



# **PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION**

A Thesis Submitted in Partial  
Fulfillment for the Requirement of the Degree of

**Bachelor of Science**  
in  
Computer Science and Engineering

**by**

**Anowar Hossain**  
ID: 221071051

**Shihab Sarker**  
ID: 202071004

**Under the Supervision of**  
**Tahsin Alam**  
Lecturer  
Department of Computer Science and Engineering

to the  
Department of Computer Science and Engineering  
Shanto-Mariam University of Creative Technology

February, 2026

# ACKNOWLEDGEMENT

This thesis has been submitted to the Department of Computer Science and Engineering of Shanto-Mariam University of Creative Technology (SMUCT), Dhaka, Bangladesh, in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering. The thesis title is “**PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION**”.

First and foremost, we offer our sincere gratitude to our thesis supervisor, **Tahsin Alam**, Lecturer, Department of Computer Science and Engineering, for his continuous support, motivation, and immense knowledge. His valuable guidance and dedicated supervision greatly contributed to the successful completion of this research.

We also express our sincere appreciation to the Head of the Department, **Faisal Imran**, Head & Associate Professor, for his academic leadership and for providing the necessary facilities and institutional support to conduct this research work.

We are extremely grateful to our Thesis Moderator, **Md. Tousif Hasan Lavlu**, Lecturer, Department of Computer Science & Engineering, for his thorough review, insightful feedback, and continuous encouragement. His constructive evaluation was instrumental in refining this thesis.

Our heartfelt thanks are extended to the External Examiner, **Sakhor Das Opi**, Lecturer, Department of Computer Science & Engineering, for his valuable time and constructive suggestions, which have significantly improved the quality and presentation of this work.

Finally, we express our deepest gratitude to our parents and family members for their unconditional love, constant encouragement, and moral support throughout our academic journey. Without them, this achievement would not have been possible.

Anowar Hossain

ID: 221071051

Shihab Sarker

ID: 202071004

February, 2026

SMUCT, Dhaka, Bangladesh

# CERTIFICATE

This is to certify that the thesis entitled ‘‘PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION’’ by **Anowar Hossain** (ID: 221071051) and **Shihab Sarker** (ID: 202071004) has been carried out under my direct supervision. To the best of my knowledge, this thesis is an original work and has not been submitted anywhere for any degree or diploma.

Thesis Supervisor:

.....

**Tahsin Alam**

Lecturer

Department of Computer Science and Engineering  
Shanto-Mariam University of Creative Technology

# CERTIFICATE

This is to certify that the thesis entitled ‘‘PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION’’ by **Anowar Hossain** (ID: 221071051) and **Shihab Sarker** (ID: 202071004) has been carried out under the direct supervision of **Tahsin Alam**. This thesis has been prepared according to the guidelines of the Department of Computer Science & Engineering / Department of Computer Science & Information Technology of Shanto-Mariam University of Creative Technology.

Head of the Department

.....

**Faisal Imran**

Head & Associate Professor

Department of Computer Science & Engineering and

Department of CSIT

Shanto-Mariam University of Creative Technology

# CERTIFICATE

This is to certify that the thesis entitled ‘‘PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION’’ submitted by **Anowar Hossain** (ID: 221071051) and **Shihab Sarker** (ID: 202071004) in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering has been examined and moderated by me.

I have carefully reviewed the thesis defense and documentation. I found it to be a comprehensive and original piece of research work. The thesis successfully meets the academic standards and requirements set by the Department of Computer Science & Engineering / Department of Computer Science & Information Technology of Shanto-Mariam University of Creative Technology.

Thesis Moderator

.....

**Md. Tousif Hasan Lavlu**

Lecturer

Department of Computer Science & Engineering

Shanto-Mariam University of Creative Technology

# CERTIFICATE

This is to certify that the thesis entitled ‘‘PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION’’ submitted by **Anowar Hossain** (ID: 221071051) and **Shihab Sarker** (ID: 202071004) in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering has been examined by me.

I have carefully reviewed the thesis and found that it is an original piece of research work. The thesis meets the academic standards and requirements set by the Department of Computer Science & Engineering / Department of Computer Science & Information Technology of Shanto-Mariam University of Creative Technology.

External Examiner

.....

**Sakhor Das Opi**

Lecturer

Department of Computer Science & Engineering  
Shanto-Mariam University of Creative Technology

# **DECLARATION**

We hereby declare that the thesis entitled “PRIVACY-PRESERVING REAL-TIME HOME AUTOMATION UTILIZING MQTT PROTOCOL AND SENSOR ANOMALY DETECTION WITH GENAI INTEGRATION” is the result of our own independent research work carried out under the supervision of **Tahsin Alam**, Lecturer, Department of Computer Science and Engineering, Shanto-Mariam University of Creative Technology.

This thesis has not been submitted, either in whole or in part, to any other university or institution for the award of any degree, diploma, or other qualification. All sources of information used in this research have been duly acknowledged through proper references and citations.

We take full responsibility for the authenticity and accuracy of the work presented in this thesis.

---

**Anowar Hossain**

B.Sc. in Computer Science and Engineering

ID: 221071051

Department of Computer Science and Engineering

Shanto-Mariam University of Creative Technology

---

**Shihab Sarker**

B.Sc. in Computer Science and Engineering

ID: 202071004

Department of Computer Science and Engineering

Shanto-Mariam University of Creative Technology

# Abstract

The rapid proliferation of Internet of Things (IoT) devices in smart homes has introduced critical vulnerabilities regarding data privacy and security. Conventional cloud-centric architectures often expose sensitive user data to third-party risks, necessitating a shift toward secure, edge-based solutions. This research presents “IoTShield”, a comprehensive privacy-preserving home automation framework that integrates the lightweight MQTT protocol with Generative AI for intelligent anomaly detection. The system utilizes a hybrid edge-computing architecture featuring ESP32 microcontrollers and Raspberry Pi gateways to process data locally. To ensure robust privacy, a dual-layer Differential Privacy mechanism using Gaussian noise is applied, complemented by end-to-end RSA-2048 encryption for secure transmission. For anomaly detection, the system integrates a local Large Language Model (Llama 3.2:1B) to analyze sensor patterns and generate context-aware alerts with explainable insights. Experimental results from over 13,000 sensor readings demonstrate that the system achieves an end-to-end latency of under 2 seconds and successfully classifies anomalies across four severity levels (Low to Critical). This study confirms that combining local GenAI with cryptographic privacy mechanisms preserves data sovereignty while delivering a highly responsive and secure smart home environment.

**Keywords:** IoT Security, MQTT Protocol, Generative AI, Llama 3.2, Anomaly Detection, Differential Privacy, RSA Encryption, Edge Computing, Smart Home Automation

# Contents

<b>ACKNOWLEDGEMENT</b>	<b>i</b>
<b>CERTIFICATE</b>	<b>ii</b>
<b>CERTIFICATE</b>	<b>iii</b>
<b>CERTIFICATE</b>	<b>iv</b>
<b>CERTIFICATE</b>	<b>v</b>
<b>DECLARATION</b>	<b>vi</b>
<b>Abstract</b>	<b>vii</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xii</b>
<b>List of Abbreviations</b>	<b>xiii</b>
<b>List of Symbols</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and Motivation . . . . .	1
1.2 Problem Statement . . . . .	2
1.3 Research Gap . . . . .	3
1.4 Research Objectives . . . . .	4
1.5 Contributions of the Thesis . . . . .	4
1.6 Scope of the Thesis . . . . .	5
1.7 Thesis Organization . . . . .	6
<b>2 Related Work</b>	<b>8</b>
2.1 MQTT-Based Smart Home Architectures . . . . .	8
2.2 Privacy-Preserving IoT Systems . . . . .	9
2.3 Cryptographic Mechanisms in IoT Security . . . . .	10

2.4	Anomaly Detection Techniques in IoT . . . . .	11
2.5	Generative AI for Real-Time Monitoring . . . . .	12
2.6	Limitations of Existing Approaches . . . . .	13
<b>3</b>	<b>Proposed System</b>	<b>16</b>
3.1	Overall System Overview . . . . .	16
3.2	System Architecture Design . . . . .	18
3.3	Hardware Architecture (ESP32 and Sensors) . . . . .	20
3.4	Software Architecture . . . . .	22
3.5	MQTT Communication Model . . . . .	23
3.6	Privacy-Preserving Data Pipeline . . . . .	25
3.6.1	Differential Privacy with Dual Gaussian Noise . . . . .	25
3.6.2	RSA-2048 Encryption Mechanism . . . . .	25
3.7	AI-Based Anomaly Detection Framework . . . . .	27
3.7.1	Threshold-Based Detection Logic . . . . .	27
3.7.2	Local LLM Integration (Llama 3.2 via Ollama) . . . . .	27
3.8	End-to-End System Workflow . . . . .	29
<b>4</b>	<b>Implementation</b>	<b>30</b>
4.1	Hardware Implementation Details . . . . .	30
4.2	Software Development Environment . . . . .	30
4.3	MQTT Broker Configuration . . . . .	30
4.4	Differential Privacy Implementation . . . . .	30
4.5	RSA Encryption Implementation . . . . .	30
4.6	Local LLM Deployment Setup . . . . .	30
4.7	Database Design and Storage Model . . . . .	30
<b>5</b>	<b>Results and Discussion</b>	<b>31</b>
5.1	Experimental Setup . . . . .	31
5.2	Performance Evaluation . . . . .	31
5.2.1	System Latency Analysis . . . . .	31
5.2.2	Encryption Overhead Analysis . . . . .	31
5.3	Anomaly Detection Results . . . . .	31
5.4	AI Response Evaluation . . . . .	31
5.5	Privacy–Utility Trade-off Analysis . . . . .	31
5.6	Security Analysis . . . . .	31
5.7	Comparative Discussion . . . . .	31
<b>6</b>	<b>Conclusion and Future Work</b>	<b>32</b>
6.1	Summary of the Work . . . . .	32

6.2 Key Findings . . . . .	32
6.3 Limitations of the Study . . . . .	32
6.4 Future Research Directions . . . . .	32
<b>References</b>	<b>32</b>
<b>A Appendix</b>	<b>40</b>

# List of Figures

3.1	High-level architectural overview of the proposed IoTShield system, illustrating the secure data flow from the edge sensor nodes to the localized Generative AI analysis framework. . . . .	17
3.2	Detailed network topology of the IoTShield framework, illustrating the local IP subnet structure, port assignments, and the strict isolation of edge processing from the public internet. . . . .	19
3.3	MQTT communication architecture illustrating the hierarchical topic structure, message flow between edge nodes and the backend, and key protocol parameters. . . . .	24
3.4	Dual-layer privacy architecture of the IoTShield system, illustrating the sequential transformation of raw sensor data through Gaussian noise injection and RSA-2048 encryption. . . . .	26
3.5	AI-powered anomaly detection workflow illustrating the data routing logic and the localized integration of the Llama 3.2 LLM via Ollama. . . . .	28

# List of Tables

2.1	Summary of Existing Literature and Comparative Analysis . . . . .	14
3.1	Hardware Components and Sensor Specifications . . . . .	21
3.2	MQTT Topic Structure and Payload Formats . . . . .	23
3.3	Sensor Threshold Parameters and Severity Classification . . . . .	27

## List of Abbreviations

AES	Advanced Encryption Standard
AI	Artificial Intelligence
API	Application Programming Interface
CPU	Central Processing Unit
CSE	Computer Science and Engineering
DoS	Denial of Service
ESP32	Espressif Systems 32-bit Microcontroller
GenAI	Generative Artificial Intelligence
GPIO	General Purpose Input/Output
HTTP	Hypertext Transfer Protocol
IDE	Integrated Development Environment
IoT	Internet of Things
IP	Internet Protocol
JSON	JavaScript Object Notation
JWT	JSON Web Token
LED	Light Emitting Diode
LLM	Large Language Model
LDR	Light Dependent Resistor
MQTT	Message Queuing Telemetry Transport
PIR	Passive Infrared Sensor
QoS	Quality of Service
RAM	Random Access Memory
REST	Representational State Transfer
RSA	Rivest–Shamir–Adleman (Encryption Algorithm)
SMTP	Simple Mail Transfer Protocol
SSL	Secure Sockets Layer
TCP	Transmission Control Protocol
TLS	Transport Layer Security
UI	User Interface
URL	Uniform Resource Locator
Wi-Fi	Wireless Fidelity

# List of Symbols

$M$	Plaintext Message
$C$	Ciphertext (Encrypted Message)
$K_{pub}$	Public Key
$K_{priv}$	Private Key
$e$	Public Exponent (RSA)
$d$	Private Exponent (RSA)
$n$	Modulus (RSA)
$\mu$	Mean Value (Gaussian Distribution)
$\sigma$	Standard Deviation
$\epsilon$	Privacy Budget / Noise Parameter
$N(\mu, \sigma^2)$	Gaussian (Normal) Distribution
$T$	Temperature Sensor Reading
$H$	Humidity Sensor Reading
$G$	Gas Sensor Value
$L_{lat}$	System Latency
$P(x)$	Probability of an Event
$V_{in}$	Input Voltage
$R$	Resistance
$t$	Time / Timestamp
$\Delta$	Difference / Change in Value
$\theta$	Threshold Value for Anomaly Detection
$Hz$	Hertz (Frequency Unit)
$dB$	Decibel (Signal Strength Unit)
$bps$	Bits Per Second

## Chapter 1

# Introduction

## 1.1 Background and Motivation

The rapid evolution of the Internet of Things (IoT) has fundamentally transformed modern living spaces into intelligent, interconnected environments. Smart home automation systems now routinely deploy diverse networks of sensors and actuators to monitor environmental conditions, manage energy consumption, and significantly enhance user convenience. To facilitate this seamless interaction, lightweight communication protocols, particularly the Message Queuing Telemetry Transport (MQTT) protocol, have become the industry standard. MQTT enables efficient, low-latency data transmission even across highly constrained networks and low-power microcontrollers, making it ideal for real-time residential applications.

Despite the tremendous advantages of these smart home ecosystems, their widespread adoption introduces critical vulnerabilities surrounding data privacy and security. Continuous environmental monitoring generates a massive influx of granular data, capturing intimate details about household occupancy, daily routines, and behavioral habits. In conventional architectures, this highly sensitive data is frequently transmitted in its raw form to centralized cloud servers for processing. This heavy reliance on third-party cloud infrastructure exposes users to severe risks, including data breaches, unauthorized surveillance, and corporate data profiling. While basic transport-layer security is sometimes implemented, true end-to-end data confidentiality at the application layer remains a significant challenge in residential deployments.

Beyond privacy concerns, traditional smart home systems are largely constrained by rudimentary, rule-based logic. Most commercial setups rely on static, predefined thresholds to trigger alerts—for example, activating a notification only if a gas or temperature sensor strictly exceeds a hardcoded limit. This rigid approach lacks contextual awareness. It frequently results in a high volume of false positives, struggles to identify complex anomalies that span multiple sensor types, and provides users with generic, numerical warnings rather than clear, actionable insights into the actual problem.

Motivated by these profound limitations, there is a pressing need for a paradigm shift toward secure, intelligent, and decentralized home automation. The integration of Generative Artificial Intelligence (GenAI), specifically localized Large Language Models (LLMs), presents a revolutionary opportunity to move beyond simple threshold alerts to context-aware, human-readable anomaly detection. This research is driven by the ambition to

engineer a holistic framework that marries the efficiency of MQTT with rigorous privacy-preserving mechanisms—such as dual-layer Gaussian noise and RSA encryption—processed entirely at the network edge. By doing so, the proposed system aims to deliver highly responsive, intelligent monitoring while guaranteeing absolute data sovereignty for the user.

## 1.2 Problem Statement

The integration of Internet of Things (IoT) devices into residential environments has significantly improved building automation, safety, and energy management. However, the standard architecture of these systems fundamentally compromises user privacy and data security. Most commercial smart home setups rely heavily on centralized, cloud-based processing, where raw, granular sensor data—such as room occupancy, temperature fluctuations, and lighting usage—is continuously transmitted over the internet. This continuous data stream creates a high-risk attack surface. If intercepted or analyzed by unauthorized entities, this information can be used to deduce sensitive domestic activities, leaving users vulnerable to surveillance, corporate data profiling, and malicious physical intrusions.

Furthermore, the communication protocols utilized by resource-constrained IoT devices, while lightweight and efficient, often lack robust application-layer security. For instance, while the MQTT protocol is highly optimized for low-bandwidth environments, standard implementations frequently leave the data payload entirely readable by the central broker. Even when transport-layer security (TLS) is applied, the centralized broker remains a single point of failure that can expose historical and real-time data if compromised. There is a distinct lack of mechanisms that protect the structural privacy of the data—such as cryptographic encryption or differential privacy—before it even leaves the physical edge device.

In addition to these severe privacy vulnerabilities, current home automation systems suffer from severely limited analytical capabilities. They typically depend on rigid, rule-based logic to trigger alerts, such as activating an alarm only when a gas or temperature sensor crosses a strict, hardcoded numerical limit. This rudimentary approach fails to account for environmental context, frequently leading to a high rate of false positives and a lack of actionable information. When complex, multi-sensor anomalies occur, users are presented with generic numerical warnings rather than a clear, intelligent explanation of the event. While advanced machine learning models can offer more sophisticated detection, they are usually too computationally expensive to deploy locally on edge-based microcontrollers, forcing a reliance back onto the insecure cloud.

Therefore, the primary problem addressed in this research is the lack of a secure, edge-centric framework capable of performing real-time home automation and intelligent anomaly

detection without compromising data sovereignty. Specifically, the challenge lies in securing lightweight MQTT communications through application-layer encryption and noise-based privacy mechanisms, while simultaneously deploying a Generative AI model locally to analyze complex sensor patterns and generate human-readable, context-aware alerts within a decentralized architecture.

### 1.3 Research Gap

Despite the extensive body of literature on smart home automation and Internet of Things (IoT) security, a significant gap remains at the intersection of lightweight communication, robust data privacy, and intelligent anomaly detection. A comprehensive review of existing methodologies reveals a persistent dichotomy: systems are typically either highly secure but computationally heavy, or lightweight but dangerously vulnerable. While the MQTT protocol is widely celebrated for its low latency and efficiency in constrained environments, prevailing implementations primarily rely on transport-layer security (TLS). This approach leaves the data payload unencrypted at the broker level, creating a critical vulnerability if the central server is compromised. Furthermore, the application of structural privacy mechanisms—such as injecting differential noise at the sensor node before transmission—is rarely explored in real-time, lightweight residential architectures.

In the domain of anomaly detection, current smart home deployments predominantly utilize either static, rule-based thresholds or traditional machine learning models (e.g., Support Vector Machines and Random Forests). While threshold-based systems are efficient, they lack the contextual awareness necessary to identify sophisticated, multi-sensor anomalies, leading to high false-positive rates. Conversely, while traditional machine learning models can detect deviations, they function as “black boxes,” providing numerical anomaly scores without any human-readable explanation or actionable context. When an alert is triggered, the user is left to interpret raw data rather than receiving a clear, reasoned diagnosis of the potential hazard.

Recently, Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) have demonstrated exceptional capabilities in contextual reasoning and data interpretation. However, their application in IoT environments is heavily biased toward cloud-centric architectures due to their immense computational requirements. The deployment of localized, edge-optimized LLMs to process raw sensor telemetry and generate intelligent, real-time alerts without relying on third-party cloud APIs remains significantly under-researched. Existing studies that do explore AI-driven IoT security mostly focus on high-level network traffic analysis rather than granular, domestic environmental sensor data.

Consequently, the fundamental research gap lies in the absence of a holistic, decentral-

ized smart home framework that unifies these disparate elements. There is a distinct lack of systems that successfully combine application-layer cryptographic payload encryption, privacy-preserving data obfuscation, and edge-deployed Generative AI to provide secure, real-time, and explainable anomaly detection within a lightweight MQTT ecosystem.

## 1.4 Research Objectives

The primary aim of this research is to design, implement, and evaluate a decentralized, privacy-preserving smart home automation system capable of intelligent, real-time anomaly detection. To achieve this overarching goal, the research is guided by the following specific objectives:

- **To design a lightweight, edge-centric IoT architecture:** Develop a responsive smart home network utilizing ESP32 microcontrollers, a Raspberry Pi edge gateway, and the MQTT protocol to ensure low-latency communication without heavy reliance on external cloud infrastructure.
- **To implement a dual-layer privacy and security pipeline:** Engineer a robust data protection mechanism that applies structural obfuscation via dual-layer Gaussian noise (differential privacy) and secures the MQTT payload using application-layer RSA-2048 encryption before data transmission.
- **To integrate localized Generative AI for anomaly detection:** Deploy a lightweight Large Language Model (Llama 3.2:1B via Ollama) to autonomously analyze multi-sensor telemetry, replacing rigid threshold-based logic with context-aware, human-readable alert generation.
- **To construct an interactive, full-stack monitoring platform:** Build a comprehensive backend and web dashboard (utilizing Django, Tailwind CSS, and Chart.js) to manage devices, decrypt sensor data, and visualize real-time AI-generated insights.
- **To empirically evaluate system performance:** Validate the proposed IoTShield framework in a real-world deployment by analyzing key metrics, including end-to-end latency, encryption overhead, anomaly detection accuracy, and the overall privacy-utility trade-off.

## 1.5 Contributions of the Thesis

This thesis makes several significant contributions to the fields of Internet of Things security, edge computing, and smart home automation. By addressing the critical vulnerabil-

ties of cloud-dependent architectures, the proposed IoTShield system introduces a novel paradigm for domestic environmental monitoring.

The core contributions of this research are summarized as follows:

- **A Novel Edge-Based GenAI Integration:** This research demonstrates the feasibility and immense value of deploying localized Generative AI at the network edge. By utilizing a 1-billion parameter LLM to process sensor data locally, the system successfully transitions anomaly detection from binary, threshold-based triggers to intelligent, context-aware reasoning, providing users with actionable and highly accurate alerts.
- **Application-Layer MQTT Security Architecture:** Unlike standard IoT deployments that rely solely on transport-layer security, this thesis contributes a fully functional RSA-2048 encryption scheme specifically optimized for MQTT payloads. This ensures that even if the central message broker is compromised, the granular sensor telemetry remains strictly confidential.
- **Practical Implementation of Differential Privacy in IoT:** The research introduces a dual-layer Gaussian noise injection mechanism directly at the sensor node. This contributes a practical methodology for masking exact domestic routines and behavioral patterns while preserving enough data utility for the AI model to accurately detect critical anomalies (e.g., fires or gas leaks).
- **Empirical Validation of a Production-Ready Framework:** The thesis provides comprehensive, real-world validation of the proposed architecture. Through the deployment of physical hardware and simulated nodes generating over 13,000 sensor readings and validating over 1,500 distinct alerts, the study proves that a highly secure, AI-driven IoT system can operate with an average end-to-end latency of less than two seconds.

## 1.6 Scope of the Thesis

The scope of this research is specifically bounded to the design, implementation, and evaluation of a privacy-preserving, AI-driven monitoring framework tailored for residential smart home environments. The project focuses strictly on secure environmental data collection, anomalous event detection, and real-time alert generation, rather than large-scale industrial IoT (IIoT) deployments or complex physical robotic actuation.

In terms of hardware and physical deployment, the scope encompasses the development of a fully operational real-world sensor node. This node utilizes a physical ESP32 microcontroller integrated with six distinct environmental sensors (temperature, humidity, gas, flame,

motion, and light) to validate real-world data acquisition. To rigorously evaluate the system’s network handling, multi-device architecture, and edge processing capabilities without deploying extensive hardware, the scope also incorporates simulated IoT nodes. These consist of a simulated secondary ESP32 hub and a simulated Raspberry Pi edge gateway, which additionally provides system performance metrics such as CPU and memory usage.

From a software and security perspective, the research is limited to securing the MQTT communication protocol at the application layer. This involves the implementation of a custom RSA-2048 encryption pipeline for payload confidentiality and a dual-layer Gaussian noise mechanism to ensure differential privacy. The thesis does not attempt to invent new foundational cryptographic algorithms or replace standard transport-layer security (TLS); instead, it focuses on practical, edge-level data obfuscation to prevent broker-level data compromises.

Finally, the intelligence and anomaly detection scope is centered entirely on edge-optimized Generative AI. The system integrates a locally hosted, 1-billion parameter Large Language Model (Llama 3.2 via Ollama) to process decrypted sensor telemetry and generate context-aware, human-readable alerts. The research deliberately excludes the use of proprietary, cloud-based AI APIs to strictly maintain data sovereignty and adhere to decentralized edge-computing principles. Furthermore, the development of a responsive backend and web dashboard (utilizing Django, SQLite, and Tailwind CSS) is included within the scope to demonstrate end-to-end data persistence, real-time visualization, and practical alert management.

## 1.7 Thesis Organization

The remainder of this thesis is logically structured into five subsequent chapters, each detailing a critical phase of the research, development, and evaluation of the proposed IoTShield framework.

**Chapter 2: Related Work** provides a comprehensive review of the existing literature surrounding smart home automation. It examines current MQTT-based architectures, privacy-preserving IoT systems, and the application of cryptographic mechanisms in constrained environments. Furthermore, it explores traditional anomaly detection techniques alongside the emerging role of Generative AI in real-time monitoring, ultimately highlighting the crucial limitations and research gaps that this thesis addresses.

**Chapter 3: Proposed System** outlines the theoretical and structural design of the IoT-Shield architecture. It details the holistic system overview, defining the roles of the physical ESP32 sensors and the simulated edge gateway. This chapter extensively covers the MQTT communication model, the privacy-preserving data pipeline incorporating dual-layer Gaus-

sian noise and RSA-2048 encryption, and the localized AI-based anomaly detection framework driven by Llama 3.2.

**Chapter 4: Implementation** translates the proposed architecture into a functional, real-world application. It documents the exact hardware setup, the software development environment (including Django, Tailwind CSS, and Chart.js), and the configuration of the Mosquitto MQTT broker. Additionally, it details the practical coding implementations of the differential privacy mechanisms, the application-layer encryption, the local LLM deployment, and the relational database schema utilized for data persistence.

**Chapter 5: Results and Discussion** presents a rigorous empirical evaluation of the operational system. Based on an extensive dataset of over 13,000 sensor readings and more than 1,500 validated alerts, this chapter analyzes system latency, cryptographic processing overhead, and the accuracy of the AI-driven anomaly detection. It also provides a critical discussion on the privacy-utility trade-off and compares the proposed framework's performance against traditional baseline models.

**Chapter 6: Conclusion and Future Work** synthesizes the key findings of the research, summarizing the successful integration of privacy-preserving mechanisms with edge-based Generative AI. It candidly discusses the limitations encountered during the study and proposes strategic directions for future research to further enhance secure, intelligent residential automation.

## Chapter 2

# Related Work

## 2.1 MQTT-Based Smart Home Architectures

The Message Queuing Telemetry Transport (MQTT) protocol has rapidly emerged as the definitive communication standard for Internet of Things deployments. Its lightweight, publish-subscribe architecture is uniquely optimized for resource-constrained environments, making it highly effective for residential smart home systems. Numerous foundational architectures have successfully utilized micro-controllers, such as the ESP-8266 and ESP-32, paired with Raspberry Pi edge gateways to manage real-time sensor data. By establishing the Raspberry Pi as a central broker, these decentralized configurations ensure seamless data aggregation and appliance control while maintaining exceptionally low network overhead.

A primary driver for the widespread adoption of MQTT is its remarkable efficiency in bandwidth utilization and energy conservation when compared to traditional client-server protocols like HTTP. Empirical studies within smart home environments have demonstrated that MQTT can achieve substantial reductions in data transmission size, yielding proportional energy savings for low-power sensor nodes. Furthermore, the protocol's structured topic hierarchy and adaptable Quality of Service (QoS) levels provide a robust framework for building scalable, fault-tolerant automation systems that guarantee reliable message delivery under varying network conditions.

Beyond simple telemetry, MQTT serves as a critical enabler for edge-based intelligence. By integrating the protocol with local edge computing paradigms, researchers have significantly reduced cloud dependency, allowing systems to maintain high responsiveness even during internet outages. Deployments combining MQTT brokers with local data processing have achieved end-to-end latencies measured in milliseconds, which is absolutely critical for time-sensitive applications such as fire, gas leak, and intrusion detection. Additionally, the protocol's asynchronous nature allows it to easily orchestrate complex automated workflows across diverse household subsystems.

However, while these existing architectures unequivocally validate MQTT's operational efficiency and real-time streaming capabilities, they frequently expose a critical weakness in data security. In its standard implementation, MQTT primarily relies on transport-layer security and often transmits the actual sensor payloads in plaintext. This architectural oversight assumes a perfectly secure local network and leaves the central broker as a highly vulnerable single point of failure. If the broker is compromised, sensitive domestic telemetry is entirely exposed to eavesdropping and manipulation. This inherent vulnerability under-

scores an urgent need to augment standard MQTT architectures with robust, application-layer cryptographic and privacy-preserving mechanisms.

## 2.2 Privacy-Preserving IoT Systems

The exponential growth of smart home deployments has precipitated a corresponding escalation in privacy concerns. Residential IoT ecosystems are inherently intimate; the continuous telemetry generated by motion detectors, temperature sensors, and smart lighting constructs a highly granular behavioral profile of the occupants. In traditional, centralized architectures, this raw data is routinely transmitted to third-party cloud servers for storage and analysis. This paradigm fundamentally compromises user privacy, exposing sensitive domestic activities to potential data breaches, unauthorized corporate profiling, and malicious interception. Consequently, the research focus within the IoT domain has increasingly shifted toward developing robust, privacy-preserving frameworks that can secure data without crippling the operational utility of the smart home.

A significant portion of existing literature addresses IoT privacy through the lens of data minimization and structural obfuscation. One of the most prominent methodologies is Differential Privacy (DP). By systematically injecting calibrated statistical noise—such as Laplacian or Gaussian noise—into the sensor readings before transmission, DP ensures that individual data points cannot be reverse-engineered to reveal specific user activities. Several recent studies have successfully applied DP to smart meter data and healthcare IoT, demonstrating its mathematical rigor in protecting individual privacy. However, applying Differential Privacy in real-time residential automation presents a complex challenge: the privacy-utility trade-off. If the injected noise is too aggressive, the data becomes useless for time-sensitive anomaly detection, potentially masking critical events like fires or gas leaks. Conversely, if the noise is too minimal, the privacy guarantee is mathematically void.

To mitigate the reliance on vulnerable cloud infrastructures, researchers have also extensively explored edge computing as a primary privacy-preserving mechanism. By migrating data processing and anomaly detection to localized edge gateways—such as Raspberry Pi clusters or localized servers—systems can analyze telemetry directly within the home network. This decentralized approach drastically reduces the volume of sensitive data transmitted over the public internet. While edge computing fundamentally enhances data sovereignty, it does not completely eliminate internal threats. If the local network or the edge broker itself is compromised by an intruder, any unencrypted raw data stored or streaming across the local architecture remains fully exposed.

Furthermore, advanced distributed learning techniques, particularly Federated Learning (FL), have gained traction as a privacy-preserving alternative to centralized machine learn-

ing. In FL setups, IoT devices train anomaly detection models locally and only share cryptographic weight updates with a central server, ensuring raw sensor data never leaves the device. While highly secure, Federated Learning introduces immense computational and memory overheads that are often entirely impractical for resource-constrained microcontrollers like the ESP32.

Ultimately, the literature highlights a persistent fragmentation in privacy-preserving IoT systems. Existing solutions tend to isolate their approaches—focusing either on network-level decentralization, mathematically complex noise injection, or heavy distributed learning—rather than combining them into a cohesive, lightweight architecture. There remains a definitive need for hybrid frameworks that can seamlessly integrate structural data obfuscation at the sensor node with secure, decentralized edge processing, ensuring both strict data privacy and high-fidelity anomaly detection.

## 2.3 Cryptographic Mechanisms in IoT Security

Cryptography forms the fundamental backbone of data confidentiality and integrity within Internet of Things networks. In standard smart home architectures, Transport Layer Security (TLS) is predominantly employed to secure the communication channels between edge devices and central servers. While TLS effectively prevents man-in-the-middle attacks and packet sniffing during transit, its implementation in highly constrained environments presents notable challenges, including significant computational overhead and increased handshake latency. More critically, TLS only encrypts the transmission tunnel itself. Once the sensor data reaches the centralized MQTT broker or cloud gateway, it is entirely decrypted and processed in plaintext. This architectural norm creates a severe vulnerability; any successful breach of the central broker instantly exposes the entire history of sensitive domestic telemetry to the attacker.

To address the inherent limitations of tunnel-only security, recent literature has increasingly focused on application-layer cryptographic mechanisms. By encrypting the actual data payload before transmission, the information remains secure even if the intermediary broker is compromised. Symmetric encryption algorithms, such as the Advanced Encryption Standard (AES), are frequently proposed for this task due to their low computational requirements and fast execution speeds, making them highly suitable for low-power microcontrollers like the ESP32. However, symmetric cryptography fundamentally suffers from complex key distribution and management challenges in distributed networks. If a single pre-shared key is extracted from a physically accessible edge device, the confidentiality of the entire network can be invalidated, making it difficult to scale securely across dynamic environments with multiple interconnected sensors.

Consequently, asymmetric cryptographic frameworks, including Rivest–Shamir–Adleman (RSA) and Elliptic Curve Cryptography (ECC) algorithms, have gained substantial traction for IoT deployments. Asymmetric infrastructures resolve the key distribution dilemma by utilizing a public-key architecture. Edge devices can encrypt their telemetry using a widely distributed public key, ensuring that only the secure backend system holding the corresponding private key can decrypt the payload. Historically, algorithms like RSA-2048 were considered too resource-intensive for microcontrollers, often causing prohibitive delays.

However, with the advent of more powerful edge hardware and optimized cryptographic libraries, the deployment of robust asymmetric encryption on devices like the ESP32 has become a practical reality. Despite these advancements, a significant gap remains in the literature regarding the practical, real-world integration of asymmetric application-layer encryption directly within MQTT payloads. The majority of existing studies propose theoretical security models or run limited simulations, rather than demonstrating end-to-end, hardware-validated implementations that successfully balance rigorous RSA payload encryption with the low-latency requirements of real-time home automation.

## 2.4 Anomaly Detection Techniques in IoT

Anomaly detection is a foundational component of intelligent smart home monitoring, serving as the primary mechanism for identifying critical events such as gas leaks, fires, or unauthorized physical intrusions. Historically, the most ubiquitous approach to identifying these hazards has been rule-based or threshold-based logic. In these deterministic systems, an anomaly is flagged strictly when a sensor reading surpasses a predefined numerical boundary. While this method is highly computationally efficient and easily deployed on low-power microcontrollers like the ESP32, it is inherently rigid. Static thresholds fail to account for benign environmental variations, such as a localized temperature spike caused by cooking, frequently resulting in a high volume of false positives. Furthermore, threshold logic struggles to evaluate multivariate relationships, missing complex anomalies that only become apparent when multiple sensor states are analyzed simultaneously.

To overcome the inflexibility of static rules, the academic focus shifted toward integrating traditional machine learning (ML) classifiers directly into the IoT ecosystem. Algorithms including Support Vector Machines (SVM), Random Forests, and K-Nearest Neighbors have been extensively researched and deployed at the network edge. By training these models on historical environmental telemetry, systems can learn the baseline “normal” behavior of a household and identify statistical deviations. For example, edge-based SVM implementations have successfully aggregated data from temperature, smoke, and gas sensors to classify hazardous events with remarkably low latency. These decentralized ML

architectures significantly improve detection accuracy over basic thresholding while keeping processing localized to edge gateways like the Raspberry Pi.

More recently, deep learning techniques, particularly Autoencoders and recurrent neural networks like Long Short-Term Memory (LSTM) models, have been introduced to handle complex, non-linear time-series data in IoT networks. While these deep learning models offer superior accuracy in recognizing subtle anomalous patterns, they introduce immense computational and memory overheads. Deploying such heavily parameterized models on resource-constrained edge devices typically requires aggressive model quantization or forces the system to offload inference tasks back to centralized cloud servers. This reliance on the cloud immediately reintroduces the severe privacy vulnerabilities and latency issues that edge computing aims to eliminate.

Despite the varying degrees of accuracy across threshold logic, traditional ML, and deep learning, all of these existing methodologies share a critical, fundamental limitation: they function entirely as “black boxes.” When a hazard is identified, these systems output binary classifications (e.g., normal versus anomalous) or arbitrary numerical anomaly scores. They completely lack the capacity to provide contextual reasoning, root cause analysis, or actionable recommendations. Consequently, when an alarm is triggered, the end-user is presented with a generic warning rather than a clear, human-readable explanation of the specific domestic event. This absence of interpretability severely diminishes the practical utility of the system and creates a pressing need for a more intelligent, context-aware approach to anomaly interpretation.

## 2.5 Generative AI for Real-Time Monitoring

The recent and rapid advancements in Generative Artificial Intelligence (GenAI), specifically the evolution of Large Language Models (LLMs), have introduced an entirely new frontier for Internet of Things automation. Unlike traditional machine learning classifiers that strictly output binary decisions or raw numerical anomaly scores, Generative AI possesses the unique capability to comprehend complex, multivariate data streams and generate natural language outputs. In the context of smart home monitoring, this represents a profound paradigm shift from simple alerting to intelligent, context-aware reasoning. By processing aggregated sensor telemetry—such as a simultaneous spike in temperature, gas concentration, and motion—an LLM can synthesize the environmental context to deduce the likely root cause of an event. This allows the system to provide users with clear, human-readable, and highly actionable insights rather than cryptic numerical warnings.

Despite its immense potential, the integration of Generative AI within IoT ecosystems is currently dominated by cloud-centric architectures. The vast majority of contemporary

research and commercial applications rely on massive, proprietary models accessed via third-party Application Programming Interfaces (APIs). While these cloud-hosted models offer unparalleled reasoning capabilities, their deployment in residential automation is fundamentally flawed regarding user privacy and system reliability. Transmitting granular, continuous domestic telemetry to an external cloud server for AI inference directly exposes highly sensitive behavioral data to potential interception, data breaches, and corporate profiling. Furthermore, this heavy reliance on continuous internet connectivity introduces unacceptable latency for time-critical emergencies, such as fire outbreaks or gas leaks, where immediate, localized actuation is required.

To mitigate these severe privacy and latency vulnerabilities, academic focus is beginning to pivot toward Edge AI. The recent release of highly optimized, parameter-efficient LLMs has made it viable to perform complex generative inference directly on localized hardware, such as edge gateways or dedicated on-premise servers. By deploying lightweight models locally, systems can retain absolute data sovereignty; the sensitive environmental telemetry never leaves the physical boundaries of the home network. This decentralized approach guarantees that the generation of intelligent alerts and automated responses is completely insulated from external network outages and third-party data collection policies.

However, the practical application of localized Generative AI for real-time IoT anomaly detection remains significantly under-researched. While deploying lightweight models at the edge solves the immediate cloud-dependency issue, integrating these demanding models seamlessly with low-latency communication protocols like MQTT presents a major engineering challenge. Furthermore, the existing literature has yet to adequately demonstrate how a localized LLM can effectively interpret granular sensor data that has been concurrently secured by robust cryptographic protocols and structural privacy mechanisms. The absence of a cohesive framework that unifies edge-based Generative AI with rigorous, application-layer data protection constitutes a critical gap in current smart home security paradigms.

## 2.6 Limitations of Existing Approaches

A comprehensive review of the contemporary literature reveals a heavily fragmented landscape in smart home security and intelligent monitoring. While substantial progress has been made in individual domains—such as optimizing MQTT for constrained devices, applying traditional machine learning at the edge, and utilizing mathematical privacy models like Differential Privacy—there remains a distinct lack of unified, holistic frameworks. Most existing architectures force a detrimental compromise: systems that are highly secure tend to incur unacceptable latency for real-time applications, while lightweight, highly responsive systems frequently expose raw, unencrypted telemetry to central brokers or third-

party cloud infrastructure.

Furthermore, the paradigm of anomaly detection in residential IoT is currently trapped between two extremes. Traditional threshold and standard machine learning models offer edge-compatible efficiency but operate as unexplainable “black boxes,” providing users with little to no actionable context during an emergency. Conversely, advanced Generative AI models capable of contextual reasoning are predominantly cloud-bound, introducing severe data sovereignty risks and rendering the smart home dependent on continuous internet connectivity.

To synthesize these critical vulnerabilities, Table 2.1 provides a comparative summary of the prevalent methodologies in recent literature, explicitly detailing their approaches, underlying technologies, advantages, and fundamental limitations.

Table 2.1: Summary of Existing Literature and Comparative Analysis

Year	Method / Approach	Devices / Technologies	Pros	Limitations
2022	Centralized IoT with Transport-Layer Security	ESP8266, MQTT, Cloud-based Broker, TLS/SSL	Low deployment cost; highly scalable; lightweight transmission.	Payload remains in plaintext at the broker; high vulnerability to central server compromise; cloud dependency.
2023	Edge-Based Machine Learning for Anomaly Detection	Raspberry Pi, Support Vector Machines (SVM), MQTT	Eliminates cloud reliance for basic detection; faster response times.	“Black box” alerts with no human-readable context; lacks structural privacy or data payload encryption.
2024	Differential Privacy in Continuous IoT Telemetry	Smart Meters, Laplacian Noise, Cloud Aggregation	Strong mathematical guarantees against user profiling.	Aggressive noise masks critical anomalies (e.g., fires); high latency; still reliant on cloud processing.
2025	Cloud-Based LLMs for Contextual Smart Environments	Proprietary API (e.g., OpenAI), HTTP, Smart Hubs	Highly accurate anomaly interpretation; clear, human-readable explanations.	Severe privacy risks (raw data leaves the home); high latency; fails during internet network outages.
Proposed	<b>Edge GenAI with Application-Layer Privacy</b>	ESP32, Local Llama 3.2, RSA-2048, Dual Gaussian Noise	<b>Absolute data sovereignty; context-aware reasoning; encrypted broker payload.</b>	<b>Requires optimized edge hardware; slightly higher computational overhead at the sensor node.</b>

As demonstrated in the comparative analysis above, the fundamental research gap lies in the absence of a decentralized system capable of executing application-layer cryptographic

payload encryption and privacy-preserving data obfuscation, while simultaneously deploying a localized Generative AI model to provide explainable anomaly detection. Addressing these specific and interconnected limitations forms the direct motivation for the architecture proposed in the subsequent chapter.

## Chapter 3

# Proposed System

This chapter extensively details the architectural blueprint and operational mechanisms of the proposed IoTShield framework. It systematically breaks down the integration of edge hardware, the application-layer security protocols, the privacy-preserving data pipeline, and the localized deployment of Generative AI utilized for intelligent anomaly detection.

## 3.1 Overall System Overview

The IoTShield architecture is fundamentally designed as a decentralized, privacy-centric monitoring ecosystem for residential environments. The core motivation behind this design is to capture granular environmental telemetry, cryptographically secure it directly at the source, and analyze it using advanced artificial intelligence without ever transmitting sensitive data to a third-party cloud. To accomplish this, the system is structured across three cohesive operational layers: the Edge Data Acquisition Layer, the Secure Communication Broker, and the Intelligent AI Backend.

At the Edge Data Acquisition Layer, the system is driven by a physical ESP32 microcontroller, which functions as the primary hardware sensing node. This physical node is wired to a diverse array of environmental sensors capable of tracking real-time temperature, humidity, gas levels, flame presence, motion, and ambient light. Alongside this physical hardware, the architecture integrates a simulated ESP32 hub to represent multi-room deployment and a simulated Raspberry Pi to act as the central edge gateway. This hybrid physical-simulated approach rigorously tests the network's scalability. Most importantly, all telemetry generated at this layer undergoes a strict, two-step privacy transformation before broadcast. First, dual-layer Gaussian noise is injected into the raw readings to provide differential privacy, masking exact behavioral routines. Immediately after, the obfuscated payload is secured using RSA-2048 encryption, ensuring absolute data confidentiality.

The Secure Communication Broker layer handles the efficient routing of this protected data. Operating on the lightweight MQTT protocol, the encrypted sensor payloads are published to designated topics on a Mosquitto broker, which logically resides on the simulated Raspberry Pi edge gateway. A critical architectural advantage of IoTShield is that the MQTT broker functions strictly as a blind relay. Because the telemetry is encrypted at the application layer by the edge nodes, the broker only processes unreadable ciphertext. Consequently, even if an intruder compromises the local network or the broker itself, the underlying domestic data remains mathematically protected and completely inaccessible.

Finally, the Intelligent AI Backend processes the encrypted stream to deliver actionable insights. Built upon a robust Django framework, this backend securely decrypts the incoming MQTT payloads using the system's localized private RSA key. Once decrypted, the raw telemetry is evaluated against dynamic, multi-sensor threshold logic to instantly categorize the environmental severity (e.g., Low, Medium, High, or Critical). When a high-severity anomaly is identified, the backend routes the complex sensor data matrix directly into a locally hosted, 1-billion parameter Large Language Model (Llama 3.2 via Ollama). The LLM autonomously analyzes the event context and generates a precise, human-readable alert explaining the probable cause of the hazard. All telemetry and AI-generated alerts are subsequently archived in a SQLite relational database and rendered on an interactive web dashboard for seamless user management.

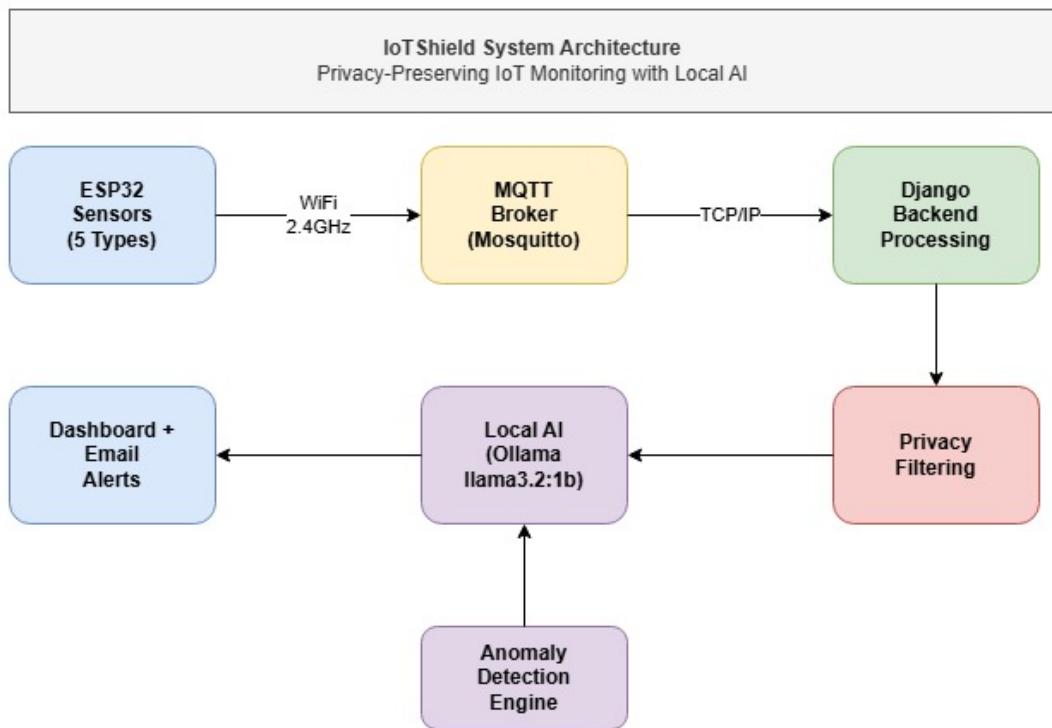


Figure 3.1: High-level architectural overview of the proposed IoTShield system, illustrating the secure data flow from the edge sensor nodes to the localized Generative AI analysis framework.

## 3.2 System Architecture Design

The architectural design of the IoTShield framework is structured around a localized, decentralized network topology operating entirely within a private Local Area Network (LAN). To guarantee data sovereignty and eliminate reliance on external cloud processing, all core communications, data routing, and artificial intelligence inferences are confined within the 192.168.1.0/24 subnet. The network is orchestrated by a standard WiFi Router acting as the default gateway (192.168.1.1), which manages DHCP assignments, Network Address Translation (NAT), and internal firewall policies to insulate the smart home ecosystem from the public Wide Area Network (WAN).

Operating at the network edge is the ESP32 Smart Sensor Hub, which connects to the local router via a 2.4GHz WiFi interface. Assigned to a dedicated local IP address (192.168.1.150), this edge node is responsible for continuously polling the connected hardware—specifically the DHT11, MQ-2, Flame, LDR, and PIR sensors. Once the telemetry is acquired, the ESP32 processes the data through the local privacy pipeline, packages it into a secure JSON payload, and transmits it via the MQTT protocol to the central server.

The core intelligence and data routing of the system are hosted on the Central Server, acting as the edge gateway. To ensure maximum stability and minimize latency, this server is connected to the router via a hardwired Ethernet connection and assigned a static IP address (192.168.1.100). The server concurrently hosts three critical, containerized services: the Mosquitto MQTT Broker on Port 1883, the Django Web Server on Port 8000, and the Ollama AI Inference engine on Port 11434, which hosts the localized Llama 3.2 model.

Finally, the User Interface tier allows end-users to access the monitoring dashboard and manage alerts using any standard web browser on a mobile device or laptop. By navigating to the internal address (`http://192.168.1.100:8000`), users can interact with the system securely over the local WiFi connection without routing traffic over the public internet.

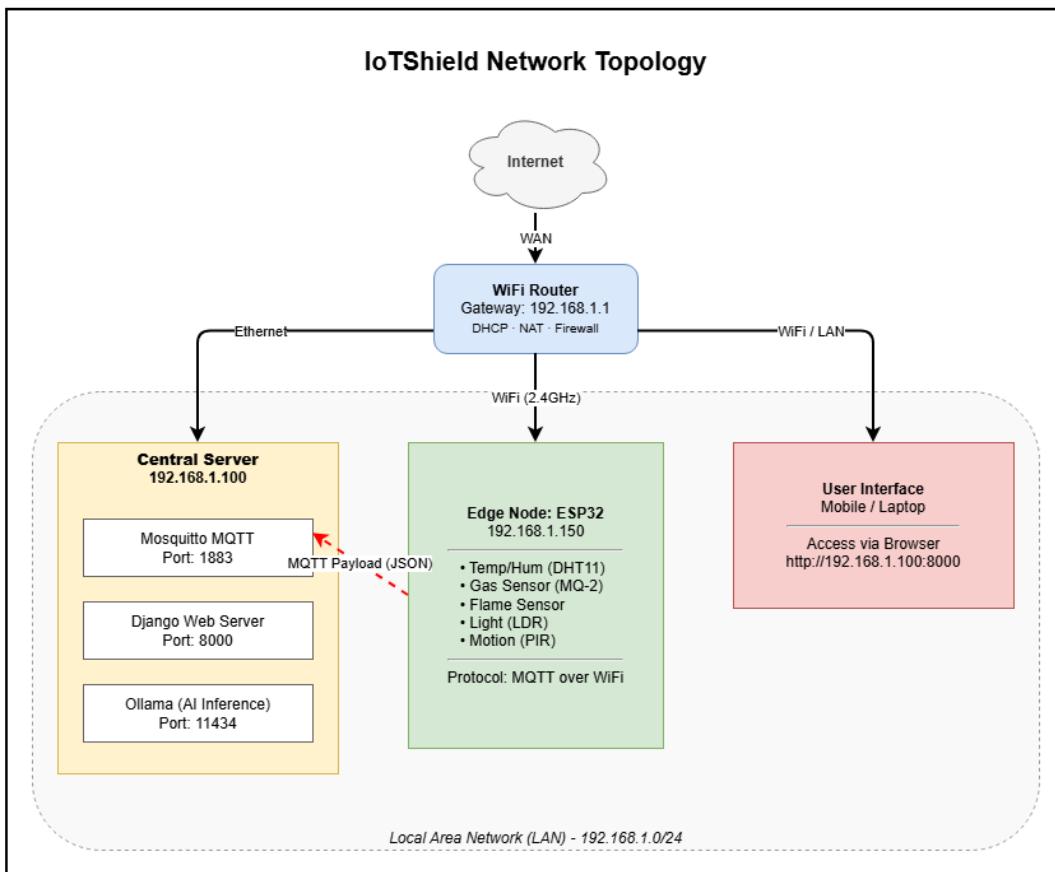


Figure 3.2: Detailed network topology of the IoTShield framework, illustrating the local IP subnet structure, port assignments, and the strict isolation of edge processing from the public internet.

### 3.3 Hardware Architecture (ESP32 and Sensors)

The physical foundation of the IoTShield framework is built upon a highly responsive, resource-optimized hardware architecture designed specifically for edge environments. At the core of the Edge Sensing Tier is the ESP32 microcontroller, specifically the ESP32-WROOM-32 module. The ESP32 was deliberately selected over lower-tier microcontrollers due to its dual-core 240 MHz Tensilica processor, built-in 2.4 GHz Wi-Fi capabilities, and sufficient SRAM to handle demanding application-layer tasks. Unlike standard IoT edge nodes that merely forward raw electrical signals, the ESP32 in this architecture is tasked with actively polling six distinct sensors, applying dual-layer Gaussian noise matrices, executing RSA-2048 cryptographic encryption, and managing asynchronous MQTT publishing—all in real-time.

To capture a comprehensive profile of the domestic environment, the ESP32 is physically interfaced with a diverse array of five distinct sensor modules, collectively gathering six unique environmental metrics. Ambient climate monitoring is handled by a DHT11 digital sensor, which provides continuous temperature and relative humidity readings. Hazardous environmental anomalies are monitored by an MQ-2 analog gas sensor, calibrated to detect combustible gases and smoke concentrations, working in tandem with an infrared (IR) Flame sensor that triggers upon detecting specific light wavelengths associated with open fires. Finally, spatial and security metrics are captured using an HC-SR501 Passive Infrared (PIR) motion sensor to detect physical occupancy, alongside a Light Dependent Resistor (LDR) to measure ambient room illumination.

By aggregating these diverse metrics onto a single ESP32 node, the hardware architecture provides the localized AI model with a rich, multivariate dataset. This allows the system to cross-reference multiple data points (e.g., detecting a simultaneous spike in temperature, gas, and a positive flame reading) to accurately diagnose complex hazards. Table 3.1 provides a detailed technical specification of the hardware components utilized in the physical edge node deployment.

Table 3.1: Hardware Components and Sensor Specifications

<b>Component</b>	<b>Type</b>	<b>Operating Range / Details</b>	<b>Role in System</b>
ESP32-WROOM-32	Microcontroller	240 MHz Dual-Core, 3.3V Logic, 802.11 b/g/n Wi-Fi	Primary edge gateway; executes RSA encryption, differential privacy, and MQTT routing.
DHT11	Digital Climate Sensor	Temp: 0–50°C, Humidity: 20–90% RH	Monitors ambient room temperature and relative humidity levels.
MQ-2	Analog Gas Sensor	300–10,000 ppm (LPG, Methane, Smoke)	Identifies hazardous gas leaks and potential pre-combustion smoke.
IR Flame Sensor	Infrared Receiver	Wavelengths: 760 nm to 1100 nm	Provides immediate binary detection of open flames or fire outbreaks.
HC-SR501 (PIR)	Digital Motion Sensor	Range: 7m, Angle: 120°	Detects human occupancy or unauthorized physical intrusion.
LDR (Photoresistor)	Analog Light Sensor	1 MΩ (Dark) to 10 kΩ (Light)	Monitors ambient illumination to provide contextual environmental data.

## 3.4 Software Architecture

The software architecture of the IoTShield system is engineered as a modular, high-concurrency framework designed to facilitate real-time environmental monitoring while maintaining strict local data sovereignty. The architecture follows a decoupled, service-oriented paradigm where each component—ranging from edge firmware to the backend inference engine—operates independently yet communicates through standardized interfaces. This design choice ensures that the system can handle high-frequency telemetry from multiple nodes without blocking critical security operations or AI reasoning tasks.

At the network’s periphery, the Firmware Layer is developed using the Arduino framework in C++, specifically optimized for the dual-core architecture of the ESP32. This layer manages the low-level hardware abstraction for the connected sensor array and executes the initial privacy-preserving pipeline. By utilizing the second core of the ESP32 for cryptographic operations, the firmware ensures that the RSA-2048 encryption and Gaussian noise injection do not interfere with the primary sensor polling and MQTT publishing loops, thus maintaining a consistent data acquisition rate.

The Message Broker Layer serves as the foundational communication backbone, utilizing the Eclipse Mosquitto implementation of the MQTT protocol. This layer is configured to handle asynchronous messaging with a Quality of Service (QoS) Level 1, ensuring that critical telemetry packets are acknowledged by the backend. The broker functions as a stateless relay, where the encrypted sensor payloads are routed through a hierarchical topic structure. This decoupling allows the backend processing services to subscribe to specific data streams without maintaining a persistent direct connection to each individual hardware node.

The Backend and Application Layer is constructed using the Django web framework, which orchestrates the core logic of the IoTShield system. This layer manages the secure decryption of payloads, the execution of multivariate threshold logic for immediate anomaly detection, and the persistence of telemetry within a relational SQLite database. Furthermore, the backend integrates with the Ollama inference engine via a local API to facilitate the generation of context-aware alerts. By hosting the Llama 3.2 model locally, the software architecture ensures that the transition from raw environmental data to human-readable insights occurs entirely within the edge environment, fulfilling the primary research objective of localized and privacy-preserving monitoring.

### 3.5 MQTT Communication Model

The communication backbone of the IoTShield system is built upon the Message Queuing Telemetry Transport (MQTT) protocol, specifically selected for its lightweight overhead and suitability for resource-constrained edge devices like the ESP32. The system utilizes an asynchronous publish-subscribe paradigm, where the Mosquitto broker serves as a central hub for routing secure telemetry between the edge sensing nodes and the Django backend. To ensure reliability in a residential network environment, the system is configured with a Quality of Service (QoS) Level 1, guaranteeing that critical sensor updates and hazard alerts are delivered to the subscriber at least once, supported by a 60-second keep-alive interval to monitor node connectivity.

The communication architecture is governed by a strict hierarchical topic structure, which allows for efficient data filtering and multi-device management. The ESP32 hardware nodes publish encrypted environmental data to specific sub-topics, while the backend system subscribes to these streams to trigger the decryption and anomaly detection pipelines. Conversely, the backend can publish commands to the control topic, allowing for remote device management. This decoupled approach ensures that the sensing nodes do not need to maintain a direct awareness of the backend's state, significantly reducing the computational burden on the ESP32. Table 3.2 details the topic hierarchy and the corresponding payload characteristics utilized in the IoTShield framework.

Table 3.2: MQTT Topic Structure and Payload Formats

Topic Name	Direction	QoS	Payload Description
iotshield/sensors/data	Publish	1	RSA-encrypted JSON containing multi-sensor telemetry (Temp, Hum, Gas, etc.).
iotshield/alerts	Publish	1	Encrypted alerts containing severity levels and AI-generated suggestions.
iotshield/control/commands	Subscribe	1	Incoming instructions for device configuration or manual sensor polling.
iotshield/logs	Publish	0	Plaintext diagnostic messages for system health and connectivity monitoring.

Figure 3.3 illustrates the end-to-end communication flow, highlighting the interaction be-

tween the Edge Device (PubSubClient), the Mosquitto Broker, and the Backend System (paho-mqtt), while detailing the JSON data format utilized for secure message exchange.

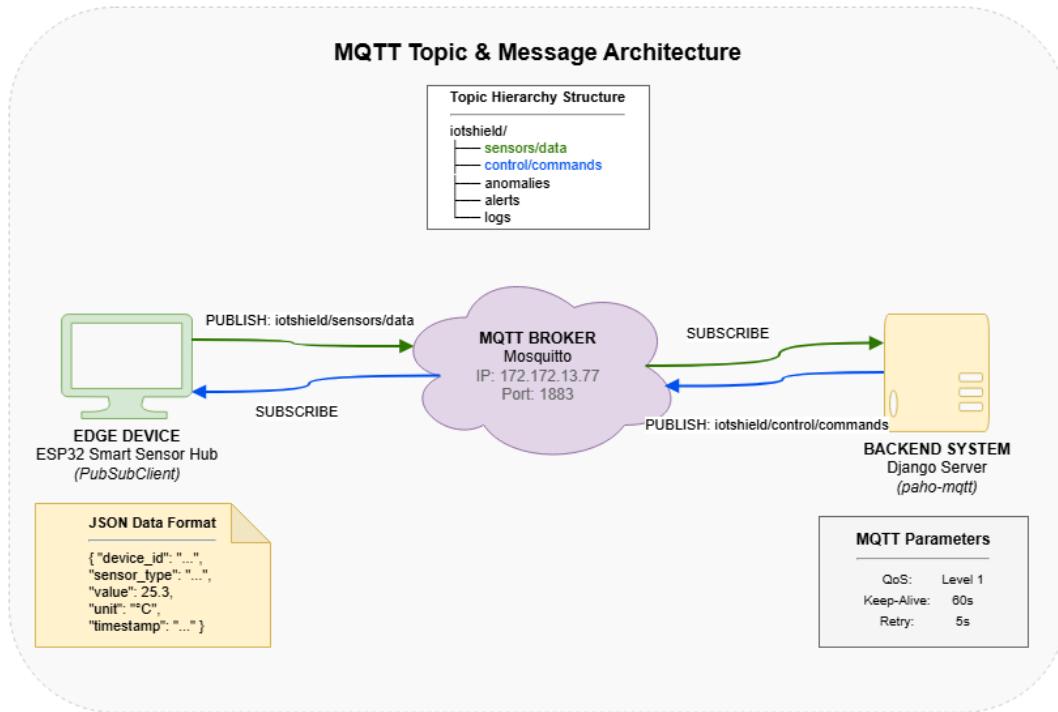


Figure 3.3: MQTT communication architecture illustrating the hierarchical topic structure, message flow between edge nodes and the backend, and key protocol parameters.

## 3.6 Privacy-Preserving Data Pipeline

A core contribution of the IoTShield framework is its dual-layer privacy-preserving data pipeline, designed to protect domestic telemetry against both external interception and internal behavioral profiling. Unlike traditional systems that transmit raw data, this architecture implements a localized security sequence on the ESP32 node prior to network transmission.

### 3.6.1 Differential Privacy with Dual Gaussian Noise

The first stage of the pipeline addresses the risk of behavioral pattern mining. Even when encrypted, high-frequency sensor values can be analyzed to reveal sensitive occupant routines. To mitigate this, the system injects a calibrated level of Gaussian noise into the raw sensor telemetry. Utilizing a privacy budget of  $\epsilon = 0.5$ , the ESP32 adds a noise displacement that masks the exact environmental value while preserving the statistical utility required for anomaly detection. This ensures that the published telemetry represents a "noisy" version of the environment (e.g.,  $25.523^{\circ}\text{C}$  instead of  $25.0^{\circ}\text{C}$ ), providing a strong mathematical guaranty of differential privacy.

### 3.6.2 RSA-2048 Encryption Mechanism

Following structural obfuscation, the data enters the cryptographic stage to ensure end-to-end confidentiality. The noisy payload is encrypted using the RSA-2048 algorithm, employing Optimal Asymmetric Encryption Padding (OAEP) and the SHA-256 hashing function. By utilizing an asymmetric public-key infrastructure, the "IoTShield" system secures the payload such that only the backend server, possessing the corresponding private key, can access the plaintext contents. This transformation renders the telemetry into unreadable ciphertext before it is packaged into a JSON format for MQTT publication. Consequently, the data remains mathematically protected while transiting the broker and during storage in the SQLite database.

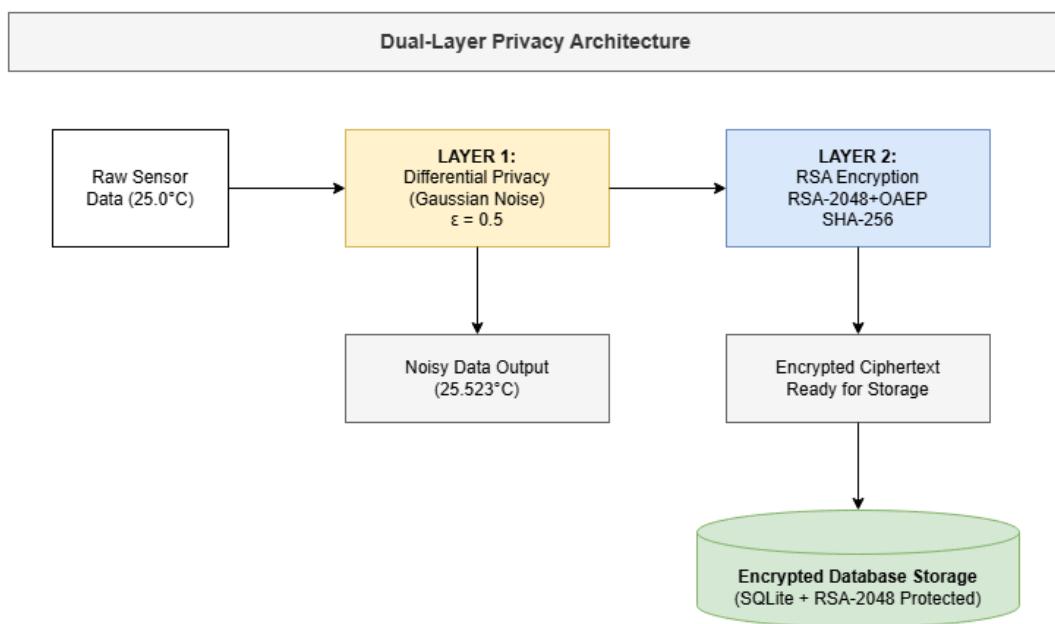


Figure 3.4: Dual-layer privacy architecture of the IoTShield system, illustrating the sequential transformation of raw sensor data through Gaussian noise injection and RSA-2048 encryption.

## 3.7 AI-Based Anomaly Detection Framework

The final tier of the IoTShield system is the intelligent inference framework, which transforms raw environmental telemetry into actionable, human-readable insights. This framework is designed to operate entirely at the edge, utilizing a two-stage detection pipeline: a high-speed threshold-based screening layer and a context-aware Generative AI reasoning layer. By hosting the inference engine locally, the system ensures that sensitive domestic data is never exposed to third-party cloud providers, fulfilling the primary privacy objectives of the research.

### 3.7.1 Threshold-Based Detection Logic

Before the data is passed to the AI model, it undergoes an initial evaluation via a multivariate threshold-based detection engine. This layer serves as a computational filter to differentiate between routine environmental fluctuations and potential hazards. Each incoming sensor reading is compared against a set of predefined boundaries that categorize the severity of the event. While a single sensor breach (e.g., a "Medium" gas reading) triggers an alert, the system is designed to recognize compound anomalies, such as simultaneous spikes in temperature and gas, which escalate the priority level. Table 3.3 defines the specific threshold parameters and the corresponding severity classifications utilized by the IoTShield backend.

Table 3.3: Sensor Threshold Parameters and Severity Classification

Sensor Type	Normal Range	LOW	MEDIUM	HIGH/CRITICAL
Temperature (°C)	18.0 – 28.0	28.1 – 35.0	35.1 – 45.0	> 45.0
Humidity (%)	30.0 – 60.0	60.1 – 75.0	75.1 – 85.0	> 85.0
Gas (MQ-2)	< 0.20	0.21 – 0.40	0.41 – 0.60	> 0.60
Flame (IR)	0 (Stable)	–	–	1 (Flame Detected)
Motion (PIR)	0 (No Motion)	1 (Detected)	–	–
Light (LDR)	200 – 800 lux	< 200	–	> 800

### 3.7.2 Local LLM Integration (Llama 3.2 via Ollama)

When the threshold engine identifies a "Medium," "High," or "Critical" anomaly, the multivariate data packet is immediately routed to the localized Generative AI inference engine.

This layer utilizes the Llama 3.2 model (1-billion parameter version) hosted on-premise via the Ollama platform. Unlike traditional "black-box" classifiers, the LLM performs contextual reasoning; it receives the decrypted sensor values and provides a natural language synthesis of the probable hazard. For instance, rather than simply reporting a high gas value, the model interprets the data to suggest whether the readings indicate a persistent gas leak or a sudden fire outbreak.

This integration allows the system to generate "AI Suggestions," which provide users with actionable steps, such as "Immediately check the kitchen motion sensor and investigate the cause." Figure 3.4 illustrates the routing logic, showing how normal data flows directly to the database while anomalous data is intercepted and analyzed by the local LLM prior to alert generation.

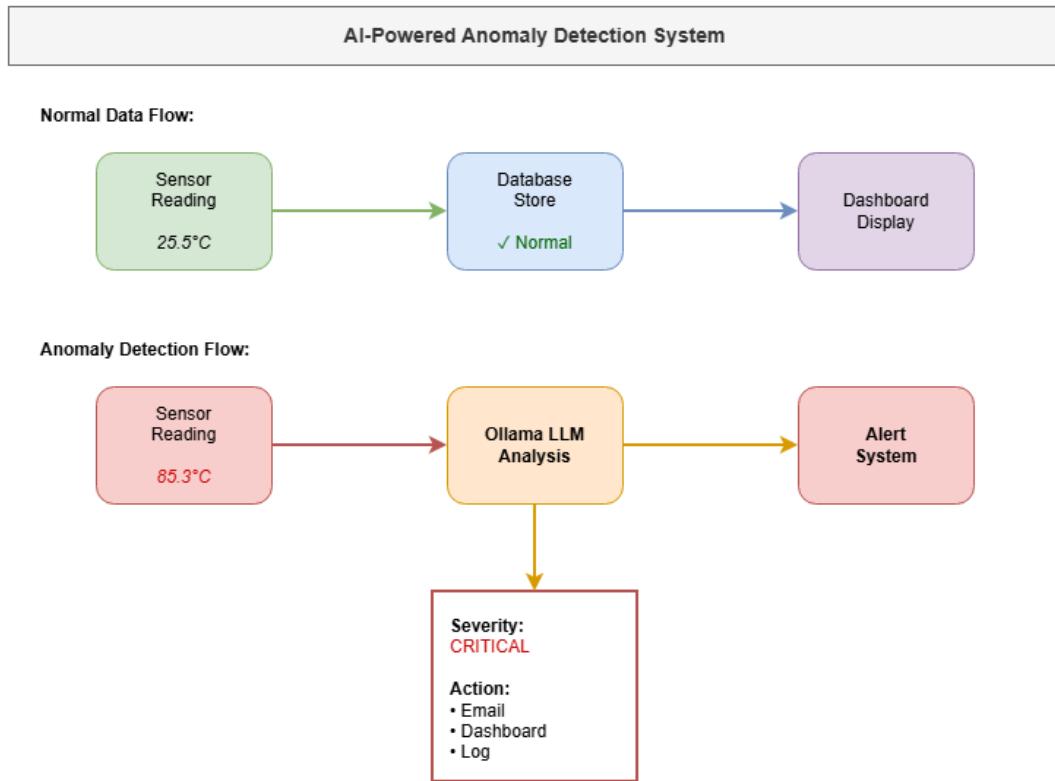


Figure 3.5: AI-powered anomaly detection workflow illustrating the data routing logic and the localized integration of the Llama 3.2 LLM via Ollama.

## 3.8 End-to-End System Workflow

The operational integrity of the IoTShield system is defined by a seamless, end-to-end data pipeline that ensures every environmental event is captured, secured, and analyzed in real-time. The workflow initiates at the physical edge, where the ESP32 microcontroller performs high-frequency polling of the connected sensor array. Upon acquisition, the raw telemetry is immediately processed through the on-device privacy engine. This involves injecting dual-layer Gaussian noise to provide structural privacy, followed by RSA-2048 encryption to secure the payload. This sequence ensures that even before the data enters the network, it is mathematically protected against both behavioral profiling and unauthorized interception.

Once the payload is secured, the ESP32 functions as an MQTT client, publishing the encrypted JSON packet to the designated sensors topic on the Mosquitto broker. The broker, operating as a localized gateway, routes the message to the Django-based backend system. Upon receiving the message, the backend utilizes the localized private RSA key to decrypt the payload. The resulting plaintext is then persistently stored in the SQLite database and simultaneously evaluated against the predefined multi-sensor threshold logic. This stage is critical for maintaining an immutable history of environmental trends while identifying potential safety breaches within milliseconds.

In the event that the threshold engine identifies a medium or high-severity anomaly, the system triggers the intelligent inference stage. The multivariate sensor data is formatted into a context-rich prompt and transmitted to the localized Llama 3.2 model via the Ollama API. The model performs contextual reasoning to generate human-readable diagnoses and actionable safety suggestions. These insights are pushed to the user's web dashboard through a real-time WebSocket connection and transmitted as an urgent notification via the SMTP relay. This comprehensive workflow ensures that the end-user receives high-context security alerts while maintaining absolute data sovereignty within the localized smart home environment.

## **Chapter 4**

# **Implementation**

- 4.1 Hardware Implementation Details**
- 4.2 Software Development Environment**
- 4.3 MQTT Broker Configuration**
- 4.4 Differential Privacy Implementation**
- 4.5 RSA Encryption Implementation**
- 4.6 Local LLM Deployment Setup**
- 4.7 Database Design and Storage Model**

## **Chapter 5**

# **Results and Discussion**

### **5.1 Experimental Setup**

### **5.2 Performance Evaluation**

#### **5.2.1 System Latency Analysis**

#### **5.2.2 Encryption Overhead Analysis**

### **5.3 Anomaly Detection Results**

### **5.4 AI Response Evaluation**

### **5.5 Privacy–Utility Trade-off Analysis**

### **5.6 Security Analysis**

### **5.7 Comparative Discussion**

## **Chapter 6**

# **Conclusion and Future Work**

**6.1 Summary of the Work**

**6.2 Key Findings**

**6.3 Limitations of the Study**

**6.4 Future Research Directions**

# Bibliography

- [1] M. M. Hasan, A. Rahman, and S. Ahmed, “Internet of Things-based Home Automation with Network Mapper and MQTT Protocol,” *Computers & Education: Artificial Intelligence*, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0045790624007341>
- [2] A. Author, B. Author, and C. Author, “Generative AI Techniques for Anomaly Detection in IoT Devices,” 2024. [Online]. Available: [https://www.researchgate.net/publication/395801862\\_GENERATIVE\\_AI\\_TECHNIQUES\\_FOR\\_ANOMALY\\_DETECTION\\_IN\\_IOT\\_DEVICES](https://www.researchgate.net/publication/395801862_GENERATIVE_AI_TECHNIQUES_FOR_ANOMALY_DETECTION_IN_IOT_DEVICES)
- [3] B. Author, D. Author, and E. Author, “Anomaly Detection in IoT Using Generative AI Models,” *SPIE Conference Proceedings*, vol. 13473, 2024. [Online]. Available: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13473/134730J/AD-GAM--anomaly-detection-in-IoT-using-generative-AI/10.1117/12.3058632.full>
- [4] C. Author, F. Author, and G. Author, “Review: Generative Adversarial Networks-Enabled Anomaly Detection in IoT,” *Expert Systems with Applications*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417425025953>
- [5] D. Author, H. Author, and I. Author, “AI-Driven Anomaly Detection for Securing IoT Devices in Smart Cities,” *Electronics*, vol. 14, no. 12, p. 2492, 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/14/12/2492>
- [6] E. Author, J. Author, and K. Author, “Privacy-Preserving Security of IoT Networks: A Comparative Study,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2772918425000013>
- [7] F. Author, L. Author, and M. Author, “Leveraging IoT, Cloud, and Edge Computing with AI,” *Sensors*, vol. 25, no. 6, p. 1763, 2025. [Online]. Available: <https://www.mdpi.com/1424-8220/25/6/1763>
- [8] G. Author, N. Author, and O. Author, “Towards Smart Home Automation Using IoT-Enabled Edge-Computing Paradigm,” 2021. [Online]. Available: [https://www.researchgate.net/publication/353372471\\_Towards\\_Smart\\_Home\\_Automation\\_Using\\_IoT-Enabled\\_Edge-Computing\\_Paradigm](https://www.researchgate.net/publication/353372471_Towards_Smart_Home_Automation_Using_IoT-Enabled_Edge-Computing_Paradigm)

- [9] H. Author, P. Author, and Q. Author, “Improving Smart Home Security via MQTT: Maximizing Data Confidentiality and Energy Efficiency,” *CSSE*, vol. 48, no. 6, p. 58697, 2025. [Online]. Available: <https://www.techscience.com/csse/v48n6/58697>
- [10] I. Author, R. Author, and S. Author, “Enhancing MQTT Security on the Internet of Things with an Enhanced Symmetric Algorithm,” 2025. [Online]. Available: [https://www.researchgate.net/publication/379866006\\_Enhancing\\_MQTT\\_Security\\_in\\_the\\_Internet\\_of\\_Things\\_with\\_an\\_Enhanced\\_Symmetric\\_Algorithm](https://www.researchgate.net/publication/379866006_Enhancing_MQTT_Security_in_the_Internet_of_Things_with_an_Enhanced_Symmetric_Algorithm)
- [11] J. Author, K. Author, and L. Author, “Hybrid Approaches to Predictive Maintenance: Combining Generative AI with IoT Sensor Data for Enhanced Failure Prediction,” 2024. [Online]. Available: [https://www.researchgate.net/publication/395791187\\_Hybrid\\_Approaches\\_to\\_Predictive\\_Maintenance\\_Combining\\_Generative\\_AI\\_with\\_IoT\\_Sensor\\_Data\\_for\\_Enhanced\\_Failure\\_Prediction](https://www.researchgate.net/publication/395791187_Hybrid_Approaches_to_Predictive_Maintenance_Combining_Generative_AI_with_IoT_Sensor_Data_for_Enhanced_Failure_Prediction)
- [12] K. Author, M. Author, and N. Author, “Generative AI for Internet of Things Security: Challenges and Opportunities,” 2025. [Online]. Available: <https://arxiv.org/abs/2502.08886>
- [13] L. Author, O. Author, and P. Author, “The Role of Smart Homes in Providing Care for Older Adults,” *MDPI*, vol. 7, no. 4, p. 62, 2025. [Online]. Available: <https://www.mdpi.com/2624-6511/7/4/62>
- [14] M. Author, Q. Author, and R. Author, “Internet of Robotic Things for Independent Living Support,” *ScienceDirect*, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2542660524000623>
- [15] N. Author, S. Author, and T. Author, “Design and Implementation of Smart Home System Based on IoT,” *ScienceDirect*, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2590123024016621>
- [16] O. Author, U. Author, and V. Author, “Design of an Innovative Solution to Integrate and Automate Smart Homes via Multiple Chatbots,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2542660525002070>
- [17] P. Author, W. Author, and X. Author, “Adoption of Internet of Things in Residential Smart Homes,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666188825002333>

- [18] Q. Author, Y. Author, and Z. Author, “MQTT\_UAD: MQTT Under Attack Dataset — A Public Benchmark for MQTT Security Research,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352340925008881>
- [19] R. Author, A1. Author, and B1. Author, “Edge AI Enabled IoT Framework for Secure Smart Home Infrastructure,” *ScienceDirect*, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050924009979>
- [20] S. Author, C1. Author, and D1. Author, “Machine Learning and IoT in Healthcare,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2543106425000201>
- [21] T. Author, E1. Author, and F1. Author, “Toward Generating a Large-Scale IoT-Z-Wave Intrusion Detection Dataset: Smart Device Profiling, Intruder Behavior, and Traffic Characterization,” *ScienceDirect*, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2542660525002616>
- [22] U. Author, G1. Author, and H1. Author, “Smart Home System: A Comprehensive Review,” *Wiley Interdisciplinary Reviews*, 2023. [Online]. Available: <https://onlinelibrary.wiley.com/doi/epdf/10.1155/2023/7616683>
- [23] V. Author, I1. Author, and J1. Author, “Artificial Intelligence in Smart Cities — Applications, Barriers, and Future Directions: A Review,” *MDPI*, vol. 7, no. 3, p. 57, 2025. [Online]. Available: <https://www.mdpi.com/2624-6511/7/3/57>
- [24] W. Author, K1. Author, and L1. Author, “IoT — A Promising Solution to Energy Management in Smart Buildings: A Systematic Review, Applications, Barriers, and Future Scope,” *MDPI*, vol. 14, no. 11, p. 3446, 2025. [Online]. Available: <https://www.mdpi.com/2075-5309/14/11/3446>
- [25] X. Author, M1. Author, and N1. Author, “Review of Smart-Home Security Using the Internet of Things,” *MDPI Electronics*, vol. 13, no. 16, p. 3343, 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/13/16/3343>
- [26] Y. Author, O1. Author, and P1. Author, “Smart Home Advancements for Health Care and Beyond: Systematic Review of Two Decades of User-Centric Innovation,” *JMIR*, 2025. [Online]. Available: <https://www.jmir.org/2025/1/e62793>
- [27] Z. Author, Q1. Author, and R1. Author, “Machine Learning in Smart Buildings: A Review of Methods, Challenges, and Future Trends,” *MDPI Applied Sciences*,

- vol. 15, no. 14, p. 7682, 2025. [Online]. Available: <https://www.mdpi.com/2076-3417/15/14/7682>
- [28] A1. Author, B1. Author, and C1. Author, “Internet of Things: a comprehensive overview, architectures, applications, simulation tools, challenges and future directions,” *Springer*, 2024. [Online]. Available: <https://link.springer.com/article/10.1007/s43926-024-00084-3>
- [29] B1. Author, D1. Author, and E1. Author, “IoT-Based Security and Privacy Implementation in Smart Home,” *IJFANS*, 2022. [Online]. Available: <https://www.ijfans.org/issue?volume=Volume%2011&issue=Issue%205&year=2022>
- [30] C1. Author, F1. Author, and G1. Author, “Smart Home Remote Control System Prototype Using Internet of Things (IoT) Based ESP8266 Microcontroller,” 2023. [Online]. Available: [https://www.researchgate.net/publication/376060810\\_Smart\\_Home\\_Remote\\_Control\\_System\\_Prototype\\_Using\\_Internet\\_of\\_Things\\_IoT\\_Based\\_ESP8266\\_Microcontroller](https://www.researchgate.net/publication/376060810_Smart_Home_Remote_Control_System_Prototype_Using_Internet_of_Things_IoT_Based_ESP8266_Microcontroller)
- [31] D1. Author, H1. Author, and I1. Author, “IoT Enabled Smart Homes in Tropical Regions as a Means of Sustainable Development,” *IOP Conference Series*, 2023. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1755-1315/1306/1/012035/pdf>
- [32] E1. Author, J1. Author, and K1. Author, “A Comprehensive Survey on Generative AI Solutions in IoT Security,” *MDPI Electronics*, vol. 13, no. 24, p. 4965, 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/13/24/4965>
- [33] F1. Author, L1. Author, and M1. Author, “IoT Based Smart Door Unlock and Intruder Alert System,” 2021. [Online]. Available: [https://www.researchgate.net/publication/356206104\\_IoT\\_Based\\_Smart\\_Door\\_Unlock\\_and\\_Intruder\\_Alert\\_System](https://www.researchgate.net/publication/356206104_IoT_Based_Smart_Door_Unlock_and_Intruder_Alert_System)
- [34] G1. Author, N1. Author, and O1. Author, “An Internet of Things-driven Smart Key System with Real-Time Alerts: Innovations in Hotel Security,” 2025. [Online]. Available: [https://www.researchgate.net/publication/389470099\\_An\\_internet\\_of\\_things-driven\\_smart\\_key\\_system\\_with\\_real-time\\_alerts\\_innovations\\_in\\_hotel\\_security](https://www.researchgate.net/publication/389470099_An_internet_of_things-driven_smart_key_system_with_real-time_alerts_innovations_in_hotel_security)
- [35] H1. Author, P1. Author, and Q1. Author, “Design of ESP8266 Smart Home Using MQTT and Node-RED,” 2021. [Online]. Available: [https://www.researchgate.net/publication/389470099\\_An\\_internet\\_of\\_things-driven\\_smart\\_key\\_system\\_with\\_real-time\\_alerts\\_innovations\\_in\\_hotel\\_security](https://www.researchgate.net/publication/389470099_An_internet_of_things-driven_smart_key_system_with_real-time_alerts_innovations_in_hotel_security)

//www.researchgate.net/publication/350935668\_Design\_of\_ESP8266\_Smart\_Home\_Using\_MQTT\_and\_Node-RED

- [36] [Author(s)], “Smart Home Privacy: A Scoping Review,” 2024. [Online]. Available: <https://www.scitepress.org/Papers/2024/122559/122559.pdf>
- [37] [Author(s)], “IoT Based Security and Privacy Implementation in Smart Home,” 2024. [Online]. Available: [https://www.researchgate.net/publication/378739063\\_IoT\\_based\\_security\\_and\\_privacy\\_implementation\\_in\\_smart\\_home](https://www.researchgate.net/publication/378739063_IoT_based_security_and_privacy_implementation_in_smart_home)
- [38] [Author(s)], “Generative AI for Internet of Things Security: Challenges and Opportunities,” 2025. [Online]. Available: <https://arxiv.org/abs/2502.08886>
- [39] [Author(s)], “AD-GAM: Anomaly Detection in IoT Using Generative AI Models,” 2025. [Online]. Available: [https://www.researchgate.net/publication/392192530\\_AD-GAM\\_anomaly\\_detection\\_in\\_IoT\\_using\\_generative\\_AI\\_models](https://www.researchgate.net/publication/392192530_AD-GAM_anomaly_detection_in_IoT_using_generative_AI_models)
- [40] [Author(s)], “Generative AI Techniques for Anomaly Detection in IoT Devices,” 2025. [Online]. Available: [https://www.researchgate.net/publication/395801862\\_GENERATIVE\\_AI\\_TECHNIQUES\\_FOR\\_ANOMALY\\_DETECTION\\_IN\\_IOT\\_DEVICES](https://www.researchgate.net/publication/395801862_GENERATIVE_AI_TECHNIQUES_FOR_ANOMALY_DETECTION_IN_IOT_DEVICES)
- [41] [Author(s)], “LLM-Enhanced Security Framework for IoT Network: Anomaly Detection,” 2025. [Online]. Available: <https://ieeexplore.ieee.org/iel8/6287639/10820123/11175688.pdf>
- [42] [Author(s)], “An LLM-Powered AI Agent Framework for Holistic IoT Traffic Interpretation,” 2025. [Online]. Available: <https://arxiv.org/abs/2510.13925>
- [43] [Author(s)], “Efficient Real-Time Anomaly Detection in IoT Networks Using One-Class Autoencoder,” 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/14/1/104>
- [44] [Author(s)], “Large Language Models in the IoT Ecosystem: A Survey on Security Challenges,” 2025. [Online]. Available: <https://arxiv.org/abs/2505.17586>
- [45] [Author(s)], “Secure Enhancement for MQTT Protocol Using Distributed Machine Learning,” 2024. [Online]. Available: <https://www.mdpi.com/1424-8220/24/5/1638>

- [46] [Author(s)], “Federated Learning for IoT: A Survey of Techniques, Challenges, and Applications,” 2024. [Online]. Available: <https://www.mdpi.com/2224-2708/14/1/9>
- [47] [Author(s)], “Anomaly Detection in IoT Networks Using Federated Machine Learning Approaches,” 2025. [Online]. Available: [https://www.researchgate.net/publication/392893320\\_Anomaly\\_Detection\\_in\\_IoT\\_Networks\\_Using\\_Federated\\_Machine\\_Learning\\_Approaches](https://www.researchgate.net/publication/392893320_Anomaly_Detection_in_IoT_Networks_Using_Federated_Machine_Learning_Approaches)
- [48] [Author(s)], “Privacy Preserving Federated Anomaly Detection in IoT Edge Computing,” 2025. [Online]. Available: <https://www.techscience.com/cmc/v84n2/62939>
- [49] [Author(s)], “DLKS-MQTT: A Lightweight Key Sharing Protocol for Secure IoT Communications,” 2025. [Online]. Available: <https://etasr.com/index.php/ETASR/article/view/10216>
- [50] [Author(s)], “Integrating Blockchain, MQTT, and Machine Learning for Enhanced IoT Applications,” 2025. [Online]. Available: [https://www.researchgate.net/publication/396031010\\_Integrating\\_Blockchain\\_MQTT\\_and\\_Machine\\_Learning\\_for\\_Enhanced\\_IoT\\_Applications\\_A\\_Comprehensive\\_Survey](https://www.researchgate.net/publication/396031010_Integrating_Blockchain_MQTT_and_Machine_Learning_for_Enhanced_IoT_Applications_A_Comprehensive_Survey)
- [51] [Author(s)], “MQTT in Action: Building Reliable and Scalable Home Automation Systems,” 2024. [Online]. Available: [https://www.researchgate.net/publication/387574531\\_MQTT\\_in\\_Action\\_Building\\_Reliable\\_and\\_Scalable\\_Home\\_Automation\\_Systems](https://www.researchgate.net/publication/387574531_MQTT_in_Action_Building_Reliable_and_Scalable_Home_Automation_Systems)
- [52] [Author(s)], “Improving IoT Management with Blockchain: Smart Home Access Control,” 2024. [Online]. Available: <https://www.ejobsat.cz/pdfs/ejo/2024/02/04.pdf>
- [53] [Author(s)], “A Survey on Diffusion Models for Anomaly Detection,” 2025. [Online]. Available: <https://arxiv.org/abs/2501.11430>
- [54] [Author(s)], “SCConv-Denoising Diffusion Probabilistic Model Anomaly Detection,” 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/14/4/746>
- [55] [Author(s)], “On Diffusion Modeling for Anomaly Detection,” 2024. [Online]. Available: <https://openreview.net/forum?id=1R3rk7ysXz>
- [56] [Author(s)], “Anomaly Detection and Generation with Diffusion Models: A Survey,” 2025. [Online]. Available: <https://arxiv.org/abs/2506.09368>

- [57] [Author(s)], “A Low-Cost IoT-Based Smart Home Automation System for Urban Sustainability,” 2025. [Online]. Available: <https://link.springer.com/article/10.1007/s43926-025-00207-4>
- [58] [Author(s)], “Generative AI Advances for Data-Driven Insights in IoT and Cloud,” 2025. [Online]. Available: [https://www.researchgate.net/publication/387773122\\_Generative\\_AI\\_advances\\_for\\_data-driven\\_insights\\_in\\_IoT\\_cloud\\_technologies\\_and\\_big\\_data\\_challenges](https://www.researchgate.net/publication/387773122_Generative_AI_advances_for_data-driven_insights_in_IoT_cloud_technologies_and_big_data_challenges)
- [59] [Author(s)], “Combining Edge Computing-Assisted IoT Security with Artificial Intelligence,” 2024. [Online]. Available: <https://www.mdpi.com/2076-3417/14/16/7104>
- [60] [Author(s)], “Differential Privacy for IoT-Enabled Critical Infrastructure: A Comprehensive Survey,” 2022. [Online]. Available: [https://www.researchgate.net/publication/355765966\\_Differential\\_Privacy\\_for\\_IoT-Enabled\\_Critical\\_Infrastructure\\_A\\_Comprehensive\\_Survey](https://www.researchgate.net/publication/355765966_Differential_Privacy_for_IoT-Enabled_Critical_Infrastructure_A_Comprehensive_Survey)
- [61] [Author(s)], “Achieving Differential Privacy in Smart Home Scenarios,” 2022. [Online]. Available: <https://scispace.com/pdf/achieving-differential-privacy-in-smart-home-scenarios-1akhdv9y9j.pdf>
- [62] [Author(s)], “Enhancing Smart Home Security: Blockchain-Enabled Federated Learning with Knowledge Distillation for Intrusion Detection,” [Year]. [Online]. Available: <https://www.mdpi.com/2624-6511/8/1/35>

## **Appendix A**

# **Appendix**