Automated microsegmentation for lateral movement prevention in industrial Internet of Things (IIot).

ARIFEEN, M., PETROVSKI, A. and PETROVSKI, S.

2021

Automated Microsegmentation for Lateral Movement Prevention in Industrial Internet of Things (IIoT)

1st Murshedul Arifeen
School of Computing
Robert Gordon University
Aberdeen, Scotland
d.arifeen@rgu.ac.uk

2nd Andrei Petrovski
School of Computing
Robert Gordon University
Aberdeen, Scotland
a.petrovski@rgu.ac.uk

3rd Sergei Petrovski
School of Electric Stations
Samara State Technical University
Samara, Russian Federation
petrovski.sv@samgtu.ru

***Abstract*—The integration of the IoT network with the Operational Technology (OT) network is increasing rapidly. However, this incorporation of IoT devices into the OT network makes the industrial control system vulnerable to various cyber threats. Hacking an IoT device at the network edge, an attacker can move laterally to compromise the main control server and manipulate the whole control system of the industrial infrastructure. In this paper, we have proposed an automated Micro-segmentation (MS) model based on Machine Learning (ML) algorithms to reduce the lateral movement of an attacker or malware. The proposed model generates the micro-segments based on network traffic and blocks the malicious traffic at each segment. We have taken UNSW-NB15 and IoTID20 datasets for our experiments. Experimental results show that after generating micro-segments and separating the normal traffic, the model limits redundant links and blocks malicious traffic. Limiting the usage of redundant links reduces the lateral movement or spreading of malware. We also considered the deterministic epidemic model to analyze the device infection rate due to lateral movement or malware propagation.**

***Index Terms*—Internet of Things, Micro-segmentation, Security, Lateral Movement, Machine learning**

I. Introduction

Information technology (IT) is the application of computers, networking devices, communication technologies for collecting, processing, storing, and communicating digital data [1]. On the contrary, OT involves industrial infrastructures, which use SCADA or control system networks for direct monitoring and controlling the industrial equipment [2]. Unlike the IT, the OT network, which includes devices like Programmable Logic Controllers (PLC), has power issues, slower processing capability, low memory and a much longer upgrade cycle [3]. The integration of the Industrial Internet of Things (IIoT) in the industrial manufacturing environment converges the IT and OT networks.The convergence of IT and OT offers various benefits including improved safety, increased productivity, efficiency, and predictive maintenance [4].

Along with these benefits, the convergence of IT and OT networks faces severe security risks. Due to the connection with the IIoT network, the OT network becomes accessible throughout the Internet [5]. Moreover, the OT devices like PLCs or other controlling devices were not designed with the consideration of security vulnerabilities [5]. An attacker can gain access to the OT network by bypassing the IoT network using lateral movement. Lateral movement or east-west traffic enables an attacker to compromise the entire network, including internal servers and other devices [6]. This compromisation of controlling devices may result in massive damage in the industrial domain. For instance, an attacker took control over the main server of Oldsmar's water treatment plant, Florida, the USA, in February 2021 [7]. By taking control at the water treatment plant, the attacker abnormally increased the amount of NaOH in the water, which may cause vision problems, pain, shock if consumed. Securing the IoT devices may prevent lateral movement.

However, the IoT network is vulnerable to various security threats [8], and these vulnerabilities create loopholes for lateral movement. Moreover, replacing the cloud network with edge devices cut the centralized control over the IoT devices. An attacker can hack or snitch the IoT devices at the network edge and inject malware. Without special security measures, any device in the network can access any other device like in Mesh topology [9], which enables the malware to reach anywhere in the network. This malware enables an attacker to compromise the internal servers. Therefore, securing the IoT network for preventing lateral movement has become indispensable. But according to a survey, 99% of security professionals are struggling to secure the IoT devices and facing challenges to update security patches using firmware update [10].

Network MS is a promising way to prevent lateral movement throughout the IoT network. MS prevents lateral movement and reduces the attack surface by splitting a large network into several smaller network segments [11]. Then, the access control of each device in a micro-segment is restricted within the segment perimeter by imposing specific security rules. Therefore, the devices within a micro-segment cannot communicate with other devices outside of its restricted perimeter. Restricting the access can confine a malware or

an attacker within the segment and reduce further movement outside the compromised device's segment. Although MS is widely applied to secure the cloud and workloads of servers [11], it is challenging for the IoT networks due to several reasons. First of all, the IoT network is large and dynamic, which creates difficulty in identifying proper segments. Secondly, it is difficult to maintain and update a large number of micro-segments with the security rules periodically. Intelligent algorithms can be used to overcome these tedious jobs of maintaining MS and security policies for the IoT networks.

In this work, we have proposed an automated MS procedure and security rules generation for each segment based on ML algorithms. The micro-segments are generated through the OPTICS clustering algorithm. Then a Decision Tree (DT) classification algorithm is used to separate the malicious network traffic from the legitimate traffic data. These traffic data are then used to generate packet filtering policy.

In section II we have discussed the related works. Section III presents the system model including, the network model, threat model, and proposed MS process. Section IV demonstrates the experimental results. In section V we have analyzed the security enhancement by MS, and finally, section VI concludes this paper with future works.

II. Literature Review

A very few research works have been conducted for preventing lateral movement in IoT network domain - some related studies are discussed in this section. The authors in [6] proposed a micro-segmentation technique based on edge cloud architecture for smart home IoT networks, using Open flow rules. The proposed model blocks attackers from accessing the LAN and WAN of the smart home IoT network. However, the open flow rules are static and need to be updated manually. Also, the approach applied for smart homes is not suitable for large scale and dynamic IIoT networks.

The authors in [12] proposed an evidence reasoning lateral movement detection technique for the cloud-edge environment. The authors also introduced vulnerability correlation process in lateral movement detection. However, this model is not appropriate for networks which replace the cloud architecture with only edge computing devices.

A micro-segmentation technique is proposed in [13] based on K-means clustering algorithm for enterprise network. However, it is required to define the number of clusters initially for the K-means algorithm, which is not effective for a large scale network like industrial IoT or other sensor networks.

The MITRE ATT&CK framework also takes into consideration lateral movements. MITRE ATT&CK can be defined as the set of individual techniques performed by an attacker to accomplish malicious tasks. It was shown in [14] that MITRE ATT&CK encompasses 440 attack techniques belonging to 27 different tactics. These malicious activities may include gaining access to the IoT network through the use of phishing links that may compromise other devices through lateral movements. Furthermore, a public repository (referred to as the MITRE ATT&CK Framework) is available which contains

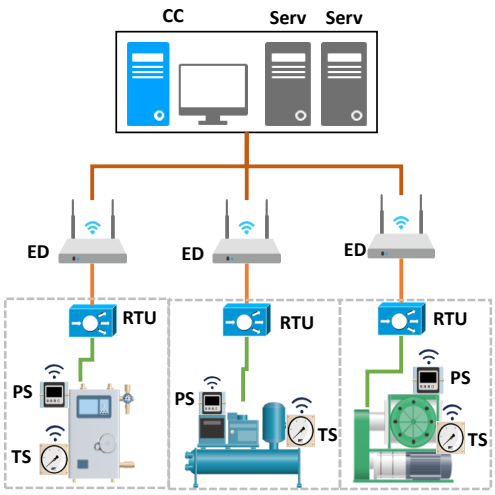

Fig. 1: Edge enabled IIoT network where different types of sensors are connected with the industrial machines. The elements of the network are PS– Pressure sensor, TS– Temperature sensor, RTU– Remote terminal unit, ED– Edge device, CC– Control centre and Serv– Server.

adversary tactics, techniques and procedures based on real-world observations [15]. This publicly available knowledge base provides a rich resource for the development of specific attack detection, prediction and mitigation models.

III. System Model

A. Network Model

In the traditional OT network like SCADA, all the data are collected and analyzed in the centralized server. However, the IIoT network improves the SCADA network by introducing edge devices at the network edge. Figure 1 shows the IIoT and edge enabled SCADA network, where edge devices are connected with the RTUs. These edge devices then receive data from the sensors, which are connected with the industrial equipment. After receiving data, the edge devices process and provide a real-time decision for maintaining the industrial machinery. An administrator can control the whole network from the control centre and send commands through the RTUs. Also, the data are stored in the central servers for future analysis and optimization [16]. This integration of IIoT and edge devices enable the administrators to monitor and control the industrial control system remotely.

B. Threat Model

In this work, we considered the threat due to lateral movement by an attacker. Advanced Persistent Threats (APT) [17] are severe and long-lasting cyber attacks, where lateral movement is an attack phase in which the attacker moves from the compromised devices to other devices [18] [19]. APT can be defined as the theft of intellectual property or espionage as opposed to achieving immediate financial gain and are

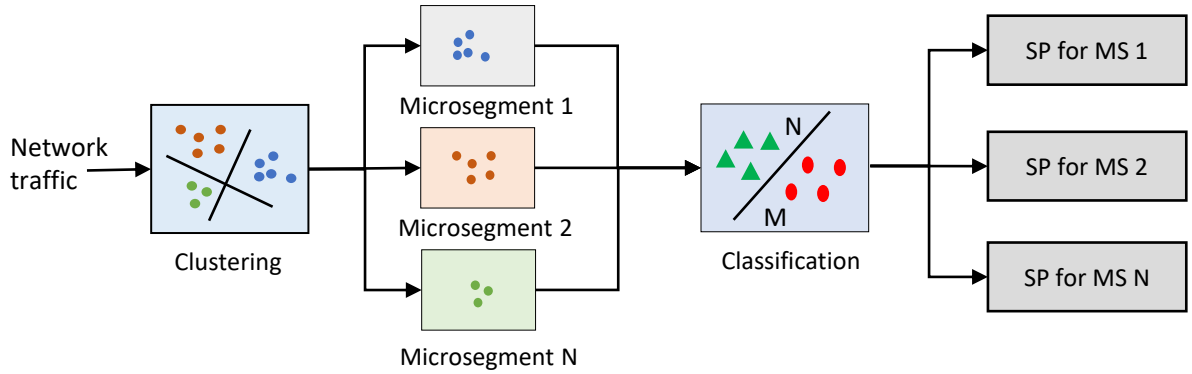

Fig. 2: Proposed Micro-segmentation generation model based on ML where, the clusters or micro-segments are generated through OPTICS algorithm and DT is used for traffic data classification. Security policies (SP) are used to restrict the redundant links.

prolonged, stealthy attacks [20] [21] . For taking control of the main server of the industrial control system, the attacker moves deeper inside the network after hacking an IoT device. Therefore, the attacker can gradually compromise the whole network. This compromisation may result in a devastating situation. Moreover, an internal employee may intentionally try to compromise a device and achieve a malicious goal.

C. Background on ML algorithm

1) OPTICS Clustering algorithm: OPTICS is the upgraded version of the DBSCAN algorithm. It was demonstrated in [22] that DBSCAN performs well in clustering network traffic compared to other models. However, unlike DBSCAN, OPTICS is better suited for large scale dataset [23] and do not require the epsilon parameter (the domain knowledge). For these reasons, here we chose the OPTICS clustering algorithm.

2) DT algorithm: A DT is a supervised classification technique that includes internal nodes, which represent the features of the traffic data (for instance, IP address, Flow ID); branches represent the decision rules, and the leaf nodes represent the outcomes (Malicious or Normal). This algorithm uses various feature selection measures like information gain or Gini index to select the best features as the root node or the internal nodes. Information gain (IG) can be defined as in equation (1) [24], which tells us how much a feature provides information about a class.

$$Gain(S, A) = E(S) - \sum_{i}^{n} \frac{|s_i|}{|S|} E(S_i) \tag{1}$$

where, n = number of attributes A, $|Si|$ = number of cases in partition S_i, $|S|$ = total cases and E is the Entropy as defined below:

$$E(S) = -\sum_{i=1}^{c} p_i \log_2(p_i)$$

D. Microsegmentation

In this subsection, we will discuss the MS generation process using ML algorithms discussed in the previous subsection.

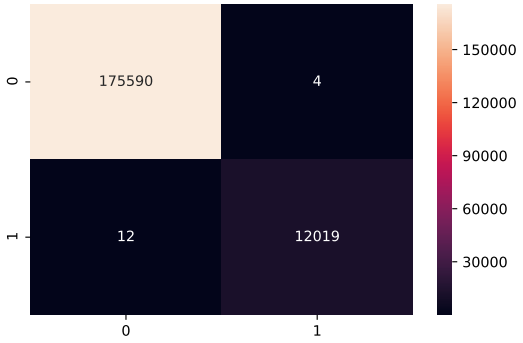

Fig. 3: Confusion matrix of IoTID20 dataset. Here, 0 denoted the Anomaly class and 1 denotes the Normal class

Figure 2 shows the proposed MS creation model based on ML algorithms. As shown in [11], MS implementation consists of several steps.

Firstly, we need to identify and group the devices which show similar functionalities or behavior. Here, we have chosen the similarity of traffic data to group the IoT devices through the OPTICS clustering algorithm. Each group of IoT devices will then work as a micro-segment.

After generating the groups of similar devices, the traffic information of each group of devices will be classified as malicious or normal for creating the security policies. For classification tasks, we have considered the DT classifier algorithm. After classifying, the algorithm will look for multiple connection of each IoT device and restrict the access of redundant links except one link for each IoT nodes. Upon failure of the current link, the algorithm will make one of the restricted link available for use. This will result in blocking the malicious traffics as well.

IV. Experiments

The experimental environment comprises of Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz, 16GB

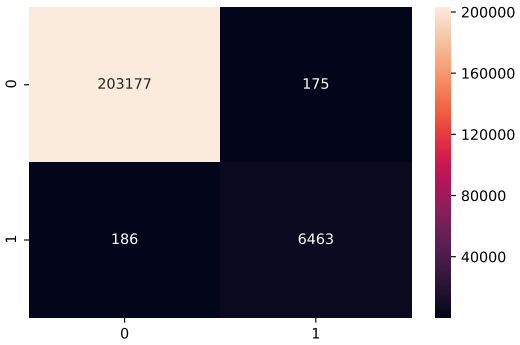

Fig. 4: Confusion matrix of UNSW dataset. Here, 0 denoted the Normal class and 1 denotes the Anomaly class

RAM and Windows 10 OS. We have considered Python's Scikit learn ML library for importing the OPTICS clustering algorithm and the DT classifier. Also, other functions are called from the library for instance train test split function and StandardScalar() for normalizing data.

A. Dataset

For the experiment, we have taken the UNSW-NB15 [25] and IoTID20 [26] datasets. These datasets contain various features of network traffic, including anomalous and normal data. The UNSW-NB15 dataset contains 48 features of the network traffic. The last feature of this dataset is the class label that is either 0 for normal and 1 for malicious traffic. The IoTID20 dataset comprises 80 network features including, three class labels. The Normal and Anomaly class are subdivided based on various cyber attacks.

B. Data Pre-processing

Before applying the ML models on the datasets, we performed data preprocessing. First, the categorical features are encoded using LabelEncoder() function and normalized using StandardScaler() function. Both of the datasets are high dimensional. Therefore, we conducted a correlation analysis and found that 8 pairs of features are highly correlated with each other in the UNSW-NB15 dataset. However, among the 8 pairs of features, only ('swin', 'dwin'), ('Stime', 'Ltime') pairs of feature showed 100% correlation. Therefore, from UNSW-NB15 dataset 'swin' and 'Stime' features are dropped. On the contrary, from the IoTID20 dataset 21 highly correlated features which showed 100% correlation are dropped from the dataset. We choose the correlation threshold as 0.95.

To further reduce the dimensions of the datasets, we applied Principle Component Analysis (PCA). From PCA, we found it is sufficient to consider only the first 30 principal components to represent the overall information of the UNSW-NB15 dataset. For the IoTID20 dataset, the first 20 principal components are adequate. However, for the DT classifier, we did not conduct the PCA procedure.

C. Training and testing

The OPTICS clustering algorithm and the DT classifier are implemented using the pythons sci-kit learn library. We took 1000 samples from each dataset randomly to conduct OPTICS clustering operations since our experimental configuration fails to do clustering for the entire dataset. We set min_samples=2, max_eps=np.inf, metric='chebyshev', cluster_method ='xi' for the OPTICS clustering method's parameters. We found that for the 'chebyshev' distance metric the OPTICS yields good results.

On the other hand, for classification algorithm, our environment supported the entire dataset. We split the entire dataset to a $70:30$ ratio for training and testing the DT classifier.

D. Results

After performing the OPTICS clustering algorithm, we got 178 clusters (micro-segments) for UNSW-NB15 dataset and 295 clusters (micro-segments) for IoTID20 dataset based on the random 1000 samples of each dataset. Table I shows the clustering results.

TABLE I: Number of clusters generated for each dataset

Dataset	Number of Clusters
UNSW-NB15	178
IoTID20	295

TABLE II: Classification performance of DT over UNSW and IoTID20 dataset

Dataset	Accuracy	Sensitivity	Specificity
UNSW-NB15	99.82%	97.20%	99.91%
IoTID20	99.99%	99.90%	99.99%

After training and testing the DT classifier on both of the dataset, we have computed the Accuracy, Sensitivity and Specificity metrics. Table II shows the evaluation results. From this table, we can see that the DT classifier performed similarly on both datasets in terms of Accuracy and Specificity, but the Sensitivity for UNSW-NB dataset is slightly lower than the one for the IoTID20 dataset. Figures 3 and 4 show the confusion matrices of the DT classifier for IoTID20 dataset and UNSW-NB15 dataset respectively. Then, we have used this trained DT classifier to differentiate between the normal and malicious traffic in each cluster or micro-segment (as depicted in Figure 2).

Table III shows the security policies for a security group generated by a clustering algorithm. The MS model with the DT classifier will block the traffics generated outside of a security perimeter from entering into the micro network bestowed by that perimeter. Also, any malicious traffics will be blocked. From table III we can see that the IoT device with the IP address 149.171.126.6 is only allowed to communicate with 59.166.0.4. But other devices with IP addresses 59.166.0.1 and 59.166.0.5 are blocked from communicating with 149.171.126.6. Hence, a single device is restricted to access the redundant links (Section V explains in more detail). The malicious traffics will be blocked automatically.

TABLE III: Security rules for a Micro-segment

DID	SIP	SPort	DIP	DPort	Proto	Action
436047	59.166.0.3	10138	149.171.126.0	42769	udp	Block
31392	59.166.0.4	8515	149.171.126.6	53	udp	Allow
345291	59.166.0.7	11498	149.171.126.0	5190	tcp	Block
632042	59.166.0.1	9310	149.171.126.6	53	udp	Block
505483	59.166.0.5	24688	149.171.126.6	58616	tcp	Block
381348	59.166.0.8	50725	149.171.126.4	53	udp	Block

Legend: DID– Device ID; SIP– Source IP; SPort– Source Port; DIP– Destination IP; DPort– Destination Port; Proto– Protocol

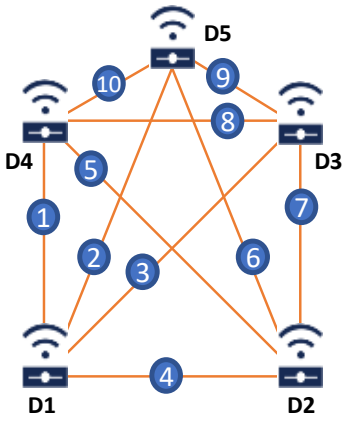

Fig. 5: A security group of IoT devices connected through mesh network

V. Security Analysis

If any restriction is not imposed explicitly, an IoT device can communicate with multiple other devices like in Mesh topology [9]. Therefore, multiple links help malware to spread more rapidly within a network. The attackers get more paths to move laterally within the network and compromise the devices. It is not acceptable to block the redundant links of an IoT device since IoT devices must communicate through other available links if the current link fails. However, we can control the number of links to reduce the spreading of malware through lateral movement. MS has the potential to minimize the spreading of malware over the network by imposing specific security policies. In this section, we will theoretically analyze the effectiveness of MS in terms of reducing malware dissemination. As an example, let us consider a segment of the IoT network shown in Figure 5, where $D5$ is the gateway node and $D1$, $D2$, $D3$ and $D4$ are the sensor nodes. The devices are connected in a Mesh topology. Without any specific security measures, the malware may spread through all the links. The number of links of this mesh topology is

$$^5C_2 = \frac{5(5-2)}{2} = 10 \tag{2}$$

Through MS, we can restrict the access to communicate with the gateway node $D5$. For instance, initially the device $D1$ could communicate with $D5$ through device $D4$, $D3$ and $D2$. But using MS, we can restrict node $D1$ from accessing $D5$ through $D4$, $D3$ and $D2$ except the direct link numbered 2. Therefore, MS restricts the access of the links numbered 1, 3 and 4 for device $D1$. Similarly, after restricting all the redundant links for other devices, the total number of allowed links in this security group will be reduced from 10 to 4.

Now, we can use the deterministic epidemic model to figure out the IoT device infection rate for MS and for without MS. The epidemic model can be defined as [27]-

$$\frac{dI(t)}{dt} = \beta I(t)[N - I(t)] \tag{3}$$

where β is the infection rate and it is constant for specific malware, N is the total number of devices, and $I(t)$ is the number of infected devices at time t. However, from the above analysis, we can see that the parameter β is proportional to the number of links in the IoT network for any malware. As the number of links increases, the probability of device infection also raises. Therefore, we have considered β as the link parameter. A close form equation of the epidemic model is also shown in [27] as,

$$I(t) = I(0)e^{\beta N t} \tag{4}$$

where, $I(0)$ is the number of devices infected at $t = 0$ unit of time. Figure 6 (Log plot) shows the device infection rate of the Mesh network shown in figure 5 with $I(0) = 1$, $t = 15$ time unit, $\beta = 10$ for without MS and $\beta = 4$ for with MS, and finally $N = 5$. We can see device infection rate is higher without applying MS than the infection rate after applying MS. Therefore, it is evident that, MS reduces device infection rate by declining lateral movements (at $t = 14$ almost 3 devices are infected without MS but only 2 devices are infected with MS). The device infection rate increases exponentially according to equation 4. Therefore, if we consider a large network instead of the simple Mesh network depicted in Figure 5, the difference between the two lines shown in Figure 6 will increase. After applying MS, it will take more time to move from the compromised device to the internal nodes. Therefore, the administrator will be able to identify and revoke the compromised devices before the attacker takes control of the main server or device.

VI. Conclusion

In this work, we have proposed an automated MS model based on the OPTICS clustering algorithm and a DT classifier for preventing lateral movement in IIoT. We have considered ML algorithms to automate the micro-segmentation process

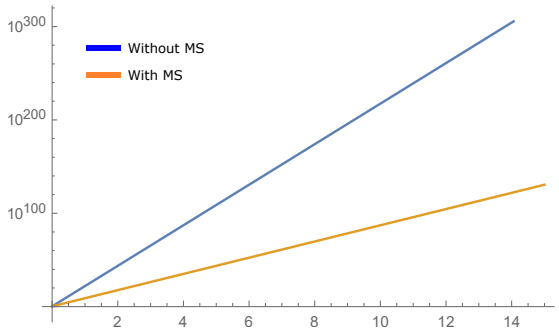

Fig. 6: Plot of equation (3). Where, the blue line shows the device infection rate before applying MS with $\beta = 10$ and the orange line shows the device infection rate after applying MS for $\beta = 4$.

since it is difficult and tedious to maintain micro-segmentation for large-scale IoT networks. We have considered the network traffic to find and group similar IoT devices using the OPTICS clustering algorithm. The IoT devices which produce similar traffic information can be grouped together. Then, we have trained a DT classifier and used the DT model obtained to separate the normal traffic from the malicious one. The model will restrict accessing the redundant links of each IoT device, which will reduce the spreading of malware. MS will also reduce the lateral movement of an attacker or malware over the entire IIoT network by imposing security rules. Furthermore, we have analyzed the effectiveness of MS in the IIoT network and showed MS reduces device infection rate.

However, in the security analysis section only a static Mesh topology of IoT devices is considered. In reality, the IoT network is more complex, heterogeneous, and dynamic. Therefore, in future work, we will apply statistical distribution for modeling the dynamic nature of large scale IIoT networks. Also, we intend to integrate a malware detection model with the MS process to identify and revoke the infected device before an extensive portion of the network becomes compromised through lateral movement. We also believe our work will open the door to further experiments of lateral movement prevention using ML in IoT networks.

REFERENCES

[1] R. Castagna and S. J. Bigelow, "What is information technology? definition and examples," Aug 2021. [Online]. Available: https://searchdatacenter.techtarget.com/definition/IT

[2] A. Hahn, "Operational technology and information technology in industrial control systems," in *Cyber-security of SCADA and other industrial control systems*. Springer, 2016, pp. 51–68.

[3] E. R. Alphonsus and M. O. Abdullah, "A review on the applications of programmable logic controllers (plcs)," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1185–1205, 2016.

[4] S. Kamal, S. Al Mubarak, B. Scodova, P. Naik, P. Flichy, G. Coffin *et al.*, "It and ot convergence-opportunities and challenges," in *SPE Intelligent Energy International Conference and Exhibition*. Society of Petroleum Engineers, 2016.

[5] S. Cheruvu, A. Kumar, N. Smith, and D. M. Wheeler, *Demystifying internet of things security: successful iot device/edge and platform security deployment*. Springer Nature, 2020.

[6] A. Osman, A. Wasicek, S. Köpsell, and T. Strufe, "Transparent microsegmentation in smart home iot networks," in *3rd {USENIX} Workshop on Hot Topics in Edge Computing (HotEdge 20)*, 2020.

[7] D. Bisson, "Poison in the water: The physical repercussions of iot security threats," Jun 2021. [Online]. Available: https://securityintelligence.com/articles/florida-water-supply-poison-iot-security/

[8] F. Majdani, L. Batik, A. Petrovski, and S. Petrovski, "Detecting malicious signal manipulation in smart grids using intelligent analysis of contextual data," in *Proceedings of the 13th International Conference on Security of Information and Networks*, 2020, pp. 1–8.

[9] Y. Liu, K.-F. Tong, X. Qiu, Y. Liu, and X. Ding, "Wireless mesh networks in iot networks," in *2017 International workshop on electromagnetics: applications and student innovation competition*. IEEE, 2017, pp. 183–185.

[10] J. Glackin, "Survey: 99% of security pros struggling to secure their iot and iiot devices," Mar 2021. [Online]. Available: https://www.securityinfowatch.com/cybersecurity/information-security/article/21216554/tripwire-survey-99-of-security-pros-struggling-to-secure-their-iot-and-iiot-devices

[11] D. Klein, "Micro-segmentation: securing complex cloud environments," *Network Security*, vol. 2019, no. 3, pp. 6–10, 2019.

[12] Z. Tian, W. Shi, Y. Wang, C. Zhu, X. Du, S. Su, Y. Sun, and N. Guizani, "Real-time lateral movement detection based on evidence reasoning network for edge computing environment," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4285–4294, 2019.

[13] M. Yousefi-Azar, M.-A. Kaafar, and A. Walker, "Unsupervised learning for security of enterprise networks by micro-segmentation," *arXiv preprint arXiv:2003.11231*, 2020.

[14] R. Al-Shaer, J. M. Spring, and E. Christou, "Learning the associations of mitre att&ck adversarial techniques," *arXiv preprint arXiv:2005.01654*, 2020.

[15] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "Mitre att&ck: Design and philosophy," *Technical report*, 2018.

[16] U. Otokwala, A. Petrovski, and H. Kalutarage, "Effective detection of cyber-attack in a cyber-physical power grid system," *Advances in Intelligent System and Computing, Kohei Arai (Ed.): Proceedings of the Future of Information and Communications conference*, vol. 1, 2021.

[17] H. N. Eke, A. Petrovski, and H. Ahriz, "The use of machine learning algorithms for detecting advanced persistent threats," in *Proceedings of the 12th International Conference on Security of Information and Networks*, 2019, pp. 1–8.

[18] A. Fawaz, A. Bohara, C. Cheh, and W. H. Sanders, "Lateral movement detection using distributed data fusion," in *2016 IEEE 35th Symposium on Reliable Distributed Systems (SRDS)*. IEEE, 2016, pp. 21–30.

[19] T. Bai, H. Bian, M. A. Salahuddin, A. Abou Daya, N. Limam, and R. Boutaba, "Rdp-based lateral movement detection using machine learning," *Computer Communications*, vol. 165, pp. 9–19, 2021.

[20] C. Five, "Advanced persistent threats: A decade in review," *Command Five PTY LTD*, pp. 1–13, 2011.

[21] ISACA, "Advanced persistent threat awareness study results: Information systems auditing manual," *Information Systems Audit and Control Association*, 2014, accessed on 2019-02-20.

[22] J. Erman, M. Arlitt, and A. Mahanti, "Traffic classification using clustering algorithms," in *Proceedings of the 2006 SIGCOMM workshop on Mining network data*, 2006, pp. 281–286.

[23] "sklearn.cluster.optics." [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.OPTICS.html

[24] P. Gulati, A. Sharma, and M. Gupta, "Theoretical study of decision tree algorithms to identify pivotal factors for performance improvement: A review," *International Journal of Computer Applications*, vol. 141, no. 14, 2016.

[25] N. Moustafa and J. Slay, "Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set)," in *2015 military communications and information systems conference (MilCIS)*. IEEE, 2015, pp. 1–6.

[26] I. Ullah and Q. H. Mahmoud, "A scheme for generating a dataset for anomalous activity detection in iot networks," in *Canadian Conference on Artificial Intelligence*. Springer, 2020, pp. 508–520.

[27] S. Yu, G. Gu, A. Barnawi, S. Guo, and I. Stojmenovic, "Malware propagation in large-scale networks," *IEEE Transactions on Knowledge and data engineering*, vol. 27, no. 1, pp. 170–179, 2014.