CHAPTER 1

1.1 Introduction

In this era of advanced technology, the continuously evolving Internet of Things is prominently reflected in our daily lives through smart homes, intelligent scenarios, and smart spaces. IoT, which was often known as the "Internet of Things," was first used in the context of supply chain management in 1999 by Kevin Ashton (Ashton, 2009).IoT can communicate and interact over the Internet or other communication networks(Gubbi et al., 2013).According to statistics, the number of IOT devices worldwide is expected to nearly double from 15.9 billion in 2023 to 32.1 billion by 2030(Lionel, 2024).Smart homes equipped with a large number of continuously operating IoT devices provide us with great convenience and comfort.However,Rapid expansion of the Internet of Things brings innovation and serious security challenges, including malicious attacks and attempts by external network environments that threaten the user privacy, security, and health (Xu et al., 2014).Therefore, it is essential to ensure that smart home devices are protected from malicious attacks and attempts by external network environments, thus safeguarding user privacy, security, and health. Machine learning, which detects normal and abnormal traffic of IoT devices, plays a crucial role in identifying such malicious attacks and attempts. This study tests and evaluates the performance of smart home networks using a machine learning method for detecting abnormal device traffic based on different classifiers. The aim of this study is to enhance the security of IoT devices through machine learning in identifying malicious traffic, thereby protecting user privacy, security, and health.

1.2 Problem Background

The Internet of Things (IoT) aims to facilitate effective communication between the real world and digital equivalents (also known as digital transformation or cyber-physical systems)(Klötzer.2017).So the rapid growth of the Internet of Things -IoT has revolutionized how we interact with our homes, leading to the rise of smart homes. These environments are equipped with a wide array of connected devices—such as smart thermostats, cameras, door locks, lighting systems, and appliances(Alam et al., 2012)—which can provide remote monitoring, automatic control of lighting and heating, and intelligent monitoring(Weber, R. H. ,2010).These devices communicate over networks to enhance convenience, energy efficiency, and security(Alam et al., 2012). However, this growing interconnectedness also introduces significant challenges, particularly in ensuring the security , reliability , and privacy of these systems. One critical challenge is the detection of network anomalies, which can indicate malicious activities, device malfunctions, or other irregularities.

1.3 Problem statement

Smart homes are becoming increasingly popular due to their ability to automate daily tasks and improve quality of life. According to recent studies, the global smart home market is expected to grow significantly in the coming years, with billions of IoT devices being deployed worldwide. These devices rely on network connectivity to function effectively, communicating with each other and with external services (e.g., cloud platforms, mobile apps).

However, this reliance on networking introduces vulnerabilities. Unlike traditional computing devices such as laptops or servers, IoT devices often have limited computational power , minimal security features , and less stringent software update mechanisms . As a result, they are attractive targets for attackers looking to exploit weaknesses in the network.

Smart home IoT networks are increasingly targeted by cyberattacks, device malfunctions, and environmental disruptions, leading to network anomalies (unexpected or malicious deviations from normal behavior). However, existing detection methods struggle to address the heterogeneity , resource constraints , and dynamic communication patterns of IoT devices, leaving smart homes vulnerable to security breaches, privacy leaks, and service disruptions.

1.4 Research Questions

The research question of the research are as follows:

1）Which data sets can be more suitable, more accurate, more representative, and more comprehensive as experimental materials?

2）How many machine learning methods are selected and how many algorithms are used to experiment with the data set?

3）How to evaluate the various performance of machine learning? What methods are used for comparison?

1.5 Research Aim

The purpose of this project is to detect abnormal and normal behavior of IoT device traffic through machine learning based on a smart home anomaly detection method,in order to identify malicious activities on the network,cluding external attacks and attempts .

1.6 Research Objective

The research objectives of this research are follows:

1）To collect more accurate, representative, comprehensive and widely used data sets;

2）To preprocess the data set. By cleaning, balancing and feature selection in advance, the data set can remove noise data and missing information；

3）To transform the data set into a more appropriate feature format for the classification model;

4）To choose a variety of machine learning methods to experiment with the data set to get more experimental results；

5）To analyze the experimental results in various characteristics to determine the machine learning model with the best abnormal performance of network traffic.

1.7 Research Scope (Current Work)

1）The data set uses a more refined and widely used UNSW BoT IoT data set;

2）The data is converted into feature vectors by using unique heat coding and label coding techniques;

3）A variety of machine learning methods are adopted, including AdaBoost, decision tree, random forest, autoencoder and artificial neural network;

4）The performance parameters of the experimental results evaluation include confusion matrix, accuracy, precision, recall and F1 score.

1.8 Expected Research Contribution

This research is a significant contribution to identify malicious patterns in traffic by using machine learning methods based on feature selection. In this way, user privacy, security, and protection can be achieved. The solution of this project shows a path to integrate secure design into existing IoT installations.

REFERENCES

Ashton, Kevin. 2009.“Internet of Things”’ Thing.” RFiD Journal, 1

Gubbi, Jayavardhana, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. 2013. “Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions.” Future Generation Computer Systems 29 (7): 1645–60.

Lionel, Sujay Vailshery. 2022. “Number of IoT Connected Devices Worldwide 2019-2021, with Forecasts to 2030.” August 22, 2022.

Xu, Li Da, Wu He, and Shancang Li. 2014. “Internet of Things in Industries: A Survey.” IEEE Transactions on Industrial Informatics 10 (4): 2233–43.

Klötzer, J. Weißenborn, and A. Pflaum,.2017.“The evolution of cyberphysical systems as a driving force behind digital transformation.”

Alam, M. R., Reaz, M. B. I., & Ali, M. A. M. 2012. A review of smart homes—Past, present, and future . IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) , 42(6), 1190–1203.

Weber, R. H. (2010). Internet of Things – New security and privacy challenges . Computer Law & Security Review, 26(1), 23–30.