



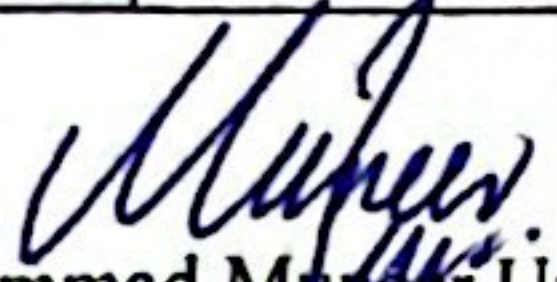
KOHAT UNIVERSITY OF SCIENCE & TECHNOLOGY

Kohat 26000, Khyber Pakhtunkhwa, Pakistan Ph # 0922-554563-554565, Fax #. 554556

Institute of Computing

CERTIFICATE REGARDING INCORPORATION OF OBSERVATIONS OF THE DGC/ASRB

Scholar Name	Research Title	Observations raised by DGC held on 22 Jan 2024	Action Taken	Remarks (if any)
Muhammad Nauman Hamed	A Real-Time Large-Scale IoT Traffic Anomalies Detection System Using Auxiliary Classifier Generative Adversarial Networks.	1. The document should be as per the approved format.	Indents are removed from the paragraphs. Headings are made as per the approved format	Please check the updated document
		2. The order of citations should be corrected. Reference number 5 comes after reference 6.	The said order has been corrected.	Please refer to page # 2
		3. Objective # 4 should be made SMART. The order of objectives should be corrected.	The objective is retyped, and important terms are added. Objective # 3 is swapped with Objective # 2 as directed during presentation	Please refer to page # 3 Objectives section
		4. How can we mark the data traffic FAKE or REAL? Appropriate words should be used.	These words are replaced with more appropriate words GENUINE and FABRICATED.	Please Refer to page # 4
		5. The figure 1 should be corrected as guided by DGC members.	The figure is redrawn, and some major changes are made as guided by DGC members.	Please Refer to page # 4
		6. Reference # 8 should be updated with more relevant work.	The said reference has been updated.	Please refer to page # 6


Dr. Muhammad Musfer Umar


Dr. Muhammad Irfan Uddin


Name & Signature of Supervisor-I

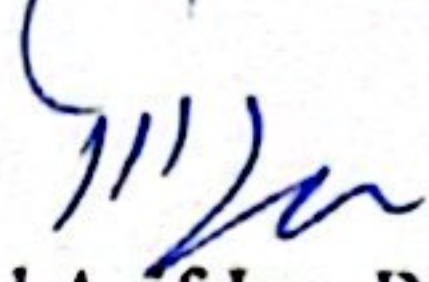
Name & Signature of Supervisor-II

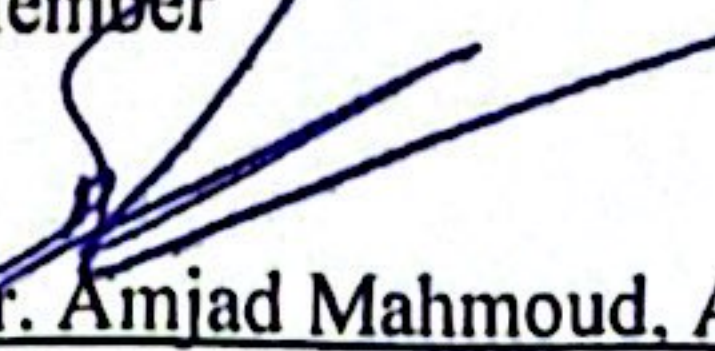
Name & Signature of Supervisor-III

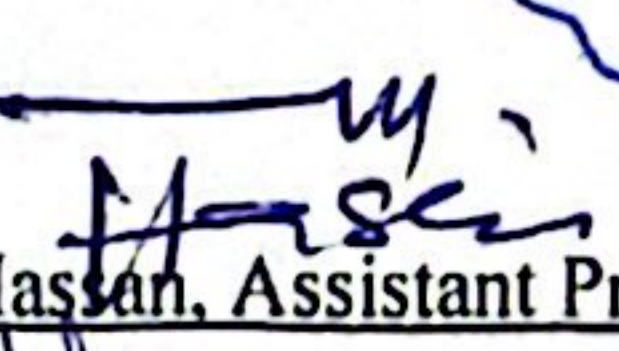
Name & Signature of Departmental Graduate Committee:

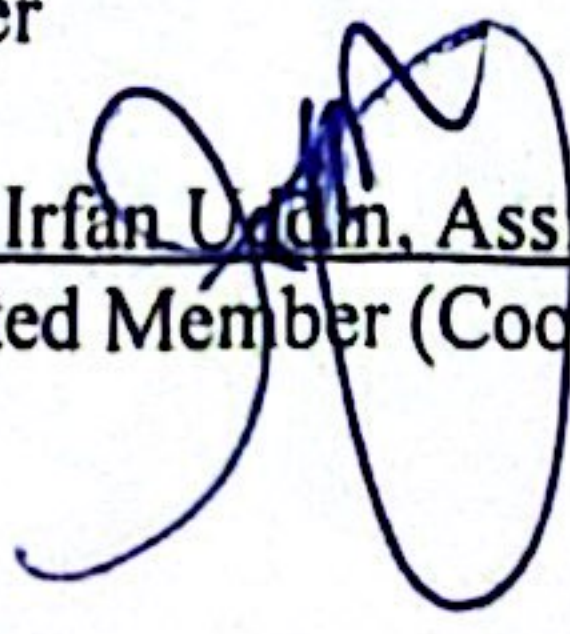
1. 
Prof. Dr. Shafiqullah Khan, IoC
Convener/Director

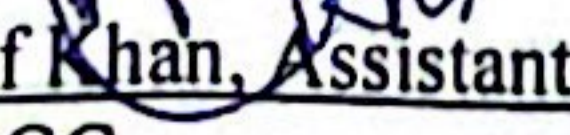
2. 
Prof. Dr. Wali Khan Mashwani, INS
Member

3. 
Prof. Dr. Muhammad Asif Jan, INS
Member

4. 
Dr. Amjad Mahmoud, Associate Professor,
IoC
Member

5. 
Dr. Saima Hassan, Assistant Professor, IoC
Member

6. 
Dr. M. Irfan Uddin, Assistant Professor, IoC
Co-Opted Member (Coordinator MS)

7. 
Dr. M. Altaf Khan, Assistant Professor, IoC
Secretary DGC

**A Real-Time Large-Scale IoT Traffic Anomalies Detection
System Using Auxiliary Classifier Generative Adversarial
Networks**

by
Muhammad Nauman Hameed

(CS320212005)

Supervisor-I Dr. M. Muneer Umar Institute of Computing,
KUST, Kohat


Signature

Supervisor-II Dr. M. Irfan Uddin Institute of Computing,
KUST, Kohat


Signature

Director Prof. Dr. Shafiullah Khan Institute of Computing,
KUST, Kohat


Signature



Institute of Computing
Kohat University of Science & Technology, Kohat-26000,
Khyber Pakhtunkhwa, Pakistan.

INTRODUCTION

The exceptional development of connected devices in the Internet of Things (IoT) era has led to an exponential increase in data flow. The enormous network of linked devices has created difficult and significant issues for the management and security of these digital networks. A major factor contributing to the fast expansion of IoT is the proliferation of gadgets, which ranges from wearables and smart homes to industrial sensors and autonomous vehicles. This spread has changed the way we interact with our environment by establishing a digital ecosystem where billions of gadgets can communicate, gather data, and help make decisions in real time [1].

Normal traffic and malicious traffic represent two distinct categories of data flows on computer networks. Normal traffic encompasses the routine, legitimate data exchanges occurring within a network. These activities include access to web-based data, sending messages, transferring files, and other control and authorized interactions among different devices and users. In contrast, malicious traffic consists of data flows with harmful intent, aiming to compromise network security, disrupt operations, or gain unauthorized access. This category encompasses a cyber threat, such as malware, denial-of-service attacks, intrusion attempts, phishing, and spam, all orchestrated by cybercriminals (malicious actors). The differentiation between normal and malicious traffic is vital in network security as it enables the deployment of appropriate security measures to detect malicious activities [2].

An important issue in the IoT landscape is the existence of unexpected and potentially harmful traffic patterns within the massive stream of data that characterizes it. These traffic patterns cover a broad range of activities, from simple data exchanges to more intricate ones. The regular data traffic produced by IoT devices, the cyclical nature of sensor data, and even the irregular and sporadic data flows that could point to odd behavior or security risks are examples of patterns. Malicious activity can seem as IoT traffic patterns, including attack types like Denial of Service (DoS), Distributed Denial of Service (DDoS), Brute Force, Mirai, and Spoofing are just a few examples [3]. Although efforts have been made to develop systems that can identify anomalous activity in IoT networks, these normal methods have downsides. They frequently encounter challenges adjusting to the constantly

changing tactics used by malevolent actors and encounter issues with the large and unbalanced datasets that are typical on the IoT context. Because of their potential difficulties in handling these diverse traffic patterns, current systems run the danger of producing false positives and false negatives during the detection process [4].

Traditional machine learning algorithms are not suitable for processing unstructured data from IoT's systems, which require strong pattern recognition tools for anomaly detection. Several data types can be used to train deep learning algorithms, ensuring secure and reliable data transmission in IoT networks. However, data-centric IDS technologies may be inefficient due to the limited scope of IoT systems and the lack of user consent for dataset sharing [5]. Generative adversarial networks (GANs) can address imbalanced datasets by generating genuine anomalous data, which can be used for anomaly detection in IoT networks. Synthetic data generation is effective when data production is expensive or when anomalies are infrequent. Deep learning methods are suitable for anomaly detection in big data IoT networks due to their ability to scale well to large datasets. Deep learning models can be adjusted to utilize GPUs and feature engineering independently, increasing accuracy and requiring subject expertise [6]. GANs are deep learning models that are getting a lot of attention in the artificial intelligence field and providing new research opportunities. The original GANs model, developed by Ian Goodfellow, generate actual pictures comparable to the original data [7-9].

The records IoT data traffic flow records show a variety of patterns, which leads to an unbalanced dataset that is vulnerable to several types of network assaults. Conventional anomaly and signature systems might not function well since they are static, and current artificial intelligence models have trouble with unbalanced data and real-time analysis. An AI model that can handle big, unbalanced datasets in real time is required. Auxiliary Classifier Generative Adversarial Networks (AC-GAN) provide improved discriminative performance, semi-supervised learning, and scalability in the face of imbalanced input. The goal of this study is to provide a strong artificial intelligence (AI) solution for effective IoT data traffic anomaly detection by utilising the advantages of AC-GAN [10] to develop a framework for detecting anomalies in imbalanced datasets seen in IoT networks. AC-GAN create data that looks and feels like actual data.

PROBLEM STATEMENT

The traffic records of IoT networks exhibit a multitude of patterns and are dynamic in nature. Conventional detection techniques, such as anomaly, and signature-based approaches, find it difficult to keep up with the changing landscape and emerging attack strategies. IoT networks data traffic with imbalanced patterns and real-time analysis presents challenges for current AI-based solutions. The amount of data traffic has increased significantly because of the expanding number of networked devices, making it challenging to spot odd traffic patterns. Massive and unbalanced datasets are difficult for current approaches to handle, which leads to false positives and false negatives during detection. To increase threat and anomaly detection accuracy in IoT networks, an AI-based methodology utilizing AC-GAN is proposed.

OBJECTIVES

- To develop an AI-based model using AC-GAN to enhance IoT traffic anomaly detection.
- To handle imbalanced and large-scale IoT traffic datasets effectively.
- To classify the normal and abnormal traffic patterns within IoT networks.
- To enhance the accuracy of anomaly and threat detection specially within the IoT ecosystem, ensuring that the research addresses relevant challenges in the context of IoT security.

MATERIAL AND METHOD

The proposed IoT data traffic detection and classification system use AC-GAN primarily. With its improved discriminative abilities, semi-supervised learning methodology, and generative power, the AC-GAN offers a chance to more successfully handle the difficulties associated with IoT network security. To gain a deeper grasp of the variety of typical traffic patterns, it can produce synthetic data, survive adversarial attacks, and adjust to unbalanced data. By utilizing the AC-GAN capabilities, we hope to create a strong and flexible AI-based model in this study that can effectively handle the intricacies of anomaly detection and IoT network security. We plan to acquire a varied dataset of IoT traffic flows as the

primary goal of the first phase of our methodology. To properly train our AI model, this dataset will include a combination of typical and abnormal traffic flow patterns. We will also carefully clean and preprocess the data to guarantee data consistency and quality. To rectify the intrinsic data imbalances that are frequently observed in IoT, we will utilize suitable methods like oversampling, under sampling, or creating artificial data. This preliminary data preparation is essential to our AI model's performance.

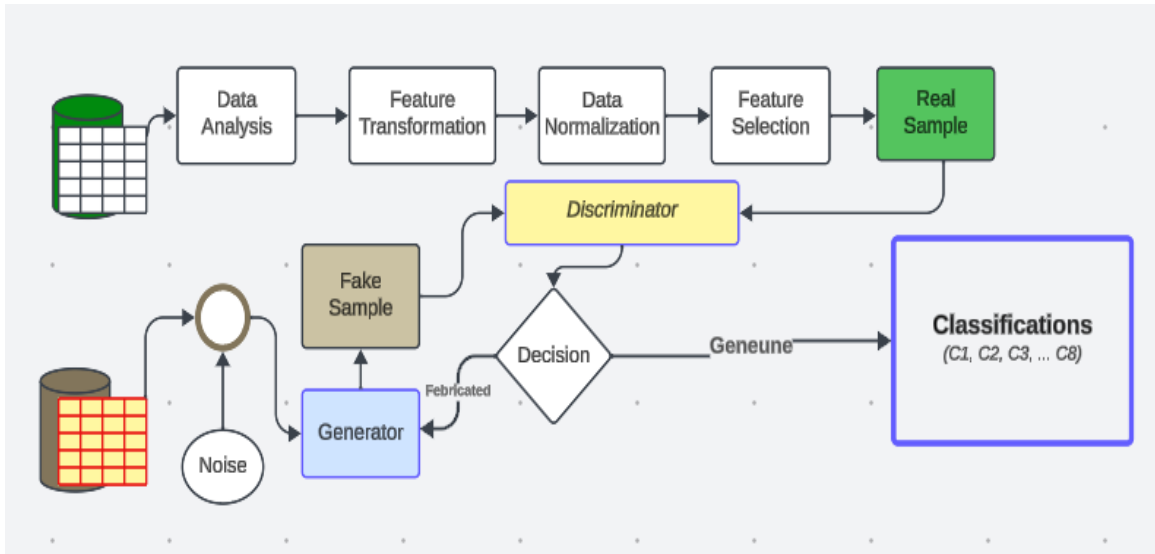


Figure 1. Workflow for detecting and classifying large scale anomalies in real-time IoT data traffic using AC-GAN.

To determine the various forms of noise present in the attacks, the data from IoT attacks is analyzed. Numerous statistical and machine learning methods can be applied to this. To determine various forms of noise, one can compute the dataset's mean, variance, and skewness, for instance. Additionally, various forms of noise in the data collection can be identified using machine learning techniques like classification and clustering. A collection of features that can be used to categories the noise is created from the attack data. To guarantee that the traits are informative and discriminative, much attention should be used when selecting them. For instance, the data set's noise can be categorized using the following features. To ensure that every feature is on the same scale, the data set is normalized. This is necessary to make sure that no feature is given undue weight by the classifier. Numerous methods, including z-score normalization and min-max scaling, can be used to normalize data. From the processed data, the most significant characteristics are

chosen. Numerous feature selection methods, including information gain, recursive feature elimination, and correlation analysis, can be used to accomplish this. This phase involves creating a collection of genuine samples—that is, samples that have noise in them. Real-world IoT attack data can be gathered for this dataset or clean data sets can be artificially noisy augmented. To discern between normal and malicious samples, a discriminator model is trained. In the following stage, the generator model creates the fictitious samples. Numerous machines learning methods, including logistic regression, support vector machines, and neural networks, can be used to train the discriminator model. A generator model is trained to produce noise filled, synthetic IoT attacks. As a feedback mechanism, the discriminator model is used to train the generator model. Many machines learning approaches, including variation auto encoders (VAEs) and GANs, can be used to train the generator model. The novel attacks are categorized as noisy or non-noisy using the learned discriminator model. An attack is categorized as noisy if the discriminator model indicates that it is fabricated. The attack is categorized as non-noisy otherwise.

TIME FRAME

S. No	Research Components	Proposed Time
1	Model Design	02 months
2	Implementation	04 months
3	Model Evaluation	03 months
4	Thesis Writing	03 months

REFERENCES

- [1] A. Abusitta, G. H. de Carvalho, O. A. Wahab, T. Halabi, B. C. Fung, and S. Al Mamoori, “Deep learning-enabled anomaly detection for iot systems,” *Internet of Things*, vol. 21, p. 100656, 2023.
- [2] C. Zhang, S. Yu, Z. Tian, and J. J. Yu, “Generative adversarial networks: A survey on attack and defense perspective,” *ACM Computing Surveys*, 2023.

- [3] F. Li, K. Liang, Z. Lin, and S. K. Katsikas, Security and Privacy in Communication Networks: 18th EAI International Conference, SecureComm 2022, Virtual Event, October 2022, Proceedings. Springer Nature, 2023, vol. 462.
- [4] O. M. Ezeme, Q. H. Mahmoud, and A. Azim, “Design and development of ad-cgan: Conditional generative adversarial networks for anomaly detection,” IEEE Access, vol. 8, pp. 177 667–177 681, 2020.
- [5] M. Salem, S. Taheri, and J. S. Yuan, “Anomaly generation using generative adversarial networks in host-based intrusion detection,” in 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2018, pp. 683–687.
- [6] A. Ferdowsi and W. Saad, “Generative adversarial networks for distributed intrusion detection in the internet of things,” in 2019 IEEE Global Communications Conference (GLOBECOM). IEEE, 2019, pp. 1–6.
- [7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” Communications of the ACM, vol. 63, no. 11, pp. 139–144, 2020.
- [8] L. Gonog and Y. Zhou, “A review: generative adversarial networks,” in 2019 14th IEEE conference on industrial electronics and applications (ICIEA). IEEE, 2019, pp. 505–510.
- [9] O. M. Ezeme, Q. H. Mahmoud, and A. Azim, “Design and development of ad-cgan: Conditional generative adversarial networks for anomaly detection,” IEEE Access, vol. 8, pp. 177 667–177 681, 2020.
- [10] P. Wang, B. Hou, S. Shao, and R. Yan, “Ecg arrhythmias detection using auxiliary classifier generative adversarial network and residual network,” Ieee Access, vol. 7, pp. 100 910–100 922, 2019.