Anomaly Detection in Cellular IoT with Machine Learning

Bernardo Santos[1(✉)], Imran Qayyrm Khan[1], Bruno Dzogovic[1], Boning Feng[1], Van Thuan Do[1,2], Niels Jacot[2], and Thanh Van Do[1,3]

[1] Oslo Metropolitan Univeristy, Pilestredet 35, 0167 Oslo, Norway
{bersan,bruno.dzogovic,boning.feng}@oslomet.no
[2] Wolffia AS, Haugerudvn. 40, 0673 Oslo, Norway
{vt.do,n.jacot}@wolffia.net
[3] Telenor ASA, Snarøyveien 30, 1331 Fornebu, Norway
thanh-van.do@telenor.com

Abstract. The number of Internet of Things (IoT) devices used in eldercare are increasing day by day and bringing big security challenges especially for health care organizations, IoT service providers and most seriously for the elderly users. Attackers launch many attacks using compromised IoT devices such as Distributed Denial of Services (DDoS), among others. To detect and prevent these types of attacks on IoT devices connected to the cellular network, it is essential to have a proper overview of the existing threats and vulnerabilities. The main objective of this work is to present and compare different machine learning algorithms for anomaly detection in the cellular IoT scenario. Five supervised machine learning algorithms, namely KNN, Naïve Bayes, Decision Tree and Logistic Regression are used and evaluated by their performance. We see that, for both normal (using a local test dataset) and attack traffic (CICDDoS2019 (**CICDDoS2019 Dataset:** https://www.unb.ca/cic/datasets/ddos-2019.html.)) datasets, the accuracy and precision of the models are in average above 90%.

Keywords: Machine learning · Anomaly detection · Mobile network security · IoT security · Cross layer security

1 Introduction

Internet of Things (IoT) is described as a "*network to connect anything with the Internet based on stipulated protocols through information sensing equipments to conduct information exchange and communications in order to achieve smart recognitions, positioning, tracing, monitoring, and administration.*" [1].

IoT is nowadays an outlet to provide applications and services such as smart healthcare, control energy, process monitoring, environmental observation and fleet management [2, 3] to companies in the industry or to end consumers at their own homes. As per 2020 forecasts in [4], 50 billion internet of things including cardiac monitors,

I. M. Pires et al. (Eds.): GOODTECHS 2021, LNICST 401, pp. 51–64, 2021.
https://doi.org/10.1007/978-3-030-91421-9_5

thermostats, smart phones, surveillance cameras, kitchen applications, cars, television everything will be connected via Internet.

Ericsson forecast report (C-IoT connections by segment and technology) states that "*The Massive IoT technologies NB-IoT and Cat-M1 continue to be rolled out around the world, but at a slightly slower pace in 2020 than previously forecasted due to the impact of COVID-19. 2G and 3G connectivity still enables the majority of IoT applications, but during 2019, the number of Massive IoT connections increased by a factor of 3, reaching close to 100 million connections at the end of the year.*" [5]. As shown in Fig. 1, Ericsson has predicted that some 29 billion IoT gadgets will be usable by 2025 [6].

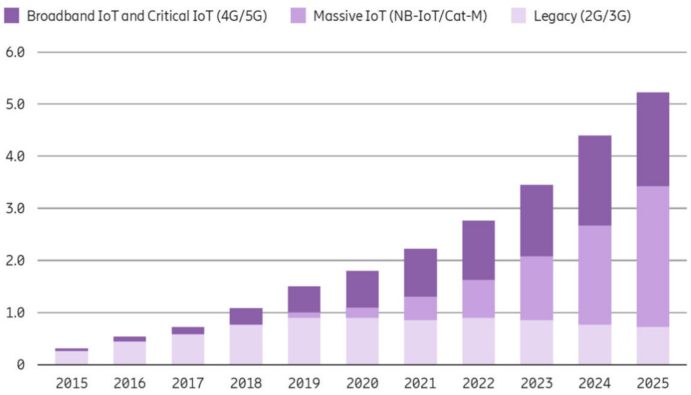

Fig. 1. Cellular IoT connections by segment and technology (billion) [6]

Due to the COVID-19 pandemic, cyber criminals gave us massive challenges specially in the health field. Due to this health crisis, they took advantage to develop their attacks on healthcare, hospitals, medical research centres and on international health public organizations. Because of this, the International Committee of Red Cross (ICRC) and other members have published a letter to various governments to do more on security and safety on these medical organizations from cyber-attacks [7].

There has been a huge interest and investment into bringing IoT capabilities to elderly care to provide senior citizens a more pleasant and lasting experience in the comfort of their own homes, even after going through some sort of incident, avoiding the need to move them to a senior home and providing the autonomy they are still accustomed to. This can be possible by deploying devices that would monitor (unobtrusively) not only the environment that the elderly person is living in, but also the elderly itself by measuring periodically its vital signs and provide immediate actions when an emergency happens (e.g., fall). As an example, the Body Sensor Network (BSN) innovation is a breakthrough that makes it possible for a physician to collect data from patients to additionally screen them via extremely compulsive devices that use lightweight protocol f or transmission of data such as CoAP [8].

The protection and security of these devices' sensors is extremely important because they hold the patient's critical data. Any unauthorized entry, leakage and capture of these devices can cause serious harm to patients. The information segment can be tampered

due to manipulation in packets that can be dangerous and life critical [8, 9]. If an intruder inflicts DoS on devices that change the value of the patient's high heartbeat, the device will not be triggered, and this will cause real problems and, in some cases, death.

IoT DDoS attacks were a main dominant attack in 2017, in line with the Arbor Security report [7], and 65% of the attacks carried out in 2016 were in majority DDoS attacks. The Mirai DDoS attack [10] was triggered by the contamination of defective IoT devices, being one of the biggest attacks ever to this segment. Consequently, DDoS attacks should be detected and mitigated. Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and DNS flooding are the most common attacks on DDoS. Protection measurements are challenging to enforce due to memory limitations, power and the heterogeneous nature of IoT devices.

To provide a way to mitigate these issues, the work presented in this paper aims at analysing different machine learning techniques that can help in detecting or even predicting an exploit targeting IoT devices connected to cellular networks. The hypothesis is as follows: If we can obtain information from the control and data plane in a cellular network, coming from IoT devices, we can use machine learning and anomaly detection algorithms in these data to see if it allows us to detect or even predict an upcoming attack.

The paper is organized as follows: In Sect. 2, we discuss about what kind of technologies and concepts are needed for this work. In Sect. 3, we showcase the planned steps that are important to consider given the problem statement. In Sect. 4, we present the outcome given on what was implemented and in Sect. 5, some take home messages are provided as to what to do with the scenarios described in this work. Finally, in Sect. 6, we make a summary by highlighting the obtained results and provide guidelines for purposes of further developing this research topic.

2 Background and Related Work

2.1 Cellular Networks

Currently there are 16 billion cellular customers from 2G to 4G and it is gradually increasing [11] to the 5G generation, with approximately 50 billion including IoT devices. 5G comes as a breakthrough for digital voice and data capacity but also for special features like IoT (Internet of Things) and AR (Augmented Reality), VR (Virtual Reality). Anything from smart cars to city grids, using different protocols such as the CISCO CCN and the MQTT protocols. Packet switching technology is used in 5G network. The latency in 5G network is only 1 ms [12].

2.2 IoT (Internet of Things)

As described in [13], IoT has a lot of security threats and challenges. According to the researchers, we need to understand the new features of IoT regarding security threats in IoT device. We define some security threats of IoT that cause attack in IoT devices as follows:

1. **Ubiquitous:** It is involved in our daily life and use all our resources. Individuals do not have an idea of the security of devices and still use them, and manufacturers

provide very little safety advice or recommendations given that the device collects sensitive data. The unsafe default configuration of these devices is one of the latest and common attacks' triggers.

2. **Diversity:** IoT has several devices that are involved in use cases and applications. IoT tracks different cloud networks through distinctive security elements and conventions. Differences in device capabilities and requirements make it difficult to create a global defence mechanism. To deploy DDoS attacks, attackers exploit these distinct qualities. The Intrusion Detection Systems (IDS), and Intrusion Prevention Systems (IPS) can provide help in preventing intrusion attacks.
3. **Privacy:** A few sensors jointly gather important any sensitive information, and it can be an easy task for a hacker to obtain it. An example of such event is described in [14], regarding a smart home activity arrangement by a home network traffic.
4. **Unattended:** Some IoT devices are special purpose devices, such as Implantable Medical Devices (IMDs). These types of devices have been operating in an unprecedented physical world for a long time without human mediation. It is extremely difficult to apply security computing and monitor if these devices are remotely hacked. In [15] authors proposed a lightweight, stable execution environment for these types of devices.
5. **Mobile:** Several IoT devices are portable and switch from network to network. As an example, a smart vehicle that collects street data when driving from one place to another. If the attacker injects the code by mobile devices, the device configuration or activity is changed. However, the change of device configuration is very difficult special when the network portability is configured on a device.

2.3 DDoS

DDoS stands for 'Distributed Denial-of-Service' and it is a kind of DoS (Denial-of-Service) where the intruder performs a attack through several locations from different sources simultaneously. DoS attacks are most driven by directing or shutting down a specific resource, and one method of operation is to exploit a system deficiency and cause failure of processing or saturation of system resources.

The authors in [16] claim that battery, computation, memory and radio transmission capability are limited in IoT devices. In this way, it is not easy to enforce security actions that involve a massive communication stack and more computing resources. Authors also suggest the usage of machine learning techniques, that is important for finding the vulnerability and security threats in IoTs.

The authors in [17] proposed a DDoS machine learning detection system that would include one pre-trained module to detect suspicious activities inside virtual machines and another online learning module to revise the pre-trained module. The structure is tested against TCP SYN, ICMP, DNS reflection and SSH brute-force attacks on nine separate machine learning algorithms and described as machine learning highlights. The finding result is the 93% accuracy by using the supervised approach in machine learning algorithm such as Naïve Bayes, SVM and Decision Tree.

Pattern discovery can be an instrument that identifies attacks by recognizing the signature of known attacks. Pattern position systems are often used as a virus detection system. Snort detecting the attack by using the attack signature is one of the good

detecting systems proposed in [18]. In sum, Payload Inspection and Machine Learning-based behaviour detection are the two feasible approaches for DDoS detection.

2.4 Machine Learning

Several machine learning techniques have been used to detect DDoS attacks. Each approach is distinguishing between the distinctive DDoS attacks and different results that are based on the data properties of the algorithm. A one-of-a-kind solution with a range of features to recognize all kinds of DDoS attacks is still not available. Due to massive amount of network data, it is difficult to recognize if the generated data is done by legitimate users or from real-time attack. Peter et al. [19] tests show that the Long Short-Term Memory Recurrent Neural Network (LSTM RNN) deep learning approach gives impressive results for detecting a DDoS attack in a network. The choice of supervised or unsupervised machine learning algorithms depends on specific parameters, such as the volume and structure of information and the form of DDoS. Five supervised machine learning approaches for detecting DDoS attack in IoT are briefly described below:

1. **K-Nearest Neighbours:** KNN [20] could be an effective and robust classification algorithm. KNN is known as an 'Instance-based Learner', which implies that the memorization of algorithm relies on continuous training experiences. KNN is a paradigm of machine learning that build on labelled dataset of the sampling data (x, y) and predicts the relationship between x and y. The main purpose is to learn the function $h : x \rightarrow y$ to predict the undetectable understanding of the target x, $h(x)$.
2. **Decision Tree:** The Decision Tree [21] is a well-known machine learning algorithm used to classify unknown data from trained data. A decision tree may be either a binary or non-binary tree that includes a root, internal and leaf node. All perceptions are placed in the root node, and each of the inner nodes holds the testing of features.
3. **Support Vector Machines (SVM):** In machine learning, support vector machines [22] are administered learning models with related learning calculations that dissect information utilized for characterization and relapse investigation. Given a lot of preparing precedents, each set apart as having a place with either of two classes, a SVM preparing calculation constructs a model that doles out new guides to one classification or the other, making it a non-probabilistic double direct classifier.
4. **Naïve Bayes Classifier:** Based on the Bayes Hypothesis, the Naïve Bayes can be a simple probabilistic classifier that is useful to large datasets [23]. When the features within the datasets are independent of each other, the Naïve Bayes model is easy to build, being a classifier that provides a speedy performance.
5. **Logistic Regression:** This model [24] is a broadly utilized statistical model that, in its fundamental shape, utilizes a logistic calculation to display a binary dependent variable. In regression analysis, logistic regression is assessing the parameters of a strategic model.

3 Approach

The aim of this work is to detect DDoS attacks through machine learning in cellular network via the packets generated by IoT devices. Here, we elaborate about the basis of

our proposed solution and describe the steps taken to be able to experiment and validate our hypothesis.

3.1 Design Phase

Our proposed method captures packets from Serving Gateway (SGW) and performs packet inspection to recognize malicious packets by extracting the features that can indicate a DDoS attack. After that, machine learning classification algorithms can segregate between normal and abnormal packets. If the packet is classified as normal traffic, it will be forward through the network and reaches the IoT application server. If it is considered abnormal and further verified as an attack, the device's info is forwarded to the Identity Management System (IDMS), which is the responsible for the temporary or permanent block of devices meaning that a device will not be able to connect to the network. For further explanation on how the proposed model detects DDoS attacks, we need to describe how the packets travel from IoT devices to IoT application servers. In a core 4G cellular network generally, the user packet transfers from the eNodeB to the SGW. The packets are then forwarded from SGW towards the Packet Gateway (PGW) which afterwards forwards these packets towards the application server. In this packet, the eNodeB attaches another IP packet that has a GTP header, which will provide information elements that can help us foresee if an attack is imminent (Fig. 2).

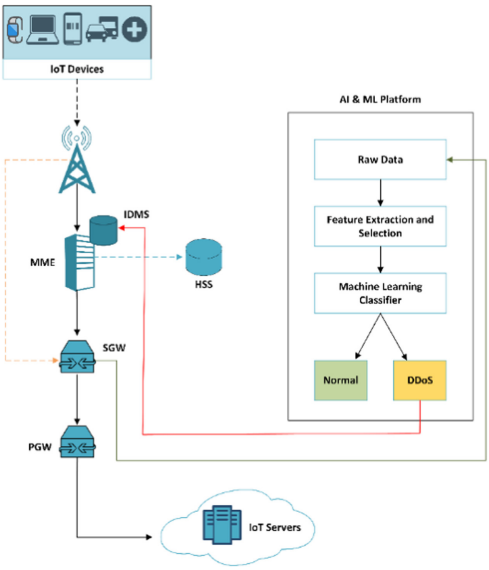

Fig. 2. Proposed method for anomaly detection

3.2 Implementation and Experimental Phase

We use different tools and packages for generating normal and DDoS traffic and analyse patterns by using machine learning technology. The tools and packages that we are using are described below:

3.2.1 Data Collection

1. **Wireshark**[1]**:** is an open-sourced and free packet analyser software that is used for analysis, network troubleshooting, communication and software protocol development. Wireshark is using the Qt widget toolkit that is implemented in the interface by using *pcap* to capture the packets. We use Wireshark in our work for capturing packets that coming from IoT devices.

Normal Dataset: For generating normal traffic, in our Secure 5G4IoT lab[2] in OsloMet, we had mobile devices (e.g., smartphones and raspberry Pi with IoT boards) connected through Wi-Fi to a mobile gateway. This gateway has a programmable SIM card that allows to connect to our test cellular network using consumer available hardware and open-source software.

DDoS Dataset: For DDoS traffic we use the CICDDoS2019[3] dataset, due to access and time restraints given the COVID-19 pandemic, available for machine learning research. This dataset is based on simulation and dated between 2016 to 2019. For this study we select this dataset as it provides a comprehensive analysis or various type of DDoS attacks.

3.2.2 Feature Extraction

To distinguish between DDoS and normal IoT traffic, we need to indicate the packet features that are selected for machine learning classification. Protocol type, port, source and destination IP and packet length have been used for recognizing most DDoS attacks. We have chosen the characteristics below to differentiate between ordinary traffic and DDoS [25, 26].

1. **Packet Size:** Under a timestamp, DDoS disperses a large number of packets, and these packets are smaller compared to an ordinary packet. Rohan et al. [26] maintain that the DDoS bundle is less than 100 bytes, while the normal operating bundle is between 100 and 1200 bytes. However, for the TCP SYN attack, the DDoS packet estimate is set at 58, 60 and 174 bytes.
2. **Packet Time Interval:** The interval between parcels in a DDoS attack is close to zero [26].
3. **Packet Size Variance:** For the most part, parcels of assault activity have the same estimate, while regular traffic has different packet measurements [26].

[1] **Wireshark:** www.wireshark.com.
[2] **Secure 5G4IoT Lab:** https://5g4iot.vlab.cs.hioa.no/.
[3] **CICDDoS2019 Dataset:** https://www.unb.ca/cic/datasets/ddos-2019.html.

4. **Protocol Type:** Two protocols (TCP and UDP) have been used for attack operations, allowing us to focus on them for our work.
5. **Destination IP:** IoT devices communicate with many expected target numbers and seldom modify their target IP over time. This highlight can also display a DDoS attacks. Inside a short timestamp, a single gadget interaction with a range of specific targets shows an attack [26].

3.2.3 Data Processing

When dealing with data pre-processing, some techniques need to be considered:

1. **Missing Values:** It is very difficult to handle the missing values in machine learning because it could create an incorrect prediction for any model. The null values and respective entries are then removed.
2. **Transformation:** The arrangement of the data collected might not be appropriate for modelling. As illustrated by the CRISP-DM method [27], in such cases, the type of data should be changed in such a way that the information can be integrated into the models at that point. Here, a few data features have been converted to numeric or float type.
3. **Labelling:** Our dataset represents the two types of classes: first packets with length below 100 packets size are represented with 1 (meaning an anomaly) and other length of packets are represented with 0 (normal traffic). Second type of data, if a packet has a length between 50 to 70 and 160 to 180 it is then represented with 1, and if the packet does not fit those intervals, it is represented with 0.
4. **Dataset Splitting:** Datasets are divided into two subsets; training and testing. The split data is divided in 70/30 ratio. The *train_test_split* helper method is used from scikit-learn library for splitting of data. With this approach, training data is divided into two parts, training and validation. The training set is used to train the model in start, then validation set is used to estimate the performance of data.

To help with this process, the following tools can be used:

1. **Python**[4]**:** It is a general purpose, high level open-source programming language. Ease of learning, efficient code and easy communication are some of the features of Python, many researchers use this programming language in this field. We use this language for machine learning experiment.
2. **Scikit-Learn**[5]**:** As open-source machine learning tool for Python programming language. It is a simple tool for data analysis and data mining. Scikit-learn consists of different algorithms for implementation for supervised and unsupervised learning.

[4] **Python:** www.python.org.
[5] **Scikit-Learn:** www.scikit-learn.org.

3.2.4 Evaluation

Evaluation is a very crucial part for understanding the performance of a chosen model. This part defines the performance of the models. Below are described the various metrics used in this study.

1. **Accuracy:** It is one way to describe your model performance by the count of correct and incorrect classifier elements. These correct and incorrect values are represented in the values of accuracy, which determines the performance of classifier.
2. **Precision:** For assessing the performance of learning model accuracy is not enough. The accuracy gives an idea that the model is trained correctly, but it does not give the detailed information of the specific application. For that reason, we use the other performance measurements, such as precision, which is the rate of correctly classified true positive or true negative.
3. **Recall:** Recall is measuring how many actual positive values are measured or recalled.

4 Results

Experiments were carried out to verify the performance and accuracy of the classifier for various combinations and sizes of data. Our DDoS detection test was based on TCP SYN attack due to time constraints. We use two different threshold scenarios in both datasets: First, we set the threshold of packets below 100 bytes. Second threshold is set between 50 and 70 or between 160 and 180 bytes.

4.1 First Threshold - Packet Length Below 100 Bytes

4.1.1 Normal Scenario

Table 1 shows how accurate they performed with normal traffic. SVM performed well in this experiment as it gives no anomaly, however rest of the algorithms show possible anomalies.

Table 1. Performance metrics: first threshold – normal scenario

Classifier name	Accuracy (%)	Precision (%)	Recall (%)
K-NN	83.70	86.72	83.70
SVM	82.55	86.15	82.55
Naïve-Bayes	75.51	82.70	75.51
Decision Tree	83.70	86.72	83.70
Logistic Regression	82.55	86.15	82.55

4.1.2 DDoS Scenario

Table 2 shows how accurate they performed with DDoS traffic. The SVM shows no anomaly in this experiment, but other classifiers show anomalies. Naïve Bayes performs well to detect the anomaly but with lower accuracy.

Table 2. Performance metrics: first threshold – DDoS scenario

Classifier name	Accuracy (%)	Precision (%)	Recall (%)
K-NN	98.21	97.54	98.21
SVM	98.19	96.42	98.19
Naïve-Bayes	97.98	97.07	97.98
Decision Tree	98.21	97.57	98.21
Logistic Regression	98.18	97.22	98.18

4.2 Second Threshold – Packet Length Between 50 and 70 Bytes and Between 160 and 180 Bytes

4.2.1 Normal Scenario

Table 3 shows how accurate they performed with normal traffic. The KNN performed well in this experiment. In the normal dataset SVM, decision tree and logistic regression give no anomaly, but KNN and Naïve Bayes classifiers show possible anomalies.

Table 3. Performance metrics: second threshold – normal scenario

Classifier name	Accuracy (%)	Precision (%)	Recall (%)
K-NN	84.52	80.85	84.52
SVM	84.25	70.98	84.25
Naïve-Bayes	82.61	78.82	82.61
Decision Tree	84.51	80.70	84.51
Logistic Regression	84.22	78.33	84.22

4.2.2 DDoS Scenario

Table 4 shows how accurate they performed with DDoS traffic. In this dataset K-NN and SVM show good results with this threshold.

Table 4. Performance metrics: second threshold – DDoS scenario

Classifier name	Accuracy (%)	Precision (%)	Recall (%)
K-NN	99.19	98.85	99.19
SVM	99.19	98.39	99.19
Naïve-Bayes	98.92	98.69	98.92
Decision Tree	99.20	98.90	99.20
Logistic Regression	99.18	98.77	99.18

5 Discussion

This study was conducted to analyse the action, performance and utilization of the machine learning algorithms in the context of intrusion detection system. Researchers and industry are working to find out good solutions in the field of machine learning and artificial intelligence for intrusion detection and prevention. However, different business partners and researchers often find it difficult to obtain excellent quality datasets to test and evaluate their machine learning models for detection of threats. This problem is the main motivation of this study, and basis for research questions. To ensure that the experiment is carried out in an appropriate manner, all classifiers were chosen based on literature review. The results were evaluated using a set of performance metrics, including precision, accuracy and recall.

5.1 First Threshold

In 1st threshold, the KNN performance metrics are fair. It achieved 83.70% accuracy with precision of 86.72%. When trained with the CICDDoS2019 dataset, KNN shows much better precision and accuracy scores averaging 98%. SVM gives 98.19% result, but it does not find any anomaly in this dataset. Naïve Bayes gives 97.98%, logistic regression gives 98.18% and decision tree gives 98.21% accuracy. Overall, for this threshold, KNN is the best classifier.

5.2 Second Threshold

If we talk about the second threshold in the CICDDoS2019 dataset, KNN also gives the good precision and accuracy scores averaging 99%. SVM gives 99.19% accuracy. Naïve Bayes gives 98.92%, logistic regression 99.18% and decision tree gives 99.20%. In this threshold, KNN also turns out to be the classifier that performs the best.

5.3 Evaluation

Throughout this work we were able to conclude that some classifiers are more sensitive hence producing results that were not the expected ones. A reason for these discrepancies is most likely due to the thresholds chosen. An establishment of more robust thresholds

that are more adequate to our studied scenario is needed to provide more reliable results. Nevertheless, we were able to detect the attack given by the supervised and labelled dataset even the with differences in performance depending on the classifier. In a real-life context and given the early stage of the implementation, the result data would have been sent to the corresponding security expert team in a telecom operator for further validation.

6 Conclusion

This work was set to look into the issues of IoT protection from the point of view of the Cellular Network in terms of the security challenges. Recognizing attacks within the cellular network is not the same as recognizing attacks in an IP network. For instance, a sudden increase in the acceptance of packets in a single node from the number of distinctive MME nodes in the case of IoT could suggest an attack, as IoT devices do not transmit packets in a very high frequency.

This work presents an overview of how other researchers discuss the issue of discovery of intrusion detection with the use of machine learning. This has provided a much better understanding that how different algorithms work and can help understand how to mitigate the propagation of DDoS attacks. In addition, it also provides an understanding of which algorithms are commonly used to deal with problems in this area.

Normal and DDoS datasets have been used and with five classification methods, such as KNN, Decision Tree, and Naïve Bayes, SVM and logistic regression, we analysed their performance as to detect possible attacks. The focus was on TCP attacks, as this protocol is commonly used to launch an attack, and due to time constraints, we just focus on the SYN attack.

Our primary focus not only for this work but also in our research is to provide ways to develop and provide a secure environment towards device-driven solutions that could enhance the quality of life of an elderly person at their own homes, but the proposal herein presented can be applicable to other verticals in which IoT can be a beneficial added factor.

The point was to identify DDoS attacks within the context of the cellular network in this proposed work, and the aim was to propose an arrangement that could lead to a specific use in the future. Subsequently, the strategy recommends a full-scale DDoS detection technique within the cellular network, and offline data has been used for training and testing of the model. We would like to recommend that this methodology be tested in a true research setting for future work. In addition, this strategy focused only on the TCP SYN flood. To secure IoT devices and services in the future, we would like to incorporate all potential DDoS attacks. We hope that this study starts as a basis to create a helping tool for telecom operators that could be used in the future to detect DDoS and other types of attacks in a more automated fashion.

Acknowledgement. This paper is a result of the H2020 Concordia project (https://www.concordia-h2020.eu) which has received funding from the EU H2020 programme under grant agreement No 830927. The CONCORDIA consortium includes 23 partners from industry and other organizations such as Telenor, Telefonica, Telecom Italia, Ericsson, Siemens, Airbus, etc. and 23 partners from academia such as CODE, university of Twente, OsloMet, etc.

References

1. Patel, K.K., Patel, S.M., et al.: Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. Int. J. Eng. Sci. Comput. **6**(5), 6122–6131 (2016)
2. Chen, S., et al.: A vision of IoT: applications, challenges, and opportunities with china perspective. IEEE Internet Things J. **1**(4), 349–359 (2014)
3. Lee, I., Lee, K.: The Internet of Things (IoT): applications, investments, and challenges for enterprises. Bus. Horizons **58**(4), 431–440 (2015)
4. Rawat, P., Singh, K.D., Bonnin, J.M.: Cognitive radio for M2M and Internet of Things: a survey. Comput. Commun. **94**, 1–29 (2016)
5. Ericsson: IoT connections outlook. https://www.ericsson.com/en/mobility-report/reports/june-2020/iot-connectionsoutlook
6. Jejdling, F. (Ericsson): Ericsson Mobility Report. https://www.ericsson.com/en/mobility-report/reports
7. Stackpole, B.: Symantec Security Summary, June 2020. COVID-19 attacks continue and new threats on the rise. https://symantec-enterprise-blogs.security.com/blogs/featurestories/symantec-security-summary-june-2020
8. Khoi, N.M., et al.: IReHMo: an efficient IoT-based remote health monitoring system for smart regions. In: 2015 17th International Conference on E-health Networking, Application and Services (Health-Com), pp. 563–568. IEEE (2015)
9. Gope, P., Hwang, T.: BSN-care: a secure IoT-based modern healthcare system using body sensor network. IEEE Sens. J. **16**(5), 1368–1376 (2015)
10. Heer, T., et al.: Security challenges in the IP-based Internet of Things. Wirel. Pers. Commun. **61**, 527–542 (2011)
11. Van der Elzen, I., van Heugten, J.: Techniques for detecting compromised IoT devices. University of Amsterdam (2017)
12. Saqlain, J.: IoT and 5G: history evolution and its architecture their compatibility and future (2018)
13. Zhou, W., et al.: The effect of IoT new features on security and privacy: new threats, existing solutions, and challenges yet to be solved. IEEE Internet Things J. **6**(2), 1606–1616 (2018)
14. Copos, B., et al.: Is anybody home? Inferring activity from smart home network traffic. In: 2016 IEEE Security and Privacy Workshops (SPW), pp. 245–251. IEEE (2016)
15. Noorman, J., et al.: Sancus: low-cost trustworthy extensible networked devices with a zero-software trusted computing base. In: 22nd fUSENIXg Security Symposium (fUSENIXg Security 2013), pp. 479–498 (2013)
16. Xiao, L., et al.: IoT security techniques based on machine learning: how do IoT devices use AI to enhance security? IEEE Signal Process. Mag. **35**(5), 41–49 (2018)
17. He, Z., Zhang, T., Lee, R.B.: Machine learning based DDoS attack detection from source side in cloud. In: 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), pp. 114–120. IEEE (2017)
18. Bakker, J.: Intelligent traffic classification for detecting DDoS attacks using SDN/OpenFlow (2017)
19. Bediako, P.K.: Long short-term memory recurrent neural network for detecting DDoS flooding attacks within TensorFlow implementation framework (2017)
20. Adeniyi, D.A., Wei, Z., Yongquan, Y.: Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. Appl. Comput. Inform. **12**(1), 90–108 (2016)
21. Tian, F., et al.: Research on flight phase division based on decision tree classifier. In: 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), pp. 372–375. IEEE (2017)

22. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
23. Patil, T.R., Sherekar, S.S.: Performance analysis of Naive Bayes and J48 classification algorithm for data classification. J. Comput. Sci. Appl. **6**(2), 256–261 (2013)
24. Wikipedia: Logistic regression. https://en.wikipedia.org/wiki/Logistic_regression
25. Oo, T.T., Phyu, T.: Analysis of DDoS detection system based on anomaly detection system. In: International Conference on Advances in Engineering and Technology (ICAET 2014), Singapore (2014)
26. Doshi, R., Apthorpe, N., Feamster, N.: Machine learning DDoS detection for consumer internet of things devices. In: 2018 IEEE Security and Privacy Workshops (SPW), pp. 29–35. IEEE (2018)
27. Cross-industry standard process for data mining. https://en.wikipedia.org/wiki/Cross-industry_standard_process_for_data_mining