



Article

A Two-Level Flow-Based Anomalous Activity Detection System for IoT Networks

Imtiaz Ullah *  and Qusay H. Mahmoud 

Department of Electrical, Computer, and Software Engineering Ontario Tech University, Oshawa, ON L1G 0C5, Canada; qusay.mahmoud@uoit.ca

* Correspondence: imtiaz.ullah@ontariotechu.net

Received: 2 March 2020; Accepted: 20 March 2020; Published: 23 March 2020



Abstract: The significant increase of the Internet of Things (IoT) devices in smart homes and other smart infrastructure, and the recent attacks on these IoT devices, are motivating factors to secure and protect IoT networks. The primary security challenge to develop a methodology to identify a malicious activity correctly and mitigate the impact of such activity promptly. In this paper, we propose a two-level anomalous activity detection model for intrusion detection system in IoT networks. The level-1 model categorizes the network flow as normal flow or abnormal flow, while the level-2 model classifies the category or subcategory of detected malicious activity. When the network flow classified as an anomaly by the level-1 model, then the level-1 model forwards the stream to the level-2 model for further investigation to find the category or subcategory of the detected anomaly. Our proposed model constructed on flow-based features of the IoT network. Flow-based detection methodologies only inspect packet headers to classify the network traffic. Flow-based features extracted from the IoT Botnet dataset and various machine learning algorithms were investigated and tested via different cross-fold validation tests to select the best algorithm. The decision tree classifier yielded the highest predictive results for level-1, and the random forest classifier produced the highest predictive results for level-2. Our proposed model Accuracy, Precision, Recall, and F score for level-1 were measured as 99.99% and 99.90% for level-2. A two-level anomalous activity detection system for IoT networks we proposed will provide a robust framework for the development of malicious activity detection system for IoT networks. It would be of interest to researchers in academia and industry.

Keywords: flow-based intrusion detection; internet of things; machine learning; cybersecurity; vulnerabilities

1. Introduction

Smart digital devices have become part of our daily lives. These systems improve the quality of life, make communication more accessible, and increase data transfer and information sharing. Security becomes significant and essential due to the development of IoT networks and the vulnerabilities that are available in these Internet of Things (IoT) devices. These vulnerabilities are technically difficult and economically very costly to remove from the existing systems. Cyber-attacks are increasing day-by-day, and their effect is becoming more devastating. Intrusion Detection System improves the cybersecurity by monitoring the network traffic for abnormal patterns. As a result, the Intrusion Detection System (IDS) became a critical aspect of protecting IoT networks. An IDS is a hardware device or software that monitors a system or network for malicious activity or policy violations. An IDS can only be useful if it generates timely, accurate alerts and provides useful, actionable information. Anomaly-based and misuse-based are typically focused and motivated detection techniques in the area of intrusion detection. Commercial products are usually preferred toward misuse detection techniques as compared to anomaly-based methods; therefore, anomaly-based approaches are still

considered to be an immature technology in the business tools. Data mining and machine learning are widely used techniques for classification and clustering in network security. The main challenge for the intrusion detection system is the capability to detect new attacks based on the previously observed events. The complexity used by attackers and the increase in zero-day attacks renders anomaly-based intrusion detection well suited to the current environment. Anomaly detection methodology can be used for fraud detection and medical diagnoses. Example of anomalies in IoT networks are alpha flows, flash crowd, port scan, network scan, outage events, worms, small packet size, or large packet size, TCP-based events (Invalid flag combinations, ACK, FIN, SYN, RST), UDP-based events, various amplification attacks, flow bomb, and mass ICMP. The capability of an IDS can be measured using multiple metrics. The IDS input and output mutual information ratio to the entropy of the input was used to develop a new evaluation matrix [1]. The matrix considers all the critical characteristics of the detection capability.

IoT devices become a pervasive part of our everyday lives. Healthcare, industrial control, retails, logistics, emergency services, traffic congestion detection, security, waste management, smart cities, smart homes, smart street lighting, and vehicle networks are the application areas of IoT devices [2]. The cybercrime destruction budget will hit \$6 trillion per year by 2020, and 50 billion IoT devices need protection by 2020, according to the cyber threat defense report 2017 [3]. A completely secure system is not possible, because there is no absolute security, human in the loop can make a mistake, most existing systems have security flaws, and misuses through the insiders and all type of intrusions are not known. Traditional security mechanisms are not appropriate for IoT networks, because the IoT platforms cover regular Internet, non-IP network, mobile network, cloud computing, fog computing, and sensor network. IoT devices also have limited processing and memory capability. Charles Mclellan [4] studied 345 security predictions from 49 various organizations and predicted IoT security threats as top rank threats for 2017. Flow-based anomaly detection is a novel methodology for detecting malicious activities. Flow-based intrusion detection only inspects the packet header to detect malicious activity. Usually, IDS uses stateful protocol analysis or in-depth packet inspection to identify abnormal activity in the network traffic [5]. Stateful protocol methods are protocol dependent and computationally very costly, while deep packet inspection becomes a bottleneck in a high-speed network. Research now is focusing on flow-based intrusion detection as an alternative to protect IP networks due to these limitations. Flow-based IDS uses a network stream to analyze the network traffic for malicious activity. Flow-based IDS inspects only packet header; therefore, flow-based anomaly detection technique faster than packet-based and state-full protocol examination. Figure 1 shows the types of flow-based intrusion detection systems.

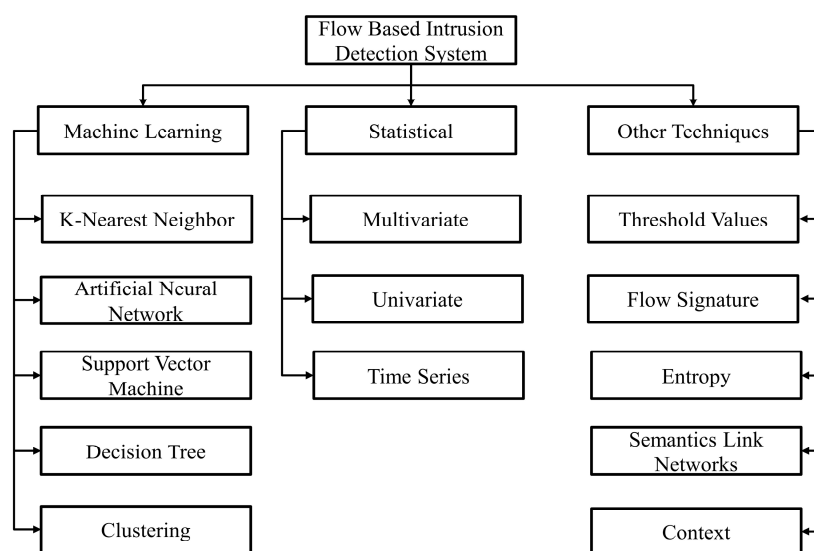


Figure 1. Classification of Flow-Based Intrusion Detection Systems.

In this paper, we propose a two-level anomalous activity detection system for IoT networks. The level-1 model classifies network traffic as normal or abnormal. If the level-1 model detects the flow as an anomaly, then the flow will be forwarded to the level-2 model for further classification to find the category or subcategory of the detected anomaly. Open issues in flow-based intrusion detection are: there are a limited number of publicly available datasets for flow-based intrusion detection. It is a critical requirement to develop a publicly available flow-based intrusion detection dataset that includes the latest regular network traffic and a variety of attacks traffic to evaluate flow intrusion detection techniques. Networks attacks embedded in packets cannot be detected via flow-based intrusion detection systems, such as XSS and SQL injection attacks. An adequate number of features are required to analyze the network flow. A solution is to compute new features from basic flow features. The flow transfer interval critically disturbs the performance of the intrusion detection system, and the detection engine might not be able to detect short-timed attacks if the flow transfer interval kept longer. We adopted a new IoT Botnet dataset and extracted the flow-based features for our proposed model.

The remainder of this paper is organized as follows: Section 2 discusses the related work. Section 3 describes the hierarchical architecture of IoT. Section 4 discusses our proposed two-level intrusion detection system for IoT networks. Section 5 presents a comparative study of the intrusion detection datasets and the IoT Botnet dataset used for the assessment of our proposed model. Section 6 presents a comparison of results, and Section 7 concludes the paper and offers ideas for future directions.

2. Related Work

Connecting IoT devices are getting more attention and significance; consequently, the frequency of cyber-attacks increased. Analyzing these attacks in network flow is a significant challenge for an intrusion detection system. Anomaly detection techniques have been a critical motivation for many researchers due to their potential in classifying new attacks. Launt et al. [6] proposed a prototype using discrete measure user login time, location, connection time, CPU time, IO activity, and protection violations to detect anomalous behavior in real-time. One of the essential requirements for IDS to detect a significant amount of intrusions [7]. Paxson [8] developed Bro, which is, in a way, an anomaly and signature-based IDS to monitor network activity in real-time to detect abnormal behavior. Handley and Paxson [9] developed several evasion techniques and countermeasures for network-based intrusion detection systems. Polymorphic attack techniques can efficiently evade a signature-based intrusion detection system. Still, anomaly-based IDS offer various protection techniques because the existing polymorphic methods cannot make the attack instance look normal [10].

IoT based systems required a secure and fast communication interface between the embedded device and the Internet. The study of intrusion detection in IoT has received a lot of attention due to the vulnerabilities and threats available in IoT networks. These vulnerabilities required new methodologies of intrusion detection to fill these gaps. Ullah and Mahmoud [11] proposed a two-level hybrid model for anomalous activity detection for IoT networks. Their model uses a flow-based methodology at level-1 and RFE for feature selection, synthetic minority over-sampling technique (SMOTE) for oversampling, and edited nearest neighbors (ENN) for cleaning the dataset at level-2. Their model achieved high accuracy and provided a robust framework for detecting malicious activity in IoT networks. Currently, the attackers using sophisticated tools to launch more dangerous attacks very quickly with a small amount of prerequisite technical knowledge of the system. An intrusion detection framework for smart-grid was proposed in [12]. The training and testing time can be significantly decreased by using a simplified set of features [13]. A filter-based feature selection model proposed [13,14] removing redundant and irrelevant features from the ISCX and NSL-KDD datasets. Umer et al. [5] deliberate techniques and challenges for flow-based intrusion detection. They categorize the available methods into a general-purpose, attack based, technique-based, and scenario-based.

Sun et al. [15] proposed a flow-based IDS using TCP flow as the main criterion to detect and classify malicious behaviors using Benford's law. Their analysis shows that each attack has a unique pattern, and using these patterns, they successfully categorize the normal flow and abnormal flow.

Many intrusion detection techniques rely on packets inspection and resource overwhelming in a high-speed network. AlEroud et al. [16] used contextual information to generate semantics links to detect suspicious network flows. Their prototype achieved promising results in detecting known and unknown attacks. These semantic links improve the detection rates for multi-steps attacks. Initially, the semantic links were static, but these semantic links can be dynamically updated to include suspicious network flow to detect novel attacks. Compromised machines are the main building blocks for various unlawful activities on the Internet. An intrusion detection system plays an essential role in detecting compromised devices. SSH is a typical application for identifying compromised machines because SSH can be used for remote server administration, and an adversary can get root access to the compromised machine via SSH service. Hofstede et al. [17] developed open-source flow-based software to detect compromised hosts and, subsequently, SSH dictionary attacks. They demonstrate their model in a lab environment and achieved satisfactory results. Satoh et al. [18] used flow features and machine learning algorithms to detect stealthy dictionary attacks over SSH. They used a campus network to evaluate their model and achieved improved accuracy with acceptable computational complexity. Zhang et al. [19] present a survey of anomaly-based intrusion detection techniques in a computer network. They divided all the methods into four categories: finite state machines, machine learning, classification, and statistical and describe available methodologies in these categories. Table 1 presents a summary of intrusion detection systems for IoT networks.

Table 1. Intrusion Detection System (IDS) for the Internet of Things (IoT).

References	Security Threat	Detection Method	Validation Strategy	Placement Strategy
[20]	Man-in-the-middle	Anomaly-based	Simulation	Centralized
[21]	Multiple Conventional attacks	Signature-based	————	————
[22]	Routing attack	Specification-based	Simulation	Hybrid
[23]	DoS	Specification-based	Simulation	————
[24]	DoS	Signature-based	Empirical	Centralized
[25]	Routing attack	Anomaly-based	Simulation	Centralized
[26]	Routing attack	Hybrid	Simulation	Hybrid
[27]	Multiple Conventional attacks	Anomaly-based	————	————
[28]	Multiple Conventional attacks	Signature-based	Hypothetical	Centralized
[29]	Multiple Conventional attacks	Specification-based	Empirical	Hybrid
[30]	Multiple Conventional attacks	Signature-based	Empirical	Distributed
[31]	DoS	Anomaly-based	Simulation	Distributed
[32]	Routing attack and MITM	Hybrid	Simulation	————
[33]	Routing attack	Hybrid	Simulation	Distributed
[34]	Conventional	Anomaly-based	Empirical	————
[35]	Multiple Conventional attacks	Anomaly-based	————	Hybrid
[36]	Routing attack	Anomaly-based	Simulation	Hybrid

A limited number of IoT based intrusion detection dataset available as a result, several researchers, used traditional network datasets for evaluating their model for IoT networks. Many researchers used the KDD, UNSW-NB15, and CICIDS2017 dataset for evaluating their model for IoT networks. A traditional network dataset is not appropriate for IoT networks, because the IoT platform covers regular Internet, non-IP network, mobile network, cloud computing, fog computing, and sensor network. IoT devices also have limited processing and memory capability. Moustafa et al. [37] developed an IoT botnet dataset via legitimate and emulated IoT networks. Smart fridge, Weather station, Motion-activated lights, Smart thermostat, and Remotely activated garage door IoT services implemented using the Node-Red tool. A typical smart home configuration was designed that contains five IoT devices operated locally and connected to the cloud infrastructure via a Node-Red tool to generate normal traffic. Meidan et al. [38] developed an IoT botnet dataset. The dataset was generated using nine commercial IoT devices. The dataset contains 115 network features that provide a reliable normal, and malicious network flow. The dataset contains separate benign network traffic for each commercial device to ensure the normal network behavior of each device for the training purpose.

Jadidi and Sheikhan [39] proposed a flow-based anomaly detection system using Multi-Layers Prescription (MLP) via a gravitational search algorithm. Their model achieved an accuracy of 99.40% for known and unknown attacks. Casas et al. [40] proposed a flow-based anomaly detection system through multiple unsupervised density-based and sub-space clustering to classify a malicious flow. Tran and Jiang [41] proposed a hardware-based flow detection engine using the block-based neural network (BBNN). Their model achieved an improved detection rate with relatively better computation time due to the hardware-based implementation. Kanda et al. [42] developed an anomaly detection technique using entropy-based principal component analysis. They used the MAWI dataset for evaluating their model and achieved better results as compared with other anomaly detectors for the same dataset. An enhanced technique for anomaly detection proposed in [43] using KNN via a probability density function. The optimization techniques that they used are particle swarm optimization, harmony search, gravitational search, and bat algorithm to determine the most excellent value of K. Duque and Omar [44] proposed an intrusion detection model via K-means clustering and achieved a better detection rate, low false-positive, and false negative. They categorized network attacks into denial of service attacks, penetration attacks, and scanning attacks. Their proposed model analyzes and identifies the signature and behaviors of these attacks. Kumar and Kumar [45] combined the genetic algorithm and neural network methodology to detect anomalies. KDD99 and ISCX2012 datasets were used for evaluation. The ISCX2012 dataset achieved a 97% highest detection rate using the Multi-Layer Perception algorithm. Decision tree and neural network techniques were used [46] with reduced features of the KDD99 dataset. The model achieved a 95% detection rate for normal traffic and 92% for the intrusion.

3. IoT Hierarchical Architecture

IoT improved human life using pervasive technology. Smart infrastructure is an application area of the Internet of Things. Figure 2 shows a hierarchical architecture of IoT services in smart infrastructure. The architecture consists of four layers: application, service, network communication, and perception layer.

3.1. Application Layer

The application layer provides personalized facilities to the end-user. The user can access these services through a smart device. The application layer delivers smart services, as requested by the user. Table 2 presents commonly used protocols used for the development and deployment of IoT applications. The standard functionality of the application layer is to send and receive data to/from the IoT service layer, smart device administration, and override the IoT lower layers.

Table 2. List of Application Layer Protocols for IoT.

1	Constrained Application Protocol (CoAP)
2	Data Distribution Service (DDS)
3	Advanced Message Queuing Protocol (AMQP)
4	Message Queue Telemetry Transport (MQTT)
5	Extensible Messaging and Presence Protocol (XMPP)

3.2. Service Layer

The service layer provides computational power to the smart infrastructure, as well as analyzing and storing large amounts of IoT devices data. There are four categories of IoT services: identity services, information aggregation services, collaborative-aware services, and ubiquitous services. Functions performed by the service layer are cloud-based administration, application-to-application communication, anomaly detection, send/receive data to/from the application layer, and network layer.

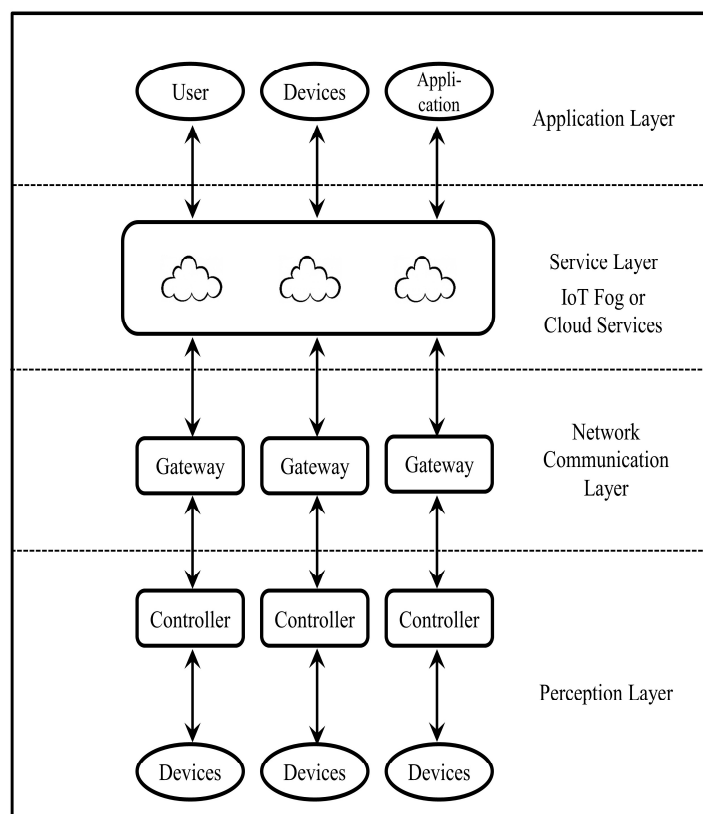


Figure 2. IoT Hierarchical Architecture.

3.3. Network Communication Layer

The network communication layer is responsible for communication between the perception layer and the service layer. The network communication layer includes communication protocols, network infrastructure, mobile networks, and the Internet. The protocols used by the network layer for communication with the IoT device are WirelessHART (Wireless Highway Addressable Remote Transducer Protocol), Bluetooth, ZigBee, WiMAX, Ethernet, IEEE 802.15.4, Wi-Fi, etc. The Protocols used by the network layer for communication with the service layer are SOAP (Simple Object Access Protocol), REST (Representational State Transfer), LWM2M (Lightweight Machine to Machine), DDS, CoAP, IBM MessagSight, Node.js, WebSocket, Ethernet, and HTTP. The network communication layer function includes device-to-device communication, anomaly detection, receiving data from the upper and lower layer, and pushing data to the upper layer and lower layer.

3.4. Perception Layer

The perception layer consists of devices that monitor, identify, and modify the physical world. The perception layer collects information from the physical world using IoT devices and provides communication with short-range IoT devices. Raspberry PI, Apple Watch, Samsung Gear, Atlas Fitness Tracker, Garmin vivofit, Gadgeteer, etc. are commonly used IoT platforms. Commonly used perception layer protocols are Wi-Fi, ZigBee, Bluetooth, and Ethernet. Standard functions performed by the perception layer includes a device to device communication, anomaly detection, receive and send data from/to network communication layer, data encryption, and decryption.

4. Two-Level Flow-Based Anomalous Activity Detection System

Anomaly detection techniques become more attractive because of their capability to detect new attacks, but the main challenge is to identify these attacks correctly and on time. Some attacks in IoT networks appear to be identical, as in traditional Internet, but the scale and simplicity of these

attacks are more significant due to the limited protection of IoT devices. In this paper, we proposed a two-level intrusion detection model for smart infrastructure. In our proposed approach, the network layer is responsible for training the level-1 model, analyze the IoT network traffic for malicious activity and forward the network flow to IoT Fog or Cloud layer to classify the category of an anomaly when the input flow detected as abnormal. The IDS at the network communication layer is responsible for classifying the network flow as normal or anomalous. The IDS at level-1 allows detecting a malicious activity near the smart infrastructure by optimizing local parameters. The level-1 IDS accelerates attack detection by getting the parameter locally. If the detected flow is anomalous, then the level-1 model transfers the flow to the level-2 model to identify the category of attack. Figure 3 shows a two-level flow-based intrusion detection model for IoT networks.

Step-1: Analyze the network flow from each IoT device. Use Wireshark or TCPdump to intercept network packets.

Step-2: Extract network features from the collected flow.

Step-3: Select flow-based features and useful general features.

Step-4: Generate training and test data. We used 60% data for training and 40% for testing.

Step-5: Train the level-1 model for binary classification. The proposed model uses a decision tree classifier at level-1 to classify the network traffic into normal traffic or anomalous traffic.

Step-6: Train the level-2 model to identify the category or subcategory of malicious flow. Random forest classifier used at level-2 to categorize the network flow received from the level-1 model. Figure 4 shows a detailed view of the two-level flow-based anomalous activity system for IoT networks.

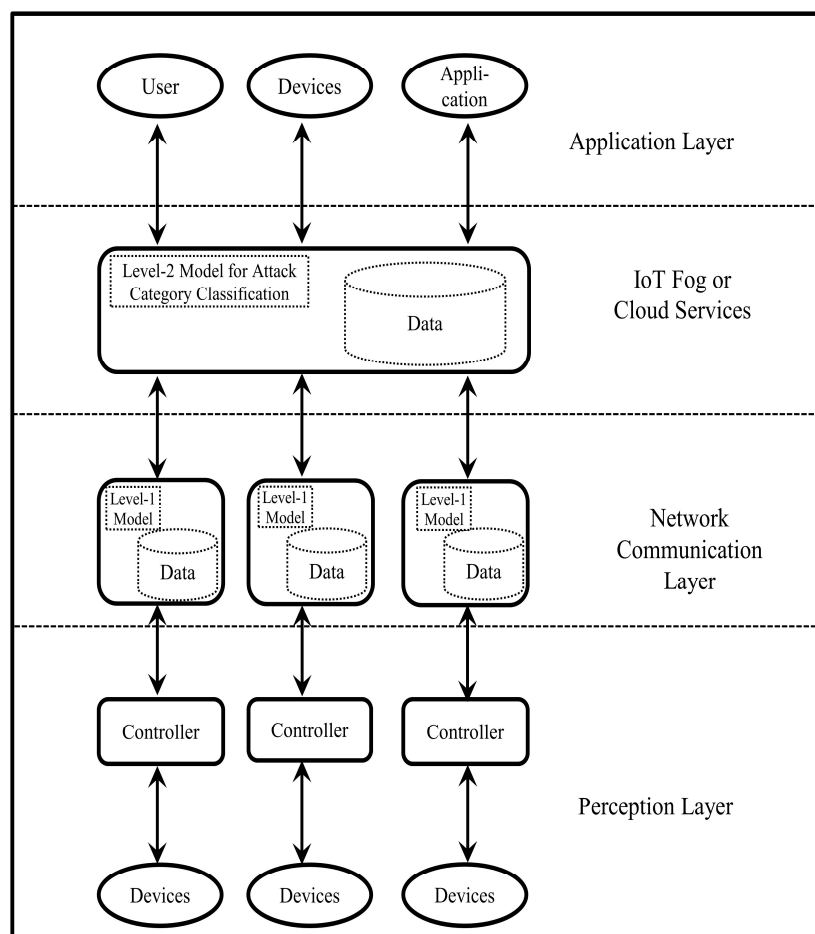


Figure 3. IoT Flow-Based Intrusion Detection Model.

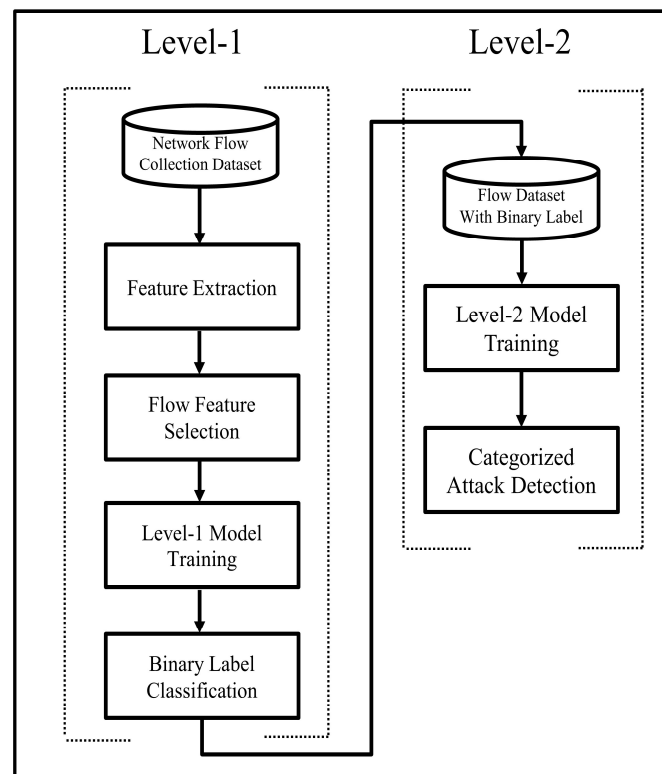


Figure 4. Two-Level Flow-Based Anomalous Activity Detection System Intrusion Detection in IoT Networks.

5. Evaluation of Datasets

Many challenges are associated with datasets used for intrusion detection. Dataset information comes from three sources: CPU/memory usage, low-level system information or user command, and network data packets. KDD99 [47] and NSL-KDD [48] are the most commonly used datasets for validating intrusion detection systems. KDD99 is a comprehensively unbalanced dataset, because 80% of the dataset instances are anomalous, while a real network contains 99.99% normal traffic. User-2-Root and Root-2-Local attacks are infrequent attack classes in the KDD99 dataset. Training and testing dataset contains duplicate records that can produce bias results for DoS and normal classes. These deficiencies were removed in the NSL-KDD dataset. The training dataset contains 78% redundant instances, while the testing dataset contains 75% redundant records [48]. These redundant records cause machine learning algorithms to prevent learning uncommon instances and are biased towards more frequent instances. The redundant records also had an impact on the testing data, which produce biased results for a technique that has improved the detection rate on frequent records.

A systematic approach was used to generate the ISCX2012 dataset for FTP, POP3, IMAP, SSH, SMTP, and HTTP protocol [49]. The ISCX2012 dataset consists of seven days of network activity. Agents were programmed to simulate user behavior. The agents effectively mimic regular user activity. Multi-stage attack scenarios conceded to generate the abnormal portion of the dataset. These attacks launched in real-time via human assistance. The essential properties of the ISCX2012 dataset are (i) Realistic network traffic: a dataset should reveal network traffic properties. For this reason, traffic must behave and look as realistic as possible. (ii) Labeled dataset: The labeled dataset plays an important role in evaluating various machine learning mechanisms. (iii) Due to the increasing number of incidences, scope, fact, and density of attacks in recent years, diverse intrusion scenarios are considered during the creation of the dataset. The ISCX dataset has a minimal number of attacks, and a small number of network features were selected, and some of these features are not usable for machine learning algorithms.

The UNSW-NB15 [50] dataset was generated by the Australian defense force academy University of New South Wales. UNSW-NB15 reflect modern attack scenario and in-depth structured network traffic information. The main properties of the UNSW-NB15 dataset are 45 unique IP addresses, three networks, nine attack families, and 49 features. The UNSW-NB15 has several advantages compared to earlier intrusion detection datasets. It contains modern malicious and normal behavior, training and testing set probability distribution are similar, and flow-based features from the packet header and payload completely replicate the network packets. The CICIDS2017 [51] dataset that was developed at the Canadian Institute for Cybersecurity (CIC), University of New Brunswick Canada uses diverse attack scenarios, e.g., Brute Force, Heartbleed, Botnet, DoS, DDoS, Web, and Infiltration attack. The dataset traffic was generated for five days. The CICIDS2017 dataset's main characteristics are complete network configuration; total traffic; labeled instances; complete interaction; complete capture; a more significant number of available protocols; attack diversity; heterogeneity; large number feature set; and, metadata. Table 3 shows a list of publicly available datasets that were used for anomaly detection systems or misuse detection systems.

Table 3. Intrusion Detection Domain Datasets.

	Abbreviation	Dataset Name
1	UNM	University of New Mexico Dataset
2	BSM98	DARPA 1998 BSM Files
3	DARPA98	DARPA 1998 TCPDump Files
4	BSM99	DARPA 1999 BSM Files
5	DARPA99	DARPA 1999 TCPDump Files
6	KDD99	KDD99 Dataset
7	UNIXDS	UNIX User Dataset
8	IES	Internet Exploration Shootout Dataset
10	NSL-KDD	NSL KDD Dataset
11	ISCX2012	Information Security Centre of Excellence 2012 evaluation dataset
12	UNSW-NB15	University of New South Wales IDS Dataset
13	CICIDS2017	Canadian Institute for Cybersecurity IDS dataset
14	BoT-IoT	BoT IoT Dataset

Malicious third parties have targeted IoT networks. A network intrusion detection system is a realistic protection mechanism for IoT networks. For this purpose, a well-structured IoT dataset is required to develop and validate the credibility of the intrusion detection system for IoT networks. An IoT Botnet dataset that was developed at the University of New South Wales Canberra, Australia [37] addresses the drawback of previously developed datasets, i.e., complete network information, diverse attack scenario, and accurate labeling. They extracted 46 network features and two label features for the IoT Botnet dataset. The IoT Botnet dataset has a limited number of flow features. The IoT Botnet dataset and the related Pcap files are publicly available [52]. We adopted a new dataset from the IoT Botnet dataset [52] using a network traffic flow analyzer [53] to increase and improve the number of flow features and the network features. The output of the traffic analyzer is the CSV file with 83 network traffic features. We select 17 flow-based and general features from our adopted IoT Botnet dataset.

New techniques and detection algorithms for intrusion detection required a well-designed dataset for IoT networks, which contains comprehensive modern normal network traffic as well as diverse intrusions scenario with an in-depth structured network traffic information. Our proposed IoT botnet dataset will provide a reference point for identifying anomalous activity across the IoT networks. The

IoT Botnet dataset can be accessed from [54]. The label feature identifies the network traffic as normal flow or anomalous flow, while the category label feature classifies the network traffic as Normal, DDoS, DoS, Reconnaissance, or Theft. The sub-category label feature identifies the network traffic as Normal, DDoS-HTTP, DDoS-TCP, DDoS-UDP, DoS-HTTP, DoS-TCP, DoS-TCP, OS-Fingerprint, Service-Scan, Keylogging, and Data-Exfiltration. Maximizing the detection rate by minimizing false positive and false negative rates is the primary objective of our proposed intrusion detection system model. Table 4 shows the flow-based features that were selected from the adopted IoT Botnet dataset.

Table 4. Selected Flow Features IoT Botnet Dataset.

Feature Name	Feature Name
Src IP	Flow IAT Min
Dst IP	Fwd IAT Tot
Dst Port	Fwd IAT Mean
Protocol	Subflow Fwd Pkts
Flow Duration	Subflow Fwd Byts
Flow Byts/s	Subflow Bwd Pkts
Flow Pkts/s	Subflow Bwd Byts
Flow IATMean	Label
Flow IAT Std	Cat
Flow IAT Max	Sub_Cat

6. Results and Discussion

Supervised learning is a pivotal step in building a computational model for intrusion detection systems. Supervised learning divides the instances into previously determined classes. In this paper, we used the IoT Botnet dataset for our proposed model. The IoT Botnet dataset required a preprocessing process because the data types and the format of some attributes are not appropriate for machine learning algorithms. Therefore, before training, the IoT Botnet dataset non-numeric features have been converted into numeric features. The IoT Botnet dataset was normalized using the column normalization.

$$\text{Column Normalization} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$F \text{ Score} = \frac{2(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (5)$$

Performance evaluation matrices for the machine learning algorithm is the critical step in measuring the success rate of a classification algorithm. However, various approaches have been widely utilized in the literature to select a better choice for prediction in specific applications. We used the Accuracy, Precision, Recall, and F score to evaluate our proposed model. Accuracy shows the number of correct predictions to the total number of predictions. Precision shows the intrusion predicted by IDS is an actual intrusion. High precision means a low false positive alarm. Precision is always measured for the intrusion class. Recall determining projected intrusion contrasted with all

intrusion present. The F score provides an enhanced measure of IDS correctness, and it is the harmonic mean of precision and recall. A high F-value means high precision and recall. After preprocessing, the IoT Botnet dataset was ready to use for machine learning algorithms. We select python sklearn package [55] to develop our proposed model. Figure 5 shows the training and testing instances for the IoT Botnet dataset. Training and validation are two essential components of machine learning. Initially, we used random sampling for generating training and testing sets, so each time the proposed model selects a random sample for training and testing. A fully-grown tree is likely to overfit the data; therefore, we split the fully-grown decision tree by using a maximum depth selection for our proposed model. Another way to reduce overfitting is to use random forest classifier, which combines multiple machine learning algorithms to obtain better predictive performance, so we used a random forest classifier at level-2 for our proposed model. In machine learning, different validation tests are used to evaluate the success rates of a proposed model. Jackknife test is efficient and reliable because of its unique results. However, the computational time of the jackknife test is an issue when using a large dataset. Therefore, we used the K-fold cross-validation test to evaluate the performance of our proposed model to minimize the running time. K-fold cross-validation test uses first k-1 folds for training and the last fold for testing. We checked the validation of the proposed model using 3, 5, 10, and 15-fold cross-validation test.

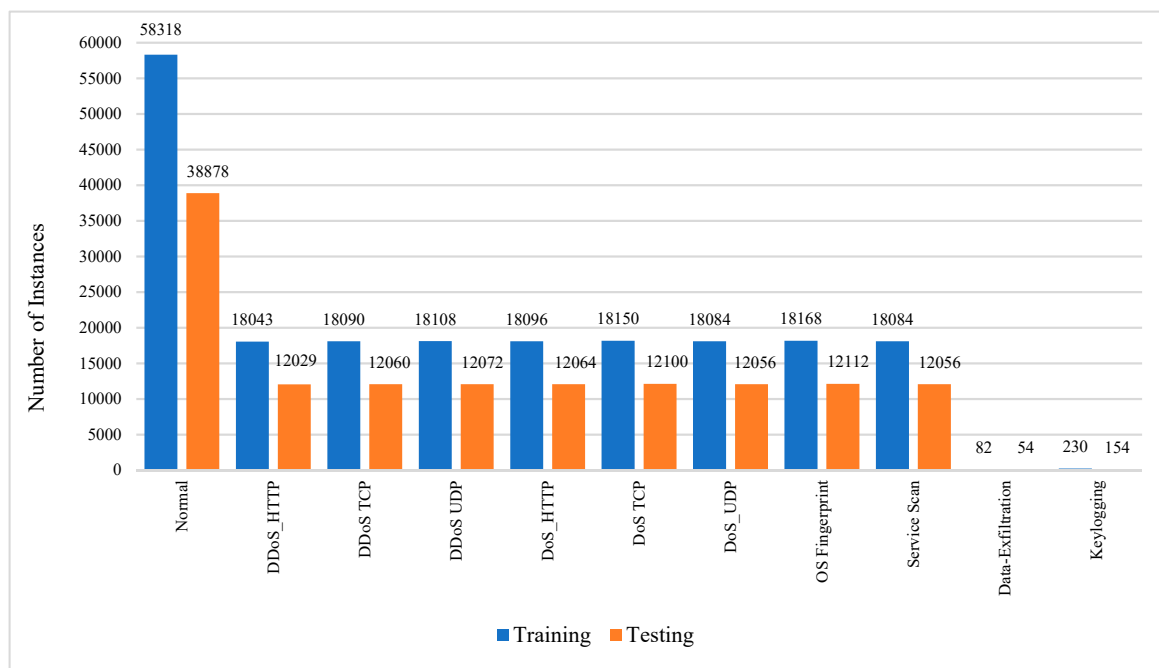


Figure 5. IoT Botnet Dataset Training and Testing Distribution.

6.1. Level-1 Model Prediction

The primary function of our proposed level-1 model is to classify the input flow correctly as normal flow or anomalous flow. A high Accuracy, Precision, Recall, and F score value will be achieved when a perfect classifier selected for intrusion detection. Therefore, a performance study accomplished to select the most excellent classification technique. We used a decision tree classifier for a level-1 model, which provides effective prediction performance. The proposed level-1 model achieved 100% Accuracy, Precision, Recall, and F score for the IoT Botnet dataset, as shown in Table 5; Table 6. We used various 3, 5, 10, and 15 folds cross-validation tests to evaluate the Accuracy, Precision, Recall, and F score of our proposed model. Table 7 shows 10-fold cross-validation tests for the IoT Botnet dataset. The result remains constant for binary classification of the IoT Botnet dataset.

Table 5. Contingency Matrix Label.

Actual Class	Predicted Class		
	Classified→	Normal	Anomaly
	Normal	29,103	0
Anomaly	0	72,265	

Table 6. Precision, Recall, and F Score.

Attack Type	Accuracy	Precision	Recall	F Score
Normal	100	100	100	100
Anomaly	100	100	100	100
Average	100	100	100	100

Table 7. Precision, Recall, and F Score for 10-Fold Cross-Validation.

No. of Fold	Precision	Recall	F Score
3-Fold	99.99	99.99	99.99
5-Fold	99.99	99.99	99.99
10-Fold	99.99	99.99	99.99
15-Fold	99.99	99.99	99.99
Average	99.99	99.99	99.99

6.2. Level-2 Model Prediction

The second function of our proposed model is to find the attack category or subcategory. We evaluate diverse classification methodologies to select an optimum classification technique for the proposed level-2 model. We used random forest classifier for the level-2 model. Figure 6 shows the Precision, Recall, and F Score of various classifiers using the IoT Botnet dataset. When the level-1 model classifies a network flow as an anomaly, then the flow will be forwarded to the level-2 model for further classification to find a category or subcategory of the detected anomaly. The Contingency Matrix for the category of IoT Botnet Dataset is presented in Table 8. The results of our prediction model for categorized and subcategorized attacks are presented in Tables 9 and 10 for the IoT Botnet dataset, respectively.

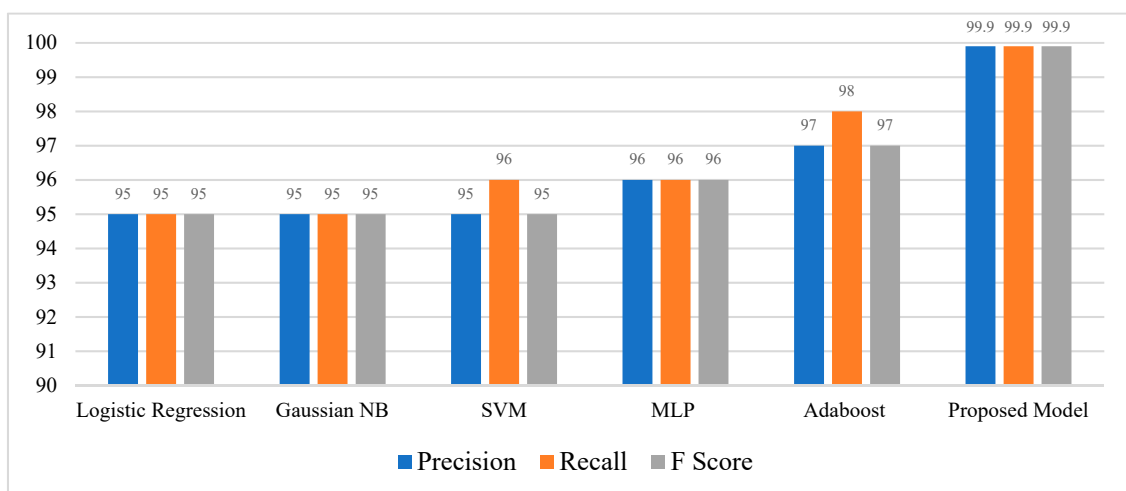


Figure 6. Precision, Recall, and F Score for Top Rank Classifier for IoT Botnet Dataset.

Table 8. Contingency Matrix for Category.

		Predicted Class				
		Classified→	Normal	DDoS	DoS	Recon-
Actual Class	Normal	29,103	0	0	0	0
	DDoS	0	26,754	8	6	1
	DoS	0	9	26,813	5	0
	Reconnaissance	0	6	0	18,137	1
	Theft	0	0	0	1	159

Table 9. Accuracy Precision, Recall, and F Score for Category.

Attack Type	Accuracy	Precision	Recall	F Score
Normal	100	100	100	100
DDoS	99.90	99.60	99.60	99.60
DoS	99.60	99.90	99.60	99.60
Reconnaissance	99.99	99.99	99.99	99.99
Theft	99.60	99.90	99.60	99.90
Average	99.80	99.90	99.75	99.80

Table 10. Accuracy, Precision, Recall, and F Score for Sub-Category.

Attack Type	Accuracy	Precision	Recall	F Score
Normal	100	100	100	100
DDoS-HTTP	99.00	99.00	99.00	99.00
DDoS-TCP	99.99	99.99	99.99	99.99
DDoS-UDP	99.99	99.99	99.99	99.99
DoS-HTTP	99.00	99.00	99.00	99.00
DoS-TCP	99.99	99.99	99.99	99.99
DoS-UDP	99.99	99.99	99.99	99.99
OS Fingerprint	99.90	99.90	99.90	99.90
Service_Scan	99.90	99.90	99.90	99.90
Keylogging	99.99	91.00	95.00	93.00
Data-Exfiltration	99.99	93.00	97.00	95.00
Average	99.80	99.00	99.00	98.80

Accuracy, Precision, Recall, and F score for normal network traffic measured as 100%, DoS, and theft as 99.90, DDoS as 99.60, and reconnaissance as 99.99, as shown in Table 9. An average Accuracy, Precision, Recall, and F score value of 99.80% was achieved for the category and 99% for the Subcategory label. We used different K-folds cross-validation tests to check the overfitting of our proposed level-2 model. Tables 11 and 12 show 10-fold cross-validation tests for category and subcategory IoT Botnet Dataset, respectively. The average Accuracy, Precision, Recall, and F score for category and subcategory of the IoT Botnet dataset remains unchanged, as shown in Tables 11 and 12. Table 13 presents a comparison of the intrusion detection system for IoT networks. Many researchers used KDD, UNSW-NB15, and CICIDS2017 dataset for evaluating their model for IoT networks, as shown in Table 13. In this paper, we empirically select 17 flow and general features from the IoT botnet dataset to validate the proposed flow-based intrusion detection model for IoT networks. We have a considerable number of instances for many attack classes, but the features are limited in number. We did not use feature selection at this stage for our proposed flow-based anomalous activity detection model for IoT networks. One way to increase the accuracy of some attack classes is to generate more flow-based features and use a feature selection algorithm to remove irrelevant and redundant attributes from the dataset to improve the predictive power of a classification algorithm.

Table 11. Accuracy, Precision, Recall, and F Score for 10-Fold Cross-Validation.

Attack Type	Accuracy	Precision	Recall	F Score
Normal	100	100	100	100
DDoS	99.60	99.50	99.50	99.50
DoS	99.60	99.80	99.70	99.80
Reconnaissance	99.99	99.99	99.99	99.99
Theft	99.70	99.80	99.70	99.70
Average	99.80	99.90	99.75	99.80

Table 12. Accuracy, Precision, Recall, and F Score for 10-Fold Cross-Validation.

Attack Type	Accuracy	Precision	Recall	F Score
Normal	100	100	100	100
DDoS-HTTP	99.20	99.10	99.10	99.10
DDoS-TCP	99.90	99.99	99.99	99.99
DDoS-UDP	99.99	99.99	99.99	99.99
DoS-HTTP	99.10	99.10	99.10	99.10
DoS-TCP	99.99	99.99	99.99	99.99
DoS-UDP	99.99	99.99	99.99	99.99
OS Fingerprint	99.90	99.90	99.90	99.90
Service_Scan	99.90	99.90	99.90	99.90
Keylogging	99.99	91.00	95.20	93.20
Data-Exfiltration	99.99	93.10	97.10	95.10
Average	99.80	99.00	99.00	98.80

Table 13. Intrusion Detection System IoT Networks.

	Attack Type	Dataset	Year	Classification	Accuracy	Precision	Recall	F-Score
[56]	DDoS	CICIDS2017	2020	Binary	99	99.20	99.30	99.30
[57]	Multi	UNSW-NB15	2020	Multi Class	98.8	—	—	—
[58]	Sybil	—	2020	Binary	94	93.5	90.1	97
[59]	Multi	Generated	2019	Multi Class	98	95.9	96.3	96.1
[60]	Multi	BoT-IoT	2019	Multi Class	95	—	—	—
[61]	Multi	NSL-KDD	2019	Multi Class	98.19	—	—	—
[62]	Multi	UNSW-NB15	2019	Multi Class	84	—	—	—
[63]	Multi	NSL-KDD	2019	Multi Class	85.81	—	—	—
[64]	Multi	Generated	2018	Multi Class	—	98	99	99
[65]	Multi	UNSW-NB15	2018	Multi Class	—	89	87	88
[66]	Multi	CICIDS2017	2019	Multi Class	99.80	98.68	92.76	95.04
[67]	Multi	NSL-KDD	2019	Multi Class	86.95	89.56	87.25	88.41
[68]	Multi	DS2OS	2019	Multi Class	99.4	—	—	—
[69]	Binary	Generated	2018	Binary	—	97.7	97.7	97.7
[70]	—	WiLab	2020	—	97.95	—	—	—
[71]	Multi	—	2020	Multi Class	74.98	74.80	77.07	—
[72]	Blackhole	Generated	2019	Binary	—	97.2	—	97
Proposed	Multi	IoT Botnet	2020	Multi Class	99.80	99	99	98.80

7. Conclusion and Future Work

In this work, we have developed a two-level flow-based anomalous activity detection system for IoT networks. The level-1 model categorizes the input flow as normal or anomalous, while the level-2 model classifies the category or subcategory of the anomalous flow. The flow-based features were empirically selected from the IoT Botnet dataset. Our proposed model achieved Accuracy, Precision, Recall, and F score of 100% for the level-1 model and 99.90 for the level-2 model. We used various K-fold cross-validation tests to evaluate the Accuracy, Precision, Recall, and F score of our proposed model. Our proposed model accomplished Accuracy, Precision, Recall, and F score of 100% for level-1 and 99.90 for the level-2 model via various K-fold cross-validation.

In future work, we plan to analyze the current model to improve the accuracy of insignificant classes of the IoT Botnet dataset.

Author Contributions: Writing—Original draft preparation, I.U.; Supervision, and Writing—review and editing, Q.H.M. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), DDG-2018-00014.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gu, O.; Fogla, P.; Dagon, D.; Lee, W.; Škorić, B. Measuring intrusion detection capability: An information-theoretic approach. In Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security, ASIACCS'06, Taipei, Taiwan, 21–24 March 2006; Volume 2006, pp. 90–101. [CrossRef]
2. Lee, K.; Kim, M.S.; Shim, P.; Han, I.; Lee, J.; Chun, J.; Cha, S. Technology advancement of laminate substrates for mobile, IoT, and automotive applications. In Proceedings of the 2017 China Semiconductor Technology International Conference, CSTIC 2017, Shanghai, China, 12–13 March 2017; pp. 1–4. [CrossRef]
3. Cyberthreat Defense Report. 2017. Available online: http://www.crn.com/sites/default/files/ckfinderimages/userfiles/images/crn/custom/Webroot_Q3_2017_CyberEdge_Cyberthreat_Defense_Report.pdf (accessed on 5 September 2019).
4. McLellan, C. Cybersecurity in 2017: A Roundup of Predictions. Available online: <https://www.techrepublic.com/article/cybersecurity-in-2017-a-roundup-of-predictions> (accessed on 20 September 2019).
5. Umer, M.F.; Sher, M.; Bi, Y. Flow-based intrusion detection: Techniques and challenges. *Comput. Secur.* **2017**, *70*, 238–254. [CrossRef]
6. Launt, T.F.; Jagannathan, R. A prototype real-time intrusion-detection expert system. In Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA, USA, 18–21 April 1988; pp. 59–66. [CrossRef]
7. Axelsson, S. The base-rate fallacy and the difficulty of intrusion detection. *ACM Trans. Inf. Syst. Secur.* **2000**, *3*, 186–205. [CrossRef]
8. Paxson, V. Bro: A system for detecting network intruders in real-time. In Proceedings of the 7th USENIX Security Symposium, San Antonio, TX, USA, 26–29 January 1998; Volume 31, pp. 2435–2463. [CrossRef]
9. Handley, M.; Paxson, V.; Kreibich, C. Network intrusion detection: Evasion, traffic normalization, and end-to-end protocol semantics. In Proceedings of the 10th USENIX Security Symposium, Washington, DC, USA, 13–17 August 2001.
10. Fogla, P.; Sharif, M.; Perdisci, R.; Kolesnikov, O.; Lee, W. Polymorphic blending attacks. In Proceedings of the 15th USENIX Security Symposium, Vancouver, BC, Canada, 31 July–4 August 2006; pp. 241–256.
11. Ullah, I.; Mahmoud, Q.H. A two-level hybrid model for anomalous activity detection in IoT networks. In Proceedings of the 16th IEEE Annual Consumer Communications & Networking Conference, CCNC 2019, Las Vegas, NV, USA, 10–14 January 2019; pp. 1–6. [CrossRef]
12. Ullah, I.; Mahmoud, Q.H. An intrusion detection framework for the smart grid. In Proceedings of the IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, Canada, 30 April–3 May 2017; pp. 1–5. [CrossRef]
13. Ullah, I.; Mahmoud, Q.H. A hybrid model for anomaly-based intrusion detection system. In Proceedings of the 2017 IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; pp. 2160–2167. [CrossRef]
14. Ullah, I.; Mahmoud, Q.H. A filter-based feature selection model for anomaly-based intrusion detection systems. In Proceedings of the 2017 IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; pp. 2151–2159. [CrossRef]
15. Sun, L.; Anthony, T.S.; Xia, H.Z.; Chen, J.; Huang, X.; Zhang, Y. Detection and classification of malicious patterns in network traffic using Benford's law. In Proceedings of the 9th Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, Malaysia, 12–15 December 2017; pp. 864–872. [CrossRef]

16. AlEroud, A.F.; Karabatis, G. Queryable semantics to detect cyber-attacks: A flow-based detection approach. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *48*, 207–223. [[CrossRef](#)]
17. Hofstede, R.; Pras, A.; Sperotto, A.; Rodosek, G.D. Flow-based compromise detection. *IEEE Secur. Priv.* **2018**, *16*, 82–89. [[CrossRef](#)]
18. Satoh, A.; Nakamura, Y.; Ikenaga, T. A flow-based detection method for stealthy dictionary attacks against Secure Shell. *J. Inf. Secur. Appl.* **2015**, *21*, 31–41. [[CrossRef](#)]
19. Zhang, W.; Yang, Q.; Geng, Y. A survey of anomaly detection methods in networks. In Proceedings of the 1st International Symposium on Computer Network and Multimedia Technology, Wuhan, China, 18–20 January 2009; pp. 1–3. [[CrossRef](#)]
20. Cho, E.J.; Kim, J.H.; Hong, C.S. Attack model and detection scheme for botnet on 6LoWPAN. In *Asia-Pacific Network Operations and Management Symposium*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 515–518. [[CrossRef](#)]
21. Liu, C.; Yang, J.; Chen, R.; Zhang, Y.; Zeng, J. Research on immunity-based intrusion detection technology for the Internet of Things. In Proceedings of the 7th International Conference on Natural Computation, Shanghai, China, 26–28 July 2011; pp. 212–216. [[CrossRef](#)]
22. Le, A.; Loo, J.; Luo, Y.; Lasebae, A. Specification-based IDS for securing RPL from topology attacks. In Proceedings of the IFIP Wireless Days, Niagara Falls, ON, Canada, 10–12 October 2011; Volume 1, pp. 1–3. [[CrossRef](#)]
23. Misra, S.; Venkata Krishna, P.; Agarwal, H.; Saxena, A.; Obaidat, M.S. A learning automata based solution for preventing distributed denial of service in internet of things. In Proceedings of the 2011 IEEE International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, Dalian, China, 19–22 October 2011; pp. 114–122. [[CrossRef](#)]
24. Kasinathan, P.; Pastrone, C.; Spirito, M.A.; Vinkovits, M. Denial-of-Service detection in 6LoWPAN based Internet of Things. In Proceedings of the IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications, Lyon, France, 7–9 October 2013; pp. 600–607. [[CrossRef](#)]
25. Wallgren, L.; Raza, S.; Voigt, T. Routing attacks and countermeasures in the RPL-based internet of things. *Int. J. Distrib. Sens. Netw.* **2013**, *2013*. [[CrossRef](#)]
26. Surendar, M.; Umamakeswari, A. InDReS: An Intrusion Detection and response system for Internet of Things with 6LoWPAN. In Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 23–25 March 2016; pp. 1903–1908. [[CrossRef](#)]
27. Gupta, A.; Pandey, O.J.; Shukla, M.; Dadhich, A.; Mathur, S.; Ingle, A. Computational intelligence based intrusion detection systems for wireless communication and pervasive computing networks. In Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research, Madurai, India, 26–28 December 2013; pp. 1–7. [[CrossRef](#)]
28. Kasinathan, P.; Costamagna, G.; Khaleel, H.; Pastrone, C.; Spirito, M.A. Demo: An IDS framework for internet of things empowered by 6LoWPAN. In Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security, Berlin, Germany, 4–8 November 2013; pp. 1337–1339. [[CrossRef](#)]
29. Amaral, J.P.; Oliveira, L.M.; Rodrigues, J.J.P.C.; Han, G.; Shu, L. Policy and network-based intrusion detection system for IPv6-enabled wireless sensor networks. In Proceedings of the 2014 IEEE International Conference on Communications, Sydney, Australia, 10–14 June 2014; pp. 1796–1801. [[CrossRef](#)]
30. Chen, J.; Chen, C. Design of complex event-processing IDS in internet of things. In Proceedings of the 6th International Conference on Measuring Technology and Mechatronics Automation, Zhangjiajie, China, 10–11 January 2014; pp. 226–229. [[CrossRef](#)]
31. Lee, T.-H.; Wen, C.-H.; Chang, L.-H.; Chiang, H.-S.; Hsieh, M.-C. A lightweight intrusion detection scheme based on energy consumption analysis in 6LoWPAN. In *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*; Springer: Dordrecht, The Netherlands, 2014; Volume 260, pp. 1205–1213. [[CrossRef](#)]
32. Krimmling, J.; Peter, S. Integration and evaluation of intrusion detection for CoAP in smart city applications. In Proceedings of the 2014 IEEE Conference on Communications and Network Security, San Francisco, CA, USA, 29–31 October 2014; pp. 73–78. [[CrossRef](#)]
33. Cervantes, C.; Poplade, D.; Nogueira, M.; Santos, A. Detection of sinkhole attacks for supporting secure routing on 6LoWPAN for Internet of Things. In Proceedings of the 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), Ottawa, ON, Canada, 11–15 May 2015; pp. 606–611. [[CrossRef](#)]

34. Summerville, D.H.; Zach, K.M.; Chen, Y. Ultra-lightweight deep packet anomaly detection for Internet of Things devices. In Proceedings of the 2015 IEEE 34th international performance computing and communications conference (IPCCC), Nanjing, China, 14–16 December 2015; pp. 1–8. [CrossRef]
35. Thanigaivelan, N.K.; Nigussie, E.; Kanth, R.K.; Virtanen, S.; Isoaho, J. Distributed internal anomaly detection system for Internet-of-Things. In Proceedings of the 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2016; pp. 319–320. [CrossRef]
36. Pongle, P.; Chavan, G. Real time intrusion and wormhole attack detection in internet of things. *Int. J. Comput. Appl.* **2015**, *121*, 1–9. [CrossRef]
37. Koroniotis, N.; Moustafa, N.; Sitnikova, E.; Turnbull, B. Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. *Future Gener. Comput. Syst.* **2019**, *100*, 779–796. [CrossRef]
38. Meidan, Y.; Bohadana, M.; Mathov, Y.; Mirsky, Y.; Shabtai, A.; Breitenbacher, D.; Elovici, Y. N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Comput.* **2018**, *17*, 12–22. [CrossRef]
39. Jadidi, Z.; Muthukumarasamy, V.; Sithirasenan, E. Metaheuristic algorithms based Flow Anomaly Detector. In Proceedings of the 2013 19th Asia-Pacific Conference on Communications (APCC), Denpasar, Indonesia, 29–31 August 2013; pp. 717–722. [CrossRef]
40. Casas, P.; Mazel, J.; Owezarski, P. Unsupervised network intrusion detection systems: Detecting the unknown without knowledge. *Comput. Commun.* **2012**, *35*, 772–783. [CrossRef]
41. Tran, Q.A.; Jiang, F.; Hu, J. A real-time NetFlow-based intrusion detection system with improved BBNN and high-frequency field programmable gate arrays. In Proceedings of the 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, Liverpool, UK, 25–27 June 2012; pp. 201–208. [CrossRef]
42. Kanda, Y.; Fontugne, R.; Fukuda, K.; Sugawara, T. ADMIRE: Anomaly detection method using entropy-based PCA with three-step sketches. *Comput. Commun.* **2013**, *36*, 575–588. [CrossRef]
43. Costa, K.A.P.; Pereira, L.A.M.; Nakamura, R.Y.M.; Pereira, C.R.; Papa, J.P.; Xavier Falcão, A. A nature-inspired approach to speed up optimum-path forest clustering and its application to intrusion detection in computer networks. *Inf. Sci.* **2015**, *294*, 95–108. [CrossRef]
44. Duque, S.; Bin Omar, M.N. Using data mining algorithms for developing a model for Intrusion Detection System (IDS). *Proced. Comput. Sci.* **2015**, *61*, 46–51. [CrossRef]
45. Kumar, G.; Kumar, K. A multi-objective genetic algorithm based approach for effective intrusion detection using neural networks. In *Intelligent Methods for Cyber Warfare*; Springer: Cham, Switzerland, 2015.
46. De Campos, L.M.L.; De Oliveira, R.C.L.; Roisenberg, M. Network intrusion detection system using data mining. *Commun. Comput. Inf. Sci.* **2012**, *311*, 104–113. [CrossRef]
47. Lee, W.; Stolfo, S.J. A framework for constructing features and models for intrusion detection systems. *ACM Trans. Inf. Syst. Secur.* **2000**, *3*, 227–261. [CrossRef]
48. Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE symposium on computational intelligence for security and defense applications, Ottawa, ON, Canada, 8–10 July 2009; pp. 1–6. [CrossRef]
49. Shiravi, A.; Shiravi, H.; Tavallaee, M.; Ghorbani, A.A. Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *Comput. Secur.* **2012**, *31*, 357–374. [CrossRef]
50. Moustafa, N.; Slay, J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Proceedings of the 2015 military communications and information systems conference (MilCIS), Canberra, Australia, 10–12 November 2015; pp. 1–6. [CrossRef]
51. Sharafaldin, I.; Lashkari, A.H.; Ghorbani, A.A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. In Proceedings of the 4th International Conference on Information Systems Security and Privacy, ICISPP 2018, Funchal, Portugal, 22–24 January 2018; pp. 108–116. [CrossRef]
52. Nickolaos, K.; Nour, M.; Sitnikova, E.; Benjamin, T. The BoT-IoT Dataset. 2018. Available online: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/bot_iot.php (accessed on 5 September 2019).
53. Lashkari, A.H.; Gil, G.D.; Mamun, M.S.I.; Ghorbani, A.A. Characterization of tor traffic using time based features. In Proceedings of the 3rd International Conference on Information Systems Security and Privacy, ICISPP 2017, Porto, Portugal, 19–21 February 2017; pp. 253–262. [CrossRef]

54. Ullah, I.; Mahmoud, Q.H. IoT-Botnet Dataset. 2020. Available online: <https://sites.google.com/view/iotbotnetdataset> (accessed on 1 March 2020).
55. Pedregosa, F.; Varoquaux, G.; Gramfort, A. Scikit-learn: Machine Learning in Python. *J Mach Learn Res.* **2011**, *12*, 2825–2830.
56. Roopak, M.; Tian, G.Y.; Chambers, J. An intrusion detection system against DDoS attacks in IoT networks. In Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January 2020; pp. 0562–0567. [[CrossRef](#)]
57. Lawal, M.A.; Shaikh, R.A.; Hassan, S.R. Security analysis of network anomalies mitigation schemes in IoT networks. *IEEE Access* **2020**, *8*, 43355–43374. [[CrossRef](#)]
58. Murali, S.; Jamalipour, A. A lightweight intrusion detection for sybil attack under mobile RPL in the internet of things. *IEEE Internet of Things J.* **2020**, *7*, 379–388. [[CrossRef](#)]
59. Van Huong, P.; Thuan, L.D.; Hong Van, L.T.; Hung, D.V. Intrusion detection in IoT systems based on deep learning using convolutional neural network. In Proceedings of the 6th NAFOSTED Conference on Information and Computer Science (NICS), Hanoi, Vietnam, 12–13 December 2019; pp. 448–453. [[CrossRef](#)]
60. Ibitoye, O.; Shafiq, O.; Matrawy, A. Analyzing adversarial attacks against deep learning for intrusion detection in IoT networks. In Proceedings of the 2019 IEEE Global Communications Conference, Waikoloa, HI, USA, 9–13 December 2019; pp. 1–6.
61. Alrashdi, I.; Alqazzaz, A.; Alharthi, R.; Aloufi, E.; Zohdy, M.A.; Ming, H. FBAD: Fog-based attack detection for IoT healthcare in smart cities. In Proceedings of the 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 10–12 October 2019; pp. 515–522. [[CrossRef](#)]
62. Hanif, S.; Ilyas, T.; Zeeshan, M. Intrusion detection in IoT using artificial neural networks on UNSW-15 dataset. In Proceedings of the 2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT), Charlotte, NC, USA, 6–9 October 2019; pp. 152–156. [[CrossRef](#)]
63. Illy, P.; Kaddoum, G.; Moreira, C.M.; Kaur, K.; Garg, S. Securing fog-to-things environment using intrusion detection system based on ensemble learning. In Proceedings of the 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 15–18 April 2019; pp. 1–7. [[CrossRef](#)]
64. Anthi, E.; Williams, L.; Słowi, M.; Theodorakopoulos, G.; Burnap, P. A supervised intrusion detection system for smart home IoT devices. *IEEE Internet Things J.* **2019**, *6*, 9042–9053. [[CrossRef](#)]
65. Alomari, K.M.; Shaalan, K. Towards machine learning based IoT intrusion detection service. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*; Springer: Cham, Switzerland, 2018; pp. 580–585.
66. Lee, J.H.; Park, K.H. GAN-based imbalanced data intrusion detection system. *Pers. Ubiquitous Comput.* **2019**. [[CrossRef](#)]
67. Li, Y.; Xu, Y.; Liu, Z.; Hou, H.; Zheng, Y.; Xin, Y.; Zhao, Y.; Cui, L. Robust detection for network intrusion of industrial IoT based on multi-CNN fusion. *Meas. J. Int. Meas. Confed.* **2020**, *154*, 107450. [[CrossRef](#)]
68. Hasan, M.; Islam, M.M.; Zarif, M.I.I.; Hashem, M.M.A. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. *Internet Things* **2019**, *7*, 100059. [[CrossRef](#)]
69. Anthi, E.; Williams, L.; Burnap, P. Pulse: An adaptive intrusion detection for the internet of things. In *2018 Living in the Internet of Things; Cybersecurity of the IoT*; London, UK, 2018; pp. 1–4. [[CrossRef](#)]
70. Kumar, R. A binary classification approach for time granular traffic in SDWMN based IoT networks. In Proceedings of the 2020 International Conference on Communication Systems & NETWORKS (COMSNETS), Bangalore, India, 7–11 January 2020; pp. 531–534.
71. Chaabouni, N.; Mosbah, M.; Zemmari, A.; Sauvignac, C. An intrusion detection system for the One M2M service layer based on edge machine learning. In *International Conference on Ad-Hoc Networks and Wireless*; Springer: Cham, Switzerland, 2019; Volume 11803.
72. Thamilarasu, G.; Chawla, S. Towards deep-learning-driven intrusion detection for the internet of things. *Sensors* **2019**, *19*, 1977. [[CrossRef](#)] [[PubMed](#)]

