

**LEVERAGING IOT SECURITY TO SUPPORT ENHANCED
SMART CITY**

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

INFORMATION SYSTEMS

Submitted by:

Abhishek Singh

2K21/ISY/03

Under the supervision of

Anamika Chauhan

Assistant Professor



DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi

June 2023

DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi

CANDIDATE'S DECLARATION

I, Abhishek Singh, Roll No. 2K21/ISY/03 of M.Tech. (Information Systems), Hereby declare that the dissertation report titled “LEVERAGING IOT SECURITY TO SUPPORT ENHANCED SMART CITY” which is submitted by me to the Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship, or other similar title or recognition.

Place: Delhi

Date:

Abhishek Singh

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi

CERTIFICATE

I, hereby certify that the dissertation which is submitted by Abhishek Singh, Roll No. 2K21/ISY/03 (Information Systems), Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge, this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place : Delhi

Anamika Chauhan

Date:

SUPERVISOR

ACKNOWLEDGEMENT

I am grateful to Prof. Dinesh Kumar Vishwakarma, HOD (Department of Information and Technology), Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, and all other faculty members of our department for their astute guidance, constant encouragement, and sincere support for this project work.

I would like to take this opportunity to express our profound gratitude and deep regard to our project mentor Anamika Chauhan, for his exemplary guidance, valuable feedback, and constant encouragement throughout the duration of the project. His valuable suggestions were of immense help throughout our project work. His perspective and criticism kept us working to make this project in a much better way. Working under him was an extremely knowledgeable experience for us.

We would also like to give our sincere gratitude to all our friends for their help and support.

ABHISHEK SINGH

Abstract

As the world rapidly urbanizes, smart cities have emerged as a promising solution to enhance urban environments through the integration of Internet of Things (IoT) technologies. However, the increasing reliance on interconnected devices and systems poses significant security challenges that must be addressed to ensure the sustainable development of these cities. By conducting an extensive analysis of scholarly literature and real-world case studies, this paper elucidates the multifaceted nature of IoT security within the smart city landscape. It unveils the complexities surrounding risks and vulnerabilities unique to this context, shedding light on the necessity for proactive and comprehensive security measures. Key areas of concern include safeguarding critical infrastructure, protecting citizen privacy, and bolstering overall urban resilience. With the widespread adoption of smart devices, the usage of these devices for daily tasks has surged, resulting in a significant increase in data generation. However, protecting IoT networks in the context of Smart Cities and modern IoT technologies poses a substantial challenge due to the potential compromise of user data by malicious actors. Additionally, the vulnerability of IoT system sensors to attacks is a pressing concern, given their limited resources and susceptibility to power drainage, which can have severe consequences for both infrastructure and human safety. Consequently, extensive efforts are being made to secure networks and nodes against such threats. This research focuses specifically on Distributed Denial of Service (DDoS) attacks, which are prevalent across various layers of IoT. The study presents effective predictive models for identifying and classifying these attacks, employing optimized features combined with different feature selection techniques. Notably, our work adopts a unique approach to accurately labeling and classifying the identified attack types within their respective subclasses. Our work also provides an exhaustive review of various IoT simulation tools and testbeds that can provide new functionality for designing, modeling, and analyzing the IoT problem and determining their solution. This will lead to improvised and optimal solutions for improving the quality of human life. According to the findings of the study, NS3, an open-access simulator, has been used in numerous studies and is widely used by researchers to solve IoT problems.

CONTENTS

Candidate’s Declaration	i
Certificate	ii
Acknowledgment	iii
Abstract	iv
Contents	v
List of Figures	ix
List of Tables	x
CHAPTER 1 INTRODUCTION	1
1.1. INTRODUCTION	1
1.2. SECURITY REQUIREMENT OF SMART CITY	7
1.2.1. Secure Network Infrastructure	7
1.2.2. Multi-Factor Authentication	7
1.2.3. Robust Data Protection	7
1.2.4. Comprehensive Cyber security Measures	7
1.2.5. Privacy By Design	8
1.2.6. Physical Security Measures	8
1.2.7. Incident Response and Disaster Recovery	8
1.2.8. Training and Awareness	8
1.2.9. Collaboration and Partnerships	8
1.3. PROBLEM STATEMENT	9
1.4. NEED OF THE STUDY	10

1.5. SCOPE OF THE STUDY	10
1.6. OBJECTIVES	11
1.7. CHAPTER PLAN	11
CHAPTER 2 LITERATURE REVIEW	12
2.1. LITERATURE REVIEW	12
CHAPTER 3 TECHNOLOGY ADOPTED	16
3.1. INTERNET OF THINGS	16
3.1.1. Components of IoT	17
3.1.2. IoT Architecture Layers	19
3.1.3. Benefits of Incorporating IoT in Smart City	22
3.1.4. Limitation of Incorporating IoT in Smart City	23
3.2. MACHINE LEARNING	25
CHAPTER 4 ATTACKS ON THE INTERNET OF THINGS	27
4.1. INTRODUCTION	27
4.2. ATTACKS ON THE LAYERS OF IOT	28
4.2.1. Physical Layer	28
4.2.2. Network Layer	29
4.2.3. Middleware Layer	29
4.2.4. Application Layer	29
4.2.5. Business Layer	29
4.3. DISTRIBUTED DENIAL OF SERVICE ATTACK	31
4.3.1. Various Forms of DDoS Attacks	31
4.3.2. Understanding The Working of DDoS Attack	33
4.3.3. Different Types of Attacks At Different Layers of IoT	34

CHAPTER 5 RESEARCH METHODOLOGY	37
5.1. INTRODUCTION	37
5.2. METHODOLOGY	37
5.2.1. Data Pre-Processing	38
5.2.2. Feature Selection	38
5.2.3. Optimizing Feature Count	39
5.2.4. Building the ML Model	39
5.3. AVAILABLE DATASET RELATED TO IOT SECURITY	40
5.3.1. ToN-IoT	40
5.3.2. Edge-IIoT	40
5.3.3. UNSW-NB15	41
5.3.4. BoT-IoT	41
5.3.5. CICIDS2017	42
5.3.6. KDD Cup ‘99	43
5.4. DATASET EMPLOYED IN OUR EXPERIMENTAL WORK	46
5.5. IOT TESTBEDS	49
CHAPTER 6 RESULTS	51
6.1. MACHINE LEARNING MODEL ADOPTED	51
6.1.1. Logistic Regression	51
6.1.2. XGBoost	51
6.1.3. Decision Tree	52
6.1.4. Random Forest	52
6.2. PERFORMANCE EVALUATION	52
6.2.1. Accuracy	52

6.2.2. Precision, Recall, and F1-Score	53
6.2.3. Confusion Matrix	54
CHAPTER 7 CONCLUSION AND FUTURE WORK	57
References	58

LIST OF FIGURES

Figure Number	Figure Name	Page Number
Figure 1.1.	Global megacity population projection 2035	3
Figure 1.2.	Number of connected IoT devices till 2021 and projected till (2022-2030)	7
Figure 3.1.	The IoT Evolution	17
Figure 3.2.	Layers of IoT Architecture	20
Figure 4.1.	Global DDoS attack data, Showcasing 4 major attacks, such as TCP-Connection, volumetric, Fragmentation, and Application	32
Figure 4.2.	A10 network security navigating DDoS attack	36
Figure 5.1.	Methodology adopted to conduct the research work	37
Figure 6.1.	Confusion matrix for Logistic Regression with PCC and SCC	55
Figure 6.2.	Confusion matrix for Random Forest with PCC and SCC	55
Figure 6.3.	Confusion matrix for Decision Tree with PCC and SCC	55
Figure 6.4.	Confusion matrix for XGBoost Classifier with PCC and SCC	56

LIST OF TABLES

Table Number	Table Name	Page Number
Table 1.1.	Smart City and Traditional City	5
Table 4.1.	Various attacks involved in different layers of IoT architecture	27
Table 4.2.	Types of attacks in various layers of IoT	30
Table 5.1.	KDD Cup “99	43
Table 5.2.	Various available dataset on IoT security	44
Table 5.3.	List of machines with their IPs	47
Table 5.4.	Different types of attacks recorded on two days with their timing	48
Table 5.5.	Comparison of different testbeds	50
Table 6.1.	Performance metrics of different machine learning models	53

CHAPTER 1

INTRODUCTION

1.1. INTRODUCTION

The rise of Internet of Things (IoT) technology has revolutionized the concept of smart cities, where interconnected devices and sensors gather and exchange vast amounts of data to improve efficiency, sustainability, and quality of life in urban areas. According to IBM's definition, the concept of a smart city encompasses three key characteristics: instrumented, interconnected, and intelligent.

Instrumented: This characteristic entails equipping the city with a range of devices like sensors and actuators. These devices enable the core systems of the city to access reliable and real-time information.

Interconnected: The smart city involves an extensive network of systems that collaborate to gather information from diverse locations and sources. This interconnected infrastructure facilitates the establishment of a link between the physical world and the digital realm by effectively integrating and coordinating interconnected and instrumented systems.

Intelligent: The intelligent aspect of a smart city refers to the utilization of data and information acquired from various systems and devices, such as sensors. This information is leveraged to enhance the quality of life for citizens, enabling intelligent decision-making and the implementation of innovative solutions.

While smart cities offer numerous benefits, the widespread deployment of IoT devices also introduces significant security challenges that must be addressed to protect these interconnected systems' privacy, integrity, and reliability. IoT security plays a crucial role in supporting the advancement of smart cities. It encompasses a range of measures and technologies designed to safeguard the IoT ecosystem from potential cyber threats and attacks. By implementing robust security practices, smart

cities can mitigate risks and foster trust among citizens, businesses, and government entities, fully realizing the potential offered by IoT-enabled urban environments.

According to the Statista report published on 27 Sep 2022, a projection of the population for the largest urban agglomerations globally by the year 2035 (Figure 1). According to the projection, the New York-Newark agglomeration in the United States is estimated to have a population of approximately 20.8 million individuals. This projection indicates the anticipated size and scale of urban growth and demographic trends in one of the largest urban areas in the world. In smart cities' Industrial Internet of Things (IIoT) environments, cybersecurity emerges as a paramount challenge. The focus lies on preventing unauthorized access, ensuring secure communication to preserve privacy, and safeguarding edge devices against malware attacks. These issues represent some of the prominent hurdles faced in the present landscape. The integration of digital and telecommunication technology aims to enhance the efficiency of traditional networks and services.

The concept revolves around amalgamating diverse solutions for different aspects of urban infrastructure, such as safety, parking systems, waste management, transportation, and city lighting. The objective is to leverage this combination of solutions to optimize the various assets within cities [1]. A smart city is an urban area that leverages diverse technologies to simplify people's lives, enhance city infrastructure, and ensure safety. The smart city mission aims to prioritize the following elements: security, enhanced infrastructure, advanced technology, energy efficiency, smart transportation, and improved healthcare services [2]. The concept of the Smart City (SC) emerged a few years ago, encompassing the notion of utilizing information and communication technologies to enhance urban functionality [3].

Although the term Smart City is increasingly prevalent, with more cities being labeled as smart each day, it is still regarded as an evolving concept that continues to evolve and develop. Smart cities have emerged as a result of the continuous evolution of urban development and the integration of advanced technologies. Over time, cities have transformed to embrace innovations and address the growing complexities of urban life. The notion of smart cities can be adhered to in the early 1990s when the idea of using ICT to enhance urban governance and services began to gain attention [4]. The evolution of smart cities has been influenced by several factors, with

technology playing a central role. Advancements in information technology, telecommunications, and data analytics have significantly contributed to the development of smart city infrastructure. These technological advancements have enabled the analysis, collection, and utilization of vast amounts of data, providing valuable insights for urban planning and resource management [5]. In the early stages of smart city development, emphasis was placed on deploying ICT infrastructure to enhance municipal services and improve the quality of life for citizens.

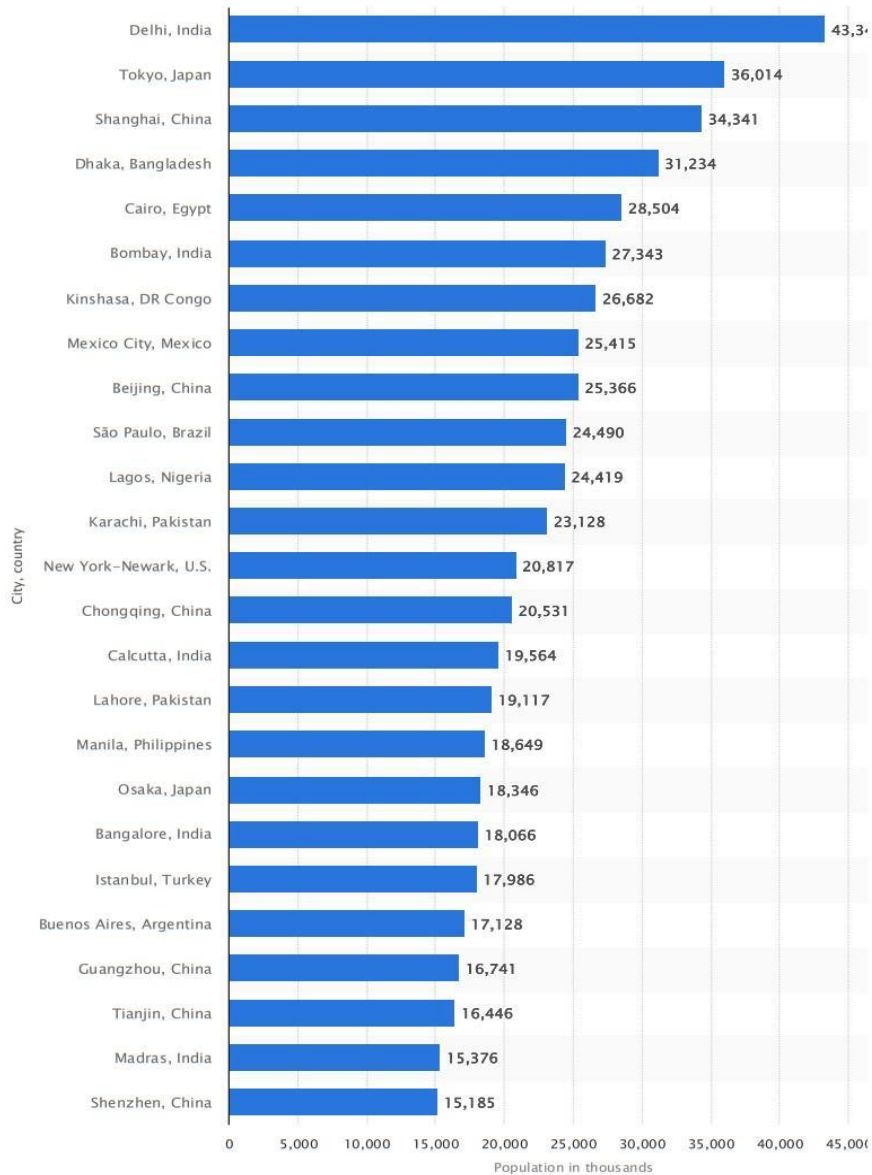


Figure 1.1: Global megacity population projection 2035

This involved the integration of sensors, networks, and data-driven systems to monitor and manage various aspects of urban life, including transportation, energy

consumption, waste management, and public safety. As technology continued to advance, smart cities evolved to embrace more sophisticated solutions [6]. The concept expanded beyond individual services to focus on creating a holistic ecosystem that promotes sustainability, efficiency, and citizen engagement. This led to the integration of smart grids, intelligent transportation systems, smart buildings, and other innovative solutions that leverage real-time data and automation. Moreover, the proliferation of Internet connectivity and the advent of IoT further propelled the evolution of smart cities. IoT devices, embedded in various urban infrastructure components, enabled seamless communication, data sharing, and interconnectivity [7]. This connectivity facilitated the optimization of urban services, improved resource allocation, and enabled the implementation of personalized solutions tailored to citizens' needs.

The role of technology in shaping smart city infrastructure cannot be overstated. It has revolutionized the way cities are planned, operated, and experienced. The integration of advanced technologies has allowed for the development of sustainable and intelligent urban environments, where data-driven decision-making, automation, and connectivity play a vital role [8]. While cities strive to become more intelligent and connected through smart city applications, concerns and challenges related to security and privacy arise. The smart city paradigm, as an information and networking framework, must effectively safeguard the involved information from unauthorized access, disclosure, disruption, modification, inspection, and destruction. To ensure a secure and private environment, various underlying requirements such as confidentiality, integrity, non-repudiation, availability, access control, and privacy [9] need to be fulfilled across the realms of information, communication, and physical infrastructure. However, securing a smart city presents unique challenges. On one hand, the collection of detailed and privacy-sensitive data from individuals' lives and surroundings takes place, while on the other hand, this information is processed, manipulated, and impacts people's lives. These distinctive characteristics make security and privacy issues particularly challenging, hindering the widespread adoption and utilization of smart city technologies.

In their study, Natalia Moch and Wioletta Wereda [10] conducted a comprehensive analysis of both traditional cities and smart cities, aiming to provide readers with a clear understanding of the distinctions between the two. Their insightful analysis offers valuable insights into the contrasting characteristics and features

exhibited by these urban models. By examining various aspects such as city organization, focus on inhabitants' needs, implementation of tasks, provision of services, interoperability of systems and services, the collaboration of individuals, and openness to innovation, the authors provide a comprehensive overview of the disparities between traditional cities and smart cities. Their expert analysis sheds light on the fundamental differences in structure, functionality, and approach, enabling readers to gain a deeper understanding of the contrasting nature of these two urban paradigms.

Table 1.1: Smart City and traditional city [10]

<i>Category</i>	<i>Traditional City</i>	<i>Smart City</i>
City organizations	lack integration among the cells within the city.	Collaboration among various entities within the city, such as residents and entrepreneurs, is fostered.
City and inhabitants	limited attention is given to addressing the needs of the residents	Emphasis is placed on addressing the requirements and enhancing the quality of life for the residents
Implementation of tasks	The execution of statutory obligations in the most straightforward manner is prioritized.	Continual enhancement of the quality of task execution is pursued.
Provision of services	The services are delivered with a focus on the convenience of the service provider rather than prioritizing the highest quality and convenience for the residents.	The integration of service management, daily operations, technology, and digital assets is implemented.
Interoperability of systems and services	Less than normal	High
Collaboration of individuals	Little	Permanent
Openness to innovation	Closed	Open

To establish a smart city that enhances the lives of its residents and promotes resource efficiency, the incorporation of cutting-edge technology and big data solutions is crucial. However, while Internet of Things (IoT) technology holds tremendous potential for rapid expansion, it also introduces security risks due to the proliferation of connected devices and participants within the network ecosystem. The provided statistical data, illustrated in Figure 2, clearly demonstrates the exponential growth of internet-connected devices. This growth inevitably leads to potential security vulnerabilities and exploitation opportunities, primarily attributed to the

insufficient implementation of safety measures in the IoT infrastructure and the inherent characteristics of the diverse network nodes. The implications of these risks are significant, as these interconnected devices frequently store sensitive personal information, including health monitoring data and security footage.

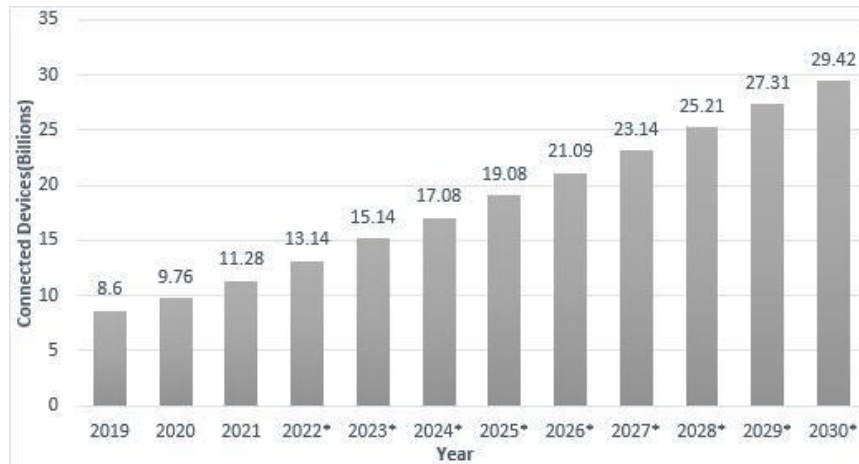


Figure 1.2: Number of connected IoT devices till 2021 and projected till (2022-2030)

The establishment of a smart city necessitates stringent security and privacy measures to safeguard the integrity of its infrastructure and protect the sensitive information of its residents. As innovative technologies and interconnected systems become integral components of smart cities, ensuring the security of these systems becomes paramount. One of the primary security needs is the implementation of robust authentication and access control mechanisms. With numerous interconnected devices and participants in the smart city ecosystem, it is imperative to verify the identity and authorize the access of each entity [11]. This prevents unauthorized access and mitigates the risk of malicious activities. To counteract potential threats, continuous monitoring, and threat detection systems should be in place. These systems employ advanced analytics and machine learning techniques to detect anomalous behavior, suspicious activities, or potential security breaches in real-time. Rapid identification and response to security incidents help mitigate the impact and prevent further damage. Another crucial aspect is the establishment of secure communication channels [12]. By employing secure protocols and implementing secure network architectures, the smart city infrastructure can mitigate the risk of unauthorized interception, eavesdropping, and tampering of data during transmission. Smart cities must also

prioritize privacy protection. Clear policies and regulations need to be established to govern the collection, use, storage, and sharing of personal data. Anonymization techniques, such as data aggregation and pseudonymization, should be employed to protect the privacy of individuals while still allowing for valuable data analysis.

1.2. SECURITY REQUIREMENT OF SMART CITY

The security requirements of a smart city are critical considerations in establishing a safe and resilient urban environment. As a smart city integrates numerous interconnected systems and devices, it becomes imperative to implement robust security measures that protect the city's infrastructure, data, and the privacy of its citizens. Expertise in this area is essential to address the unique challenges associated with securing a complex and interconnected ecosystem.

1.2.1. Secure Network Infrastructure

A smart city requires a secure network infrastructure that employs strong encryption protocols, secure communication channels, and strict access controls. This ensures that data transmitted within the network remains confidential and protected from unauthorized access.

1.2.2. Multi-Factor Authentication

Implementing multi-factor authentication mechanisms is crucial for verifying the identities of individuals, devices, and systems accessing the smart city's network. This helps prevent unauthorized access and protects critical systems and data from compromise.

1.2.3. Robust Data Protection

The protection of sensitive data is paramount in a smart city environment. Encryption techniques should be employed to secure data at rest and in transit. Additionally, data anonymization methods can help protect the privacy of individuals while still enabling valuable data analysis.

1.2.4. Comprehensive Cybersecurity Measures

A smart city must establish a multi-layered cybersecurity framework, including firewalls, intrusion detection systems, and security monitoring tools. These measures continuously monitor the network for potential threats, promptly detect and respond to security incidents, and mitigate risks in real time.

1.2.5. Privacy by Design

Privacy considerations should be integrated into the design and implementation of smart city technologies and services. Privacy impact assessments, consent mechanisms, and strict data handling policies should be in place to ensure compliance with privacy regulations and protect the personal information of citizens.

1.2.6. Physical Security Measures

Physical security is an essential aspect of securing a smart city. Implementing surveillance systems, access control mechanisms, and physical barriers around critical infrastructure locations protects against unauthorized physical access and potential sabotage.

1.2.7. Incident Response and Disaster Recovery

A well-defined incident response plan and disaster recovery strategy are essential for addressing security incidents effectively. This includes processes for incident detection, reporting, containment, and recovery to minimize the impact of security breaches and restore normal operations efficiently.

1.2.8. Training and Awareness

Regular training programs and awareness initiatives should be conducted to educate employees, citizens, and stakeholders about cybersecurity best practices, recognize and report potential threats, and maintain secure behaviors. Increased awareness enhances the overall security posture of the smart city ecosystem.

1.2.9. Collaboration and Partnerships

Collaboration between government entities, private organizations, cybersecurity experts, and law enforcement agencies is crucial for sharing threat intelligence, exchanging best practices, and addressing emerging security challenges collectively. Partnerships foster a proactive approach to security and enable the development of innovative solutions.

1.3. PROBLEM STATEMENT

In the context of IoT-enabled smart cities, Distributed Denial of Service (DDoS) attacks pose a significant threat due to their resource-intensive nature. These attacks involve overwhelming a targeted host or network with an excessive volume of request packets, leading to a depletion of system resources and subsequent disruption of services. As smart city nodes and sensors typically operate with limited resources, it becomes crucial to address this issue by developing an efficient machine-learning model with low computational complexity. Mitigating DDoS attacks in smart cities requires the deployment of intelligent defense mechanisms capable of identifying and filtering malicious traffic while minimizing the impact on system resources. To achieve this, an expertly crafted machine learning model is needed—one that can effectively distinguish legitimate traffic from malicious requests, while being lightweight enough to operate within the constraints of resource-limited IoT devices. Designing such a model entails several key considerations. Firstly, it should be trained on large-scale datasets that encompass various traffic patterns, including both normal and DDoS attack scenarios specific to smart city environments.

By leveraging representative datasets, the model can learn to accurately detect and classify attack traffic, thereby enabling proactive defense measures. Secondly, the machine learning model should be engineered with a focus on computational efficiency, as resource-constrained IoT devices have limited processing capabilities. By adopting techniques such as feature selection, dimensionality reduction, and algorithm optimization, the model's computational complexity can be minimized without compromising its detection accuracy. Furthermore, the model should be designed to operate in real-time, enabling timely detection and mitigation of DDoS attacks within the smart city infrastructure. This necessitates the development of lightweight algorithms that can quickly process incoming network traffic and make prompt decisions regarding the presence of attack patterns. Additionally, the model should be adaptable to dynamic network conditions, allowing it to adjust its detection thresholds and update its knowledge base to accommodate evolving attack strategies.

1.4. NEED OF THE STUDY

Currently, the existing machine learning models used for DDoS attack detection in IoT-enabled smart cities tend to be highly complex and resource-consuming. These models typically operate as binary classifiers, determining only whether an attack is present or not, without providing detailed information about the specific type of attack. Hence, there is a clear need for a comprehensive study aimed at developing a low-resource consumption model capable of accurately classifying and identifying the type of DDoS attack. Addressing this need requires a multifaceted approach that encompasses data collection, feature engineering, algorithm design, and model optimization. Firstly, a diverse and representative dataset should be curated, encompassing various types of DDoS attacks that are prevalent in smart city environments. Ultimately, the study's findings will contribute to the advancement of security in IoT-enabled smart cities by offering a practical and effective solution that efficiently detects and classifies DDoS attacks while operating within the limitations of resource-constrained IoT devices.

1.5. SCOPE OF THE STUDY

The scope of this study encompasses the development and evaluation of a low-resource consumption machine learning model for the classification and identification of DDoS attacks in IoT-enabled smart cities. This study holds significant relevance in scenarios where the probability of nodes being targeted by DDoS attacks is notably high. Particularly, in situations where nodes play critical roles in making important decisions to prevent malicious activities, the findings of this study can be effectively utilized to enhance the security posture. In IoT-enabled smart cities, nodes often act as pivotal components responsible for crucial decision-making processes that directly impact the overall functioning and safety of the infrastructure. By deploying the low-resource consumption machine learning model developed in this study, smart cities can bolster their defenses against DDoS attacks and mitigate potential risks to these vital nodes.

1.6. OBJECTIVES

1. To study the state-of-art of different attacks in various layers of IoT.
2. To study the existing machine learning model used in the classification or identification of the various DDoS attacks.
3. To identify the appropriate dataset related to the IoT attacks.
4. To build a less resource-consuming and computationally less expensive model.

1.7. CHAPTER PLAN

The structure of the chapters is as follows:

Chapter 1 provides introductory information about the smart city and the requirement of the city to become smart, following the benefits and drawbacks of employing IoT in smart cities.

Chapter 2 is about the literature review done in order to gain a theoretical and experimental understanding of the topic.

Chapter 3 imparts the knowledge about technologies adopted to achieve the objective and do the study. It also gives knowledge about the employed dataset and other available datasets related to IoT security.

Chapter 4 provides detailed information about distributed denial of service attacks, their various kinds, and their working.

Chapter 5 gives information about the research methodology adopted to perform the experiment and do the study.

Chapter 6 discusses the results obtained from the experiment and the performance of various machine learning models in different settings.

Chapter 7 concludes the work done and give detail about future work.

CHAPTER 2

LITERATURE REVIEW

2.1. LITERATURE REVIEW

The paper [13] presents a comprehensive study on enhancing the security of an anonymous roaming authentication scheme with two-factor security in smart city environments. The authors address the challenges associated with secure authentication and propose an improved scheme that leverages two-factor security mechanisms. The research contributes valuable insights and practical solutions for strengthening the security of authentication schemes in the context of smart cities, thereby promoting secure and reliable communication in these dynamic environments. Paper [14] discusses rapid advancement of smart technologies has revolutionized data generation and collection, encompassing a wide range of sensitive information in smart cities, including personal, organizational, environmental, energy, transport, and economic data. The objective is to identify the guilty agent responsible for the data leakage and enhance the security of critical data. By computing the probability of an agent is guilty based on the data allocation, the model effectively identifies the information leaker and safeguards confidential information. The successful implementation of this model contributes to the efficient usage and security of sensitive data in smart city environments, promoting the overall advancement and reliability of smart city concepts.

The paper [15] presents an integrated data mining and Business Intelligence architecture for analyzing non-emergency data in a Smart City context. The approach enriches the data with additional contextual information and generates informative dashboards based on Key Performance Indicators (KPIs) and association rules. The experiments conducted in a real Smart City environment validate the effectiveness of the proposed approach. The NED (Non-Emergency Data Analyzer) system serves as a data mining and Business Intelligence environment specifically designed for analyzing non-emergency data in a Smart City. Further improvements can be made, such as extracting multiple-level rules using advanced data mining algorithms to

uncover hidden correlations. Overall, this research contributes to understanding urban security perceptions and supporting decision-making processes in Smart City management.

Paper [16] explores the concept of smart cities and the role of technology in improving various domains such as utility, healthcare, transportation, and home. While numerous IoT applications have been developed to enhance the intelligence of smart cities, there is a need to ensure the suitability of their security strategies. The paper addresses the challenges faced by system designers in implementing effective security measures and provides guidance on when, where, and how to implement security strategies in each smart city domain. It introduces the ANT-centric architecture, which adopts a data-centric viewpoint to achieve end-to-end data security in a zero-trust environment.

Paper [17] introduce an ActivityNetwork-Things (ANT)-centric security reference architecture that encompasses three distinct architectural views when studying IoT systems: semantic, internet, and device. This approach provides a comprehensive framework for addressing security concerns in IoT systems by considering various aspects of the system's architecture and operation. Whereas the paper by Rani [18] discusses the advancement of sensing and computational technologies in smart cities has led to the integration of the Internet of Things (IoT) and Cyber-Physical Systems (CPS) in various processes and infrastructure. While this has enhanced the quality of services in areas such as healthcare, transportation, and environmental management, it has also increased the vulnerability to cyber-attacks.

Contemporary urban environments are witnessing substantial advancements in the realm of smart cities, particularly concerning applications that prioritize precise location awareness, low latency, and robust security. Key examples encompass real-time manufacturing, patient health monitoring, and emergency fire events. To effectively meet these demands, the progress of smart cities relies heavily on the adoption of sophisticated computing paradigms. In this context, fog computing [19] emerges as a pivotal complement to cloud computing, offering notable advantages under its proximity to end devices. As the prevalence of smart cities continues to rise, there has been a notable increase in the utilization of emerging technologies like the Internet of Things (IoT), artificial intelligence (AI), and many more. This paper

[20][21] has conducted a comprehensive analysis to identify the potential security risks associated with each of these prominent emerging technologies employed in smart city initiatives. Regarding IoT devices, the study identified device vulnerabilities, data breaches, and Distributed Denial of Service (DDoS) attacks as the most significant threats. The study put forth several recommended countermeasures to mitigate these security risks.

Specifically, for IoT devices, it is advised to implement robust authentication and encryption protocols, regularly update the devices, employ network segmentation techniques, and closely monitor network traffic. This research paper [22] presents an innovative three-layer architecture called SafeCity, which focuses on the interconnected ecosystem of smart cities comprising cameras, sensors, and various physical devices. Within SafeCity, advanced data analysis techniques and machine learning algorithms are utilized to process and analyze the vast amounts of data generated in the pervasive environment. By adopting this approach, SafeCity achieves notable benefits such as reduced processing time, increased throughput, and enhanced efficiency in handling large-scale data ingestion.

In light of unauthorized access attempts, the integrity of cloud-based information is at risk. The primary cause of such security breaches is the Distributed Denial of Service (DDoS) attack, which poses a significant threat to cloud systems. To address this issue, a novel approach known as FACVO-based DNFN [23] has been developed for effectively detecting DDoS attacks within cloud environments. This cutting-edge solution combines the power of fractional anti-corona virus optimization techniques with deep neuro-fuzzy network models, enabling accurate identification and mitigation of DDoS attacks in cloud-based systems. The FACVO-based DNFN algorithm offers enhanced security measures and improved detection capabilities to safeguard cloud infrastructures against potential threats. The novel methodology put forward in this study demonstrated remarkable performance in terms of testing accuracy, True Positive Rate (TPR), True Negative Rate (TNR), and precision when applied to the NSL-KDD dataset without any attack instances. Specifically, the achieved values were 0.9304, 0.9088, 0.9293, and 0.8745, respectively employing the BoT-IoT dataset.

The Internet of Things (IoT) is susceptible to Application layer Denial of Service (DoS) attacks, particularly those based on protocol vulnerabilities. Such attacks have the potential to cause extensive service disruptions in traditional systems. To address this issue, [24] research presents a comprehensive framework for detecting Application layer DoS attacks specifically targeting the MQTT protocol. The proposed scheme has been thoroughly evaluated using both legitimate and protocol-compliant DoS attack scenarios. The results obtained from the experiments reveal a concerning trend: even when legitimate access to MQTT brokers is denied and resource restrictions are in place, attackers can still overpower server resources. Furthermore, researchers have identified key MQTT features that exhibit high accuracy in detecting these attacks.

The research [25] introduces a novel architecture comprising two integral components: DoS/DDoS mitigation and DoS/DDoS detection. The component of detection offers a meticulous approach to identifying and classifying attacks. The study proposes a multi-class classifier based on the "Looking-Back" concept. The effectiveness of the DoS/DDoS detection component is evaluated using the Bot-IoT dataset. The evaluation results reveal highly promising outcomes, demonstrating the potential of the classifier to achieve an exceptional accuracy rate of 99.81%.

The research [26] presents an innovative feature engineering and machine learning (ML) framework designed specifically for detecting Distributed Denial-of-Service (DDoS) attacks within the IoT-CIDDS dataset. The framework comprises two distinct phases, each contributing to the overall effectiveness of the solution. The initial phase focuses on data set enrichment, utilizing advanced algorithms to enhance the dataset. A key emphasis is placed on employing sophisticated feature engineering techniques to perform comprehensive statistical analysis, enabling a deeper understanding of the dataset's probability distribution and the correlations among its features. Moving to the second phase, an ML model specifically tailored for DDoS attack detection. To assess the complexity of the feature-engineered dataset, five different ML techniques are employed.

CHAPTER 3

TECHNOLOGY ADOPTED

3.1. INTERNET OF THINGS

The Internet of Things (IoT) represents a significant and transformative shift in the field of information technology. The term "Internet of Things" or IoT is derived from the combination of two words: "Internet," which refers to the global system of interconnected computer networks utilizing the standard Internet protocol suite (TCP/IP) to connect billions of users worldwide, and "Things," which encompasses the various physical devices and objects that are embedded with sensors, software, and connectivity capabilities. The Internet of Things (IoT) refers to the integration of physical objects, equipped with sensors and actuators, into the digital realm through wired or wireless networks. This integration enables real-time monitoring, control, and communication between the physical and digital domains. By connecting objects to computing systems, the IoT allows for seamless data exchange, enabling the digital monitoring and even remote control of physical objects in various applications and industries [27].

The phrase "Internet of Things" (IoT) was initially introduced in 1999 by Kevin Ashton, a prominent British technology innovator. Ashton used this term to describe a concept wherein physical objects could be interconnected to the Internet through sensors. He specifically employed the term to demonstrate the potential of connecting Radio-Frequency Identification (RFID) tags, commonly utilized in corporate supply chains, to the Internet. This connection allowed for automated counting and tracking of goods, eliminating the need for human intervention[28]. Since its inception, the Internet of Things has gained significant popularity as a descriptor for scenarios wherein objects, devices, sensors, and everyday items are equipped with Internet connectivity and computing capabilities. This broadens the range of possibilities, enabling seamless communication and data exchange between physical objects and the digital world. The evolution of the Internet of Things (IoT) can be depicted through

several distinct phases, as demonstrated in Figure 3. Today, the term "Internet of Things" serves as a widely recognized expression for illustrating the expansion of Internet connectivity to diverse objects and devices, revolutionizing various industries and enhancing everyday experiences.

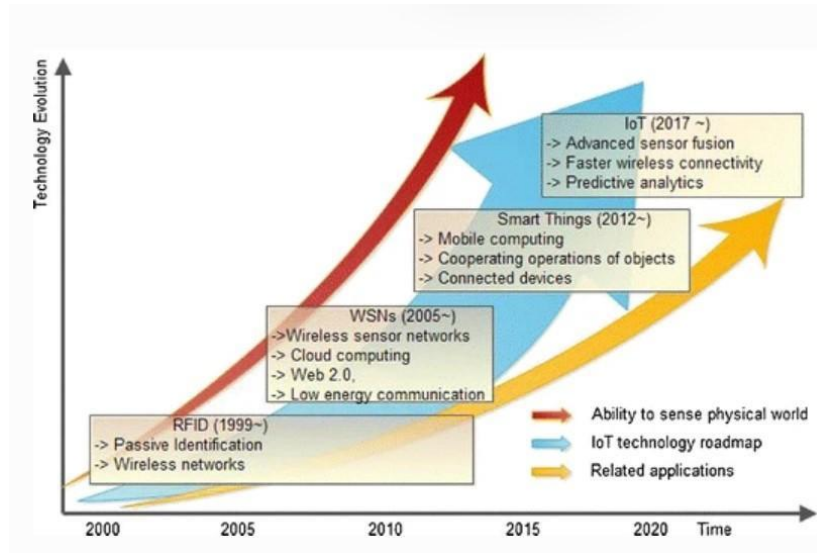


Figure 3.1: The IoT evolution

The IoT comprises several key components that work in harmony to enable its functionalities. First and foremost are the physical devices or "things" themselves, equipped with sensors, actuators, and connectivity capabilities. These devices generate vast amounts of data, which is transmitted through networks, both wired and wireless, to cloud-based platforms or edge computing systems. These platforms provide storage, processing power, and advanced analytics, allowing for data interpretation, actionable insights, and intelligent decision-making.

3.1.1. Component of IoT

The Internet of Things (IoT) is composed of several key components that work together to enable its functionalities. These components include:

1. Things/Devices: The foundation of the IoT is the network of physical objects or devices, often referred to as "things." These devices can range from everyday items such as sensors, actuators, and appliances to more complex entities like industrial machinery and smart infrastructure. Equipped with sensors, processors, and

connectivity capabilities, these devices collect data, interact with the environment, and communicate with other devices.

2. **Sensors and Actuators:** Sensors play a crucial role in the IoT ecosystem by capturing real-world data such as temperature, humidity, motion, or light intensity. These sensors convert physical parameters into electrical signals that can be processed and analyzed. Actuators, on the other hand, enable devices to take actions based on the data received. They can control physical processes, trigger responses, or adjust settings in the environment.

3. **Connectivity:** Connectivity is a fundamental aspect of the IoT, facilitating seamless communication between devices and enabling data exchange. IoT devices utilize various connectivity technologies such as Wi-Fi, Bluetooth, cellular networks, Zigbee, Z-Wave, or even satellite communication. The choice of connectivity depends on factors like range, power consumption, bandwidth requirements, and the specific application context.

4. **Networks:** IoT devices connect to networks that serve as the communication infrastructure for data transfer. These networks can be local area networks (LANs), wide area networks (WANs), or even the Internet. Depending on the scale and requirements of the IoT deployment, networks can be centralized or decentralized, and they may employ different network topologies such as star, mesh, or hybrid configurations.

5. **Data Processing and Analytics:** The enormous amount of data generated by IoT devices require efficient processing and analysis to derive valuable insights. Cloud-based platforms, edge computing systems, or a combination of both are utilized for data processing and analytics. Cloud platforms offer scalable storage, computational power, and advanced analytics capabilities, while edge computing brings data processing closer to the devices, reducing latency and enabling real-time decision-making.

6. **Applications and Services:** IoT applications and services leverage the data collected from devices to deliver specific functionalities and enhance user experiences. These applications span various domains, including smart homes, healthcare, transportation, agriculture, industrial automation, and smart cities. IoT applications

often involve user interfaces, data visualization tools, automation systems, and integration with existing software platforms to provide valuable services and insights.

7. Security and Privacy: Security and privacy are critical considerations in the IoT landscape. IoT systems must implement robust security measures to protect against unauthorized access, data breaches, and malicious activities. This includes authentication mechanisms, encryption protocols, access controls, secure data transmission, and regular software updates. Privacy aspects involve safeguarding personal data, implementing privacy-by-design principles, and complying with relevant regulations and standards.

3.1.2. IoT Architecture Layers

The architecture of the Internet of Things (IoT) is a complex framework that enables the seamless integration of devices, networks, and applications to facilitate efficient communication and data exchange. This architecture is composed of multiple layers, each serving a specific purpose and contributing to the overall functionality and effectiveness of IoT systems. Let's delve into these layers, exploring their unique characteristics and roles. The architecture of the Internet of Things (IoT) remains a topic of ongoing discussion among researchers worldwide, resulting in a variety of proposed architectures. While some researchers advocate for a three-layer IoT architecture, others support a four-layer approach, arguing that the evolving nature of IoT necessitates additional layers to meet application requirements. Furthermore, in response to the significant security and privacy challenges inherent in IoT, a five-layer architecture has been proposed as a viable solution. This recently proposed architecture is believed to effectively address the security and privacy concerns associated with IoT while fulfilling its operational demands. But, some layers are common in all the architectures of IoT [29] [30], such as:

1. Perception Layer: At the bottom of the IoT architecture, lies the Perception Layer, also known as the Sensing Layer. This layer comprises various sensors, actuators, and other devices that interact with the physical world. Sensors capture real-world data such as temperature, humidity, light intensity, motion, and more, while actuators enable physical actions in response to commands. The Perception Layer forms the bridge between the physical and digital realms, facilitating data acquisition and transforming physical events into digital signals.

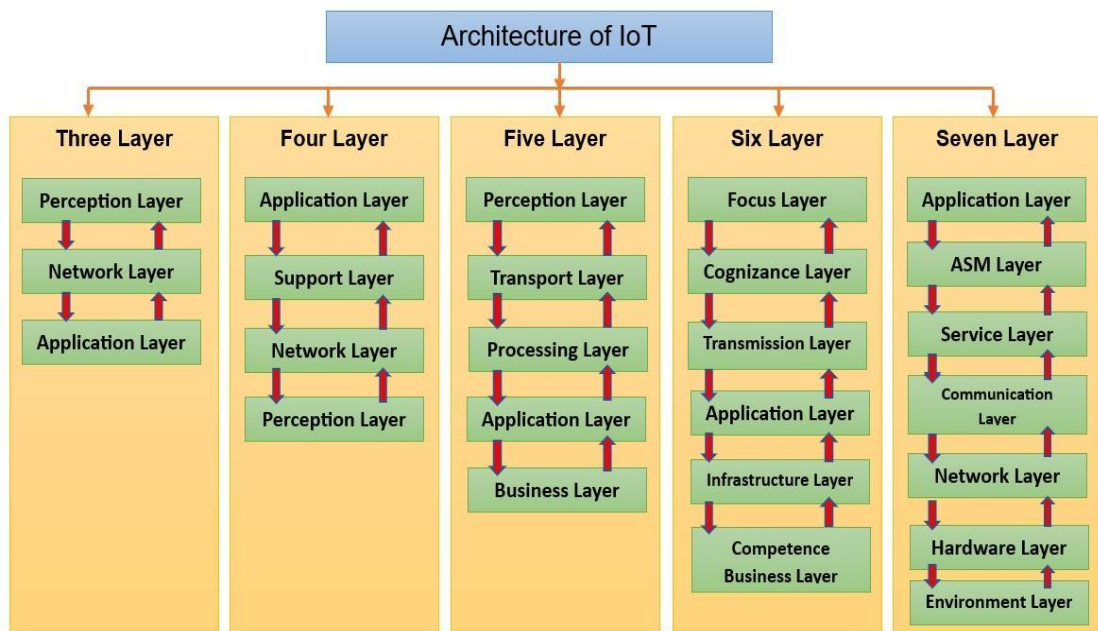


Figure 3.2: Layers of IoT Architecture

2. **Network Layer:** Sitting above the Perception Layer is the Network Layer, responsible for establishing reliable and secure connectivity between devices, gateways, and cloud platforms. This layer encompasses the communication protocols, network infrastructure, and technologies that enable data transmission over wired or wireless networks. It ensures that devices can communicate seamlessly and efficiently, forming a robust and scalable network for IoT deployments.

3. **Middleware Layer:** The Middleware Layer acts as an intermediary between the Network Layer and the Application Layer, providing essential services for data processing, integration, and management. It enables data filtering, aggregation, and transformation, ensuring that only relevant and meaningful information is passed on to higher layers. Additionally, the Middleware Layer offers functionalities such as device discovery, security, data storage, and protocol translation. It plays a crucial role in decoupling the complexity of the underlying networks from the application logic.

4. **Application Layer:** The Application Layer represents the highest level of the IoT architecture and encompasses the applications, services, and user interfaces that leverage the data collected from the devices. This layer facilitates data analysis,

visualization, and decision-making, allowing users to monitor and control IoT systems effectively. Applications can range from simple mobile apps to sophisticated enterprise solutions, catering to various domains such as healthcare, agriculture, transportation, and smart cities.

5. **Security Layer:** Embedded throughout the entire IoT architecture is the Security Layer, which ensures the confidentiality, integrity, and availability of IoT systems and data. It encompasses various security measures, including encryption, authentication, access control, and intrusion detection, to protect against cyber threats and unauthorized access. The Security Layer is of paramount importance due to the vast amount of sensitive data generated and transmitted by IoT devices.

6. **Processing Layer:** Called the middleware layer, plays a crucial role in the IoT architecture. Positioned above the transport layer, it serves as a vital component responsible for collecting and processing information received from various sources. Its primary objective is to extract relevant and meaningful data while eliminating redundant or irrelevant information. Furthermore, the processing layer addresses the challenges associated with big data in IoT by effectively managing and optimizing the vast amount of information received. The presence of significant data volumes can potentially impact the overall performance of IoT systems. Moreover, it is important to note that the processing layer is vulnerable to various threats, which can sabotage its integrity and disrupt the performance of the entire IoT infrastructure. Safeguarding the processing layer from such attacks is crucial for ensuring the seamless and efficient operation of IoT systems.

7. **Business Layer:** The business layer plays a pivotal role in the IoT architecture, acting as the system's overseer and embodying the intended behavior of the application. Comparable to a manager, this layer assumes responsibilities for the effective management and control of IoT applications, as well as the associated business and profit models. In addition, the business layer addresses the critical aspect of user privacy, ensuring that appropriate measures are in place to protect sensitive information. Furthermore, it possesses the capability to define the creation, storage, and modification of data within the system. It is essential to fortify the security of this layer, as any vulnerabilities can enable attackers to manipulate applications by bypassing or exploiting the underlying business logic. Many security issues

encountered in IoT stem from flaws or inadequacies in the application's security controls. Strengthening and implementing robust security measures at the business layer is imperative to mitigate potential risks and safeguard the integrity of the entire IoT ecosystem.

3.1.3. Benefits of Incorporating IoT in Smart City

Incorporating IoT into a smart city framework brings numerous benefits, ranging from increased efficiency and resource management to improved safety, sustainability, and quality of life. These advantages pave the way for economic growth and innovation, making IoT an indispensable component of smart city development.

3.1.3.1. Enhanced Efficiency and Resource Management

By integrating IoT technologies into a smart city infrastructure, various systems, and services can be interconnected and optimized for efficient resource management. IoT-enabled sensors and devices can collect real-time data on energy consumption, traffic patterns, waste management, and more. This data can be analyzed and utilized to make informed decisions, leading to better allocation of resources, reduced energy consumption, and improved overall efficiency.

3.1.3.2. Improved Infrastructure and Utilities

IoT applications in a smart city can significantly enhance the monitoring and management of critical infrastructure and utilities. For instance, IoT sensors embedded in buildings, bridges, and roads can provide continuous structural health monitoring, detecting signs of wear and tear in real time. This enables proactive maintenance and reduces the risk of infrastructure failures. Similarly, IoT can optimize the management of utilities such as water and electricity by enabling remote monitoring, automated control, and demand-response systems.

3.1.3.3. Enhanced Safety and Security

IoT technologies contribute to improved safety and security in a smart city. Connected devices, such as surveillance cameras, smart streetlights, and public safety sensors, can monitor public spaces, detect anomalies, and provide real-time alerts to authorities. This enables faster emergency response, crime prevention, and overall enhanced public safety. Additionally, IoT-enabled solutions like smart fire detection

systems and early warning systems for natural disasters can help minimize risks and improve emergency preparedness.

3.1.3.4. Sustainable Environmental Practices

IoT plays a vital role in promoting sustainability and environmental conservation in smart cities. By collecting and analyzing data on air quality, noise levels, waste management, and energy consumption, city officials can implement targeted measures to reduce pollution, promote recycling, and optimize energy usage. IoT-based smart grids facilitate the integration of renewable energy sources, enabling cities to transition to cleaner and more sustainable energy systems.

3.1.3.5. Enhanced Quality of Life

The integration of IoT in a smart city leads to an improved quality of life for residents. IoT applications can provide real-time information on public transportation, parking availability, and traffic conditions, allowing citizens to make informed decisions and optimize their daily routines. Smart healthcare solutions powered by IoT enable remote patient monitoring, personalized healthcare services, and timely interventions. Additionally, IoT-driven smart homes enhance comfort, convenience, and energy efficiency for residents.

3.1.3.6. Economic Growth and Innovation

Smart city initiatives leveraging IoT technologies create a conducive environment for economic growth and innovation. By attracting businesses and entrepreneurs into the IoT ecosystem, cities can foster technological advancements, job creation, and economic opportunities. The availability of real-time data and analytics enable data-driven decision-making, leading to the development of new services, business models, and sustainable urban planning.

3.1.4. Limitations of Incorporating IoT in a Smart City

3.1.4.1. Security and Privacy Concerns

One of the primary limitations of incorporating IoT in a smart city is the heightened security and privacy risks associated with interconnected devices and systems. IoT devices are susceptible to cyberattacks, unauthorized access, and data breaches. Safeguarding the vast network of IoT devices and ensuring the privacy of

citizens' personal information requires robust security measures, encryption protocols, and continuous monitoring. Failure to address these concerns can lead to significant vulnerabilities and compromise the trust of residents.

3.1.4.2. Interoperability and Standardization Challenges

The successful implementation of IoT in a smart city relies on the interoperability and seamless integration of diverse devices, platforms, and systems. However, achieving interoperability stays a question due to the lack of standardized protocols, data formats, and communication interfaces across different IoT devices and manufacturers. This fragmentation can hinder the exchange and utilization of data, limit scalability, and impede the overall effectiveness of the smart city infrastructure.

3.1.4.3. Complex Infrastructure and Maintenance

The deployment of IoT infrastructure in a smart city requires significant investment in terms of hardware, connectivity, and supporting systems. Building and maintaining a vast network of sensors, gateways, and communication infrastructure can be complex and resource-intensive. Additionally, regular maintenance, software updates, and addressing technical issues become critical to ensure the continuous operation of IoT systems. Failure to manage the infrastructure effectively can result in system downtime, reduced efficiency, and increased costs.

3.1.4.4. Data Management and Analytics

IoT generates massive volumes of data from various sensors and devices within a smart city. Effectively managing, processing, and analyzing this data to derive meaningful insights can be a significant challenge. Smart cities need robust data management frameworks, advanced analytics tools, and data governance practices to handle the velocity, variety, and veracity of IoT-generated data. Without proper data management strategies in place, the abundance of data can become overwhelming, leading to information overload and difficulty in extracting valuable insights.

3.1.4.5. Limited Citizen Adoption and Digital Divide

The success of a smart city heavily relies on citizen adoption and engagement with IoT-enabled services and technologies. However, there can be barriers to adoption, including a lack of awareness, limited digital literacy, and concerns regarding data privacy. This digital divide can create disparities among citizens, with certain groups

benefiting more from smart city initiatives than others. Addressing these challenges requires comprehensive awareness campaigns, accessible user interfaces, and inclusive strategies to ensure equitable access and participation.

3.1.4.6. Potential for Overreliance and System Dependency

Relying extensively on IoT systems for critical city functions and services can introduce a level of dependency that poses risks in the event of system failures or disruptions. Technical glitches, network outages, or cyberattacks can lead to service interruptions and impact essential services like transportation, energy distribution, or emergency response. Implementing backup systems, redundancy measures, and robust disaster recovery plans becomes crucial to mitigate the risks associated with overreliance on IoT systems.

3.2. MACHINE LEARNING

Machine learning is a branch of artificial intelligence (AI) that focuses on the development of algorithms and models capable of enabling computers to learn and make predictions or decisions without being explicitly programmed. It empowers machines to automatically analyze and interpret complex data, recognize patterns, and extract meaningful insights, thereby driving intelligent decision-making processes [31]. At its core, machine learning revolves around the concept of creating mathematical models that can learn from data and improve their performance through experience. These models are constructed using a variety of algorithms and techniques, each designed to address specific types of learning tasks and datasets [32]. In the realm of machine learning, a wide range of algorithms are employed to address diverse data problems. Data scientists emphasize that there is no universal, one-size-fits-all algorithm that can efficiently solve every problem. The selection of an algorithm depends on various factors such as the nature of the problem at hand, the number of variables involved, and the most suitable model for the specific task.

Supervised learning is one of the fundamental approaches in machine learning. In this paradigm, a model is trained using labeled data, where the input samples are accompanied by corresponding desired outputs. The model learns to map inputs to outputs by generalizing from the training examples and is subsequently capable of

predicting outputs for new, unseen inputs. This technique is widely employed in applications such as image classification, spam filtering, and speech recognition[33].

Unsupervised learning, on the other hand, deals with unlabeled data and aims to discover hidden patterns or structures within the dataset. The model learns to identify similarities or differences between data points and group them accordingly. Clustering and dimensionality reduction are common unsupervised learning techniques that find applications in customer segmentation, anomaly detection, and data visualization [34] [35].

Reinforcement learning is another prominent approach, inspired by behavioral psychology. In this framework, an agent interacts with an environment and learns to take actions that maximize a notion of cumulative reward. Through trial and error, the agent explores different actions and adjusts its behavior based on the feedback received. Reinforcement learning has gained attention in domains like robotics, game-playing, and autonomous systems [36] [37] [38].

CHAPTER 4

ATTACKS ON THE INTERNET OF THINGS

4.1. INTRODUCTION

As discussed in [39], the paper highlights the critical importance of security in IoT systems and the urgent need for more robust and sophisticated security solutions to prevent cyber-attacks and protect the confidentiality, integrity, and availability of sensitive data. The authors also discuss the limitations of existing systems for intrusion detection and emphasize the necessity for future research to build more effective and efficient mechanisms for detecting and preventing IoT attacks. Overall, this paper provides valuable insights and recommendations for researchers, industry professionals, and policymakers to address the security challenges in IoT systems and develop more secure and trustworthy IoT applications.

Table 4.1: Various Attack Involved in Layers of IoT Architecture

Layer of IoT	Components of IoT	Possible Attacks
Perception Layer	Sensors, Cameras, Microphones, etc.	Spoofing, Jamming, Tampering, Physical Damage
Network Layer	Routers, Switches, Hubs, Gateways, etc.	Eavesdropping, DDoS Attacks, DoS Attacks, MITM Attack
Middleware Layer	Protocols, APIs, Messaging Services, etc.	Buffer Overflow, Injection Attack, Broken Access Control
Application Layer	IoT Applications, Dashboards, Cloud Services, etc.	Privilege Escalation, Data Breach, SQL Injection, Cross-Site Scripting (XSS)

The use of IoT devices has grown rapidly in recent years, leading to an increase in cyber-attacks against these devices. Cyber attackers target IoT networks to exploit their weak links and compromise the connected devices. Similarly, the Industrial Internet of Things (IIoT) has gained momentum and is being used to interconnect

machines, sensors, and actuators in large manufacturing plants. IIoT adoption has helped companies reduce operational costs and increase productivity. However, as IIoT relies on the internet to operate, it is vulnerable to cyber-attacks if security measures are not taken into consideration. Industry 4.0 has emerged as a new version of smart industries, combining cloud and fog computing, cyber-physical systems (CPS), and data analytics to automate the manufacturing process. The increased use of IoT and IIoT devices in various applications has led to an increase in cyber-attacks, necessitating the need for protection mechanisms against them. This paper surveys the various types of attacks that IoT and IIoT networks may face and highlights ways to mitigate them. The paper's objective is to provide insights into IoT and IIoT devices' security issues and challenges, serving as a guide for establishing a secure network for such devices. While fully securing IoT and IIoT devices may be a long process, understanding the different types of cyber-attacks against them is crucial for developing new protection mechanisms [40].

4.2. ATTACKS ON THE LAYERS OF IOT

The IoT architecture consists of different layers, each with its own set of functions and vulnerabilities. These layers are physical, communication, middleware, application, and business layers.

4.2.1 Physical Layer

The physical layer is the lowest layer of the IoT architecture, and it includes devices such as sensors and actuators that collect and transmit data. The devices in this layer are responsible for collecting data and communicating it to the upper layers. The physical layer is essential for the industry as it helps in gathering real-time data that can be analyzed to identify potential problems in the system. However, the physical layer is prone to attacks such as tampering, eavesdropping, and cloning, which can result in the loss of sensitive data or the disruption of critical processes. These attacks can lead to significant problems and losses in industries such as manufacturing and healthcare.

4.2.2. Network Layer

The communication layer is responsible for transmitting data between devices in the IoT network. It includes protocols such as Wi-Fi, Bluetooth, and Zigbee. The communication layer is critical to the functioning of the IoT network, as it facilitates data exchange between different devices and enables the remote monitoring and control of devices. However, the communication layer is vulnerable to attacks such as man-in-the-middle (MITM) attacks, denial-of-service (DoS) attacks, and replay attacks. These attacks can result in the interception, modification, or disruption of data transmissions, leading to significant problems and losses for industries such as transportation and logistics.

4.2.3. Middleware Layer

The middleware layer is responsible for managing the data flow and interactions between devices in the IoT network. It includes software and protocols such as MQTT and CoAP. The middleware layer is critical to the functioning of the IoT network, as it ensures that data is transmitted securely and efficiently between devices. However, the middleware layer is vulnerable to attacks such as buffer overflow attacks, injection attacks, and cross-site scripting (XSS) attacks. These attacks can result in the corruption or theft of data, leading to significant problems and losses for industries such as energy and utilities.

4.2.4. Application Layer

The application layer is responsible for processing and analyzing data collected from devices in the IoT network. It includes software applications such as data analytics and artificial intelligence (AI) algorithms. The application layer is critical to the functioning of the IoT network, as it enables real-time monitoring and control of devices and facilitates decision-making based on data insights. However, the application layer is vulnerable to attacks such as malware attacks, insider attacks, and social engineering attacks. These attacks can result in the theft or destruction of data, leading to significant problems and losses for industries such as finance and banking.

4.2.5. Business Layer

The business layer is responsible for managing the business processes and services enabled by the IoT network. It includes applications such as supply chain

management and customer relationship management. The business layer is critical to the functioning of the IoT network, as it enables efficient and effective management of business operations. However, the business layer is vulnerable to attacks such as phishing attacks, ransomware attacks, and identity theft. These attacks can result in the disruption or destruction of business processes, leading to significant problems and losses for industries such as retail and e-commerce.

Table 4.2: Types of attacks in various layers of IoT

Layer	Functionality	Importance for Industry	Types of Attacks	Problems and Losses
Perception Layer	Collects data from sensors and other sources	Provides raw data for analysis and decision making	Eavesdropping, Tampering, Physical Attacks, Denial-of-Service (DoS)	Loss of data accuracy, equipment damage, disruption in data collection
Network Layer	Transfers data between devices	Enables communication and control of devices	Man-in-the-Middle (MitM), Spoofing, DoS, Network Scanning	Interference with communication, unauthorized access, loss of control over devices
Middleware Layer	Provides data processing, storage, and other services	Enables integration of multiple devices and applications	Injection, Authentication and Authorization Bypass, Configuration Exploits, Remote Code Execution	Unauthorized access, data breaches, loss of control over devices and applications
Business Layer	Enables application development and management	Facilitates business processes and decision making	Data Leaks, Advanced Persistent Threats (APTs), Malware, Phishing, Social Engineering	operational disruption, loss of sensitive data, Financial losses, reputational damage

4.3. DISTRIBUTED DENIAL OF SERVICE ATTACK

The paper titled "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms"[3] provides a comprehensive analysis of distributed denial of service (DDoS) attacks and the corresponding defense mechanisms. The authors delve into the intricacies of DDoS attacks, beginning with an explanation of denial of service (DoS) attacks as the foundation. They then present a taxonomy of DDoS attacks, categorizing them into distinct types based on their methods of execution and impact on targeted systems. By detailing various attack vectors such as volumetric attacks, TCP state-exhaustion attacks, application layer attacks, protocol attacks, reflective/amplified attacks, and resource depletion attacks, the authors provide a holistic view of the diverse tactics employed by malicious actors. Additionally, the paper emphasizes the significance of implementing robust DDoS defense mechanisms to safeguard critical infrastructures and mitigate the disruptive effects of these attacks. By consolidating a wide range of attack types and defense strategies, this systematic review serves as a valuable resource for researchers, practitioners, and organizations striving to enhance their understanding of DDoS attacks and fortify their security posture.

4.3.1. Various forms of DDoS attacks

4.3.1.1. Volumetric Attacks

These attacks inundate the target network or system with an exceedingly high volume of traffic, overwhelming its available bandwidth and resources. Noteworthy examples encompass UDP floods, ICMP floods, and DNS amplification attacks.

4.3.1.2. TCP State-Exhaustion Attacks

These attacks exploit vulnerabilities within the TCP protocol to exhaust the resources of the target server. Common tactics encompass SYN floods and ACK floods, which exploit the server's limitations on concurrent connections or available ports.

4.3.1.3. Application Layer Attacks

These attacks focus on the application layer of the network protocol stack, intending to overwhelm the target's web servers, databases, or other application-specific resources. Prominent examples include HTTP floods, Slowloris attacks, and DNS query floods.

4.3.1.4. Protocol Attacks

These attacks capitalize on weaknesses within network protocols to disrupt the target's infrastructure. An illustration of such an attack is an Internet Control Message Protocol (ICMP) flood, which overwhelms the target by bombarding it with ICMP echo request packets.

4.3.1.5. Reflective/Amplified Attacks

These attacks exploit legitimate services that can generate a disproportionately larger response to a small request, enabling the attacker to amplify the impact of the attack traffic. Notable instances include DNS amplification and NTP amplification attacks.

4.3.1.6. Resource Depletion Attacks

These attacks specifically target vital resources, such as server CPU, memory, or database connections, aiming to deplete them and induce service degradation or disruption.

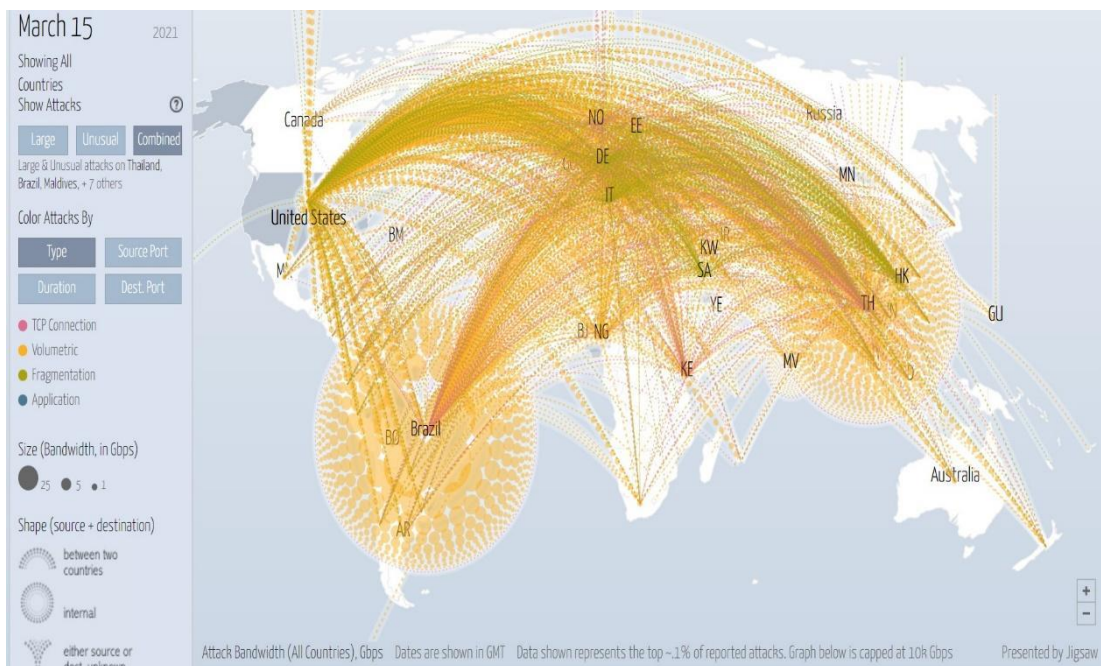


Figure 4.1: Global DDoS attack data, showcasing 4 major attacks such as TCP-connection, Volumetric, Fragmentation, and Application,

source: <https://www.digitalattackmap.com/#anim=1&color=0&country=ALL&list=0&time=18763&view=map>

The paper titled "DDoS attacks in IoT networks: a comprehensive systematic literature review"[4] offers a detailed analysis of distributed denial of service (DDoS) attacks specifically targeting Internet of Things (IoT) networks. The authors conduct a systematic literature review to gather insights from existing research and provide a comprehensive overview of the current state of knowledge in this field. They highlight the growing threat landscape surrounding IoT devices and the increasing risk of DDoS attacks targeting these networks.

The paper begins by defining DDoS attacks and their potential impact on IoT networks, emphasizing the criticality of IoT security. The authors delve into the unique characteristics and challenges posed by IoT devices, including their limited resources, heterogeneity, and vast deployment scale, which make them vulnerable to DDoS attacks. They outline the distinct phases of a DDoS attack, including reconnaissance, botnet recruitment, command and control, and the actual attack execution.

Furthermore, the paper provides a comprehensive categorization of DDoS attacks in IoT networks. It highlights various attack vectors, such as ICMP flood, SYN flood, UDP flood, HTTP flood, and DNS amplification, and discusses their specific characteristics and potential impact on IoT devices and networks. The authors also examine the motivations behind DDoS attacks in IoT networks, including financial gain, political motives, and competitive advantage.

4.3.2. Understanding the Working of DDoS Attack

In the context of IoT, a Distributed Denial of Service (DDoS) attack occurs when an attacker targets IoT devices and networks, overwhelming them with a massive influx of illegitimate traffic or requests.

1. Attackers search for vulnerable IoT devices within networks. These devices may have weak security measures or outdated firmware, making them susceptible to compromise. Common vulnerabilities include default or easily guessable passwords, unpatched software, or insecure communication protocols.

2. Botnet Formation: Once vulnerable IoT devices are identified, the attacker gains control over them. This control is often achieved through malware, such as botnet malware, which infects and takes command of the compromised devices. The attacker creates a network of these compromised devices, known as a botnet, which can be used collectively to launch the DDoS attack.

3. **Command and Control (C&C):** The attacker establishes a command and control infrastructure to orchestrate the attack. This infrastructure allows the attacker to remotely command the compromised IoT devices within the botnet and direct their actions.

4. **Attack Initiation:** The attacker instructs the compromised IoT devices to flood the target with a massive volume of traffic or requests. This flood of traffic overwhelms the target's resources, such as network bandwidth, processing power, or memory, causing disruption and rendering the IoT devices or services unavailable to legitimate users.

5. **Traffic Overload:** The targeted IoT devices or networks become overwhelmed and unable to handle the excessive traffic generated by the botnet. This overload may result in performance degradation, unresponsiveness, or even a complete service outage.

4.3.3. Different types of attacks at different layers of IoT

4.3.3.1. ICMP Flood

An ICMP flood attack targets the Internet Control Message Protocol (ICMP) by overwhelming a target network with many ICMP echo request packets. This flood of packets can consume network resources and disrupt the target's connectivity, rendering it unresponsive to legitimate traffic.

4.3.3.2. SYN Flood

An SYN flood attack exploits the TCP three-way handshake process by overwhelming the target server with a flood of SYN requests without completing the handshake. This depletes server resources and prevents legitimate connections from being established, causing service disruptions.

4.3.3.3. UDP Flood

In a UDP flood attack, the attacker floods the target with a large volume of User Datagram Protocol (UDP) packets. Since UDP is connectionless and does not require a handshake, this attack floods the target's resources and causes network congestion, leading to service unavailability.

4.3.3.4. HTTP Flood

An HTTP flood attack targets web servers by flooding them with a massive number of HTTP requests. By overwhelming the server's capacity to handle incoming

requests, the attacker can exhaust server resources, ensuing in a DoS for legitimate users.

4.3.3.5. DNS Amplification

A DNS amplification attack exploits vulnerable DNS servers to flood the target with a high volume of DNS response traffic. By using spoofed IP addresses, the attacker can amplify the volume of traffic, leading to network congestion and disruption.

4.3.3.6. NTP Amplification

An NTP amplification attack abuses vulnerable Network Time Protocol (NTP) servers to flood the target with amplified traffic. The attacker spoofs the source IP addresses and requests a large amount of data from the NTP servers, causing congestion and service degradation.

4.3.3.7. SSDP Amplification

An SSDP amplification attack exploits Simple Service Discovery Protocol (SSDP) devices to generate a high volume of response traffic to the target. By sending crafted requests to vulnerable SSDP devices, the attacker amplifies the traffic, overwhelming the target's resources.

4.3.3.8. SNMP Amplification

An SNMP amplification attack leverages Simple Network Management Protocol (SNMP) devices to generate amplified traffic toward the target. The attacker manipulates vulnerable SNMP devices to respond to large amounts of data, causing network congestion and disrupting services.

4.3.3.9. DNS Flood

A DNS flood attack involves overwhelming the target's DNS servers with a flood of DNS query traffic. By saturating the DNS infrastructure, the attacker disrupts the resolution process, rendering the target's services unreachable.

4.3.3.10. Slowloris

Slowloris is a type of application-layer DDoS attack that exploits the way web servers handle simultaneous connections. The attacker establishes numerous connections to the target server but sends HTTP requests very slowly, keeping the connections open and tying up server resources, ultimately leading to service unavailability.

4.3.3.11. Application Layer Attacks

This category includes various attacks targeting specific applications or protocols, such as HTTP-based attacks, DNS-based attacks, or attacks on specific software vulnerabilities. These attacks aim to exploit weaknesses in the application layer and disrupt the functioning of specific services or applications.



Figure 4.2: A10 network security navigating DDoS attack

Source: <https://threats.a10networks.com/>

CHAPTER 5

RESEARCH METHODOLOGY

5.1 INTRODUCTION

This section outlines the methodology employed to conduct the research study and achieve the objectives outlined in the study. This chapter provides detailed knowledge about the dataset explored regarding IoT security and the dataset that we have employed to conduct our experiment. The chapter also describes the testbed and simulator for conducting the study related to IoT projects.

5.2. METHODOLOGY

Experimentation starts with selecting the appropriate dataset. The dataset selected to experiment is discussed in section 5.4.

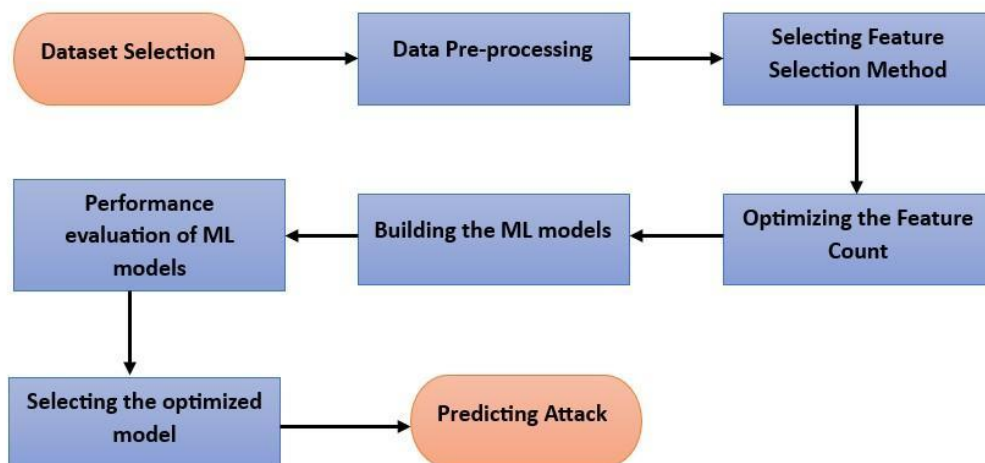


Figure 5.1: Methodology Adopted to conduct the research work

5.2.1. Data Pre-Processing

In the pre-processing stage, the dataset consists of seven separate CSV files, each containing data related to different types of DDoS attacks, NetBIOS attacks, UDP lag attacks, MSSQL attacks, Portmap, Syn attacks, UDP attacks, and LDAP attacks. The initial objective is to determine the attack category with the least amount of data to ensure a balanced representation among the different attack types. Through a comprehensive analysis, it is found that the “begin” category had the lowest volume of data.

To address this data imbalance, a combined file is created by merging the individual attack files. Subsequently, the dataset is modified to equalize the number of instances for each attack type, thereby ensuring a more balanced representation across the categories. This equalization process aimed to prevent bias and ensure that each attack type had an equal opportunity for analysis and modeling. As part of the data preprocessing, rows with missing values (N/A) were removed to ensure the dataset’s integrity. The next step involved examining the data types of the remaining columns. It is identified that several columns contained string-type data, including Timestamp, Destination IP, Source IP, Similar HTTP, and Flow ID. As these columns did not contribute to the numerical analysis and modeling, they were dropped from the dataset.

Moreover, to facilitate further analysis, the attacks were encoded as follows: MSSQL was assigned the value 11101, LDAP was assigned 11111, NetBIOS was assigned 11110, Portmap was assigned 11010, Syn was assigned 1011, UDP was assigned 1001, UDPLag was assigned 1000, and BENIGN (representing non-attack instances) was encoded as 0. This encoding scheme facilitated the classification and analysis of the attack types within the dataset.

5.2.2. Feature Selection

To identify the columns that significantly influence the results of our machine learning model, we employed feature selection techniques, specifically Spearman correlation and Pearson correlation constants. With a dataset containing 89 columns, this step aimed to identify the most relevant features while considering the constraints posed by limited computational resources. This feature selection process aimed to optimize the machine learning model’s performance while considering resource constraints. By identifying the most influential features, we could reduce the number

of columns of the dataset and alleviate the computational burden associated with analyzing all 89 columns.

5.2.3. Optimizing Feature Count

In this stage, our primary objective is to determine the optimal number of columns that would contribute to achieving higher accuracy in our machine-learning model. To accomplish this, we employed the previously discussed feature selection techniques—Spearman correlation constant (SCC) and Pearson correlation constant (PCC)—and integrated them into the machine learning model. The first step involved applying both SCC and PCC to the machine learning model and obtaining correlation scores for every column in the existing dataset. These obtained score values served as indicators of the strength and direction of the relationships between the columns and the target variable. Next, we focused on selecting a subset of columns within a range of 5 to 30 based on the obtained correlation scores. This range is determined to strike a balance between minimizing the number of columns for computational efficiency and maximizing the information captured for accurate predictions. To evaluate the impact of column selection on the accuracy, we calculated the accuracy of the machine learning model for each subset of columns within the specified range. Upon analyzing the results, it is observed that the top 15 ranked columns, as determined by both SCC and PCC, consistently demonstrated a higher accuracy compared to other subsets of columns. This convergence in accuracy for the top 15 ranked columns indicated the effectiveness of the feature selection techniques in identifying influential features that contributed significantly to the predictive power of the machine learning model.

5.2.4. Building the ML model

In this stage, we integrated four distinct machine learning models into our analysis: XGBoost, Decision Tree Classifier, Random Forest Classifier, and Logistic Regression. The primary objective was to select a model that would consume minimal computational resources, including computational power and time. This consideration arose from the fact that the deployed nodes were situated in remote locations with limited computational capabilities, making resource efficiency a critical factor.

Continuing with the methodology we present the further steps undertaken in the research methodology, along with the results obtained from the conducted experiments.

5.3. AVAILABLE DATASET RELATED TO IoT SECURITY

5.3.1. ToN-IoT

The purpose of this dataset [41][42] is to gather and investigate diverse data sources from both the Internet of Things and the Industrial Internet of Things (IIoT). It encompasses heterogeneous data collected from various origins, such as telemetry data from linked devices, system logs from Linux and Windows operating systems, and system network traffic. The dataset is constructed based on a realistic network environment. It aims to assess the accuracy and efficiency of different artificial intelligence-based cybersecurity applications. To achieve this, the ToN-IoT dataset is created to interconnect physical systems, cloud layers, multiple virtual machines, and blurred boundaries. These interactions are dynamically orchestrated using Network Function Virtualization (NFV), service coordination, and Software-Defined Networking technology.

The dataset includes continuous sets of valid and malicious events occurring in IoT services, operating systems, and network systems. Moreover, this dataset is presented in CSV format, with columns having different categories indicating whether the behavior is an attack or non-attack, along with the sub-attack type. The attack subclasses encompass nine different types of attacks, namely Cross-Site Scripting, DDoS, DoS (Denial-of-Service), password cracking attacks, reconnaissance, or verification attempts.

5.3.2. Edge-IIoT

The dataset [43] [42] aimed at cybersecurity applications in the IoT and IIoT focuses on intrusion detection systems utilizing machine learning techniques. It encompasses data from a diverse range of devices, including low-cost digital sensors for humidity and temperature sensing, flame sensors, soil moisture sensors, heart rate sensors, ultrasonic sensors, Ph sensor meters, and more. Within this database, 14 different types of attacks related to IoT and IIoT protocols are analyzed and categorized into five threat categories: Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) attacks, information gathering attacks, injection attacks, man-in-the-middle attacks, and malware attacks. Among the 1,176 features included in the dataset, 61 exhibit significant correlations. Notably, the Edge-IIoT dataset

encompasses a total of 20,952,648 instances, consisting of 11,223,940 normal records and 9,728,708 attack records.

The dataset comprises approximately 2.7 GB of data collected over a span of around 53 hours. It encompasses a big total of 1,194,500 observations, consisting of 1,108,248 benign(non-attack) samples and 88,016 malicious(attack) samples. There are 41 features within the dataset, chosen specifically based on the variation of their values during the different attack phases. The testbed used for collecting the data involved various types of attacks, including command injection, Denial-of-Service (DoS), reconnaissance, and backdoor attacks.

5.3.3. UNSW-NB15

The UNSW-NB15 dataset was published as “Cyber Range” for Cyber Security in 2015 by the Australian Center’s lab and has since become widely utilized within the research community (ACCS). The dataset [44] [41], was generated with the help of raw network packets produced by the IXIA “perfect storm program”. During the testing phase, the authors implemented 9 distinct attack scenarios: worms, exploits, shellcode, reconnaissance, generic attacks, backdoor, analysis, fuzzes, and Denial-of-Service (DoS). A total of 49 network traffic features were extracted from the dataset, employing the programs of Argus and Bro-IDS for this purpose. The dataset contains 2,540,044 streams, consisting of 321,283 aggressive (malicious) and 2,218,761 benign (non-malicious) streams.

5.3.4. Bot-IoT

The development of Bot-IoT [45][46] involved the utilization of a testbed comprising a Node-red tool, network taps, network firewalls, multiple VMs with different operating systems, [45], and the security tool named Argus Network security tool. This dataset consists of various sets and subsets that differ in terms of the number of features, size, and file format. The initial set, known as the Raw Set, includes approximately 71 GB of files of PCAP packet capture. These files capture the network data impeded in the existing testbed environment by tapping the network. Considering this set is in its raw form, it requires processing with network analysis tools like Zeek, Wireshark, or Argus before it becomes usable for traditional ML models.

Consequently, the features derived from the Raw Set may differ based on the chosen tool for PCAP processing.

The second set, called the full Set, comprises CSV files, totaling around 73M instances, obtained with the help of the security tool discussed above. Every element in this collection symbolizes a network session, and the characteristics of the session encompass the combined data of all the packets and bytes involved in a singular communication session between two hosts. Compared to other processed sets or subsets, the Full Set exhibits the lowest number of total features. It consists of 3 dependent features and 26 independent features. It is worth noting that these 26 independent features exclusively comprise the data of network flow from Argus and do not incorporate the additional 14 calculated features that were developed. Moreover, it is essential to highlight that the CSV files of the Full Set do not include a header row, and each file contains 6 columns with no interspersed data among the features. These particulars hold significant importance when attempting to convert or import the data into another format or analysis tool.

5.3.5. CICIDS 2017

The CICIDS2017 dataset [47] comprises a collection of both benign and commonly observed attacks, closely resembling real-world data captured in PCAP format. The dataset provides labeled flows obtained through network traffic analysis employing CICFlowMeter. The category of flow is based on various attributes such as the timestamp, protocols, ports of source and destination, IPs of source and destination, and the presence of an attack. The data is available in CSV files, along with a corresponding definition of the extracted features. Generating realistic background traffic was the main goal while creating the dataset. To achieve this, the B-Profile system proposed by Sharafaldin et al. (2016) was utilized. The B-Profile system profiles the abstract behavior of human interactions and generates benign background traffic that closely mimics naturalistic patterns. In the case of this dataset, the abstract 25 users' behavior was constructed, incorporating protocols like SSH, HTTP, HTTPS, FTP, and email.

Data capture for this dataset commenced on 3-07-2017, at 9 AM, and concluded on 07-07-2017, at 5 PM, spanning a duration of five days. The normal day, Monday,

solely includes benign traffic. The implemented attacks encompass Brute Force FTP, Brute Force SSH, Denial-of-Service (DoS), Heartbleed, Web Attack, Infiltration, Botnet, and Distributed Denial-of-Service (DDoS). These attacks were executed in both the morning and afternoon sessions from Tuesday to Friday.

5.3.6. KDD Cup ‘99

The KDD Cup dataset [48] [49] is widely recognized and extensively employed for conducting experiments related to anomaly detection in computer networks. This dataset was specifically curated for the Competition of the Third Knowledge Discovery and Data Mining Tools, comprising data transfers within a virtual environment. It serves as a subset of the DARPA ‘98 datasets, which was obtained through simulations emulating the operations of a typical US Air Force LAN. The DARPA dataset spanned nine weeks and encompassed TCP dump data. The data collection and distribution of the KDD Cup dataset took place at the Lincoln Laboratory of Massachusetts Institute of Technology (MIT).

The KDD Cup intrusion detection(ID) benchmark comprises three main components. Firstly, the complete KDD Cup dataset includes instances of both attacks and non-attack connections. In total, it encompasses 4,898,431 records of individual connections, each characterized by 41 features classified as either normal or indicative of attacks. Furthermore, all attacks within the KDD Cup dataset are categorized into four distinct groups, as outlined in the corresponding table.

Table 5.1: KDD Cup ‘99 dataset

Attacks Category	Name of Attacks
Probe	satan, portsweep, nmap, ipsweep
DoS	teardrop, smurf, pod, Neptune, land, back
U2R(User to Root)	rootkit, perl, loadmodule, buffer_overflow
R2L(Remote to Local)	warezmaster, warezlient, spy, phf, multihop, imap, guesspasswd, ftp_write.

Table 5.2: Various available datasets on IoT security

Dataset	Access Type	Link	Dataset Detail	
			Type of Attack	Number of Record
TON-IoT	Open	[41]	Backdoor	508116
			DoS	3375328
			DDoS	6165008
			Injection	452659
			MITM	1052
			Scanning	7140161
			Ransomware	72805
			Password	1718568
			XSS	2108944
			Normal	796380
Edge-IIoT	Open	[43]	MITM	290
			Fingerprinting	680
			Ransomware	7760
			XSS	12,060
			Backdoor	19,230
			Password	39,950
			DDoS-HTTP	38,830
			Uploading	29,450
			Vulnerability scanner	40,032
			Port-Scanning	15,992
			DDoS-TCP	40,065
			SQL-Injection	40,671
DDoS-ICMP	54,355			
DDoS-UDP	97,256			

Table 5.2: (Continued)

			Normal	1,091,199
UNSW-NB15	Open	[44]	Generic	215,481
			Exploits	44,525
			Fuzzers	24,246
			Reconnaissance	13,987
			DoS	16,353
			Worms	178
			Shellcode	1515
			Backdoor	2330
			Analysis	2680
			Normal	2,218,764
BoT-IoT	Open	[45]	OS Fingerprinting	358,278
			Service Scanning	1,463,368
			Keylogging	1469
			Data Exfiltration	118
			DoS-TCP	12,315,999
			DoS(UDP)	20,659,489
			DoS(HTTP)	29,709
			DDoS(TCP)	19,547,608
			DDoS(UDP)	18,965,109
			DDoS(HTTP)	19,790
			Normal	9,543

Table 5.2: (Continued)

CIC-IDS 2017	Open	[47]	FTP-Patator	7,945
			DoS Hulk	231,079
			SSH-Patator	5,899
			DoS Golden Eye	10,293
			DoS Slowloris	5,796
			DoS Slowhttptest	5,499
			Heartbleed	11
			Infiltration	36
			XSS	652
			SQL Injection	21
			Brute Force	1,507
			Bot	1,996
			Portscan	158,930
			DDoS	128,027
Normal	2,273,097			
KDD Cup '99	Open	[48]	Probe	41,102
			DoS	3,883,370
			U2R	52
			R2L	1,126

5.4. DATASET EMPLOYED IN OUR EXPERIMENTAL WORK

CICDDoS2019 dataset [50] consists of a combination of benign network traffic and the most up-to-date common Distributed Denial-of-Service (DDoS) attacks. The dataset aims to closely resemble real-world data captured in PCAP format. Additionally, the dataset includes labeled flows obtained through network traffic analysis using CICFlowMeter-V3.

“CICFlowMeter is a comprehensive tool designed for generating and analyzing network traffic flows. It offers the capability to generate bidirectional flows, where the direction of flow is determined by the first packet, establishing both the forward (source to destination) and backward (destination to source) directions.” This allows for the calculation of more than 80 statistical network traffic features, including metrics like Length of packets, Number of bytes, Number of packets, and Duration. These features can be separately computed for the backward and forward directions. The data is provided in CSV files.

During the creation of this dataset, generating realistic background traffic was of utmost importance. To achieve this, the authors employed the B-Profile system. This system profiles the abstract behavior of human interactions and generates realistic benign(non-attacked) background traffic within the proposed architecture of the testbed. For this specific dataset, the abstract 25 users' behavior was constructed, considering protocols like SSH, HTTP, HTTPS, FTP, and email.

Table 5.3: List of machines with their Ips

Machine	Operating System	Internet Protocols Address
Server	Ubuntu 16.04(Web Server)	192.168.50.1(first day)
		192.168.50.4(second day)
PC's(first day)	Win 10	192.168.50.7
PC's(first day) Firewall	Win 8.1	192.168.50.6
	Win Vista	192.168.50.5
	Win 7	192.168.50.8
	Fortinet	205.174.165.81
PC's(second day)	Win 10	192.168.50.8
	Win 8.1	192.168.50.7
	Win Vista	192.168.50.6
	Win 7	192.168.50.9

The dataset encompasses a variety of modern reflective DDoS attacks, including SNMP, DNS, NTP, SYN, UDP-Lag, UDP, MSSQL, LDAP, NetBIOS, and PortMap. These attacks were executed during the designated period. As indicated in Table, a total of 12 DDoS attacks were executed on the day of training, including SNMP, DNS, NTP, SYN, UDP-Lag, UDP, MSSQL, LDAP, NetBIOS, SSDP, WebDDoS, TFTP, and PortMap. On the day of testing, 7 attacks were executed, which include SYN, UDP-Lag, UDP MSSQL, LDAP, NetBIOS, and PortScan.

Table 5.4: Different types of attacks recorded on 2 days with their timing

Day Number	Attacks Types	Attack Time Duration
1	SYN	11:28 – 17:35
	UDP-Lag	11:14 – 11:24
	UDP	10:53 – 11:03
	MSSQL	10:33 – 10:42
	LDAP	10:21 – 10:30
	NetBIOS	10:00 – 10:09
	PortMap	9:43 – 9:51
2	TFTP	13:35 – 17:15
	SYN	13:29 – 13:34
	WebDDoS	13:18 – 13:29
	UDP-Lag	13:11 – 13:15
	UDP	12:45 – 13:09
	SSDP	12:27 – 12:37
	SNMP	12:12 – 12:23
	NetBIOS	11:50 – 12:00
	MSSQL	11:36 – 11:45
	LDAP	11:22 – 11:32
	DNS	10:52 – 11:05
	NTP	10:35 – 10:45

It is worth noting that the WebDDoS attack had a low traffic volume, and PortScan was only executed on the day of testing, making it unexplained for the evaluation of the proposed model. This dataset [51] addresses the limitations and deficiencies found in previous datasets. It offers a comprehensive and fully labeled collection of network traffic data. 80 network traffic features have been extracted and computed for both benign(non-attack) and DoS flows using the widely available CICFlowMeter, which can be accessed through the Cybersecurity website by Canadian Institute [50]. Furthermore, the research paper associated with the dataset performs an in-depth analysis of the generated dataset. This analysis aims to identify the most effective feature sets for detecting various types of DDoS attacks. Specifically, the study focuses on reflective DDoS attacks, such as TFTP, MSSQL, LDAP, and DNS as well as SYN, UDP-Lag, and UDP attacks. By examining and evaluating different feature sets, this research aims to provide insights into the detection and mitigation of these specific types of DDoS attacks.

5.5. IOT TESTBEDS

Simulators serve as valuable tools for researchers, enabling them to conduct feasibility studies for their proposed solutions. However, as the research progresses, a crucial stage emerges wherein it becomes necessary to validate and the proposed solution is implemented in the real-world using hardware and networks [52].

Smart Santander primarily focuses on IoT applications and services specifically designed for smart city domains. The testbed initially encompassed a diverse range of sensors, including Wireless Sensor Networks (WSN), parking sensors, RFID, and QR codes.

These sensors comprised a combination of fixed nodes and mobile nodes, the latter being integrated into vehicles such as buses and taxis. With approximately 20,000 sensor nodes, this extensive testbed serves as a robust infrastructure for conducting experiments and evaluating IoT solutions. The implementation of the Smart Santander testbed leverages the programming languages Java and JavaScript.

Table 5.5: Comparison of Different Testbeds

Testbed	Number of Sensors	Variety of nodes	Mobility	Programming language	API integration	Focused area
FIT-IoT LAB	More than 2,700	Yes(support different types of nodes)	Robot driven	JAVA	REST	Analysis of protocols and algorithms performance
Smart Santander	Approx. 20,000	Yes(support different types of nodes)	a mix of robot and vehicle driven	Java and JS	REST	Smart City application and services
JOSE	The number is huge (exact count not known)	Yes(support different types of nodes)	Not Known	C, Java, and JS	SOAP	Real-time execution of multiple IoT services.

The FIT IoT-LAB represents a prominent open-access testbed designed for conducting extensive IoT experiments. Operating across six distinct sites in France, this testbed offers researchers and developers a robust platform for exploring and advancing IoT technologies. The testbed boasts a significant number of distributed nodes, strategically positioned across the sites. These nodes collectively provide users with a diverse range of precise scientific tools to support the design, development, and optimization of IoT-related devices and systems.

CHAPTER 6

RESULTS

6.1. MACHINE LEARNING MODEL ADOPTED

Our research work is executed on 4 ML models i.e. Random forest, XGBoost, Decision Tree, and Logistic Regression. It has been observed from the experiment that XGBoost outperforms all the other models in all the parameters. The building model adopted in our research work is XGBoost (Extreme Gradient boosting) to build a multi-class classifier that can handle large data sets and also combines the predictions of weak models to make a strong learner.

6.1.1. Logistic Regression

Logistic regression is a supervised machine learning algorithm used for classification tasks, where the goal is to predict the dependent variable based on one or multiple independent features. When dealing with a single independent feature, it is referred to as logistic regression, whereas when there are multiple independent features, it is known as multiple regression.

6.1.2. XGBoost

XGBoost is a powerful ML algorithm that leverages the concept of boosting to enhance predictive accuracy. Boosting is a technique that combines multiple weak models, known as base learners, to create a stronger and more accurate model. XGBoost offers a parallel tree-boosting technique, which means it constructs an ensemble of decision trees sequentially. Each subsequent tree is built to correct the mistakes made by the previous trees, leading to a gradual improvement in the overall prediction. The boosting process involves assigning higher weights to misclassified instances, thereby allowing subsequent trees to focus on those instances and learn from their errors.

6.1.3. Decision Tree

The supervised learning algorithm you are referring to is called a decision tree. Decision trees are widely used for classification tasks, where the goal is to predict the class or category of a target variable based on learned rules derived from previous data. The decision tree starts with a root node that represents the entire dataset. The algorithm then recursively splits the data at each internal node based on the values of specific features. The splitting criteria are determined by selecting the feature that best separates the data into homogeneous classes, aiming to maximize the purity or homogeneity of the resulting subsets.

6.1.4. Random Forest

The random forest classifier is an ensemble machine-learning algorithm that combines the predictions of multiple decision trees to make accurate classifications. It operates by constructing a multitude of decision trees during training, each tree is trained on a random subset of the data and features. During prediction, the random forest aggregates the predictions of all individual trees to arrive at a final classification. This approach helps to mitigate overfitting and improve generalization by introducing randomness and diversity into the model. Furthermore, random forests can handle high-dimensional data, identify important features, and provide estimates of feature importance. Overall, the random forest classifier is renowned for its robustness, scalability, and ability to deliver accurate and reliable predictions across a wide range of classification problems.

6.2. PERFORMANCE EVALUATION

The evaluation of machine learning models involves the calculation of various metrics to assess their performance. These metrics include the accuracy of the model, the confusion matrix, precision, recall, and F-score. By applying different machine learning models, we can compare their accuracies and other performance evaluation metrics, enabling us to make informed decisions about the effectiveness of each model.

6.2.1. Accuracy

In the field of machine learning, accuracy refers to the measure of how well a predictive model can correctly classify or predict instances within a given dataset. It is

an essential evaluation metric used to assess the performance and reliability of machine learning algorithms. Accuracy is typically expressed as a percentage and is calculated by dividing the number of correctly predicted instances by the total number of instances in the dataset. The resulting value represents the proportion of correct predictions made by the model.

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{Total\ Predictions}$$

6.2.2. Precision, Recall, and F1-Score

Precision: can be described as the ratio of correctly classified attacks (TP) with total flows classified as the attack (TP+FP).

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)}$$

Table 6.1: Performance metrics of different machine learning models

Feature Selection Technique Used	Machine Learning Model	Accuracy	Precision	Recall	F1-Score
Pearson Correlation Coefficient(PCC)	Logistic Regression	78.94	69.55	69.44	68.33
	Random Forest Classifier	84.91	85.08	84.97	85.02
	Decision Tree Classifier	85.69	86.77	83.60	83.88
	XGBoost Classifier	86.13	86.84	86.20	85.89
Spearman Correlation Coefficient(SCC)	Logistic Regression	84.89	85.18	77.81	78.52
	Random Forest Classifier	87.76	87.75	87.91	87.78
	Decision Tree Classifier	88.46	89.04	83.45	84.31
	XGBoost Classifier	88.70	89.14	89.42	89.32

Recall: can be described as the ratio of correctly classified attacks (TP) with the total flows which are attacks (TP+FN).

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)}$$

F1-Score: is the harmonic mean(HM) calculated by considering precision and recall.

$$F1 - Score = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

6.2.3. Confusion Matrix

A confusion matrix is a widely used tool in machine learning for evaluating the performance of a classification model. It provides a tabular representation of the predicted and actual class labels of a dataset. Although it may sound complex, it is quite straightforward and useful for assessing the model's accuracy and identifying potential errors. A confusion matrix consists of four components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Here's what each term means:

True Positives (TP): These are the instances where the model correctly predicts the positive class. For example, if the model correctly identifies 50 out of 100 spam emails, then TP would be 50.

True Negatives (TN): These are the instances where the model correctly predicts the negative class. Continuing with the previous example, if the model correctly classifies 900 out of 1,000 non-spam emails, then TN would be 900.

False Positives (FP): These are the instances where the model incorrectly predicts the positive class. In our spam email scenario, if the model incorrectly labels 100 non-spam emails as spam, then FP would be 100.

False Negatives (FN): These are the instances where the model incorrectly predicts the negative class. For instance, if the model fails to identify 50 spam emails out of 100, then FN would be 50.

We have plotted confusion matrices for all the adopted models discussed above with both feature selection techniques.

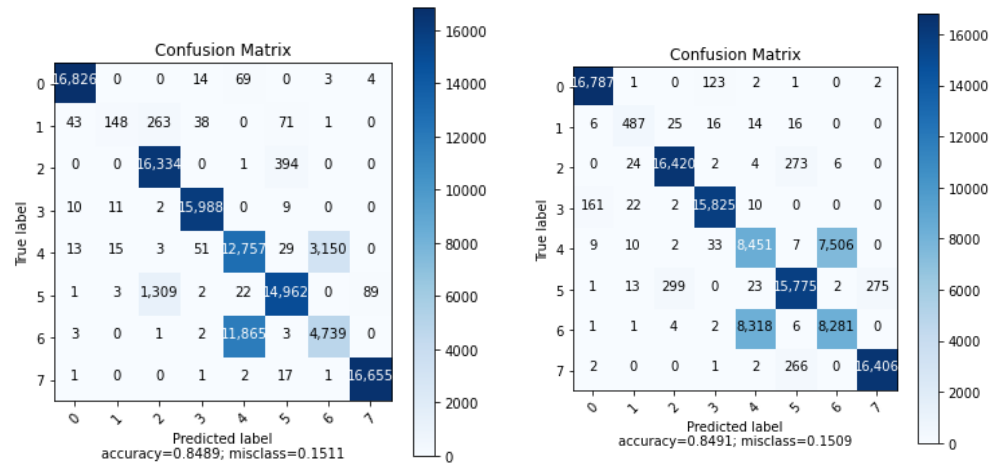


Figure 6.1: Confusion matrix for Logistic Regression with PCC and SCC

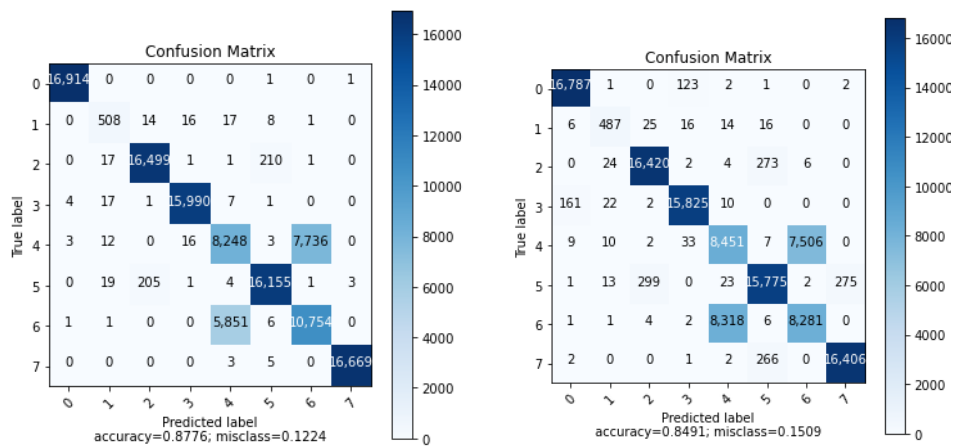


Figure 6.2: Confusion matrix for Random Forest Classifier with PCC and SCC

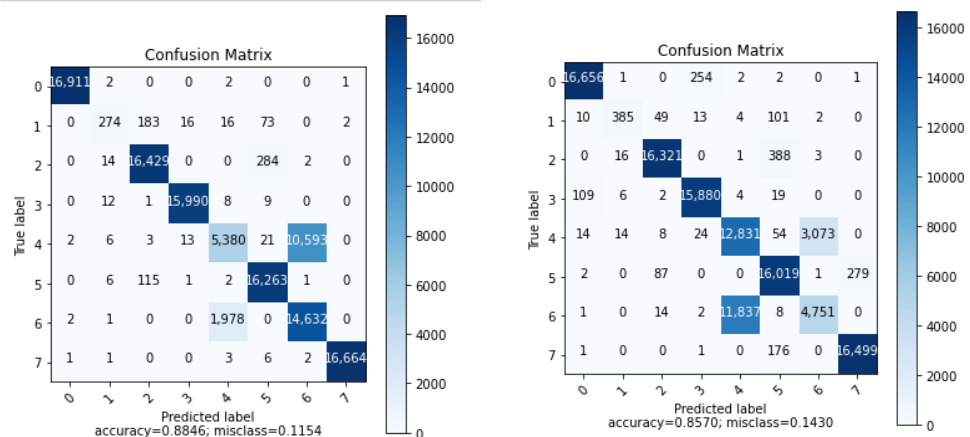


Figure 6.3: Confusion matrix for Decision Tree Classifier with PCC and SCC

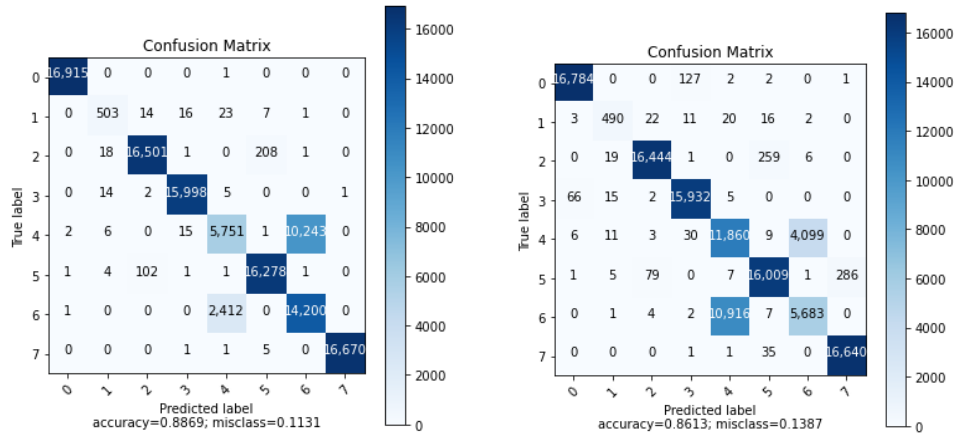


Figure 6.4: Confusion matrix for XGBoost Classifier with PCC and SCC

CHAPTER 7

CONCLUSION AND FUTURE WORK

In conclusion, this study delved into the detrimental impact of DDoS attacks on IoT devices within smart city environments. The increasing interconnectivity of devices in smart cities brings immense advantages, but it also exposes vulnerabilities that malicious actors can exploit. DDoS attacks pose a significant threat to the availability and functionality of IoT devices, thus jeopardizing critical services and essential infrastructure in smart cities. To combat this challenge, a machine learning approach was proposed in this research to predict DDoS attacks on IoT devices within smart cities. By harnessing the power of machine learning algorithms which is combined with efficient feature selection techniques, the proposed solution aims to swiftly detect with less resource consumption and mitigate DDoS attacks in real-time, significantly bolstering the security and resilience of smart city infrastructures with minimum resource requirement. The research undertaken involved the selection of efficient open datasets available which is covering all the recent DDoS attack types. Advanced feature engineering techniques were deployed to extract meaningful insights, while we used an efficient machine learning model which is XGBoost (Extreme Gradient Boosting), and compared its performance with other classical machine learning models such as Decision tree classifier, Logistic regression, Random forest classifier. The experimental results presented in this study underscore the efficacy of the XGBoost classifier with the Spearman Correlation coefficient(SCC) approach in accurately predicting DDoS attacks on IoT devices within smart cities. The model exhibited impressive accuracy, precision, recall, and F1-score, thus affirming its potential to precisely identify and mitigate DDoS attacks in real-world scenarios. The implications of this research are profound for the security and well-being of smart cities. By proactively detecting and mitigating DDoS attacks, the proposed machine learning approach plays a crucial role in thwarting service disruptions, safeguarding critical infrastructure, and ensuring uninterrupted operations of smart city systems. As future work unfolds, refining the model, incorporating additional pertinent features, and evaluating its performance in larger-scale smart city deployments are vital steps for further enhancements.

References

- [1] V. Vujović and M. Maksimović, “Raspberry Pi as a Wireless Sensor node: Performances and constraints,” 2014. doi: 10.1109/MIPRO.2014.6859717.
- [2] S. Singh, P. K. Sharma, B. Yoon, M. Shojafar, G. H. Cho, and I. H. Ra, “Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city,” *Sustain. Cities Soc.*, vol. 63, 2020, doi: 10.1016/j.scs.2020.102364.
- [3] B. N. Silva, M. Khan, and K. Han, “Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities,” *Sustainable Cities and Society*, vol. 38. 2018. doi: 10.1016/j.scs.2018.01.053.
- [4] M. Duygan, M. Fischer, R. Pärli, and K. Ingold, “Where do Smart Cities grow? The spatial and socio-economic configurations of smart city development,” *Sustain. Cities Soc.*, vol. 77, 2022, doi: 10.1016/j.scs.2021.103578.
- [5] S. Makani, R. Pittala, E. Alsayed, M. Aloqaily, and Y. Jararweh, “A survey of blockchain applications in sustainable and smart cities,” *Cluster Comput.*, vol. 25, no. 6, pp. 3915–3936, 2022, doi: 10.1007/s10586-022-03625-z.
- [6] C. S. Lai *et al.*, “A Review of Technical Standards for Smart Cities,” *Clean Technologies*, vol. 2, no. 3. 2020. doi: 10.3390/cleantechnol2030019.
- [7] H. Kumar, M. K. Singh, M. P. Gupta, and J. Madaan, “Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework,” *Technol. Forecast. Soc. Change*, vol. 153, 2020, doi: 10.1016/j.techfore.2018.04.024.
- [8] V. Albino, U. Berardi, and R. M. Dangelico, “Smart cities: Definitions, dimensions, performance, and initiatives,” *J. Urban Technol.*, vol. 22, no. 1, 2015, doi: 10.1080/10630732.2014.942092.
- [9] A. Martinez-Balleste, P. Perez-Martinez, and A. Solanas, “The pursuit of citizens’ privacy: A privacy-aware smart city is possible,” *IEEE Commun. Mag.*, vol. 51, no. 6, 2013, doi: 10.1109/MCOM.2013.6525606.
- [10] N. Moch and W. Wereda, “Smart security in the smart city,” *Sustain.*, vol. 12, no. 23, 2020, doi: 10.3390/su12239900.

- [11] M. Lacinák and J. Ristvej, "Smart City, Safety and Security," in *Procedia Engineering*, 2017, vol. 192. doi: 10.1016/j.proeng.2017.06.090.
- [12] E. Ismagilova, L. Hughes, N. P. Rana, and Y. K. Dwivedi, "Security, Privacy and Risks Within Smart Cities: Literature Review and Development of a Smart City Interaction Framework," *Inf. Syst. Front.*, vol. 24, no. 2, 2022, doi: 10.1007/s10796-020-10044-1.
- [13] Q. Xie and L. Hwang, "Security enhancement of an anonymous roaming authentication scheme with two-factor security in smart city," *Neurocomputing*, vol. 347, 2019, doi: 10.1016/j.neucom.2019.03.020.
- [14] V. Dattana, K. Gupta, and A. Kush, "A probability based model for big data security in Smart city," 2019. doi: 10.1109/ICBDSC.2019.8645607.
- [15] L. Cagliero *et al.*, "Monitoring the citizens' perception on urban security in Smart City environments," in *Proceedings - International Conference on Data Engineering*, 2015, vol. 2015-June. doi: 10.1109/ICDEW.2015.7129559.
- [16] J. Fan *et al.*, "Understanding Security in Smart City Domains From the ANT-centric Perspective," *IEEE Internet Things J.*, 2023, doi: 10.1109/JIOT.2023.3252040.
- [17] K. Y. Lam, S. Mitra, F. Gondesen, and X. Yi, "ANT-Centric IoT Security Reference Architecture - Security-by-Design for Satellite-Enabled Smart Cities," *IEEE Internet Things J.*, vol. 9, no. 8, 2022, doi: 10.1109/JIOT.2021.3073734.
- [18] S. Rani, A. Kataria, M. Chauhan, P. Rattan, R. Kumar, and A. Kumar Sivaraman, "Security and Privacy Challenges in the Deployment of Cyber-Physical Systems in Smart City Applications: State-of-Art Work," *Mater. Today Proc.*, vol. 62, 2022, doi: 10.1016/j.matpr.2022.03.123.
- [19] M. Songhorabadi, M. Rahimi, A. M. MoghadamFarid, and M. Haghi Kashani, "Fog computing approaches in IoT-enabled smart cities," *J. Netw. Comput. Appl.*, vol. 211, p. 103557, Feb. 2023, doi: 10.1016/J.JNCA.2022.103557.
- [20] J. Telo, "Smart City Security Threats and Countermeasures in the Context of Emerging Technologies," *Int. J. Intell. Autom. Comput.*, vol. 6, no. 1, pp. 31–

- 45, Feb. 2023, Accessed: May 16, 2023. [Online]. Available: <https://research.tensorgate.org/index.php/IJIAC/article/view/18>
- [21] P. M. Rao and B. D. Deebak, "Security and privacy issues in smart cities/industries: technologies, applications, and challenges," *J. Ambient Intell. Humaniz. Comput.*, 2022, doi: 10.1007/s12652-022-03707-1.
- [22] H. Zhang, M. Babar, M. U. Tariq, M. A. Jan, V. G. Menon, and X. Li, "SafeCity: Toward Safe and Secured Data Management Design for IoT-Enabled Smart City Planning," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3014622.
- [23] G. S. R. Emil Selvan, R. Ganeshan, I. D. J. Jingle, and J. P. Ananth, "FACVO-DNFN: Deep learning-based feature fusion and Distributed Denial of Service attack detection in cloud computing," *Knowledge-Based Syst.*, vol. 261, 2023, doi: 10.1016/j.knosys.2022.110132.
- [24] N. F. Syed, Z. Baig, A. Ibrahim, and C. Valli, "Denial of service attack detection through machine learning for the IoT," *J. Inf. Telecommun.*, vol. 4, no. 4, 2020, doi: 10.1080/24751839.2020.1767484.
- [25] A. Mihoub, O. Ben Fredj, O. Cheikhrouhou, A. Derhab, and M. Krichen, "Denial of service attack detection and mitigation for internet of things using looking-back-enabled machine learning techniques," *Comput. Electr. Eng.*, vol. 98, 2022, doi: 10.1016/j.compeleceng.2022.107716.
- [26] Kamaldeep, M. Malik, and D. M. Dutta, "Feature Engineering and Machine Learning Framework for DDoS Attack Detection in the Standardized Internet of Things," *IEEE Internet Things J.*, 2023, doi: 10.1109/JIOT.2023.3245153.
- [27] A. Ali Laghari, K. Wu, R. Ali Laghari, M. Ali, and A. Ayub Khan, "A Review and State of Art of Internet of Things (IoT)," *Arch. Comput. Methods Eng.*, vol. 29, pp. 1395–1413, 2022, doi: 10.1007/s11831-021-09622-6.
- [28] S. Li, & Li, D. Xu, and S. Zhao, "The Internet of Things: a survey", doi: 10.1007/s10796-014-9492-7.
- [29] M. Burhan, R. A. Rehman, B. Khan, and B. S. Kim, "IoT elements, layered architectures, and security issues: A comprehensive survey," *Sensors*

(Switzerland), vol. 18, no. 9, 2018, doi: 10.3390/s18092796.

- [30] N. M. Kumar and P. K. Mallick, “The Internet of Things: Insights into the building blocks, component interactions, and architecture layers,” in *Procedia Computer Science*, 2018, vol. 132. doi: 10.1016/j.procs.2018.05.170.
- [31] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electron. Mark.*, vol. 31, no. 3, 2021, doi: 10.1007/s12525-021-00475-2.
- [32] A. Paleyes, R. G. Urma, and N. D. Lawrence, “Challenges in Deploying Machine Learning: A Survey of Case Studies,” *ACM Comput. Surv.*, vol. 55, no. 6, 2022, doi: 10.1145/3533378.
- [33] Himanshu Chauhan, “Overview of Machine Learning,” *Int. J. Adv. Res. Sci. Commun. Technol.*, 2022, doi: 10.48175/ijarsct-4844.
- [34] J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61. 2015. doi: 10.1016/j.neunet.2014.09.003.
- [35] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, 2013, doi: 10.1109/TPAMI.2013.50.
- [36] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *J. Artif. Intell. Res.*, vol. 4, 1996, doi: 10.1613/jair.301.
- [37] V. Uc-Cetina, N. Navarro-Guerrero, A. Martin-Gonzalez, C. Weber, and S. Wermter, “Survey on reinforcement learning for language processing,” *Artif. Intell. Rev.*, vol. 56, no. 2, 2023, doi: 10.1007/s10462-022-10205-5.
- [38] T. Islam, D. M. H. Abid, T. Rahman, Z. Zaman, K. Mia, and R. Hossain, “Transfer Learning in Deep Reinforcement Learning,” in *Lecture Notes in Networks and Systems*, 2023, vol. 447. doi: 10.1007/978-981-19-1607-6_13.
- [39] N. Mishra and S. Pandya, “Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review,” *IEEE Access*, vol. 9. 2021. doi: 10.1109/ACCESS.2021.3073408.
- [40] Y. Shah and S. Sengupta, “A survey on Classification of Cyber-attacks on IoT

- and IIoT devices,” 2020. doi: 10.1109/UEMCON51285.2020.9298138.
- [41] ToN-IoT, <https://research.unsw.edu.au/projects/toniot-datasets> (accessed May 20, 2023).
- [42] I. Tareq, B. M. Elbagoury, S. El-Regaily, and E. S. M. El-Horbaty, “Analysis of ToN-IoT, UNW-NB15, and Edge-IIoT Datasets Using DL in Cybersecurity for IoT,” *Appl. Sci.*, vol. 12, no. 19, 2022, doi: 10.3390/app12199572.
- [43] EdgeIIoT, “No Title.” <https://www.kaggle.com/datasets/mohamedamineferrag/edgeiiotset-cyber-security-dataset-of-iot-iiot> (accessed May 20, 2023).
- [44] “UNSW.” <https://research.unsw.edu.au/projects/unsw-nb15-dataset> (accessed May 20, 2023).
- [45] BoT-IoT, “No Title.” <https://research.unsw.edu.au/projects/bot-iot-dataset> (accessed May 20, 2023).
- [46] J. M. Peterson, J. L. Leevy, and T. M. Khoshgoftaar, “A Review and Analysis of the Bot-IoT Dataset,” 2021. doi: 10.1109/SOSE52839.2021.00007.
- [47] CICIDS, “CICIDS 2017.” <https://www.unb.ca/cic/datasets/ids-2017.html> (accessed May 20, 2023).
- [48] K. CUP, “KDD CUP.” <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html> (accessed on May 20, 2023).
- [49] D. Protić, “Review of KDD Cup ’99, NSL-KDD and Kyoto 2006+ datasets,” *Vojnoteh. Glas.*, vol. 66, no. 3, 2018, doi: 10.5937/vojtehg66-16670.
- [50] “DDos dataset.” <https://www.unb.ca/cic/datasets/ddos-2019.html> (accessed Mar. 20, 2023).
- [51] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, “Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy,” in *Proceedings - International Carnahan Conference on Security Technology*, 2019, vol. 2019-October. doi: 10.1109/CCST.2019.8888419.
- [52] A. Singh, H. Nandanwar, and A. Chauhan, “Simulation Tools and Testbeds for

Internet of Things(IoT): ‘Comparative Insight,’” 2022 2nd Int. Conf. Comput. Sci. Eng. Appl. ICCSEA 2022, 2022, doi: 10.1109/ICCSEA54677.2022.9936302.

PAPER NAME

thesis.docx (2).pdf

WORD COUNT

15887 Words

CHARACTER COUNT

94326 Characters

PAGE COUNT

61 Pages

FILE SIZE

1.3MB

SUBMISSION DATE

May 22, 2023 10:38 AM GMT+5:30

REPORT DATE

May 22, 2023 10:39 AM GMT+5:30

● 10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 6% Internet database
- Crossref database
- 7% Submitted Works database
- 4% Publications database
- Crossref Posted Content database

● Excluded from Similarity Report

- Bibliographic material
- Cited material
- Quoted material
- Small Matches (Less than 10 words)



ICCSEA-2022



Organised by
Department of Computer Science and Engineering,
GIET University, Gunupur

Certificate

This is certified that Mr./ Ms./Prof./Dr. Abhishek Singh
affiliated to Delhi Technological University.....has participated in the “2nd International
Conference on Computer Science Engineering and Applications”(ICCSEA-2022)
held at GIET University, India on 8th September 2022. He also Presented a paper/
Delivered a keynote talk/ Co-ordinated / Organised / Chaired on Technical Session Titled
Simulation Tools and Testbeds for Internet of Things(IoT): “Comparative Insight”

.....
Conference Chair

.....
General Chair



2022 Second International Conference on Computer Science, Engineering and Applications

**(Technically Co-Sponsored by IEEE Kolkata Section & IEEE Bhubaneswar Subsection)
IEEE Conference Record No. : 54677**

Date of Conference: 8th September 2022

Organized by



**Department of Computer Science and Engineering
GIET University, Gunupur-765022, Odisha, India**

The ICCSEA proceedings of the conference will be submitted for publication on IEEE Xplore (Confirmed).

Our ICCSEA-2020 has been published by IEEE Xplore and Indexed with SCOPUS and Web of Science.

<https://www.giet.edu/>

<https://www.iccsea.in/>

We are pleased to welcome you to **2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)** which will be held in the Department of Computer Science and Engineering, GIET University, Gunupur-765022, Odisha, India on 8th September 2022.

2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA) will provide an excellent international forum for sharing knowledge and results in theory, methodology and applications of computer science, engineering and applications. The conference looks for significant contributions to all major fields of the computer science and information technology in theoretical and practical aspects. The aim of the ICCSEA conference is to provide a platform to the researchers and practitioners from both academia as well as industry to meet and share cutting-edge development in the field.

Conference Website: <http://iccsea.in/>

ICCSEA Call for Paper(s): <https://www.iccsea.in/call-for-papers>

Paper submission Link: <https://easychair.org/my/conference?conf=iccsea2022>

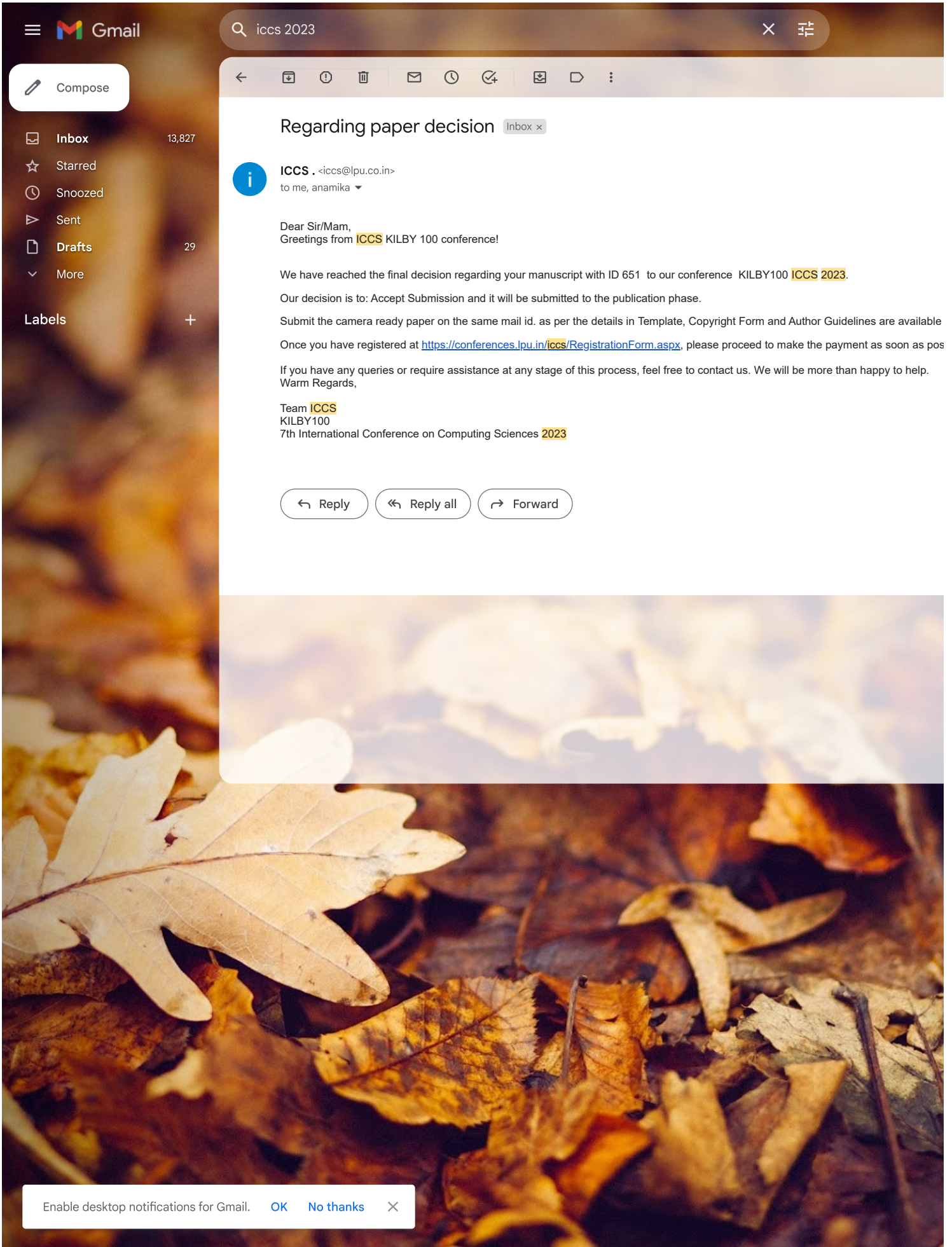
Special Session Proposal(s): <https://www.iccsea.in/special-session>

Registration fee details: <https://www.iccsea.in/registration>

Important Dates for Conference

- **Paper submission deadline: March 30, 2022**
- **Notification of acceptance: April 07, 2022**
- **Camera ready & Registration deadline: May 05, 2022**
- **Conference Date: September 08, 2022**

For any query, kindly email at iccsea@giet.edu





KILBY 100
7TH INTERNATIONAL JOINT CONFERENCE
ON COMPUTING SCIENCES (ICCS-2023)

in association with

Southern Federal University, Russia
(https://sfedu.ru/index_eng.php)

Mizan Tepi University, Ethiopia (<http://www.mtu.edu.et/>)
TH

CONFERENCE MODE:

ONLINE REGISTRATION (REGISTRATIONFORM.ASPX)

SCHEDULE (PDF/KILBY100-TECHNICAL-SESSION-SHEET.XLSX)

KILBY 100 (index.php)



All accepted papers will be published in Scopus indexed proceedings/Journals.

About KILBY100

The year 2023 marks the 100th birth anniversary of Jack Kilby who was an American electrical engineer and inventor who is best known for his contribution to the development of the microchip. He was born on November 8, 1923 in Jefferson City, Missouri, USA. Kilby's invention of the microchip in 1958 revolutionized the electronics industry, making it possible to produce integrated circuits in high volumes, and at low cost. His innovation paved the way for the modern computer and information technology revolution. Kilby received the Nobel Prize in Physics in 2000 for his work in this field. School of Computer Science and Engineering, under the aegis of Lovely Professional University, pays homage to this great engineer by hosting "KILBY100" – 7th International Conference on Computing Sciences.

Objectives of the Conference

- › The main objective of the conference is to provide a unique platform to facilitate the scientists, researchers, academicians, industrialists, and students to share the recent advancements and the challenges in the all aspects of Computer Science and Engineering.
- › To exchange Innovative Ideas among the researchers in the area of Computer Science and Engineering from all around the world.
- › To provide an opportunity for national and international experts and industry leaders to share their experiences and success stories.

THE PATRONS



Shri Ashok K Mittal

CHIEF PATRON

"Chancellor,
Lovely Professional University, Puniab, India "



Smt Rashmi Mittal

PATRON

"Pro Chancellor,
Lovely Professional University, Punjab, India "



Dr. Preeti Bajaj

HONORARY CHAIR

"Vice Chancellor,
Lovely Professional University, Punjab, India "

PREVIOUS CONFERENCES



Turing100 – 1st International Conference on