter networks try to define anomalies, i.e. behavior that is related to expected attacks. In a paper by Zhang et al. [9] a distributed analysis module is proposed that should detect malicious behavior within all sections of a smart grid. The patterns of such malicious behavior are deduced statistically from observations of the network traffic by means of SVM based machine learning algorithms. Formal definitions of anomaly detection methods can be found in [10]. A paper by Mitchell und Chen [11] describes a behavior-based intrusion detection system that aims at anomalies in the behavior of smart metering devices, where rules are defined in propositional calculus. A paper by Wei et al. [12] sets its focus on communication patterns in network traffic.

One of the very few papers that tries to define normal behavior (and everything else is classified as malicious) is [13]. It describes an intrusion detection system based on specifications of normal behavior; these specifications are mainly defined as statistical values. The system, however, works only on the two lower levels of the network, which reduces the relevance of the parameters observed significantly. Another one is [14] where the authors describe a purely statistical model based on Brown's, Holt's, and Winters' models. Realtime anomaly detection based on payload specifications is described by Düssel et al. [15]. This system defines a multidimensional feature space and detects suspicious packets by comparing them with the byte sequences of normal packets. It analyzes n-grams of the TCP payload and hence can be applied to upper levels of the protocol.

Encrypted packet classification is dealt with in a paper by Chen et al. [16], but not specifically for smart meter networks. Another paper [17] describes a classification method for TLS/SSL traffic based on a neural network. An intrusion detection system für smart meter network that partially relies on metadata of network traffic (number and size of packets sent and received) is described in [18]. A research report on intrusion detection in smart meter data [19] deals with machine learning methods, but does not go into details about the method and the infrastructure.

3 SMART METER NETWORK TRAFFIC

We restrict the analysis to smart meter networks using the common DLMS-COSEM/G3PLC protocol (which is more or less equivalent with IEC 62056). This choice is due to the protocol used in the real world example described below, but should be no real restriction as the method works with encrypted packets only and thus cannot access the contents. The packets conforming these protocols are first filtered out and all other packets are neglected. The number of uses cases executed over the network is limited. As a reference for the use cases available in smart meter networks see [20], where all use cases are described in detail. (Though the document dates from 2015 it is still valid and current as of June 2024.) The scope of these use cases covers the end-to-end encrypted communication path between a terminal device, which is both a smart meter and a load switching device, and the central system. This includes the following components:

- The WAN communication unit of the end device (smart meter and load switching device). Not included in the scope are the other parts of the end device, including e.g. the customer interface of the end device and the actual electricity meter (the metering unit).
- The last mile: This refers to the communication of the end device with the next gateway. This communication can take place on the network layer (layer 3 in the OSI model) via IPv4 or IPv6 and on the physical layer (layer 1 in the OSI model) via PLC (Power Line Carrier) or radio. For PLC, the G3-PLC protocol is used in layer 2 (data link layer), which itself secures the last mile with end-to-end encryption.
- The second mile: This is the communication between the gateway that communicates with the end device and the central system. This communication takes place on the physical layer via fiber optic cables, copper cables (e.g. Ethernet; no power cables) or radio.
- Point-to-point: This is the direct communication of a terminal device with the central system via mobile radio (GPRS, LTE, etc.) If a terminal device communicates directly with the central system via mobile radio, the subdivision into last and second mile does not apply.
- The central system: This is the end point of communication for terminal device protocols. The central system has interfaces for connecting to the IT backend systems. These systems and the interfaces of the central system to these IT backend systems are no longer part of the scope of the use cases.

The entire scope is represented on layer 4 of the OSI model as a continuous TCP connection between the end device and the central system. On layer 5, the TLS (Transport Layer Security) protocol is used to secure the payload of the higher layers end-to-end. In the higher layers, special protocols like DLMS-COSEM are used to transmit requests and responses for the defined use cases.

The above-mentioned document describes 83 use cases, which are divided into 12 groups:

1. Reading (5 use cases) – reading data from a smart meter
2. Switching off and enabling to switch on (8 use cases) – switching off the load switching device and enabling the load switching device to be switched on

3. Parameterization (22 use cases) – setting various parameters and switching various functions of the end device on or off
4. Firmware upgrade (12 use cases) – loading, installing and canceling upgrades of calibrated and non-calibrated parts of the terminal device
5. Alarms and events (2 use cases) – sending alarms and events to the end device
6. Load switching (8 use cases) – functions for controlling the switching table of the load switching device
7. Calibration / testing (2 use cases) – calibration and testing of the smart meter
8. Interface activation and deactivation (3 use cases) – activate or deactivate the maintenance and communication interface on the end device
9. Prepayment (4 use cases) – activate and deactivate prepayment, load credit
10. Registration and deregistration of the end device (2 use cases) – automated registration and re-registration of the end device
11. Gateway (6 use cases) – switch gateway function of the end device on and off, query status of the gateway
12. Security (9 use cases) – functions for managing cryptographic parameters and roles

Of these 83 use cases, 31 use cases describe physical manipulation of the device on site and do not use network communication at all. The remaining 52 use cases, involving data transmission via the network communication path, include many that are used only in exceptional situations and do not occur in everyday operation.

Each connection between a smart meter and a server uses a different encryption key. This can be used to assign packets to use cases as there is only one use case at a time executed between a specific smart meter and the server. Each use case must first establish a connection between the server and the smart meter by executing a handshake consisting of a TCP packet "hello" sent from the server to the smart meter followed by an acknowledgement packet from the smart meter to the server. Usually a use case ends by a packet terminating the connection (there is only one exception to this rule – the "reset"-use case). Between establishing and terminating the connection the number of packets sent is always greater than one and less than 21.

4 PACKET CLASSIFICATION

The general idea is to carry out a series of defined use cases in a lab environment mimicking a real-life smart meter network traffic. This results for each of the defined use cases in a set of network packets pertaining to this use case. The exploration of features extracted for each use case offers a profound insight into its operational intricacies. With the help of a data analyzing tool (the Python library SK-Learn) possible features of the packets can be calculated, limited solely to the metadata (as the packets are encrypted). Overall 29 features were extracted. Each recorded use-case cycle of the smart meter network is represented by a time series. In its current form, the acquired data is challenging to process using a classification algorithm due to factors such as the vast amount of data and numerous identical values. To mitigate this, automatic feature extraction (with the feature-importance-function of the tool used) was employed to reduce the dimensionality of the data while ensuring sufficient accuracy. Based on the extracted features from the recorded data (only metadata was used due to encryption), distinct differences between the use cases were discernible. Following fine-tuning and weighting of these features, the set was reduced to just 23 (per use case). Table 1 shows these 23 features. A subset of the final extracted features, weighted according to their correlation or stability, is illustrated in Fig. 1.

These features were used to define the normal packet distribution for each use case. In a next step network data from the real-life smart meter network was used to finetune the characterization of the use case – packet relationship by features. Features and feature values (lower and upper bounds) are used to formulate conditions that describe every use case uniquely. The set of these conditions constitutes a model for admissible network traffic: only a packet with feature values fitting into the model is admissible.

After successfully extracting these features, they were used to assign packets to use cases. The fundamental metrics of 'Count', 'Standard Deviation', 'Arithmetic Mean', 'Minimum' and 'Maximum' values, and 'Sum' lay the groundwork for comprehending the data flow dynamics within the network. These metrics provide a quantitative overview of packet and byte transmission, offering insights into the volume and variability of data exchange. Further granularity is achieved through the analysis of specific features such as 'PKT_CNT', 'BYTES_CNT', and 'TCP_LEN'. These metrics delve into the intricate details of packet and byte counts, as well as the length of TCP payloads, crucial for assessing network efficiency and throughput. Temporal dynamics play a pivotal role in network performance, as evidenced by the 'TIME_DELTA' feature. This metric encapsulates the temporal gaps between network activities, facilitating the identification of latency issues and assessing the duration of use cases. By quantifying the intervals between network events, it enables a thorough examination of the time taken for each operational phase, aiding in the optimization of network responsiveness and the refinement of use case durations. Bidirectional communication is fundamental to smart meter networks, reflected in features like 'TX_PKT_CNT', 'RX_PKT_CNT', 'TX_BYTES', and 'RX_BYTES'. These metrics gauge the volume and directionality of data transmission, essential for assessing network reliability and data integrity. In summary, the features extracted from the smart meter network serve as indispensable assets for reducing dimensionality and gaining insights for each use case.

This model is used for analyzing real-life network traffic and for detecting anomalies. During the monitoring phase the feature values of the packets are measured and compared to the patterns defined for the use cases. In order to check, whether the packet fits into one of the admissible use cases, the distance of the combined feature values to the use case classes is computed and compared to the threshold value. If it lies within the threshold of one of the allowed use cases the packet can be assigned to this use case. With the help of this method packets can be classified with respect to use cases. Packets that do not fit into the defined use cases are categorized as anomalies and reported to the operations personnel for further action. Moreover, sensible use cases (such as turning off the power supply for a smart meter) can be re-checked with the protocol data of the server. This will identify use cases not initiated by the power supply company – which are of course highly suspicious and should lead to an alert, too.

Table 1: Feature list

Feature	Explanation
stdev(IP_LEN)	Standard deviation of the length of the IP-packet header
stdev(TCP_LEN)	Standard deviation of the length of the TCP-Paket packet header
stdev(TCP_PAYLOAD_LEN)	Standard deviation of the length of the payload of the TCP-packet
avg(TIME_DELTA)	Average time delay between 2 packets
avg(IP_LEN)	Average length of the IP-packet header
avg(TCP_LEN)	Average length of the TCP-packet header
avg(TCP_PAYLOAD_LEN)	Average length of the payload of the TCP-packets
avg(TX_BYTES)	Average number of bytes transmitted
avg(RX_BYTES)	Average number of bytes received
min(IP_LEN)	Minimum length of the IP-packet header
max(IP_LEN)	Maximum length of the IP-packet header
min(TCP_PAYLOAD_LEN)	Minimum length of the payload of the TCP-packets
max(TCP_PAYLOAD_LEN)	Maximum length of the payload of the TCP-packets
min(TCP_LEN)	Minimum length of the TCP-Paket Header
max(TCP_LEN)	Maximum length of the TCP-Paket Header
sum(TX_PKT_CNT)	Sum of all packets sent
sum(RX_PKT_CNT)	Sum of all packets received
sum(RX_BYTES)	Sum of bytes received
sum(TX_BYTES)	Sum of bytes sent
sum(TIME_DELTA)	Sum of the time delays between 2 packets
sum(TCP_PAYLOAD_LEN)	Sum of the length of the payloads of the TCP-packets
pkt_Cnt	Number of the packets of the use case (TX + RX)
bytes_Cnt	Number of the bytes of the use case (TX + RX)

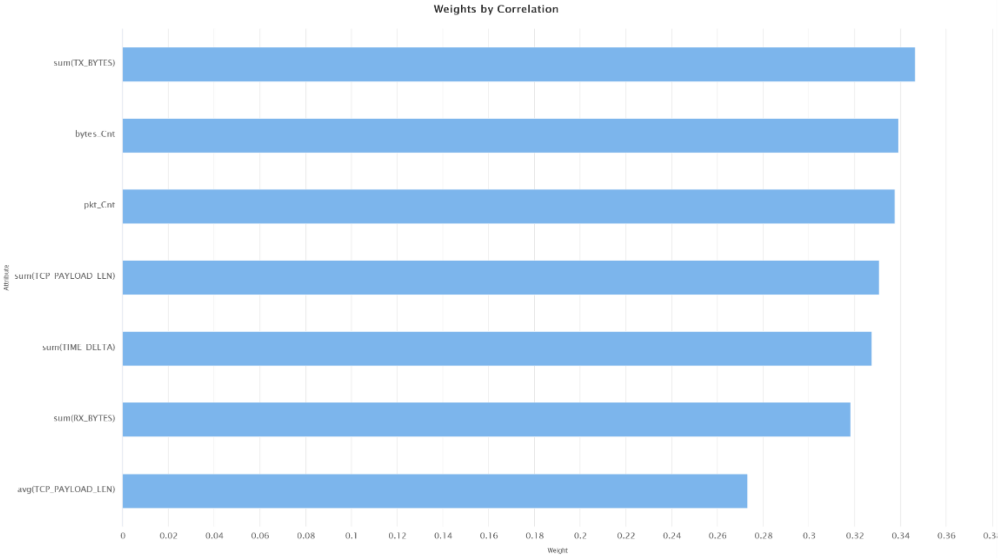


Figure 1: Subset of features with most influence for prediction

5 PRACTICAL APPLICATION

The experiments were carried out in the smart meter infrastructure of a middle-sized Austrian city (Wels, 65.000 inhabitants). The configuration of the network is shown in Fig. 2: Smart meter infrastructure.

From the use cases mentioned in [20] only a small subset is used in everyday communication. Some of the use cases – as mentioned above – are not involved in network activities and hence can be neglected. From the remaining use cases only 7 are used in day-to-day operation of the environment. Thus, only those were subject to the investigations in the real-world environment:

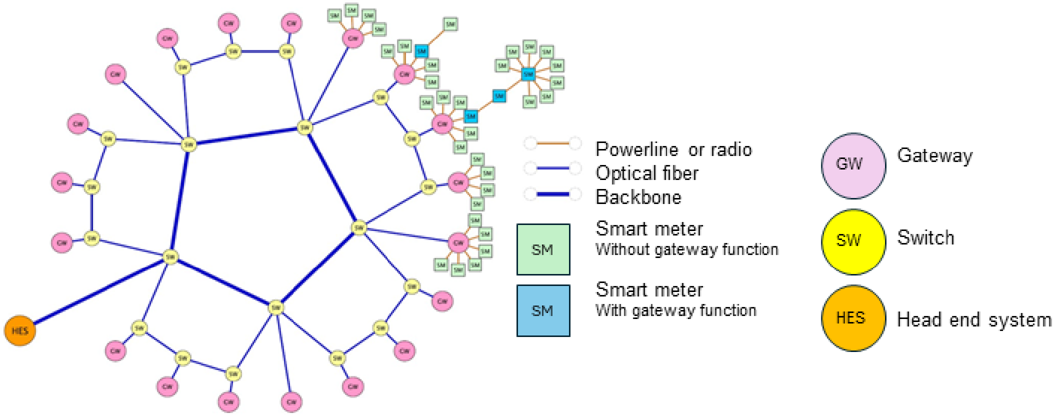


Figure 2: Smart meter infrastructure

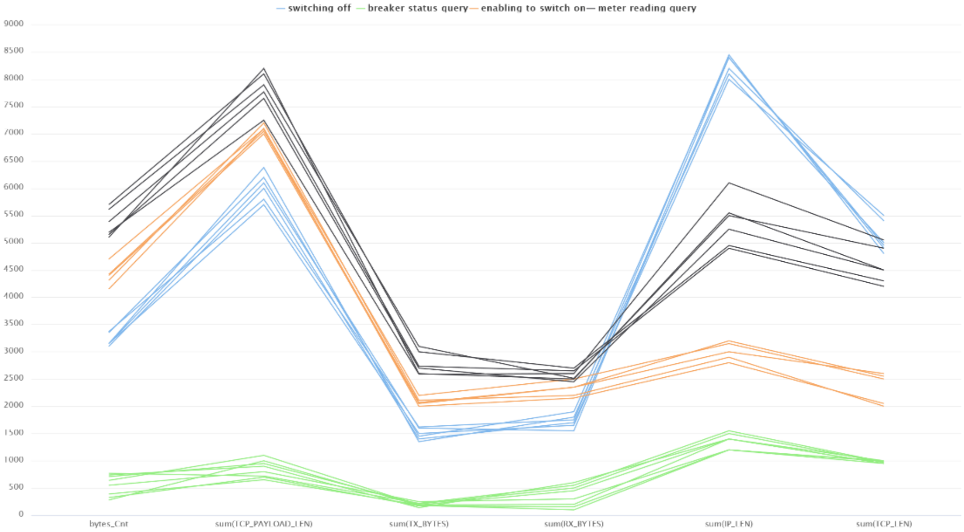


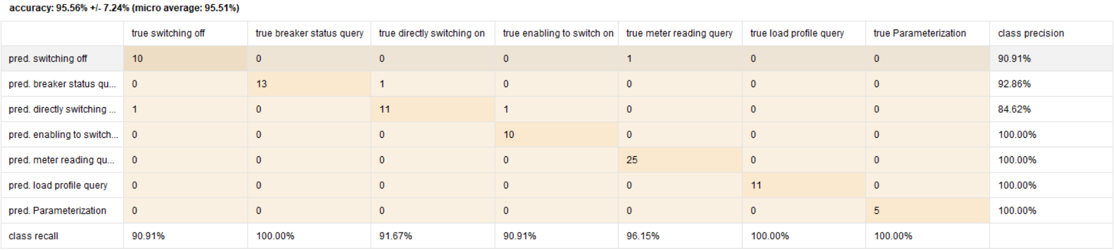
Figure 3: Feature differences between the individual use cases

1. Switching off the load switching device
2. Enabling the load switching device to switch on
3. Direct switching-on of the load switching device
4. Event query
5. Load profile query
6. Parameterization
7. Meter reading query

The classification method obtained by statistical machine learning – as described above – was applied to data from the real-world smart meter network. The data was provided by the energy supplier of the city and consisted of several recordings, each showing between 6 and 12 consecutive hours. The analysis showed clear differences in the feature values of the various use cases – see Fig. 3. The assignment of the packets to use cases was checked by comparing it to the log files of the system.

By standard methods of statistical machine learning a threshold (maximum distance of the feature set from the center of the use case) for defining a use case was calculated (by means of the tool Rapid Miner Studio) and resulted in a value of 0.62. Every set of values exceeding this threshold was considered an outlier – meaning that the data did not pertain to one of the defined use cases. The confusion matrix (Fig. 4) shows the reliability of the method.

Such outliers in the context of smart meter use case data indicate severe anomalies, which may have significant implications concerning network security. For instance, unusual patterns or unexpected spikes in features and requests, such as repeated attempts to switch off or tamper with load switching devices, could indicate potential security breaches or malicious activities. Identifying these outliers is crucial for promptly detecting and mitigating security threats to the smart meter network. Moreover, these outliers may also signal



accuracy: 95.56% +/- 7.24% (micro average: 95.51%)

	true switching off	true breaker status query	true directly switching on	true enabling to switch on	true meter reading query	true load profile query	true Parameterization	class precision
pred. switching off	10	0	0	0	1	0	0	90.91%
pred. breaker status qu...	0	13	1	0	0	0	0	92.86%
pred. directly switching ...	1	0	11	1	0	0	0	84.62%
pred. enabling to switch...	0	0	0	10	0	0	0	100.00%
pred. meter reading qu...	0	0	0	0	25	0	0	100.00%
pred. load profile query	0	0	0	0	0	11	0	100.00%
pred. Parameterization	0	0	0	0	0	0	5	100.00%
class recall	90.91%	100.00%	91.67%	90.91%	96.15%	100.00%	100.00%	

Figure 4: Confusion Matrix

operational anomalies or errors, which could impact the reliability and integrity of the network. For example, irregularities in event queries, load profile queries, or meter reading queries may indicate communication issues, device malfunctions, or unauthorized access attempts.

The system was implemented to show its viability within the context of the smart meter infrastructure of the city in question.

6 CONCLUSION AND FUTURE WORK

The work showed clearly that in a smart meter network a classification of use cases solely based on metadata is possible with high accuracy. Thus, for identifying use cases in smart meter networks with encrypted traffic it is not necessary to decrypt the network traffic. The identification of use cases can be used for detecting anomalies not only in the network traffic itself, but in the way the network is used, too. Hence it constitutes an important building block in ensuring security and safety of smart meter infrastructures. By avoiding the necessity of decryption such anomaly detection can be executed with much less overhead and may be introduced in any part of the network infrastructure without having to distribute decryption keys and algorithms.

Further possible applications of this analysis of real-world data from a smart meter network using the extracted features, could be predictive maintenance and fault detection in the power grid. By analyzing the patterns and trends in the data, such as fluctuations in power consumption, irregularities in network traffic, or anomalies in voltage levels, utility companies can anticipate potential equipment failures or malfunctions before they occur. This proactive approach to maintenance not only minimizes downtime but also helps prevent costly damages and ensures the reliability and efficiency of the power distribution system.

ACKNOWLEDGMENTS

This research was funded in whole, or in part, by the program KIRAS of the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology of the Republic of Austria, grant number 867029.

REFERENCES

[1] Kush, N., Foo, E., Ahmed, E., Ahmed, I., & Clark, A. (2011). Gap analysis of intrusion detection in smart grids. Proceedings of the 2nd International Cyber Resilience Conference.

[2] Argandeh, R., Meier, A. v., Mehrmanesh, L., & Mili, L. (2016). On the definition of cyber-physical resilience in power systems. Renewable and Sustainable Energy Reviews, 58, 1060-1069.

[3] Radoglou-Grammatikis, P.I., Sarigiannidis P.G. (2019). Securing the Smart Grid: A Comprehensive Compilation of Intrusion Detection and Prevention Systems. IEEE Access, vol. 7, pp. 46595-46620, 2019, doi: 10.1109/ACCESS.2019.2909807.

[4] Gungor, C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., & Hancke, G. (2013). A Survey on Smart Grid Potential Applications and Communication Requirements. IEEE Transactions on Industrial Informatics, 9/1, 28-42. doi:10.1109/TII.2012.2218253

[5] Sharmwey A. W. (2023). Safeguarding the future: A comprehensive analysis of security measures for smart grids. World Journal of Advanced Research and Reviews, 19(01), 847–871. Doi: 10.30574/wjarr.2023.19.1.1387

[6] Gunduz M.Z., Das R. (2020). Cyber-security on smart grid: Threats and potential solutions, Computer Networks, Volume 169, 107094, ISSN 1389-1286, https://doi.org/10.1016/j.comnet.2019.107094

[7] Kursawe, K., & Peters, C. (2015). Structural Weaknesses in the Open Smart Grid Protocol. 10th International Conference on Availability, Reliability and Security, (S. 1-10). doi:10.1109/ARES.2015.67

[8] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A Survey. ACM Computing Surveys, 41/3, 1-58. doi:https://doi.org/10.1145/1541880.1541882

[9] Zhang, Y., Wang, L., Sun, W., Green II, R., & Alam, M. (2011). Distributed Intrusion Detection System in a Multi-Layer Network Architecture of Smart Grids. IEEE Transactions on Smart Grid, 2/4, 796-808. doi:10.1109/TSG.2011.2159818

[10] Moghaddass, R., & Wang, J. (2018). A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data. IEEE Transactions on Smart Grid, 9/6, 5820-5830. doi:10.1109/TSG.2017.2697440

[11] Mitchell, R., & Chen, I. (2013). Behavior-Rule Based Intrusion Detection Systems for Safety Critical Smart Grid Applications. IEEE Transactions on Smart Grid, 4/3, 1254-1263. doi:10.1109/TSG.2013.2258948

[12] Wei, D., Lu, Y., Jafari, M., Skare, P., & Rohde, K. (2010). An Integrated Security System of Protecting Smart Grid against Cyber Attacks. Innovative Smart Grid Technologies (ISGT), (S. 1-7). doi:10.1109/ISGT.2010.5434767

[13] Jokar, P. (2012). Model-based Intrusion Detection for Home Area Networks in Smart Grids. University of British Columbia

[14] Andrysiak, T., Saganowski, Ł., & Kiedrowski, P. (2017). Anomaly Detection in Smart Metering Infrastructure with the Use of Time Series Analysis. Journal of Sensors, 2017/8782131. doi:https://doi.org/10.1155/2017/8782131

[15] Düssel, P., Gehl, C., L. P., Bußer, J., S. C., & K. J. (2009). Cyber-Critical Infrastructure Protection Using Real-time Payload-based Anomaly Detection. Critical Information Infrastructures Security. CRITIS 2009. Springer. doi:https://doi.org/10.1007/978-3-642-14379-3_8

[16] Chen, X., Yu, J., Ye, F., & Wang, P. (2018). A hierarchical approach to encrypted data packet classification in smart home gateways. IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, (S. 41-45). doi:10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00022

[17] Zhang, Y., Zhao, S., Zhang, J., Ma, X., & . Huang, F. (2019). STNN: A novel TLS/SSL encrypted traffic classification system based on stereo transform neural network. IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS). doi:10.1109/ICPADS47876.2019.00133

[18] Sun, C.C., Sebastian-Cardenas, D.J., Hahn, A., Liu, C.C. (2021). Intrusion Detection for Cybersecurity of Smart Meters. IEEE Transactions on Smart Grid, vol. 12, no. 1, pp. 612-622, doi: 10.1109/TSG.2020.3010230N3

[19] Ravinder, M., Kulkarni, Vikram (2023). Intrusion detection in smart meters data using machine learning algorithms: A research report. Frontiers in Energy Research, 11, 1147431. https://doi.org/10.3389/fenrg.2023.1147431

[20] Oesterreichsenergie (2015). Smart Metering Use Cases für das Advanced Meter Communication System (AMCS), Version 1.1. Oesterreichsenergie