# Review on Cyber Security Incidents in IoT

Priyanka More, Santosh Chavan

UG Student, Dept. of Computer Science & Engineering, ADCET, Sangli, Maharashtra, India

Assistant Professor, Dept. of Electronics & Telecommunication Engineering, Government College of Engineering, Karad, Maharashtra, India

**ABSTRACT:** As there are rapid technological changes and being given to understand industry 4.0 revolution, Cyber-Physical Systems currently become one of the main targets of hackers and any damage to them leads to high losses to a nation. The main objective understood the fundamental and theoretical concepts of security in the digital world. We propose a deep-learning-based for continuous security monitoring analysis for IoT. In our study, the Cooja IoT simulator has been utilized for the generation of high-fidelity attack data, within IoT networks ranging from up to 1000 nodes. We propose a highly scalable, deep-learning-based attack detection methodology for the detection of IoT routing attacks with high accuracy and precision.

**KEYWORDS**: Anamoly detection; Cyber security; Contiki-Cooja, Deep Learning, Threats; Incidents; Internet of Things (IoT), Machine Learning; Security;

## I. INTRODUCTION

The term IoT is a system of interconnected devices, machines, and related software services. It has been playing an important role in modern society since it enabled energy efficient automation for enhancing the quality of life.

In this paper, we propose a highly-scalable deep-learning-based attack detection method for realistic IoT scenarios. We have processed data with a size close to 64x106. We obtained a high degree of training accuracy(upto99.5%) and F1-scores (upto99%). In this study, we have focused on specific IoT routing attacks, namely, decreased rank, version number modification, and hello-flood.

There are three main intrusion detection approaches in the literature; misuse detection, anomaly detection, and specification-based detection. Additionally, a hybrid-based system is also located under the intrusion detection topic and it's, brief, a mixture of misuse detection and anomaly detection. The intrusion detection scheme is also depicted in Figure 1.1.

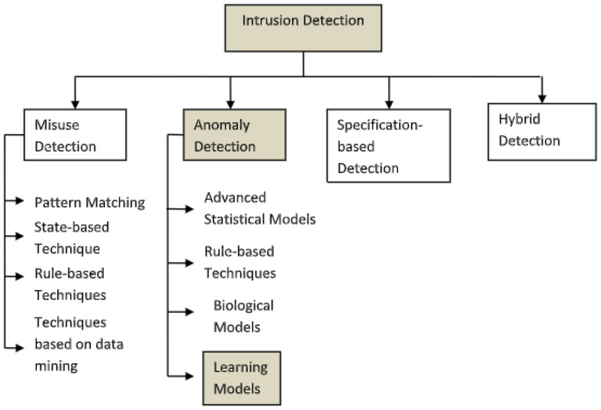


Figure 1.1: Intrusion Detection Approaches

Misuse detection is highly elective in detecting known attacks. However, it is insufficient against unknown or novel attacks because their signatures are not yet known. Additionally, any medication to the signatures can cause an increase in false alarm rate and that will decrease the electiveness and reliability of the detection system. Specific-based detection aims to set particular behavior based on the default deny principles. If the specifications are violated, the system will think there is an abnormal situation. The anomaly-based detection approach is constructed on normal activity profile and it assumes that any adversary action will conflict with the normal activity. Anomaly-based detection is examined under four subheadings; advanced statistical models, rule-based techniques, biological models, and learning models. We adopted learning models, because of that even if a misuse detection can give a faster response. Learning models have a more robust structure against unknown attacks than others [10].

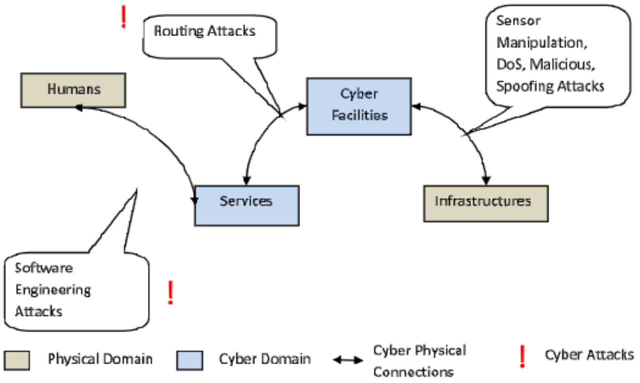

Routing Attacks
Sensor Manipulation, DoS, Malicious, Spoofing Attacks
Humans
Cyber Facilities
Services
Infrastructures
Software Engineering Attacks
Physical Domain
Cyber Domain
Cyber Physical Connections
Cyber Attacks

Figure 1.2: An Overview of IoT Attacks

In our study, we have demonstrated the viability of our methodology with simulations up to 1000 nodes whereas the existing studies, such as [24], [25], and [26], have demonstrated the methodology with the little number of nodes(upto50), which is not a realistic approach for an IoT environment. We used data generated by real-life equivalent simulations because of a lack of availability of public IoT attack datasets. The Cooja simulation generates raw packet capture files, which are first converted into Comma Separated Values (CSV) files for text-based processing. The CSV files are then inputting to the feature pre-processing module of our system. The features are calculated based on the traffic flow information in the CSV files. First, the feature conversation process is applied to some features, which is located in raw datasets. As a result of this analysis, some of the features are dropped in the pre-feature selection process. After feature pre-processing, the datasets corresponding to each scenario is labeled and mixed to produce a pre-processed dataset, consisting of a mixture of attack and benign data. These datasets are fed into the deep learning algorithm. This methodology is also depicted in Figure 1.3.

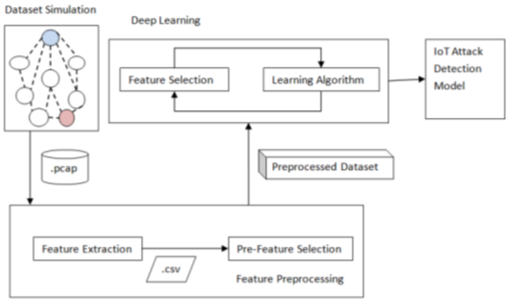


Figure 1.3: Proposed Research Methodology

Our proposal puts the minimum burden on the IoT network since it requires only the network packet traces for detection and prediction of attacks which can be collected externally by a network recording equipment or specially designated nodes. According to our knowledge, we are the first to prepare a deep learning-based methodology for routing attack detection.

The objectives of this study are:

- To deeply understand the IoT, threats, risks, and routing attacks.
- To create routing attack datasets and their preparation.
- To build a neural network by deep learning and train them by produced datasets

- To evaluate the models.

## II. BACKGROUND

### A. Internet of Things (IoT)

Internet of Things (IoT) is one of the biggest innovations of this century, considering the impact on our daily life. The areas of its usage are rapidly increasing. In 2017, the number of devices, that are called IoT, is approximately 27 billion and this number is estimated to reach 50.1 billion by 2020 [29]. In another aspect, the market size of IoT is estimated to reach approximately $9 trillion by 2020 [30]. IoT is a network that contains software, nodes, and servers. The contributions affect the living standard positively.

The list of IoT usage areas is way too long, to name a few; smart home and it's devices, wireless sensors, smart locks, smart meters, wearable devices, security cameras, smart plugs, Radio Frequency Identification (RFID), Machine to Machine(M2M), Machine to Human(M2H) devices and so on.

### 1 Challenges of IoT

IoT devices have the ability of data collecting, transferring, and processing in smart applications [36]. These data types spread to many areas such as health, transportation, military, etc. Security of the sensitive data, that's the biggest risk of IoT, comes into prominence. This risk roots in two main vulnerabilities. First; heterogeneous devices and inter-operable connections make the management of IoT systems more complex. Second; many devices have resource limitations, lack of computational capability, low latency. The second reason also makes the detection of possible and unknown attacks to IoT devices difficult [37]. These reasons make the routing protocols vulnerable. For example, WSN consist of nodes that include one or more sensors that have low-cost and limited power. These sensors' objectives are sensing the environment and communicating with the base station.

### 2 Simulation of Routing Attacks

We used the Cooja IoT simulator to simulate different IoT network communication scenarios. Cooja, coupled with the Contiki operating system, is an across-level simulation tool [2] l. Contiki makes it possible to load and unload individual programs and services to the simulated sensors[46]. We have conducted a simulation of each attack as mentioned above, by running a real sensor code in the Cooja simulator. We made the simulations on the cloud-based system. The Contiki environment includes 64-bit Java Run time Environment on top at 64-bit Ubuntu operating system and Contiki 3.0. Cooja user interface is shown in Figure 2.5.

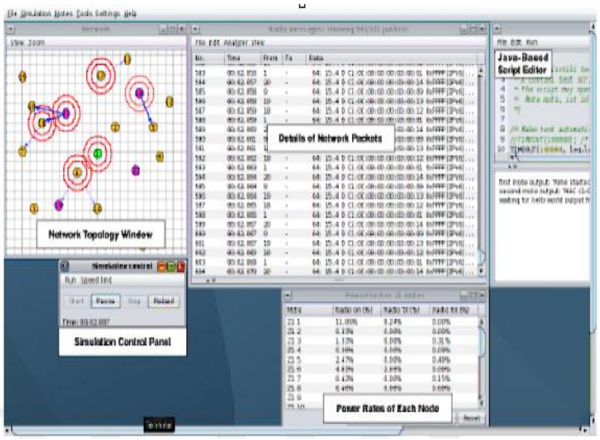


Figure 2.5: Cooja User Interface

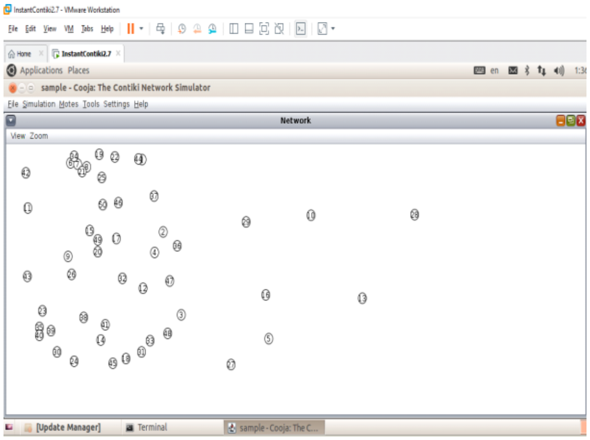
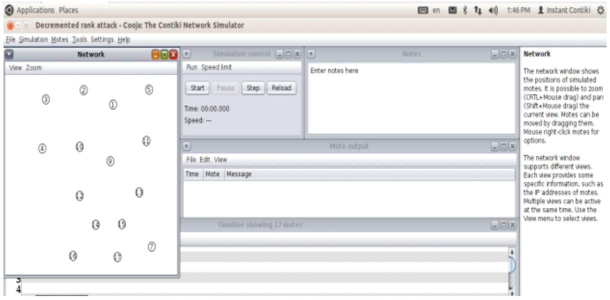

InstantContiki2.7 - VMware Workstation
sample - Cooja: The Contiki Network Simulator
Network
View Zoom
[Update Manager]
Terminal

Figure 2.7 Cooja: The Contiki Network Simulator

Routing attacks take place at the network layer. IoT systems are generally vulnerable to routing attacks. Among the most significant routing attacks are decreased rank, hello-flood, and version number attacks. The decreased rank attack is such a traffic misappropriation attack. In decreased rank attacks, malicious nodes advertise lower rank of other nodes to neighbor nodes by sending DIO packets. So the neighbor nodes change their routing path including the attacker node by sending DAO packets. The decreased rank attack can be applied to make an introduction for a black hole, eavesdropping, and sinkhole attacks. A decreased rank attack is also visualized in Figure 2.9. In this figure, node 1 is the DODAG root node and the others are normal nodes except for node 9, which is the malicious node that conducts the decreased rank attack in the network. The nodes, 3 to 8, are not effected by the attack. Nodes 10 and 11 are partially affected by the attack and their communication is partly interrupted. Some of the packets transmitted over these nodes could be taken by malicious node because the malicious node is in their routing table, in other words, they can send some packets over the malicious node to convey coming packets to the destination. The nodes, 12 to 18, are the victim nodes whose entire communication is transmitted over a malicious node

.



Network

The network window shows the positions of simulated motes. It is possible to zoom (CRTL+Mouse drag) and pan (Shift+Mouse drag) the current view. Motes can be moved by dragging them. Mouse right-click motes for options.

The network window supports different views. Each view provides some specific information, such as the IP addresses of motes. Multiple views can be active at the same time. Use the View menu to select views.

Figure 2.9 Decremented rank attacks

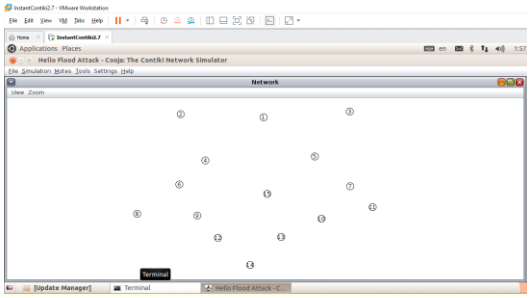

InstantContiki2.7 - VMware Workstation
File Edit View VM Tabs Help
Home
InstantContiki2.7
Applications Places
Hello Flood Attack - Cooja: The Contiki Network Simulator
File Simulation Motes Tools Settings Help
Network
View Zoom
Terminal
[Update Manager]
Terminal
Hello Flood Attack - C...

Figure 2.10: Hello Flood Attack -Cooja

It's seen that number of the received packet by the malicious node increases when the attack happens. We aim to use this anomaly as a feature. So we extracted Reception Rate(RR)(2.1), ReceptionAverageTime(RAT)(2.2), Received Packets Counts(RCP), Total Reception Time(TRT). Additionally, DIO and DAO packet-count is calculated because of that the attack is started by DIO and DAO packets.

$$RR = \frac{ReceivedPacketCountoftheNode}{1000[ms]} \qquad (2.1)$$

$$RAT = \frac{TotalReceptionTime}{ReceivedPacketCountoftheNode} \qquad (2.2)$$

In equation 2.1 and TR, 1000 is millisecond within all simulations. We have applied to the window while extracting features. The main purpose of the HELLO message is to introduce and integrate new nodes into the network. The nodes broadcast HELLO messages with their metrics such as signal power and ID number. All the other nodes create their routing table to send their messages. A malicious node sends HELLO massages by DIS packets to his victims by strong signal power and suitable routing metrics, appearing like a neighbor node.

The attacker node becomes the most favorable for the victims. This attack is called the hello-flood attack. The malicious node, Node 15, broadcasts HELLO messages to Nodes 4-13, except Nodes 8 and 11. The victim nodes change their routing table because the malicious nodes advertise high-quality metrics. After that, the effect of hello-flood attack is depicted in Figure 2.10

It's obvious that the number of transmitted packets of malicious node increase. So we extracted Transmission Rate (TR) (4.2), Transmission Average Time (TAT) (2.4), Transmitted Packets Counts (TPC), Total Transmission Time(TTT), and DIS features to identify this attack.

$$TR = \frac{TransmittedPacketCountoftheNode}{1000[ms]} \qquad (2.3)$$

$$TAT = \frac{TotalTransmissionTime}{TransmittedPacketCountoftheNode} \qquad (2.4)$$

In RPL, version numbers of nodes are changed by the root node. When the rot node changes them, each node starts the communication for reconstructing their routing table. So the network topology is changed. In a version number attack, the malicious node changes its version, then other nodes are forced to change their routing table. So the malicious node can promote itself to take better a place in other nodes routing tables. This can jeopardize the network's information security and performance due to changes in topology.

### B. Deep Learning for Cyber Security

Nowadays there is the use of deep learning (DL) for detecting routing attacks that target IoT. Before giving information about DL, ML should be explained to better understand DL. Because ML can be seen as the ancestor of DL.

#### 1. Machine Learning

Machine Learning(ML) is one of the pathways in leading Artificial Intelligence(AI) research. The popularity of ML comes from two purposes or two tasks have to be done by ML. First, the task that can be done by machines, second the task that can't be performed by humans. Learning activity comes into prominence to be intelligent what refers to a system that can keep up with changes in its environment. If a system can accord to the changes, this ability can help it to survive.

Supervised learning and unsupervised learning are the main types of ML, first of them are exist by using a fully labeled dataset whereas the other one exists a fully unlabelled dataset [11]. In supervised learning, the model receives datasets that include some features of vectors and labels that are the corresponding outputs of feature vectors. Thus the model learns to produce correct outputs as a result of a given new input. Classification and regression are the most popular product of supervised learning [12].

#### 2. Deep Learning

Deep learning (DL) is a kind of Neural Networks (NN) training and has NN architecture. Difference between 'old school' NN and deep learning is that DL has many hidden layers [12]. DL also learns the features itself, which enables the learning process to be more accurate, and also it is shown to be more efficient and accurate than shallow learning [13].

In supervised learning, there are three types of datasets. First, the training set is one of the key terms of the learning process. It is an enabler for the learning algorithm to be supervised and it contains the expected results under the label feature. They take one or more input from the previous neurons with the connection weights, sum them up, put it in the activation function and produce an output (2.6) that is, basically, fired or not. The mathematical representation of the neuron is shown in 2.5. After this addition, the activation function puts the Y into the process.

$$Y = X(\text{input}) * (\text{weight}) + \text{bias} \qquad (2.5)$$

$$\text{Output} = f(Y) \qquad (2.6)$$

’Fire' means to activate, the name is inspired by the biological working of the brain.

### C. IoT Routing Attack Dataset(IRAD)

First of all, we need the datasets that have routing attacks. In this research area, the lack of a dataset is one of the biggest challenges. So we simulated the routing attacks within different scenarios and processed raw datasets to make them ready for to detection process. Subsequently, we transform the PCAP files to CSV with Wireshark. After that, a feature extraction process is applied to the generated CSV files by using our developed Python data pre-processing script. Finally, we concatenate the same attack datasets to make a comprehensive dataset.

In the simulation, if there is no malicious node, all nodes are normal, so the benign scenarios have the same values. We tried to give our best during the dataset simulation process as in the whole thesis. Because the simulation process is getting harder when the number of nodes increases in the network topology. For these reasons, the values of Total Packet Count are different from each other.

### 1. Feature Extraction

The ML algorithms need some attributes about data for learning which are obtained by feature extraction. After the scenarios are simulated, the datasets are produced as PCAP files. We dissected the PCAP file to CSV by using Wireshark and the pre-processing section is fed by these CSV files. The data pre-processing step involves extracting useful features from the data for preventing over-fitting and to obtain problem-oriented attributes.

Cooja exports PCAP and CSV files after the end of the simulation. However, the raw data files aren't sufficient to be the input to learning algorithms because the raw dataset includes information such as source/destination nodes address and packet length, which causes noise and overfitting in the learning algorithm. We opted not to calculate global statistics over total simulated time or total packet count since this kind of calculation could decrease the importance of the main extracted features. So we have divided all the simulation to time frames, or windows of 1000 ms duration. Before this process, it is necessary to sort the datasets by simulation time, because the sequence of packet simulation time is highly important for feature extraction and Cooja extracts PCAP files in the wrong time sequence. It happens especially for wide range network topologies and long simulation times. The pseudo-code of our data pre-processing algorithm is also shown in Algorithm 1.

```
Algorithm 1 Enrichment of IoT Raw Dataset
Function
array ← RAWdataset.csv
Sorted array                                          Sorting by time
Feature conversion
Feature Extraction:
    1000ms ← Windowing Size
    Calculating Feature values within windowing size
    Labeling the dataset
    End of the Feature Extraction
End the function.
```

Raw datasets involve both quantitative and qualitative features. However, our learning algorithm accepts just quantitative values. So we applied feature conversion to qualitative features to transform their unified format.

We first convert the source and destination address from IPv6 format to Node id. For example:

$$\text{fe80 :: c30c : 0 : 0 : 12} \Longrightarrow 12 \qquad (3.1)$$

The broadcast packets are handled as follows. In the raw dataset, if the destination address is ff02::1a, that means the source node sends broadcast packets. This value is converted to 9999 to avoid any coincidence with another node:

$$\text{ff02 :: 1a} \Longrightarrow 9999 \qquad (3.2)$$

We also encoded the information of the packets: DAO is used in RPL for unicasting the destination information due to the selected parents. DIO is the most important message type in RPL. It keeps the current rank of the node and determines the best route through the base node by using specific metrics as distance or hop-count. Another message type is DIS. Nodes use DIS for joining to WSN. Ack is an acknowledgment message type for using to give responses by nodes [48]. These are encoded respectively: 1, 2 3, 4.Other types in our datasets are Protocol Data Unit (PDU) and UDP packets which are simulated data packets.

First, we calculated the Transmitted and Received Packets Counts (TPC and RCP) for each node in 1000ms in a specified time frame. Then, we divide these values to 1000ms and get Transmission Rate and Reception Rate for each node, TR (4.2), and RR (2.1) respectively, for all time frames. The duration time for each packet transmission and reception are calculated. Total Transmission Time (TTT) and Total Reception Time (TRT) is calculated by adding up the duration time of each transmission and reception packet in 1000ms. Then Transmission and Reception Average Time for each node, TAT (2.4), and RAT (2.2), are calculated. Our last features are about control packets; DAO, DIO, and DIS. The number of transmitted control packets of each node is calculated within the windowing size, 1000 ms.

The labeling process is also important. In our datasets, attack packets are labeled 1 and benign packets are labeled 0. We labeled the datasets, which has malicious node and activity, 1 and labeled the benign datasets 0. Because of malicious nodes aspect the entire network activity and influence normal node communication. For instance, most of the routing attacks may change the network topology. Accordingly, the routing path of normal nodes is changed.

### 2. Feature Normalization

Feature normalization is a pre-processing method for scaling all values of each feature into a certain range. It makes the data smoother and cleans the bias from data, ensuring a high accuracy rate [15].

We get different datasets from different scenarios for each routing attack. The datasets have different data values in different ranges due to their network topology. In this situation, datasets don't give relevant results to us and the learning algorithm couldn't work effectively. So we performed feature normalization for pulling all datasets in the same range by using our data normalization algorithm. We applied quantile transform and min-max scaling to datasets, respectively [16]. Each feature is enforced the normal quantile transform, separately.

```
Algorithm 2 Data Normalization Algorithm
Function
        Mixed Dataset ← Benign, Malicious Dataset
Feature Normalization:
        Quantile Transform Function ← Mixed Dataset
        Min Max Scale Function ← Transformed Dataset
End of the Feature Normalization
IoT Dataset ← Mixed Datasets                    // concatenating the datasets
End the function.
```

### 3. Feature Importance and Selection

Feature selection is a key step in ML. Feature selection is generally applied to the dataset before running the ML algorithm because it eliminates the irrelevant, weakly relevant features and selects the optimal subset of all features. It identifies the proper subset of all data features and makes the data serviceable. There are two main challenges; the large size of data and its inconvenient form. A dataset has two dimensions; the number of instances and number of features that are usually way too large. This huge volume also brings complexity. On the other side, datasets are created without features or attributes. Particularly, the network datasets are captured from the Internet or closed networks as PCAP form.We used a combination of random decision trees, histograms, and Pearson coefficient correlation [28] for the feature selection process.

## III. CONCLUSION

The wide uses of IoT nowadays bring some risks and means for cybercriminals to use in their attacks against governments, organizations, or individuals. The concept that deep learning can successfully deal with IoT security. The routing attacks under consideration were (decreased rank attack, hello-flood attack, and version number attack) are easily detected by the proposed attack detection models. In this paper, the issue of routing attack detection for IoT is handled. A further challenge is to create as there is a lack of datasets and also the data is not preferable when it has unrealistic content. Our attack datasets were simulated by using real-codes of simulation tools. The datasets, which we produced and also per-processed. The biggest effort of this work was generating and processing the attack datasets.

## IV. FUTURE WORK

The dataset has three routing attacks; decreased rank attack, hello flood attack, and version number attacks. By adding new routing attacks, one can plan to enhance our IoT attack dataset. We aim to increase the model prediction performance of three routing attacks and include more routing attacks in the study. We are also planning to create one deep neural network model to detect multiple attacks. By the way, we will diversify the scenarios with scenarios that have a different rate of malicious and normal nodes.

## REFERENCES

1. I. Fried berg, K. McLaughlin, P. Smith, D. Laverty, and S. Sezer, "STPASafeSec: Safety and security analysis for cyber-physical systems," J. Inf. Secure. Appl., vol. 34, pp. 183–196, 2017.
2. Wang and Z. Lu, “Cybersecurity in the Smart Grid: Survey and challenges,” Comput. Networks, vol. 57, no. 5, pp. 1344–1371, 2013.
3. W. Ten, G. Manimaran, and C. C. Liu, "Cybersecurity for critical infrastructures: Attack and defense modeling," IEEE Trans. Syst. Man, Cybern. Part systems Humans, vol. 40, no. 4, pp. 853–865, 2010.
4. J. Walker, B. J. Williams, and G. W. Skelton, “Cybersecurity for emergency management,” Technol. Homel. Secur. HST 2010 IEEE Int. Conf., pp. 476–480, 2010.
5. J. J. Walker, T. Jones, M. Mortazavi, and R. Blount, “CyberSecurity Concerns for Ubiquitous/Pervasive Computing Environments,” 2011 Int. Conf. Cyber-Enabled Distrib. Comput.Knowl.Discov., pp. 274–278, 2011.
6. N. S. Ali, “A four-phase methodology for protecting web applications using an effective real-time technique,” Int. J. Internet Technol. Secur. Trans., vol. 6, no. 4, p. 303, 2016.
7. Al-Mhiqani, M.N., Ahmad R., Abdulkareem K. H., Ali N.S., "Investigation Study of Cyber-Physical Systems: Characteristics, Application Domains, and Security Challenges, "ARPN Journal of Engineering and Applied Sciences, Vol. 12, No. 22, pp. 6557-6567, 2017
8. Ali, N. S., & Shibghatullah, A. S., "Protection Web Applications using Real-Time Technique to Detect Structured Query Language Injection Attacks," International Journal of Computer Applications, Vol. 149, No. 6, pp. 0975-8887,2016.
9. Sridhar, S., Hahn, A., & Govindarasu, M."Cyber-physical system security for the electric power grid". Proceedings of the IEEE, 100(1), 210-224.
10. Ten, C. W., Liu, C. C., & Manimaran, G. ."Vulnerability assessment of cybersecurity for SCADA systems".IEEE Transactions on Power Systems, 23(4), 1836-1846, 2008.
11. B. Miller and D. Rowe, "A survey SCADA of and critical infrastructure incidents," in Proceedings of the 1st Annual Conference on Research in information technology - RIIT ‟12, 2012, p. 51.
12. M. Kjaerland, “A taxonomy and comparison of computer security incidents from the commercial and government sectors,” Comput.Secur., vol. 25, no. 7, pp. 522–538, Oct. 2006.
13. S. Hansman and R. Hunt, “A taxonomy of network and computer attacks,” Comput.Secur., vol. 24, no. 1, pp. 31–43, Feb. 2005.
14. A. Rabiah, Y. Zahari, “A Dynamic Cyber Terrorism Framework,” Int. J. Comput. Sci. Inf. Secur., vol. 10, no. Xxx, 2012.
15. C. Blackwell, “A security ontology for incident analysis,” in Proceedings of the Sixth Annual Workshop on CyberSecurity and Information Intelligence Research - CSIIRW ‟10, p. 1, 2010.
16. J. Giraldo, E. Sarkar, et al., "Security and privacy in cyber-physical systems: A survey of surveys," IEEE Design & Test, 2017.
17. J. R. Klinefelter and T. A. Klinefelter, Minimalist Investor Maximum Profits, 1st editio. Page Publishing Inc, 2015.
18. W. B. Miller, D. C. Rowe, and R. Woodside, “A Comprehensive and Open Framework for Classifying Incidents Involving Cyber-Physical Systems,” in IAJC/ISAM Joint International Conference, 2014.
19. Furkan Yusuf Yavuz 'Deep Learning in Cyber Security for the Internet of things Thesis
20. Mohammed Nasser and ETL 'Cyber-Security Incidents: A Review Cases in Cyber-Physical
21. Systems',(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No.1, 2018

## BIOGRAPHY

**Priyanka More** is a Student in the Computer Science and Engineering Department, Annasaheb Dange College of Engineering and Technology, Ashta, India. She received a Diploma in Computer Engineering in 2018 from LESP, Sangli, India. Her research interests are IoT, Cyber Security, Machine Learning etc

**Santosh Chavan** is an Assistant Professor in the Electronics & Telecommunication Department, Government College of Engineering, Karad, M.S.Autonomus Institute, and India. He received a Master of Communication Networks (ME) degree in 2013 from SPPU, Pune, MS, India His research interests are Computer Networks (wireless sensor Networks), IoT, Cyber-Security, Data science etc.