Defense Scheme to Protect IoT from Cyber Attacks using AI Principles

T. Ahamed

Tariq Ahamed Ahanger*

College of Computer Engineering & Sciences, Prince Sattam Bin Abdulaziz University, KSA *Corresponding author: t.ahanger@psau.edu.sa

Abstract: Even in its infancy, the internet of things (IoT) has enticed most of the modern industrial areas like smart cities, automobiles, medical technology. Since IoT connects everything together, it is vulnerable to a variety of devastating intrusion attacks. Being the internet of different devices makes it easy for attackers to launch their attacks. Thus, to combat all these attacks, an attack analysis is presented in this article using the basic principles of Artificial Neural Networks. Internet packet traces are used to train to the supervised ANN (Multilevel Perceptron) and evaluated after the training to decline the DDoS Attacks. This research article mainly focuses on the categorization of traffic patterns into legitimate traffic and attack traffic patterns in IoT network. The ANN processes are evaluated and tested in a simulated IoT network. The experimental results show a greater accuracy in detection of various DDoS attacks.

Keywords: ANN, IoT, DDoS, Security, IDS, AI.

1 Introduction

In the current world more objects are associated with the Internet than the people. And this difference will pursue to grow further, as the objects increases in their capability to directly interact with the Internet or evolves into the physical objects of data which can be accessible through the Internet systems. This scenario approaching towards greater independence in object interaction with the Internet is generally interpreted as the (IoT) Internet of Things [4,26]. The greater number of devices which avails the internet services is growing faster and to have them connected every time by wire or otherwise will establish a mighty source of data information at anyone's finger tips [17]. To enable communication between smart machines is an advanced technology, but the technologies which compose the IoT are not fresh for us. IoT, as understood by its name, is the method of converting data to any virtual-platform on current internet infra, which is retrieved from diverse class of things [3]. The main concept of IoT is to grant independent exchange of important information between unseen embedded distinctly identifiable devices in the real world around us, overwhelmed by the dominant technologies like RFID (Radio-Frequency Identification) and WSNs (Wireless Sensor Networks) sensed by the sensor nodes and processed further for decision building [1], in order to perform an automated action.

2 Security threats

The complexity in the nature of IoT security rotates around the reality that, since, it is a great challenge to combine several technologies into one; the system tries to connect devices securely which have limited computation capability, storage, and power [2]. Few of the devices utilized by IoT can hold only a little basic mechanism of security measures, some of which are not capable to maintain the confidentiality and integrity of the users' information data [19]. Furthermore, the devices-for example, RFIDs, and sensors has inadequate user interfaces, such as an On-Off button or status signal, hence represents a limited psychological visual for the owners when it approaches towards trusting these type of devices [13]. Presently, privacy is considered as a major concern; which slows down the advancement of many upcoming technologies. Moreover, it has been proved that the technology which doesn't provide enough trust and exposes the individual's identity diminishes [7]; in the recent past, many new technologies failed to maintain adequate privacy and security mechanisms, which causes pain and tremendous sufferings to the distressed. For the sake of gaining trust of a common man in the IoT, the technologists need to make sure the same kind of failures with regard to security and privacy do not appear in the system, by safeguarding the required mechanisms to ensure such things existed from the beginning.

3 Sources of threats

Subsequently, discussing the IoT features and how they can be used in several scenarios, it is now time to discuss and to identify the potential threats confronting the communication mechanism in IoT. There are three primary entities which poses threats to the privacy and security in IoT:

- (i) Malicious User: It is the possessor of the IoT-device with power to carry out attacks to acquire the secrets of the device manufacturer, and also to acquire access to secret functionality. The Malicious user reveals the shortcomings in the system to retrieve information, selling secrets to 3 rd parties, even launches an attack on the systems [12].
- (ii) Bad Manufacturer: It is the builder of the device with the capability to explore the underlying technology to retrieve the information of the IoT devices, or users. Aforementioned, a manufacturer can intendedly introduce holes in the security design which can be exploited subsequently to access the user's secret data and also revealing it to 3rd parties. Similarly, the manufacturing of badly secured things results in the compromise of their users' privacy. Adding to that, in the context of IoT where different devices connect one another, a device manufacturer just to harm the reputation of their competitors; can attack their devices [6].
- (iii) External Adversary: It is an external entity that is not considered as a member of the system and has no authorization to access it. An adversary aims to gather the information about the users' of the system with the malicious intentions for example, creating financial losses and subverting the user's reliability. It also, causes flaw in the system by maneuvering the data sensed and transmitted [11].

4 Classes of attacks

To understand the risks attached, classification of attacks is essential to every system. We opted to focus on many categories to determine threats.

• Device Tampering: As we know IoT devices are smaller devices which are integrated in many other systems, for example; light, switches, TVs, cars, ovens, and many more. Few of the IoT devices remains unattended most of the time, therefore, they could be stolen easily irrespective being noticed by anyone. If a device got into the wrong hands, several types of attacks can be executed like software manipulation, hardware tampering, and secret stealing [20]. It's essential to indicate that an attacker could tamper with the vulnerable device and utilize it to inject fake data into the system, utilize the device for attacking purpose or deviate it to its expected functionality.

- Information Disclosure: It's an act of disclosing the information to an object which lacks the permission to access it. This comprises targeted attack, accidental exposure, and correlation or inference [18]. An adversary can retrieve information from the network links by eavesdropping, physically accessing the device.
- Privacy Breach: Contradictory to Information Disclosure, an attacker doesn't essentially require access to the secret information of the user to learn about him. The attacker could ascertain confidential information from several other sources like traffic analysis and meta data [5].
- Denial-of-Service: DoS is associated to the characteristic of being not accessible when an authorized user requests. The underlying system should be capable enough to continue operating even though some unintended action is being carried out by some adversary. DoS attacks can be carried out by device stealing, disrupting the communication links, manipulating the software [14].
- Spoofing: It is used to steal the credentials which belong to others so as to gain access to not accessible services. These credentials could be retrieved form a device directly, eavesdropping a channel of communication, or phishing [15].

5 Intrusion detection system

Any kind of unapproved or unauthorized activities in a network or a system are called intrusions. An IDS (Intrusion Detection System) is a group of the tools, mechanism, resource to identify, assessment which describes intrusions [8]. Intrusion detection is usually a part of a comprehensive protection system which is installed in a device or around a system and it can't be taken as a stand-alone measure of protection. Intrusion can be defined as: "any type of activities which tries to manipulate the truth, secrecy, or the resource availability" [9]. In the vice-versa intrusion prevention techniques are referred to as the first line of defense against these intrusions. Nevertheless, as in any type of security system, total prevention of intrusion is impossible. The node compromise and intrusion heads to secret information like security keys being disclosed to the intruders in the system, which results collapse of the security mechanism [10]. Consequently, IDSs are developed and designed to make intrusions public, before disclosing the system resources which are secured. IDSs are consistently acknowledged as a second line of defense from the view point of security. IDSs are considered as web equivalent of the robbers' alarms that are being utilized in physical security systems currently [16]. The normal functional requirement of IDSs are; "high true positive rate, measured as the percentage of anomalies detected, and low false positive rate, measured as the percentage of normalcy deviations detected as anomalies".

According to architecture, IDS can be divided into two groups:

- Network Based IDS: TheIDS which is network based is responsible for the protection of the whole network environment from any type of intrusions. This type of IDS architecture requires comprehensive knowledge of the system status and also monitors the different components of the network as well as the transactions which carry out between them [11]. A network IDS watches the actions on an entire network and carries the traffic analysis for possible security breaches or threats. Agent technology performs a primary role in this type of IDS architecture.
- Host Based IDS: The IDS which is host based watches all the ongoing activities on an individual information system host. It makes sure that not a single security policy of information

system is being disrupted [12]. The host based IDS are installed only on a single terminal/host and are given the responsibility of observing the status of that specific host(or server) only.

6 IDS patterns

The primary purpose of developing an Intrusion Detection system is to detect and identify the potential threats efficiently. There IDS system can be classified into two categories:

- Misuse detection: or signature-based detection. The behavior of the system is compared with previously known types of attack patterns (or signatures). Those action patterns which may present a security threat should be described and stored in the system. Afterwards, the misuse detection technique attempts to recognize any type of bad behavior with respect to these stored patterns in the system [18].
- Anomaly detection: targets on normal behaviors, instead of attack behaviors. Firstly, these types of systems define what comprises a normal behavior which is normally carried out by an automated training and then intrusion activities are flagged that differ from this normal behavior by a specified threshold [22].

7 Intrusion detection techniques

There is a diverse class of IDS methods which are established on different formats, compositions and schemes. Following are the description of those frequently used techniques:

- Evolutionary Algorithm: The algo produces an application path that contributes standard operational models. This algorithm is designed to detect standard operational behavior, attempted intrusion, and error state by categorizing archetype which depends upon Non-Identical conditions [13].
- Rule Based: This technique compares data against signatures with state transition analysis. Every data packet is practiced to FSM (Finite State Machine) and follows transitions till the ultimate state reaches, resulted in the detection an attack [8,10].
- Statistical Analysis: This technique involves the comparison of present set of data pattern with the predefined set of basic criteria. The normal data behavior is compared with the deviations over a temporal period. In Anomaly Detection system this technique is used [21].
- Protocol Verification: This technique is based upon extensively checking the protocol behavior and their fields in comparison with the pre- defined set of standards. The data which violates the defined standards is considered to be malicious. This approach has good success rate in lucrative systems but the flaws are producing erroneous positives for undefined protocols [10].

8 Artificial neural network (ANN)

ANN neurons are utilized to model complicated hypotheses. Complexity of the hypotheses depends upon the number of neurons involved. The hypotheses evaluation is carried out by the setup of input nodes in the feedback process and the propagation of event data across the network towards output nodes where it is categorized as legitimate or suspicious [3]. Gradient descents are used at this stage to thrust the error back from the output node across to the network using BPA(Back Propagation Algorithm) to scan the hidden nodes for error measurement. Thus the cost of function in terms of gradient descent can be calculated. To assimilate the structure and sequence developed in the system, neural network systems go through the training phase. A typical structure of an ANN neuron is shown in Fig. 1.

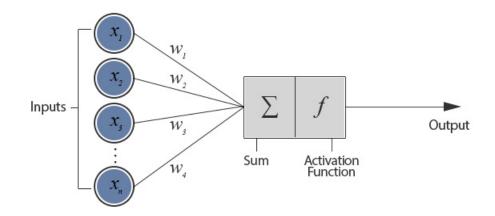


Figure 1: ANN neuron

9 ANN learning process

Two types of learning procedures for ANN's are as follows:

• Unsupervised Learning Procedure: The ANN in this learning has an input

$$g = \{a_i\}_{i=1}^N$$

i.e., unlabeled data set; need to search certain patterns in that data.

SOM is a category of Artificial Neural Network which is trained by unsupervised learning method to generate less dimensional, discrete representation input space of samples of training called map.

• Supervised Learning Procedure: In this type of learning, the ANN is presented with labelled set of training data which learns the mapping from data inputs *a* to outputs *b*, provided a labelled input-output sets of pairs,

$$g = \{(a_i, b_i)\}_{i=1}^N$$

where g is called the training set and N is the total number of training samples. It is presumed that b_i is a group variable from an infinite set

$$b_i \in \{1, \ldots, X\}$$

The MLP (Multi-layer Perceptron) is a category of artificial neural network trained by supervised learning procedure. The MLP is utilized for the detection of intrusions which were established on an offline analysis approach. Also, in a diverse outlook, it was used to detect intrusion in a data on a network by comparing it with the SOM Self-Organizing Maps.

10 MLP architecture

A three layer MLP architecture with feed-forward ANN is shown in Fig. 2. The network holds a function of unipolar transfer sigmoid in every output and hidden layer neurons. Also, an algorithm with stochastic learning and a function of mean square error is utilized. The input nodes which represent neural network with labels x_1, \ldots, x_3 have been implemented, and +1 as bias units.

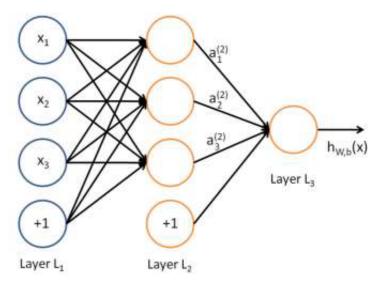


Figure 2: Feed-forward neural network

The ANN architecture has total of three input units (L_1) , three hidden units (L_2) and finally one output (L_3) composed of the three layers of neural network.

Feed forward Learning Algorithm: The feed forward algorithm is the simplest form of devised artificial neural network. The information moves in only forward direction, which starts from the input layer, the hidden layer and finally to the output layer respectively. The algorithm states as follows; Let V_l is the number of total elements which excludes the bias elements. Hence, the network parameters are

$$\{W, b\} = \{W^1, b^1, W^2, b^2\}$$

where W_{ij}^{lr} represents the combined parameter connected with the element j in layer lrand element i in lr + 1. Likewise, b_i^{lr} is the bias connected with I unit in the layer lr + 1. Therefore above mentioned elements and functions reflect $W^1 \in O^{3*3}$ and also, $W^2 \in O^{3*1}$. Suppose, a_i^1 represent the output element in layer lr.

For lr = 1, we just supposed that $a_i^1 = x_i$ represent the *i*th input. Thus, the ANN model will derive hypotheses $h_{(w,b)}(x)$ which give a real number as output. Thus we can represent

the above model in a mathematical form as:

$$a_1^{(2)} = f\{W_{11}^1 x_1 + W_{12}^1 x_2 + W_{13}^1 x_3 + b_1^1\}$$
(1)

$$a_2^{(2)} = f\{W_{21}^1 x_1 + W_{22}^1 x_2 + W_{23}^1 x_3 + b_2^1\}$$
(2)

$$a_3^{(2)} = f\{W_{31}^1 x_1 + W_{32}^1 x_2 + W_{33}^1 x_3 + b_3^1\}$$
(3)

$$h_{(w,b)}(x) = f\{W_{11}^2 a_1^{(2)} + W_{12}^2 a_2^{(2)} + W_{13}^2 a_3^{(2)} + b_1^2\}$$
(4)

Eq. 4, weighted total sum of the inputs from I in lr is the feedforward algorithm.

Backward Learning Algorithm: The process of learning in involves 4 stages, which are as follows:

- For all the layers in the neural network, the feed forward algorithm calculates the activation.
- The output from layer L_3 computes the measure of error in the output:

$$\Delta_{i}^{(L_{3})} = \frac{\Delta}{\Delta_{y}^{(L_{3})}} \frac{1}{2} \| z \cdot h_{(w,b)}(x) \|^{2}$$
$$= -\{z_{i} - a_{i}^{(L_{3})}\} \cdot f\left(y_{i}^{(L_{3})}\right)$$
(5)

Here, $a_i^{(3)} = f\left(y_i^{(L_3)}\right)$ is the sigmoid function.

• To calculate the errors in lr = 2, lr = 3 for every node I in layer lr:

$$\Delta_{i}^{(lr)} \left(\sum_{k=1}^{V_{t}+1} W_{ki}^{(lr)} \Delta_{k}^{(l+1)} \right) f' \left(y_{i}^{(L_{3})} \right)$$
(6)

• Lastly, the intended partial derivatives are computed as:

$$\begin{split} \frac{\Delta_i^{(lr)}}{\Delta W_{kj}^{(l)}} & M(W,b;z,x) = a_j^{(l)} \Delta_k^{(l+1)} \\ \frac{\Delta}{\Delta b_k^{(l)}} & M(W,b;z,x) = \Delta_k^{(l+1)} \end{split}$$

11 Proposed system and evaluation

We have used eight node sensors to compose an IoT network. Seven client nodes and one server relay node for data analytical purpose. Network tap is used to capture traffic to avoid any kind of hindrance or change in the live traffic. The sensor node sends data towards server node and server node acknowledges by sending receive data reply that is based on data itself. This process enables the sensor nodes as shown in Fig. 3, to acclimatize their behavior and respond to the ongoing phenomenon.

In our research setup, we have an external intruder who attacks the IoT network as shown in Fig. 4. Server node is the only target of target of attackers because the server node analyzes, keeps the record and responds to the sensor nodes. The attacker launches the DoS attack by

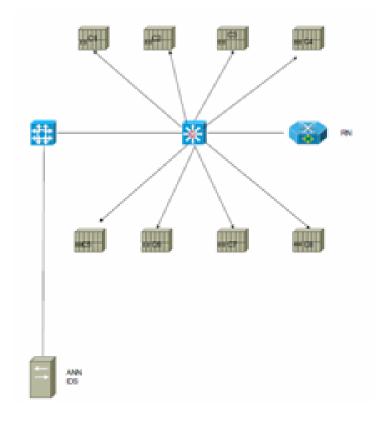


Figure 3: Normal condition

sending 10^7 packets towards the target node from a single host and the same amount of attack traffic was launched from four hosts to launch DDoS towards the target node. We used a custom C script to craft UDP packets to use them as attack packets. Therefore, the server node comes to a halt and do not respond. The sensor nodes are not able to accommodate their changed behavior and finally cause the halt in the monitored system.

Therefore the detection of these attacks at the right time is very important to allow the uninterrupted working of sensor network and assure the reliability of the network.

No. of sample patters used for training	2575
No. of sample patters used for Validation	499
No. of sample patters used for test	499

	Table 2:	Amount	of Pattern	Traffic	Used	For	Categorization.
--	----------	--------	------------	---------	------	-----	-----------------

Traffic	Size	%
DDoS	2294	63.98
DoS	2294	63.98
Normal	1279	36.02

In this section we will evaluate the performance of designed Artificial Neural Network intrusion detection method as discussed in the previous section. We trained the network with the

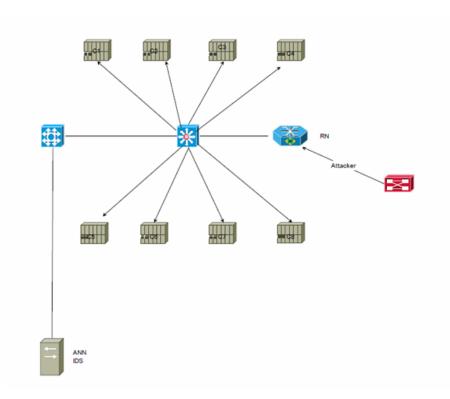


Figure 4: Attack situation

following parameters as shown in Table 1 and the amount of sample patterns used for categorization is shown in Table 2.

Neural network confusion matrix is shown in Fig. 5, is used to plot

- Training Set
- Testing Set
- Validation Set
- The Final All Confusion Matrix

The output of the system is of two types:

- 1. True Positive
- 2. False Positive

The attack traffic that is correctly categorized is measured as True positive as shown in green block and the correctly categorized normal traffic is measured as False positive as shown in red block. After all the evaluation and tests as shown in 4th final confusion matrix in Fig. 5, our proposed system showed the detection accuracy greater than 99% in the categorization of network traffic. Our method proved that the artificial neural network algorithm used is capable of detecting DDoS and Dos attack traffic during the flow of genuine IoT network traffic. It also improves the stability and reliability of the IoT network by signaling the response system at the right time to avoid the network disruptions, thus improving and enhancing the performance of the IoT network.



Figure 5: Confusion matrix

12 Conclusion

A DDoS detection method is presented in this research article using ANN for IoT network. The detection method is mainly based on categorization of legitimate traffic patterns and attack traffic patterns. The proposed system is simulated and tested in an organized simulated IoT network and the obtained results proved more than 99% detection accuracy. The system successfully identified the attack traffic and performed well in true and false negative accuracy. In the future, the system will be trained with latest threat patterns and will be tested to check its reliability with the modern world technology.

Acknowledgement

This research was funded and conducted at Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia during the academic year 2017–2018 under research number 2017/01/8029.

Bibliography

- [1] Ahamad, T. (2016). Detection and Defense Against Packet Drop Attack in MANET, International Journal of Advanced Computer Science and Applications (IJACSA), 7(2), 2016.
- [2] Ahamad, T.; Aljumah, A. (2015). Detection and defense mechanism against DDoS in MANET, Indian Journal of Science and Technology, 8(33), 2015.
- [3] Alan, S.; Overill, R.E.; Radzik, T. (2016). Detection of known and unknown DDoS attacks using Artificial Neural Networks, *Neurocomputing*, 172, 385–393, 2016.

- [4] Aldaej, A.; Ahamad, T. (2016). AAODV (Aggrandized Ad Hoc on Demand Vector): A Detection and Prevention Technique for Manets, *International Journal of Advanced Computer Science and Applications(IJACSA)*, 7(10), 2016.
- [5] Aljumah, A.; Ahamad, T. (2016). Black Hole and Mobile Ad Hoc Network (MANET): A Simple Logical Solution, In: 11th International Conference on Cyber Warfare and Security: ICCWS2016, 1-9, 2016.
- [6] Aljumah, A.; Ahamad, T. (2016). A Novel Approach for Detecting DDoS using Artificial Neural Networks, International Journal of Computer Science and Network Security, 16(12), 132-138, 2016.
- [7] Ahmed, E.; Yaqoob, I.; Gani, A.; Imran, M.; Guizani, M. (2016). Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges, *IEEE Wireless Communications*, 23(5), 10–16, 2016.
- [8] Alrajeh, N. A.; Khan, S.; Shams, B. (2013). Intrusion detection systems in wireless sensor networks: a review, *International Journal of Distributed Sensor Networks*, 1-7, 2013.
- [9] Alshehri, A.; Sandhu, R. (2016). Access Control Models for Cloud-Enabled Internet of Things: A Proposed Architecture and Research Agenda. In: Collaboration and Internet Computing (CIC), 2016 IEEE 2nd International Conference on, 530–538, 2016.
- [10] Bucerzan, D.; Cayrel, P.-L.; Dragoi, V.; Richmond, T. (2017). Improved Timing Attacks against the Secret Permutation in the McEliece PKC, *International Journal of Computers Communications & Control*, 12(1), 7-25, 2017.
- [11] Butun, I.; Morgera, S. D.; Sankar, R. (2014). A survey of intrusion detection systems in wireless sensor networks, *IEEE Communications Surveys & Tutorials*, 16(1), 266–282, 2014.
- [12] Creech, G.; Hu, J. (2014). A semantic approach to host-based intrusion detection systems using contiguousand discontiguous system call patterns, *IEEE Transactions on Computers*, 63(4), 807–819, 2014.
- [13] Elhag, S.; Ferná ndez, A.; Bawakid, A.; Alshomrani, S.; Herrera, F. (2015). On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems, *Expert Systems with Applications*, 42(1), 193–202, 2015.
- [14] Elkhodr, M.; Shahrestani, S.; Cheung, H. (2016). The internet of things: new interoperability, management and security challenges. arXiv preprint arXiv:1604.04824.
- [15] Han, G.; Shu, L.; Chan, S.; Hu, J. (2016). Security and privacy in Internet of things: methods, architectures, and solutions. *Security and Communication Networks*, 9(15), 2641– 2642, 2016.
- [16] Gong, W. (2016). The Internet of Things (IoT): what is the potential of the internet of things (IoT) as a marketing tool?, Bachelor's Thesis, University of Twente, 2016.
- [17] Gunasekaran, A.; Subramanian, N.; Tiwari, M.K. (2016). Information technology governance in Internet of Things supply chain networks, *Industrial Management & Data Systems*, 116.7, 2016.
- [18] Kim, G.; Lee, S.; Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, 41(4), 1690– 1700, 2014.

- [19] Madakam, S.; Date, H. (2016). Security Mechanisms for Connectivity of Smart Devices in the Internet of Things, In *Connectivity Frameworks for Smart Devices* (pp. 23–41). Springer International Publishing.
- [20] McKelvey, B.; Tanriverdi, H.; Yoo, Y. (2016). Complexity and Information Systems Research in the Emerging Digital World. *MIS Quarterly*.
- [21] Mitchell, R.; Chen, I. R. (2014). A survey of intrusion detection techniques for cyber-physical systems. ACM Computing Surveys (CSUR), 46(4), 55, 2014.
- [22] Moshtaghi, M.; Erfani, S. M.; Leckie, C.; Bezdek, J. C. (2017). Exponentially Weighted Ellipsoidal Model for Anomaly Detection. *International Journal of Intelligent Systems*, 32(9), 881-899, 2017.
- [23] Niu, J.; Jin, Y.; Lee, A.J.; Sandhu, R.; Xu, W.; Zhang, X. (2016). Panel Security and Privacy in the Age of Internet of Things: Opportunities and Challenges. In Proceedings of the 21st ACM on Symposium on Access Control Models and Technologies, 49–50, 2016.
- [24] Samaila, M. G.; Neto, M.; Fernandes, D. A.; Freire, M. M.; Iná cio, P. R. (2017). Security Challenges of the Internet of Things, *Beyond the Internet of Things*, 53–82, 2017.
- [25] Singh, M.; Rajan, M. A.; Shivraj, V. L.; Balamuralidhar, P. (2015). Secure mqtt for internet of things (iot). In: Communication Systems and Network Technologies (CSNT), 2015 Fifth International Conference on, IEEE, 746–751, 2015.
- [26] Tariq, U.; Aldaej A. (2018). Outlook of Coordinated Transmission Control in 5G Networks for IoTs, International Journal of Computers Communications & Control, 13(2), 280-293, 2018.
- [27] Tellez, M.; El-Tawab, S.; Heydari, H. M. (2016). Improving the security of wireless sensor networks in an IoT environmental monitoring system. In Systems and Information Engineering Design Symposium (SIEDS), 2016 IEEE, 72–77, 2016.
- [28] Xu, K.; Qu, Y.; Yang, K. (2016). A tutorial on the internet of things: from a heterogeneous network integration perspective, *IEEE Network*, 30(2), 102–108, 2016.
- [29] Zheng, Z.; Xie, S.; Dai, H. N.; Wang, H. (2016). Blockchain Challenges and Opportunities: A Survey, Int. J. Web and Grid Services, 14(4), 2018.
- [30] Zhao, S.; Cheng, B.; Yu, L.; Hou, S. L.; Zhang, Y.; Chen, J. L. (2016). Internet of Things Service Provisioning Platform for Cross-Application Cooperation, *International Journal of Web Services Research (IJWSR)*, 13(1), 1–22, 2016.