www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 830-836

e-ISSN : 2583-1062
Impact Factor : 7.001

INTRUSION DETECTION SYSTEM IN IOT ENVIRONMENT USING MACHINE LEARNING TCHNIQUES

Shackhhi Sharma1, Mohammad Mudassar Khan[2]

[1]Department of Information Technology, Mahakal Institute of Technology, Ujjain, Madhya Pradesh, India.

[2]Assistant Professor, Department of Computer Science and Engineering, Mahakal Institute of Technology, Ujjain, Madhya Pradesh, India.

ABSTRACT

Internet of Things (IoT) and its applications are the most popular research areas at todays enviroment. IoT characteristics, one side make it easily applicable to real-life applications, whereas on the other side expose it to cyber threats. Many academics are increasingly interested in enhancing the security of IoT systems. Machine learning (ML) approaches were employed to function as intrusion detection systems (IDSs) to provide better security capabilities. This work proposed a intrusion detection system based on machine ML approaches to detect attacks in IoT. A comprehensive study is carried on the classifiers which can advance the development of anomaly intrusion detection systems (IDSs). Institutions choose wise testing and verification techniques by comparing the highest rates of accuracy. IoT use has been accelerating recently across a variety of industries, including health care, smart homes, intelligent transportation, smart cities, and smart grids. The main goals of this study are to motivate IoT security researchers for developing IDSs using machine learning techniques, and suggesting appropriate methods.

Keywords: Internet of Things, Intrusion detection, Machine Learning, Security, Privacy, Vulnerability, threats, Attack.

1. INTRODUCTION

Internet of Things (IoT) is an innovative concept of connecting thousands of low-power embedded devices with each other and with the global internet. Some of the major applications of IoT include: 1) smart cities - monitoring parking spaces, traffic congestion, managing street lighting; 2) smart environment - forest fire detection, control of CO2 emissions; and 3) smart industrial control - monitoring air quality, temperature; and 4) eHealth - monitoring patient conditions in hospitals and elderly homes, UV radiation detection, etc. With the above applications, the IoT exchanges vast amounts of critical and private data, whose security if compromised can lead to severe economic consequences as well as threaten human lives.

we'll take a glance at a few of the fundamental building blocks that go into the construction of the Internet of Things. We need technologies that can sense data, identify devices, and assign unique Internet Protocol (IP) addresses to the many IoT gadgets in order for the Internet of Things to function properly. Additionally, we need technologies that enable smooth two-way communication between devices that have a wide range of capabilities. These technologies are essential to the Internet of Things and serve as its core components. Sensors, controllers, network infrastructure, and application software make up the IoT's underlying structure. Let's quickly go through each of these. It is shown in figure 1

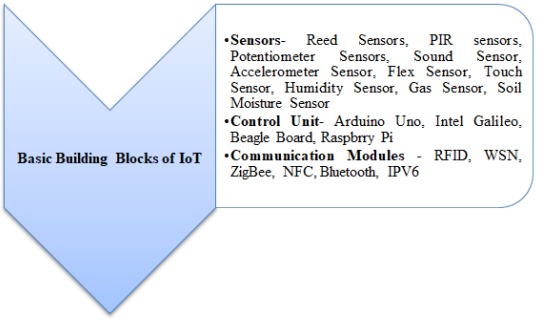

Figure 1 Basic Building Blocks of IOT

The field of machine learning is concerned with the question of how to build computer programs that automatically improve with experience, that is, acquire new knowledge, skills and techniques for organizing knowledge [5]. According to [6], since the computers were invented, the question of whether they could learn was present, because if it were possible to understand how to program them to automatically improve with the experience, this would be a great advance. Unsupervised techniques are those that the algorithm receives as samples without labels, so, by means of data grouping strategies, the input values are organized, in order to generate groups belonging to the same category

[7]. So, your goal is to organize samples based on your attributes, without the aid of an oracle. In the latter type as the semi-supervised name suggests, the input data set has samples with and without labels [8]. It is known that the process of labeling databases is expensive, for this reason there is an interest in developing strategies that meet the prerequisite of labeled databases to induce models using supervised techniques for machine learning [9]

An intrusion detection system (IDS) evaluates whether or not certain events are suggestive of an attack or are simply the result of regular system functioning by dynamically monitoring the system in question. Debar and colleagues (1999). In below figure depicts the IDS, with solid lines representing data/control flow and dotted lines representing responses to intrusive procedures.

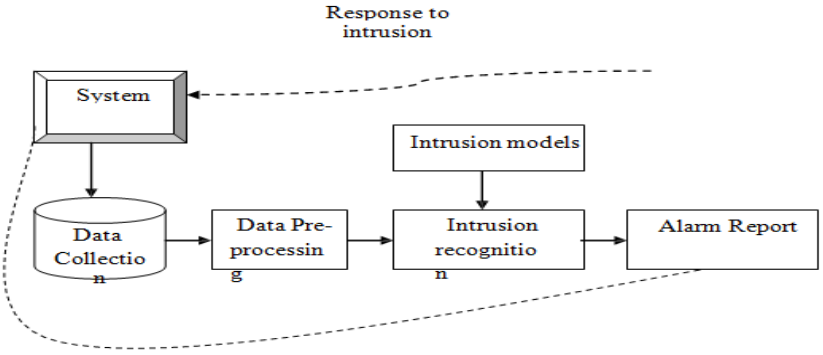

Figure 2. Intrusion Detection System

2. RELATED WORKS

IoT network is a promising technology, IoT implementation is growing rapidly but cybersecurity is still a loophole, detection of attacks in IOT infrastructures is a growing concern in the field of IoT. With the increased use of Internet of Things in different areas, cyber-attacks are also increasing proportionately and can cause failures in the system. IDS become the leading security solution. Anomaly based network intrusion detection (IDS) detection plays a major role in protecting networks against various malicious activities. Improving the security of loT networks has become one of the most critical issues. This is due to the large-scale development and deployment of loT devices and the insufficiency of Intrusion Detection Systems (IDS) to be deployed for the use of special purpose networks.[10]

The development in Machine Learning has enabled the improvement of different incredible analytical strategies that could be utilized to upgrade IoT security. IoT privacy and security were fundamental significance and assume a critical role in the commercialization of IoT innovation.[11] Conventional security and privacy arrangements affects from various issues that are identified with the dynamic quality of the IoT networks.

The field of machine learning is concerned with the question of how to build computer programs that automatically improve with experience, that is, acquire new knowledge, skills and techniques for organizing knowledge [6]. According to [7], since the computers were invented, the question of whether they could learn was present, because if it were possible to understand how to program them to automatically improve with the experience, this would be a great advance. Learning consists in the development of methods that are capable of extracting concepts from data samples, therefore, it is a process by which the computer, through descriptors and characteristics, defines concepts and based on them is able to predict labels for new entries. For learning to be possible, initially techniques are used to train a classifier [8].

Unsupervised techniques are those that the algorithm receives as samples without labels, so, by means of data grouping strategies, the input values are organized, in order to generate groups belonging to the same category [10]. So, your goal is to organize samples based on your attributes, without the aid of an oracle. In the latter type as the semi-supervised name suggests, the input data set has samples with and without labels [9]. Thus, active learning is a machine learning approach in which the algorithm has certain autonomy, as one of the steps of the algorithm is to select samples to be labeled. Samples are chosen from certain metrics, known as sampling strategies. Thus, there is optimization of the model's convergence process, since there is a reduction in the number of labeled samples present in the database [11].

Smart cities, in particular, are heavily dependent on IoT networks for various critical services such as transportation, energy management, and public safety. This increased dependence makes smart cities more vulnerable to cyber-attacks. Security is considered the main problem with IoT based networks. Attacker tries to hijack IoT network by sending false data packets. Due to that network becomes vulnerable and intruder easily unbalances the entire system. In smart cities there is extensive use of IoT networks which makes its security a huge concern. Attackers can exploit vulnerabilities in IoT devices to gain unauthorized access to sensitive information, disrupt city services, or cause

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 830-836
e-ISSN : 2583-1062
Impact Factor : 7.001

physical damage. This makes intrusion detection a critical requirement for ensuring the security of smart cities [13]. Tele-medicine is a new concept in smart cities where doctor will operate and consult the patient remotely. Therefore, intruder can disturb the process by deploying DoS/DDoS, Sybil, spoofing, wormholes and man-in-the middle attack [14]. This can cause some serious life threatening situation for the patient. Figure 2.1, Depicts the concept of IoT network of multiple devices using an IDS and its advantage in case of any intrusion.

However, to make smart city environment safe from such attacks, threat detection is very much necessary. Therefore, Intrusion detection system (IDS) plays an important role in identification of various attacks on IoT-networks [12].

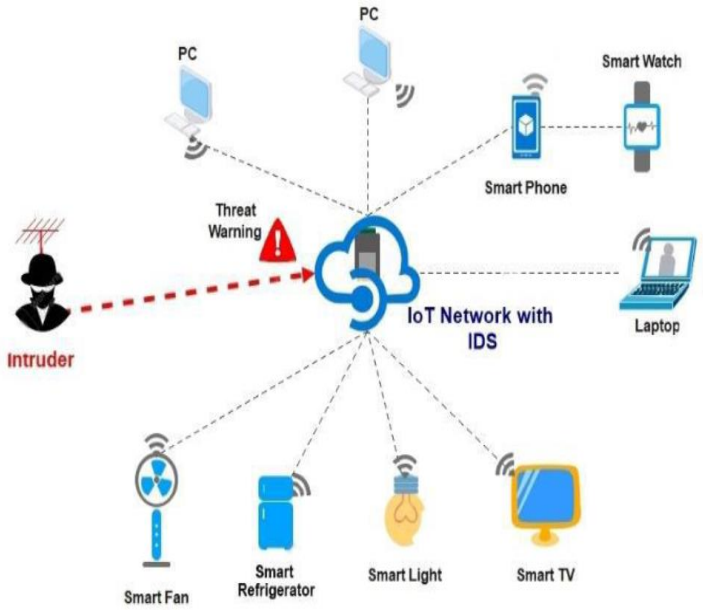

Figure 3 IoT Network with IDS

3. THEORETICAL BACKGROUND

K-means clustering based IDS (KM-IDS)

Algorithm 1 K-means clustering based KM-IDS algorithm

1: Input: test data of n nodes Y1, Y2, ..., Yn

2: Determine optimal K (number of safe zones) by minimizing the distortion function, J

3: Randomly initialize K safe zone centroids: C1, C2, ...,Ck

4: repeat

5: for i = 1 to n do

6: S(i) = index (from 1 to K) of safe zone centroid closest to Yi

7: end for

8: for k = 1 to K do

9: C(k) = average of points assigned to safe zone k

10: end for

11: until p iterations

12: Let Ni = Number of nodes in S(i)

13: attacksDetected = 0

14: for i = 1 to K do

15: for j = i+1 to K do

16: attacksDetected+ = Ni.Nj

17: end for

18: end for

Algorithm 1 shows the steps involved in KM-IDS. In Steps 1-2, the network is first initialized and each nodes x,y coordinates are entered into the 6BR. Next step for 6BR is to divide the network in K safe zones. 6BR determines K by first plotting a graph of distortion cost function for various potential K values and then using the Elbow-method to choose

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 830-836
e-ISSN : 2583-1062
Impact Factor : 7.001

the optimal value of K. In Step 3, after optimal K is selected, 6BR randomly initializes K safe zone centroids (C1, C2, , Ck). Steps 5-7 are safe zone assignment steps where each node is assigned to the safe zone with the closest centroid to the node. In Steps 8-10, 6BR moves the centroids: the coordinates of centroids of each safe zone are reassigned to the mean of Coordinates of the its assigned nodes. 6BR repeats the safe zone assignment and moving the centroid steps 30 times (or, p iterations) to get the optimal centroids for each safe zone. Steps 14-18 show that in KM-IDS, any request from a node to become a neighbor of another node located in a different safe zone is considered as a wormhole attack. In this case, 6BR denies the request and flags an attack.

Decision Tree based IDS (DT-IDS)

Algorithm 2 Decision Tree based DT-IDS algorithm

Input: input data of n nodes X1,X2, ...,Xn and corresponding training data

2: Create adjacency matrix, A of size nxn

for i = 1 to n **do**

4: **for** j = i to n **do**

A[i][j].dist = dist(Xi,Xj)

6: if(i=j) A[i][j].connected=true

else A[i][j].connected=false

8: A[j][i] = A[i][j]

end for

10: **end for**

distance = 0

12: **for** each unique tuple (Xi, Xj, true) in training data **do**

A[i][j]. connected = true

14: A[j][i]. connected = true

distance += A[i][j].dist

16: **end for**

threshold = mean(distance)

18: Input: test data of n nodes Y1,Y2,.Yn

Create adjacency matrix, B of size nxn

20: Initialize B following Steps 3 to 10

attacksDetected = 0

22: **for** i = 1 to n **do**

for j= i+1 to n **do**

24: if(B[i][j] greater than threshold)

attacksDetected++

26: **end for**

end for

Algorithm 2 shows the steps involved in DT-IDS. In Step1, 6BR is first trained in DT-IDS where the training network is entered in the 6BR. First, coordinates of each node are entered into the 6BR and an adjacency matrix is initialized with the nodes and the distances between each node, as shown in Steps 2-10. In Steps 12-16, the adjacency matrix is updated with the training network connections. A threshold distance is determined in Step 17 by taking the mean of all the connections in the training network. In Steps 19-20, a similarly distributed network is entered into 6BR and the coordinates of each node are stored in an adjacency matrix. In Steps 22-27, the new adjacency matrix is scanned and wormhole attack detections are made. In DT-IDS, any request from a node to become a neighbor of another node located at a distance greater than threshold is considered as a wormhole attack. In this case, 6BR denies the request and flags an attack.

4. DOMAIN DESCRIPTION AND PROPOSED SYSTEM

Hybrid-IDS. In this approach, every request for a direct connection made to 6BR is approved or rejected using a twostage IDS system. We first train the IDS with a known data set to determine a threshold for direct connections using DT-IDS. In a new dataset, which is similar to the trained dataset, the IDS first divides the network into an optimal number of safe zones using KM-IDS. In the native KM-IDS implementation, if the 6BR receives request to

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 830-836
e-ISSN : 2583-1062
Impact Factor : 7.001

update neighbors from two nodes belonging to separate safe zones (e.g. nodes 9 and 6) then the IDS of 6BR would detect a attack and deny the request. However, in Hybrid-IDS, if the distance between the nodes 9 and 6 is within the threshold that was determined using DTIDS, the request will be granted. By using KM-IDS in Step 1 and DT-IDS in the Step 2, we are introducing a filter to eliminate false positives that were generated by KM-IDS.

In this algorithm shows the steps involved in Hybrid-IDS. In Steps 1-2, 6BR is first trained with a training network to determine a threshold using DT-IDS. In Step 3, a similarly distributed network is initialized and each nodes coordinates are entered into the 6BR. In Step 4, 6BR divides the new network into K safe zones using KM-IDS. Steps 6-15 is the 2-step process to determine wormhole attacks in Hybrid IDS. In the inner for loops in Steps 8-13, a request from a node to become a neighbor of another node located in a different cluster is considered as a candidate for wormhole attack. Final decision is taken in the outer for loops (Steps 6, 7, 14, 15) in which if the distance between the two nodes is greater than the threshold, 6BR denies the request and flags an attack.

Proposed Algorithm:

Algorithm 3 Hybrid-IDS algorithm

Input: input data of n nodes X1,X2, ...,Xn and corresponding training data

Determine threshold using DT-IDS

3: Input: test data of n nodes Y1, Y2, ..., Yn

Determine K safe zones using KM-IDS

attacksDetected = 0

6: **for** i = 1 to K **do**

for t in S(i) **do**

for j = i+1 to K **do**

9: **for** u in S(j) **do**

if dist(t,u) greater than threshold

attacksDetected++

12: **end for**

end for

end for

15: **end for**)

We are selecting different data set and apply above algorithm to find accurate result with minimum execution time and memory uses.

5. RESULT ANALYSIS

Table 1 Comparsion between two IDS algorithm

Parameter/Algorithm	Previous IDS algorithm	Proposed Hybrid-IDS algorithm
Data Set	173	173
Execution Time	0.0025	0.002
Accuracy	94.99	99.99
Memory Uses	7.78	0.5926

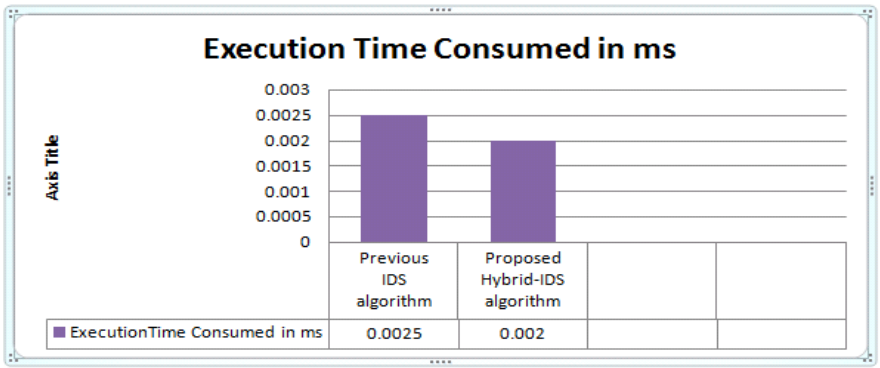

Figure 4 Comparison graph between time response

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 830-836
e-ISSN : 2583-1062
Impact Factor : 7.001

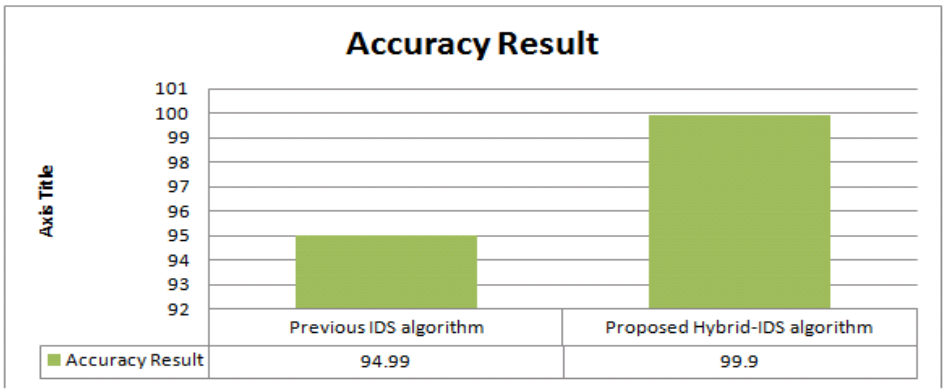

Figure 5 Comparsion graph between result of accuracy

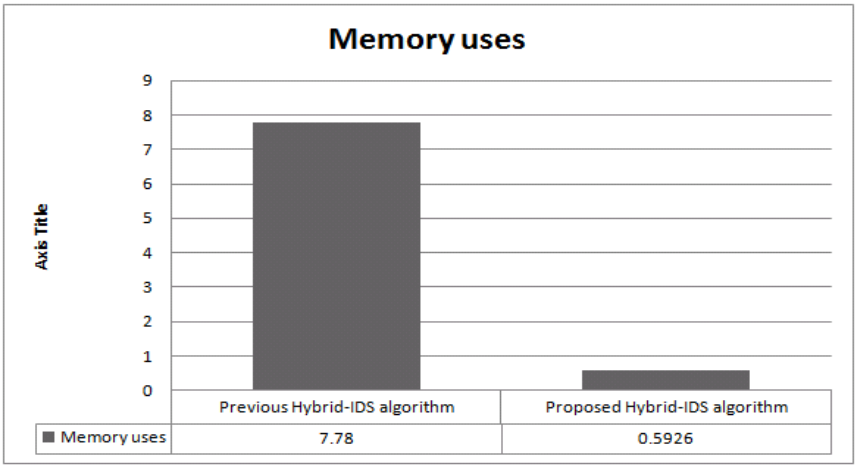

Figure 6 Comparsion graph between memory uses

6. CONCLUSION AND FUTURE WORK

The efficiency of intrusion detection system is critical to the protection of computer systems(IDS). Customer of intrusion detection systems (IDS) want their IDSes to provide a trustworthy and continuous detection services in order to protect their computer and networks. This IDS system are used in various IoT application with the help of different machine learning algorithm.In future try this Hybrid-IDS algorithm in multiple different topologies in network.

7. REFERENCES

[1] Prachi Shukla, "ML-IDS: A Machine Learning Approach to Detect Wormhole Attacks in Internet of Things," Intelligent Systems Conference 2017 7-8 September 2017 | London, UK.

[2] Eman Ashraf 1,2, *, Nihal F. F. Areed 2,3, Hanaa Salem 1, Ehab H. Abdelhady 2 and Ahmed Farouk 4, "IoT Based Intrusion Detection Systems from The Perspective of Machine and Deep Learning: A Survey and Comparative Study," Delta University Scientific Journal Vol.05-Iss.02 (2022) 367-386.

[3] Pongle, P.; Chavan, G. A survey: Attacks on RPL and 6LoWPAN in IoT. In Proceedings of the 2015, International conference on pervasive computing (ICPC), Pune, India 8–10 January 2015; pp. 1–6.

[4] Catarinucci, L.; De Donno, D.; Mainetti, L.; Palano, L.; Patrono, L.; Stefanizzi, M.L.; Tarricone, L. An IoT-aware architecture for smart healthcare systems. IEEE Internet Things J. 2015, 2, 515–526.

[5] Abdulrahman Salim A.Alrahman, Dr. Abdullahi Abdu Ibrahim, "INTRUSION DETECTION SYSTEM IN IOT NETWORK USING MACHINE LEARNING," 978-1-7281-9090-7/20/$31.00 ©2020 IEEE

[6] Abdulrahman Salim A.Alrahman, Dr. Abdullahi Abdu Ibrahim, "INTRUSION DETECTION SYSTEM IN IOT NETWORK USING MACHINE LEARNING," 978-1-7281-9090-7/20/$31.00 ©2020 IEEE

[7] Jiao B., Lian Z. and Gu X., "A dynamic inertia weight particle swarm optimization algorithm," 2008. Chaos, Solitons and Fractals 37, 698-705

[8] Yang X., Yuan Jinsha, Yuan Jiangye and Mao H., "A modified particle swarm optimizer with dynamic adaptation," 2007. Applied Mathematics and Computation 189, 1205-1213

[9] Dong C., Wang G., Chen Z. and Yu Z., "A Method Of Self-Adaptive Inertia Weight For PSO," 2008. 2008 International Conference on Computer Science and Software Engineering

[10] Amine Khatib1,2,3, Mohamed Hamlich1, Denis Hamad2 "Machine Learning based Intrusion Detection for Cyber-Security in IoT Networks ," E3S Web of Conferences **297**, 01057 (2021)

[11] Umar Albalawi," A Comprehensive Analysis On Intrusion Detection In Iot Based Smart Environments Using Machine Learning Approaches," INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 04, APRIL 2020 ISSN 2277-8616

[12] Amine Khatib1,2,3, Mohamed Hamlich1, Denis Hamad2 "Machine Learning based Intrusion Detection for Cyber-Security in IoT Networks ," E3S Web of Conferences **297**, 01057 (2021)

[13] Pooja G, Sundar R, Harshini R, Arjuna S. Recent Trends and Challenges in Smart Cities. EAI Endorsed Transactions on Smart Cities. 2022 Sep 21;6(3).

[14] Nawir M, Amir A, Yaakob N, Lynn OB. Internet of Things (IoT): Taxonomy of security attacks. In2016 3rd international conference on electronic design (ICED) 2016 Aug 11 (pp. 321-326). IEEE.