# Securing Smart Homes: Threat & Vulnerability Analysis with AI Defense

A Thesis Report Submitted to the
Department of Computer Science and Engineering,
Bangabandhu Sheikh Mujibur Rahman Science and Technology University
in Partial Fulfillment of the Requirements for the Degree of
B.Sc. in Computer Science and Engineering

By
Md. Juwel Mallick
ID: 18CSE018
4th year 2nd semester
Session: 2018 - 2019

Supervised by
Md. Monowar Hossain
Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
BANGABANDHU SHEIKH MUJIBUR RAHMAN SCIENCE AND TECHNOLOGY UNIVERSITY

May 2024

# Abstract

Smart homes are increasingly ubiquitous, offering enhanced convenience, comfort, and efficiency through automated control of various household systems. These integrated technologies provide users with remote access and real-time monitoring capabilities, revolutionizing daily living experiences and paving the way for a more interconnected and digitally-enabled lifestyle. At the same time, various types of threats and security risks are being created due to the weakness of the system. This thesis focuses on the comprehensive risk analysis of smart home automation systems, emphasizing identifying and analyzing potential threats. The research aims to catalog vulnerabilities, quantify security risks, and develop AI models for attack detection. The anticipated outcomes include actionable recommendations to enhance smart home security and contribute valuable insights for homeowners, manufacturers, and policymakers. Ethical considerations guide the research, ensuring a responsible approach to security testing. The thesis addresses unique cybersecurity challenges in smart homes, offering practical solutions for a more secure and resilient ecosystem.

# Declaration

The research work entitled "**Securing Smart Homes: Threat & Vulnerability Analysis with AI Defense**" has been carried out in the Department of Computer Science and Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University is original and conforms to the regulations of this University.

I understand the University's policy on plagiarism and declare that no part of this project has been copied from other sources or been previously submitted elsewhere for the award of any degree or diploma.

**(Supervisor)**

………………………………………………………………

Md. Monowar Hossain

Assistant Professor

Department of Computer Science & Engineering

BSMRSTU

**(Candidate)**

………………………………………………………………

Md. Juwel Mallick

ID:18CSE018

4th Year 2nd Semester.

Department of Computer Science & Engineering

# Acknowledgment

I want sincere gratitude to my Academic Supervisor **Md. Monowar Hossain**, Assistant Professor of the Department of ComputerScience & Engineering, Bangabandhu Sheikh Mujibur Rahman Science & Technology University. Without his kind direction and proper guidance this study would have been a little success. In every phase of the thesis, his supervision and guidance shaped this report to be completed perfectly

I perceive this as a good opportunity for my career development.

I am very grateful to the great creator for completing the thesis proposal.

Md. Juwel Mallick
18CSE018
May, 2024

# Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 Introduction

The integration of diverse smart devices, such as thermostats, security cameras, and voice assistants, has significantly transformed how we engage with and oversee our residences. However, alongside the convenience they offer, smart homes introduce potential vulnerabilities, posing security risks. This thesis endeavors to systematically scrutinize these risks within modern smart home environments. By meticulously assessing various facets of security, including the likelihood of breaches and their potential impacts, this study aims to provide actionable insights for enhancing the resilience and safeguarding of smart home ecosystems against emerging threats.

## 1.2 Background and Motivation

In the modern world, many ecosystems are present such as smart homes, industrial automation, smart cities, eHealth, and cloud computing these are becoming increasingly prevalent due to their ability to integrate Internet of Things (IoT) devices, offering convenience, security, and energy efficiency. [1]

Figure: 1.1 Home automation [2]

Smart homes are one of the major ones equipped with interconnected devices and automated systems. Integrating Internet of Things (IoT) devices into home environments offers unexampled levels of monitoring and control. However, this rise in smart homes has also brought about a significant number of security challenges, threats, and vulnerabilities. These technologies become more deeply embedded in our daily lives, and understanding and preventing risks has become compulsory. This thesis aims to determine security concerns in smart home automation by identifying and analyzing threats, conducting a comprehensive risk analysis, and designing Artificial intelligence models for attack detection. Expected outcomes include providing actionable recommendations for enhancing smart home security and valuable insights for manufacturers, homeowners, and policymakers. The thesis offers practical solutions for a more resilient and secure system and addresses unique cybersecurity challenges in smart homes.

## 1.3 Objectives and Research Questions

Here clearly defines the objectives guiding the research, providing a roadmap for the study. Objectives include the identification of vulnerabilities in smart home systems, the quantification of security risks, and the development of AI models for real-time threat detection.

The main objectives are-

- ✓ Vulnerability Identification
- ✓ Threat Analysis
- ✓ Machine Learning Model Development for IDS
- ✓ Actionable Insights for Homeowners, Manufacturers, and Policymakers
- ✓ Security Awareness Increase

The research questions address specific aspects of smart home security, guiding the investigation into potential risks and their implications.

## 1.4 Significance and Relevance of the Research

The significance of this research lies in its contribution to addressing security concerns in smart home automation. As these technologies become deeply embedded in our daily lives, understanding and mitigating risks become imperative.

Significance or outcomes are-

- Security Awareness Increase
- Vulnerability Identification
- AI Model Development
- Contribution to Cybersecurity
- Stakeholder Impact

Figure: 1.2 Home automation [3]

The study holds relevance in ensuring the sustained trust and adoption of smart home systems in the evolving landscape of modern living. The outcomes of this research are expected to provide actionable insights for homeowners, manufacturers, and policymakers, fostering a more secure and resilient smart home ecosystem.

# Chapter 2

## Methodology

### 2.1 Literature Review

Here analysis is an in-depth examination of existing literature related to smart home security. It encompasses a comprehensive review of relevant research studies. The literature review serves as the foundation for understanding the current state of smart home security, prevalent vulnerabilities, and existing mitigation techniques.

1. The Security Awareness of Smart Home: This foundational thesis extensively examines vulnerabilities in smart home environments, addressing risks in devices, applications, and communication protocols. [4]

2. Detecting Cybersecurity Attacks in IoT Using AI (Machine Learning) Methods: This online thesis explores machine learning in IoT security, offering insights into AI applications for identifying and mitigating security threats, although not specific to smart homes. [5]

3. Ethical Hacking and Penetration Testing in Smart Home Systems: Investigating ethical hacking methodologies, here focuses on smart home environments, providing essential insights for ethical research practices in simulated attack scenarios. [6]

| **Thesis Name** | **Devices** | **Applications** | **Protocols** | **IoT Security** | **AI Applications** | **Ethical Hacking** |
|---|---|---|---|---|---|---|
| **The Security Awareness of Smart Home** | ✓ | ✓ | ✓ | ✓ | | |
| **Detecting Cybersecurity Attacks in IoT** | | | | ✓ | ✓ | |
| **Ethical Hacking and Penetration Testing in IoT** | ✓ | ✓ | | | | ✓ |

Figure: 2.1 some thesis analysis

## 2.2 Risk & Thread Analysis

The risk analysis component also delves into the examination of potential threats and vulnerabilities stemming from both external and internal sources. Through comprehensive vulnerability assessments, the analysis identifies weaknesses in the smart home system's architecture, protocols, and configurations. Additionally, it evaluates the effectiveness of existing security measures and protocols in mitigating these risks. This phase aims to provide actionable insights for strengthening the security posture of smart home automation systems and minimizing the likelihood of successful cyber-attacks.

Figure: 2.2 Risk analysis [7]

### 2.2.1 Environment of Smart Home System

The smart home system environment encompasses a network of interconnected devices and systems like thermostats, lighting, security cameras, locks, and appliances.  These devices

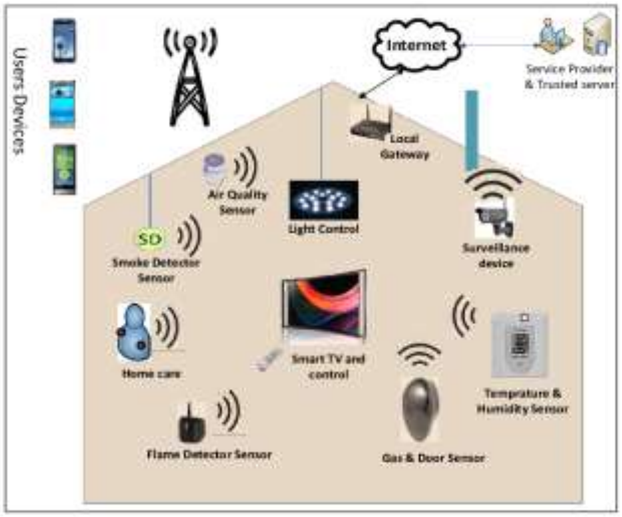

Figure: 2.3 Environment of smart homes [8] communicate through protocols like Wi-Fi, Bluetooth, Zigbee, and Z-Wave, facilitated by a central control hub. Sensors collect data on conditions and user preferences, enabling automated actions and intelligent decision-making. Integration with mobile apps and cloud platforms allows remote access and management.

### 2.2.2 Assets of Stakeholders

Stakeholders in smart home systems, including users, manufacturers, and policymakers, own critical information and physical assets necessitating robust protection measures against cyber threats.

| Information Assets | Physical Assets |
|---|---|
| User credentials | IoT smart devices |
| Information collected by smart devices | IoT hubs |
| Smart home status information | IoT gateways |
| Information about the installed assets | Sensors/Actuators |
| Log information | Cloud server |
| Video, Picture, Voice Information | |
| Location tracking information | |
| Personal information (e.g., health data) | |

Figure: 2.4 Assets of stakeholders [9]

### 2.2.3 Risk & Threat Identify

The main risks that are generally found in smart home automation systems are-

- Unauthorized Access
- Data Leakage
- Device Manipulation
- Insecure Devices and Protocols

- Privacy Concerns

## 2.3 Machine Learning Model Designed for IDS

The IDS machine learning model utilizes various algorithms and features like protocol type and service to accurately detect anomalies in network traffic within smart home environments.

### 2.3.1 Dataset Description

The auditable dataset [2] consists of comprises several intrusions simulated in a military network environment, mimicking a typical US Air Force LAN. It contains TCP/IP dump data that shows connections that have been classified as normal or as attacks with particular attack types. A connection record has about approximately 100 bytes in it. From normal and attack data, 41 quantitative and qualitative features (38 quantitative features and 3 qualitative) are derived for each TCP/IP connection.

2 categories in the class variable:

- Normal
- Anomalous

Database file (2.88 MB) with 25192 rows and 42 columns.

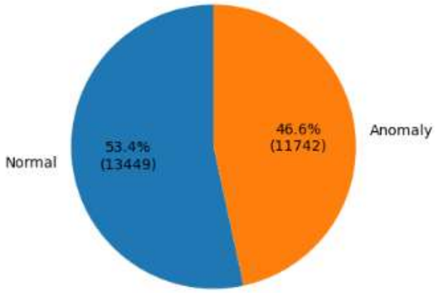

Figure: 2.5 Distribution of Normal and Anomaly Data

### 2.3.2 Data Preprocessing

1) Data Import: The Pandas library was used to import the dataset, which contained network traffic data, into the Python environment.
2) Initial Exploration: To guarantee correct import and comprehension of the data, the initial exploration process includes analyzing the dataset's structure and summary statistics. (Include the table data description)

| | *protocol_type* | *service* | *flag* | *class* |
|---|---|---|---|---|
| **count** | 25192 | 25192 | 25192 | 25192 |
| **unique** | 3 | 66 | 11 | 2 |
| **top** | tcp | http | SF | normal |
| **freq** | 20526 | 8003 | 14973 | 13449 |

Figure: 2.6 All object(string) type features

3) Data Cleaning: Data cleaning steps were taken to address any issues such as duplicates, missing values, or irrelevant features. This guaranteed the quality and consistency of the dataset for further examination.
4) Correlation Analysis: Correlation analysis was used to evaluate multicollinearity and find correlations between variables, which helped to improve feature selection and model construction.[3]
5) Feature Engineering**:** Feature engineering techniques were applied to retrieve appropriate features from the dataset and get it ready for training models.
6) Label encoding: To ensure that machine learning algorithms could work with categorical variables (object type), label encoding was used to convert them to numerical representations. This simplified the modeling procedure and standardized the dataset for analysis.
7) Data Splitting: To ensure there was enough data for both model training and evaluation, the dataset was divided into training and testing sets using a 70-30 ratio.

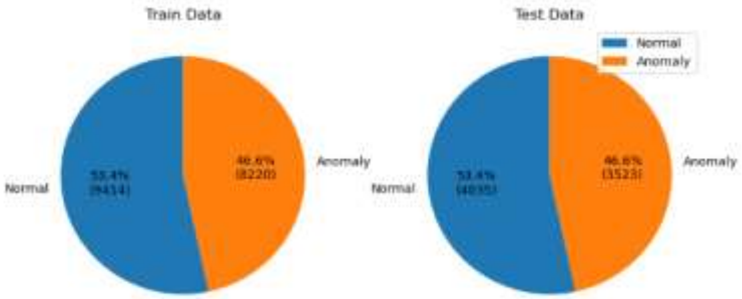

Figure: 2.7 Train data 70% and test 30%

8) Normalization and Standardization: To improve model performance, the characteristics were scaled to a similar range and centered around zero using data normalization and standardization: procedures.

### 2.3.3 Model Building

1) Model Initialization: Initialized K-Nearest Neighbors (KNN), Decision Tree, and XGBoost (XGB) machine learning models for intrusion detection.
*2)* Model Training and Evaluation: Using the training data, the initialized models were trained to identify patterns and correlations in the dataset. Their generalization capacity was then assessed by evaluating their performance using the testing data.
*3)* Performance Metrics Calculation: Several performance indicators were calculated to measure each model's efficacy, including accuracy, precision, recall, and F1-score. While precision evaluated the percentage of true positive predictions among all positive predictions, accuracy tested the total correctness of the model's predictions. The recall

function calculated the percentage of real positive cases that were true positive forecasts. Furthermore, the model's balance between accuracy and recall was assessed using the F1-score, which is computed as the harmonic mean of precision and recall.

4) Feature Importance Analysis: To identify the most important features for distinguishing between normal and anomalous network traffic, a comprehensive analysis of feature importance for both decision tree and XGBoost models was conducted. The goal of this investigation was to pinpoint the critical characteristics that significantly impact the models' ability to predict outcomes and classify data accurately.

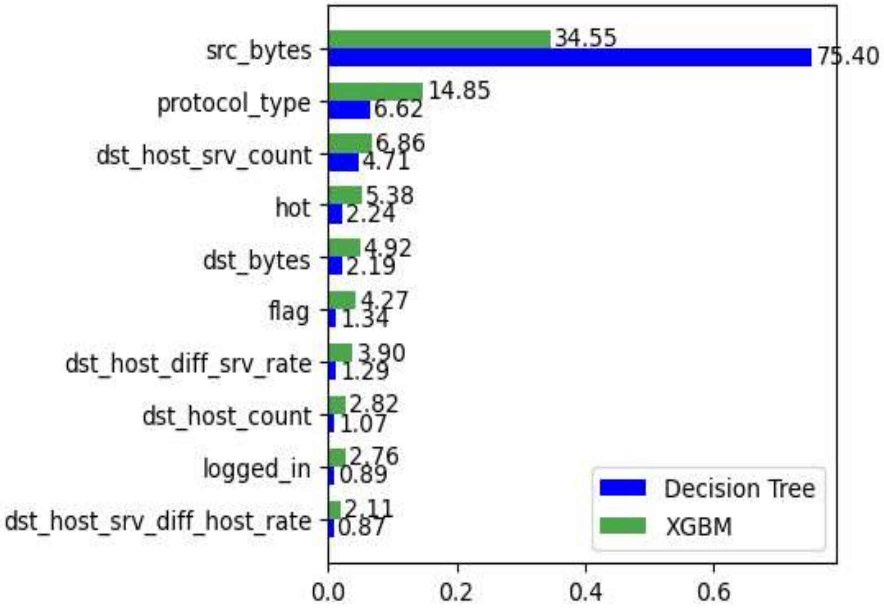

Figure: 2.8 Comparison of top 10 features importance

### 2.3.4 Integration with Network Infrastructure

The incorporation of the Intrusion Detection System (IDS) into the smart home network infrastructure marked an important step in enhancing cybersecurity protocols in home environments. The IDS achieved pervasive visibility by carefully placing sensors and monitoring agents throughout the network topology. This allowed for real-time monitoring and threat detection across network segments and associated devices.

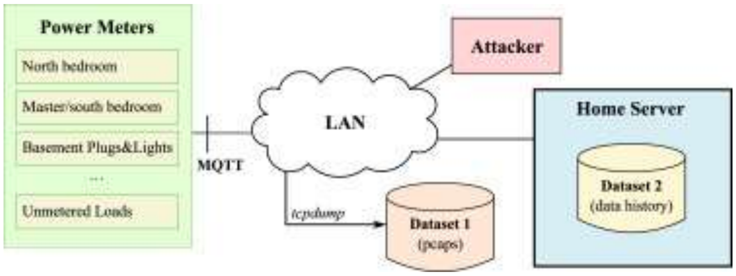

Figure: 2.9 Intrusion & detection in smart homes [10]

The IDS's adaptive resilience against complex cyberattacks was guaranteed by routine performance assessments and iterative enhancements based on changing threat environments. This proactive strategy promotes a culture of security awareness and proactive risk mitigation among homeowners as well as protecting the integrity of smart home ecosystems.

# Chapter 3

## Results and Analysis

### 3.1 Train and Test Accuracy

The train and test accuracy graph show cases the performance of K-Nearest Neighbors (KNN), Decision Tree, and XGBoost (XGB) models. It demonstrates the accuracy achieved by each model during the training and testing phases. KNN exhibits a training accuracy of 99.69% and a testing accuracy of 99.10%, while Decision Tree achieves 99.99% and 99.54% accuracy, and XGB attains 99.99% and 99.67% accuracy, respectively.

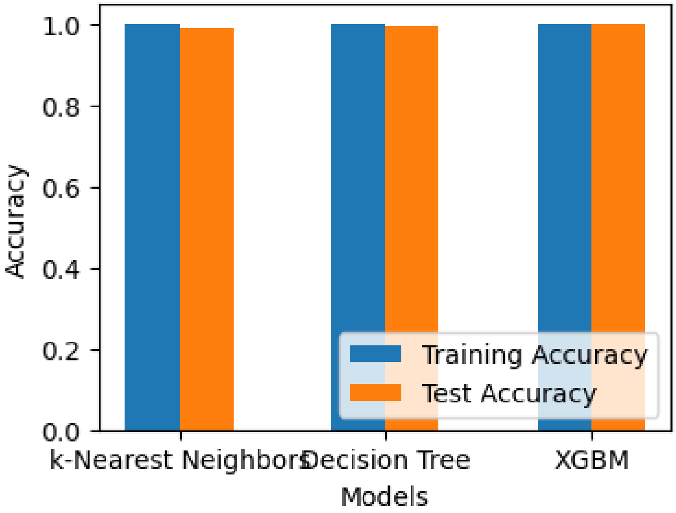

Figure: 3.1 Training and Test Accuracy of Models

### 3.2 Metrics of All Models

The metrics graph compares precision, recall, and F1-score across KNN, Decision Tree, and XGB models. KNN shows a precision of 99.11%, recall of 99.09%, and F1-score of 99.10%. Decision Tree demonstrates precision of 99.54%, recall of 99.53%, and F1-score of 99.54%. XGB exhibits a precision of 99.67%, recall of 99.66%, and F1-score of 99.67%. These metrics aid in assessing the models' effectiveness in accurately detecting and classifying network traffic.

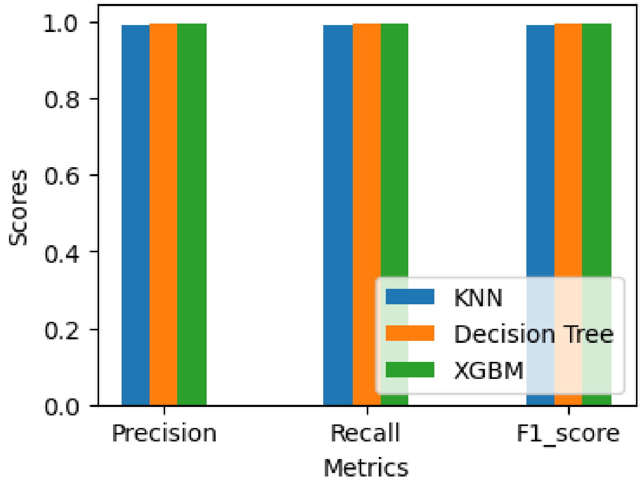

Figure: 3.2 Precision, Recall, and F-1 Score for Different Models.

In the results analysis, it was observed that among the three machine learning models—K-Nearest Neighbors (KNN), Decision Tree, and XGBoost (XGB)—XGBoost consistently demonstrated the highest accuracy across both the training and testing datasets. This indicates that XGBoost outperformed the other models in accurately classifying network traffic and detecting anomalies within the smart home system. Therefore, based on the achieved accuracy rates, XGBoost emerges as the most effective model for intrusion detection in the context of smart home automation systems.

## 3.3 Actions after Suspicious Activity Detection

Following the detection of suspicious activity, proactive measures were implemented to mitigate potential threats and safeguard the integrity of the smart home ecosystem. Infected machines were

Figure: 3.3 Network Traffic Analysis [11] promptly isolated to contain any further network damage, while the blocking of malicious traffic bolstered data and device security. Additionally, any identified security breaches were promptly reported to network administrators or homeowners, facilitating further investigation and resolution.

- Isolation of infected machines
- Blocking of malicious traffic
- Reporting of security breaches to administrators or homeowners

# Chapter 4

## Recommendations, Conclusion, and Future Work

## 4.1 Recommendations

In light of the increasing prevalence of smart home technologies and the associated security risks, it is imperative to take proactive measures to bolster the security posture of smart home automation systems. As smart devices become more integrated into our daily lives, ensuring their security and privacy is paramount to safeguarding users' homes and personal data. To address these concerns, the following recommendations are proposed to mitigate potential security vulnerabilities and enhance the overall security resilience of smart home ecosystems.

### 4.1.1 User Awareness and Education

Provide users with comprehensive training on smart home security best practices, including strong password management, regular software updates, and the use of multi-factor authentication. Educate users about the potential risks associated with smart devices and the importance of maintaining a secure home network.

Figure: 4.1 Multi-factor authentication [12]

### 4.1.2 Manufacturers' Responsibilities

Encourage manufacturers to prioritize security in the design and development of smart home devices, implementing robust encryption protocols and secure authentication mechanisms. Advocate for standardized security protocols across all smart home products to ensure interoperability and compatibility while maintaining high-security standards.

### 4.1.3 Policy Measures

Advocate for the implementation of regulatory frameworks and industry standards that mandate security requirements for smart home devices, including rigorous testing and certification processes. Collaborate with policymakers to develop privacy laws and regulations that protect consumers' data privacy rights and establish clear guidelines for data collection, storage, and sharing practices.

### 4.1.4 Continuous Monitoring and Maintenance

Encourage users to regularly monitor their smart home networks for any unusual activities or unauthorized access attempts, using intrusion detection systems or network monitoring tools. Establish protocols for routine maintenance and updates of smart home devices, ensuring that security patches and firmware updates are promptly installed to address newly discovered vulnerabilities.

### 4.1.5 Collaboration and Information Sharing

Foster collaboration among stakeholders, including users, manufacturers, cybersecurity experts, and policymakers, to share threat intelligence, best practices, and lessons learned in mitigating smart home security risks. Establish community forums or online platforms where users can share their experiences, report security incidents, and seek assistance in addressing security concerns.

Figure: 4.2 Collaboration stakeholder & information sharing [13]

By implementing these recommendations, stakeholders can work together to create a more secure and resilient smart home ecosystem, safeguarding against potential cybersecurity threats and protecting the privacy and safety of users' homes and data.

## 4.2 Challenges and Limitations

This research encountered diverse challenges. Maintaining ethical standards in simulated attacks was crucial. Managing the variety of smart home devices for testing posed a notable challenge. Ensuring timely security patching for identified vulnerabilities demanded precision. Balancing privacy concerns during data collection required a nuanced approach. Resource constraints in time, budget, and tools influenced the study's scope. Despite these challenges, the research yielded valuable insights into the intricate dynamics of smart home security.

The main challenges are-

- Ethical Considerations
- Device Diversity
- Complexity of IoT Ecosystem
- Privacy Concerns
- Resource Limitations

## 4.3 Future Research Directions

To improve the efficacy and accuracy of intrusion detection systems, future research in the field of smart home security may explore advanced anomaly detection methods, such as deep learning and reinforcement learning. In addition, a promising direction for future research is to examine how blockchain technology may be used to protect data integrity and strengthen smart home networks. Some important topics requiring more study include:

- Integration of deep learning and reinforcement learning algorithms for more robust anomaly detection.
- Exploration of blockchain technology to enhance security and integrity in smart home networks.
- Investigation of scalable and resource-efficient solutions for real-time threat detection and response in smart home environments.
- Evaluation of novel encryption and authentication mechanisms to safeguard sensitive data transmitted within smart home networks.
- Design a small-scale smart home system prototype incorporating IoT devices such as smart locks, thermostats, cameras, and motion sensors.

## 4.4 Enhancing Security Measures for Smart Home Devices

In enhancing security for smart home devices, key measures include regular firmware updates, strong authentication, encryption for data protection, access control, intrusion detection systems, secure configurations, privacy settings, vendor accountability, security audits, and user education. These strategies collectively bolster the resilience of smart home ecosystems against cyber threats.

## 4.4 Conclusion

This research project aims to conduct a comprehensive risk analysis of smart home automation systems, focusing on the identification and classification of potential attacks and threats using AI and machine learning to learn security vulnerabilities. The expected results include actionable recommendations for improving smart home security, with significant implications for homeowners, manufacturers, and policymakers

# References

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989

[3] Theien, F.L., 2020. The Security Awareness of Smart Home Users in Norway (Master's thesis, NTNU).

[4] Panda, Mrutyunjaya, A. Mousa Abd Allah, and Aboul Ella Hassanien. "Developing an efficient feature engineering and machine learning model for detecting IoT-botnet cyber-attacks." *IEEE Access* 9 (2021): 91038-91052.

[5] Karie, N.M., Sahri, N.M., Yang, W., Valli, C. and Kebande, V.R., 2021. A review of security standards and frameworks for IoT-based smart environments. *IEEE Access*, *9*, pp.121975-121995.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] Elsayed, Nelly et al. “Intrusion Detection System in Smart Home Network Using Bidirectional LSTM and Convolutional Neural Networks Hybrid Model.” 2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS) (2021): 55-58.

[8] Kavallieratos, G., Chowdhury, N., Katsikas, S., Gkioulos, V. and Wolthusen, S., 2019. Threat analysis for smart homes. Future Internet, 11(10), p.207.

[9] A. J. Alam Majumder, C. B. Veilleux and J. D. Miller, "A Cyber-Physical System to Detect IoT Security Threats of a Smart Home Heterogeneous Wireless Sensor Node," in IEEE Access, vol. 8, pp. 205989-206002, 2020, doi: 10.1109/ACCESS.2020.3037032

[10] Agustín Lara, Vicente Mayor, Rafael Estepa, Antonio Estepa, Jesús E. Díaz-Verdejo, Smart home anomaly-based IDS: Architecture proposal and case study, Internet of Things,Volume 22,2023,100773, ISSN 2542-6605, https://doi.org/10.1016/j.iot.2023.100773.

[11] Alasmari, R. and Alhogail, A.A., 2024. Protecting Smart-Home IoT Devices From MQTT Attacks: An Empirical Study of ML-Based IDS. IEEE Access, 12, pp.25993-26004.

[12] https://helpy.io/blog/why-do-you-need-password-multi-factor-authentication/

[13] https://www.linkedin.com/pulse/quiet-stakeholders-arent-always-happy-paul-slater