APPLIED COMPUTER SCIENCE

Applied Computer Science 2025, vol. 21, no. 2, pp. 96–110
https://doi.org/10.35784/acs_7397

Submitted: 2025-03-08 | Revised: 2025-04-24 | Accepted: 2025-04-30

Keywords: Belief-Desire-Intention model, internet of things, penetration testing, cybersecurity attacks, machine learning

Mohammed J. BAWANEH [1], *Obaida M. AL-HAZAIMEH* [1, 2*], *Malek M. AL-NAWASHI* [1], *Monther H. AL-BSOOL* [1], *Essam HANANDAH* [2]

[1] Al-Balqa Applied University, Jordan, dr_mjab@bau.edu.jo, dr_obaida@bau.edu.jo, nawashi@bau.edu.jo, monther.bsool@bau.edu.jo

[2] Philadelphia University, Jordan, ehanandeh@philadelphia.edu.jo

[*] Corresponding author: dr_obaida@bau.edu.jo

Enhanced IoT cybersecurity through machine learning - based penetration testing

Abstract

The Internet of Things (IoT) is a new technology that builds on the old Internet. A network connects all objects using technologies such as Radio Frequency Identification (RFID), sensors, GPS, or Machine-to-Machine (M2M) communication. The development of IoT has been negatively impacted by security concerns, which has led to a significant increase in research interest. However, very few methods look at the security of IoT from the attacker's point of view. As of today, penetration testing is a common way to check the security of traditional internet or systems. It usually takes a lot of time and money. In this paper, we look at the security problems of the Internet of Things (IoT) and suggest a way to test for them. This way is based on a combination of the belief-desire intention (BDI) model and machine learning. The results of the experiments showed that they were very good at detecting and defending against cyberattacks on IoT devices. The proposed BDI-based recall method provided 85% of the results. The 90% precision suggests that the measurements are very accurate. The F1-score was 87.4%, and the accuracy was 95%. The proposed BDI is of exceptional quality in every part of the penetration-testing model. Therefore, it is possible to create a system that can detect and defend against cyberattacks based on the proposed BDI model.

1. INTRODUCTION

The Internet of Things (IoT) was a big deal when it came out in 1999. MIT came up with the idea, and it's been a huge part of the next generation of information technology ever since (Yalli et al., 2024). The "Internet of Things" is like an expansion of the original Internet. It links physical objects to the web so that they can transfer data, identify things, find locations, track things, monitor things, and manage them. It uses technologies like Machine to Machine (M2M) communication, Radio Frequency Identification (RFID), and sensors (Mphale et al., 2024; Santos et al., 2014). As the current literature says, the application, network, and perception layers make up the framework of the Internet of Things (IoT), as shown in Figure 1 (Gokhale et al., 2018). Users get a bunch of different services from the application layer in all sorts of situations. Data processing and transmission happen at the network layer. At the end of the day, it's the perception layer's job to gather data and identify physical things using various hardware terminals like RFID, sensors, GPS, and more. Smart grids, intelligent traffic, smart cities, smart homes, intelligent healthcare, physical activity, and smart buildings are just a few of the current areas that have made use of IoT technology. But security has been a big worry because of the growing number of attacks (Abu-Ein et al., 2025; Al-Hazaimeh et al., 2022; Cao et al., 2022; Tahat et al., 2020).

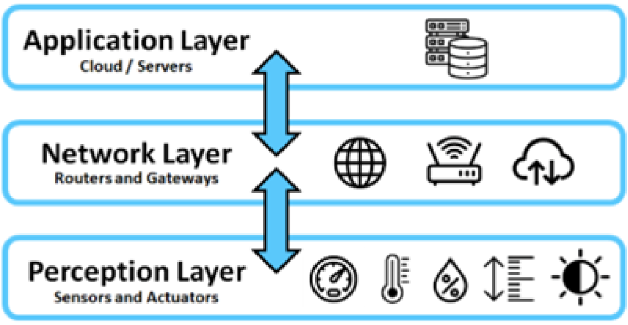

Fig. 1. Architecture of IoT layers

Penetration testing is a prevalent method that emulates authentic attacks to evaluate the security of traditional Internet or systems (Hu et al., 2020). Penetration testing execution standard (PTES) (Safitra et al., 2023) defines penetration testing as a process that includes pre-engagement interaction, The process of data collection, threat modeling, vulnerability analysis, exploitation, post-exploitation, and reporting is of paramount importance in this field. The distinction between an "attacker" and a "penetration tester" is primarily a matter of legal interpretation. Penetration testing is a method of assessing the security of a system against unauthorized access or destruction. It does not, however, impact the availability of the system under test. In the 1970s, the U.S. military discovered vulnerabilities through penetration testing and subsequently hired hackers to attack mirror targets. This enabled software engineers to reinforce computer networks. Penetration testing is a highly effective method for enhancing the security of a target system. A considerable number of firms and organizations are implementing this strategy for the purpose of detecting and addressing system vulnerabilities, thereby preventing future harm (Abu-Dabaseh & Alshammari, 2018). The prevailing focus of IoT security research is on the analysis, defense, or attack of particular devices. The security of the Internet of Things (IoT) is becoming increasingly critical. However, there is a paucity of comprehensive options that are focused on the capabilities of attackers. Despite the extensive research conducted on threat modeling, vulnerability assessment, and intrusion detection, few have systematically employed the attacker's perspective to assess IoT security across the attack surface (Bella et al., 2023). Penetration testing is a prevalent practice, but it can be costly and time-consuming (Ujjwal & Chodorowski, 2019). Automation has the potential to enhance the efficiency of penetration testing significantly. In this research, we explore issues related to IoT security and privacy, and we offer a methodology for penetration testing that is based on a combination of the Belief-Desire Intention (BDI) model and machine learning. The subsequent sections of the paper are structured as follows: In Section 2, an analysis of the security challenges associated with the Internet of Things (IoT) is conducted. In Section 3, a pertinent literature review is presented. The fourth section of this text delineates the recommended technique. In Section 5, we validate the automation of penetration testing for IoT through a simulated experiment. The ensuing discourse and prospective endeavors are delineated in Section 6 and culminate in Section 7.

2. THE SECURITY CHALLENGES OF IOT

The security of Internet of Things devices has emerged as a topic of significant interest in the twenty-first century. While the Internet of Things brings everything closer together and connects the entire globe, it also creates countless points of access that can be exploited by a variety of different types of attacks. Although the Internet of Things (IoT) is a very short phrase, it encompasses the entire world with all of its smart technologies and services (Shaukat et al., 2021). The Internet of Things connects the real and virtual worlds using intelligent devices and associated services over various communication protocols.

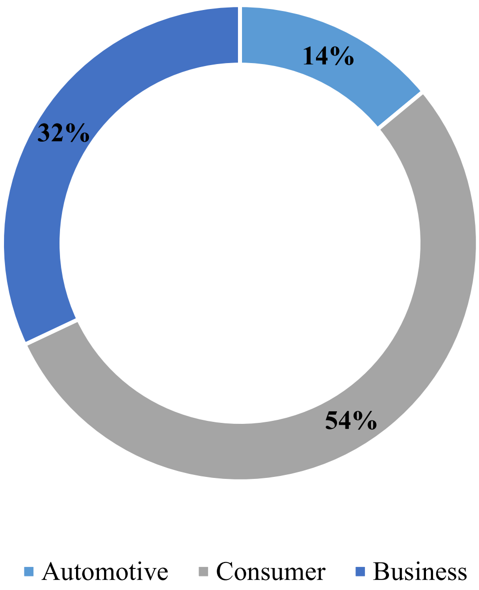

Fig. 2. Users of IoT devices by 2024 – estimated

The Internet of Things is turning a 25-year-old fantasy into reality. Intelligent technology, especially the Internet of Things, permeates today's society. People are unable to think autonomously without Internet of Things devices and services. One survey predicts that 50 billion devices will be online by 2020, and that number will continue to grow (Tawalbeh et al., 2020). Figure 2 shows the expected usage of IoT devices in 2020 (Prince et al., 2024). The Internet of Things industry is expected to reach $3,911.1 trillion by 2025. Figure 3 shows the global IoT market, the number of connected devices, and projections through 2030 (Al-Sarawi et al., 2020). For example, electrical and computer engineering researchers have focused on the Internet of Things, its development, and security over the past few decades.

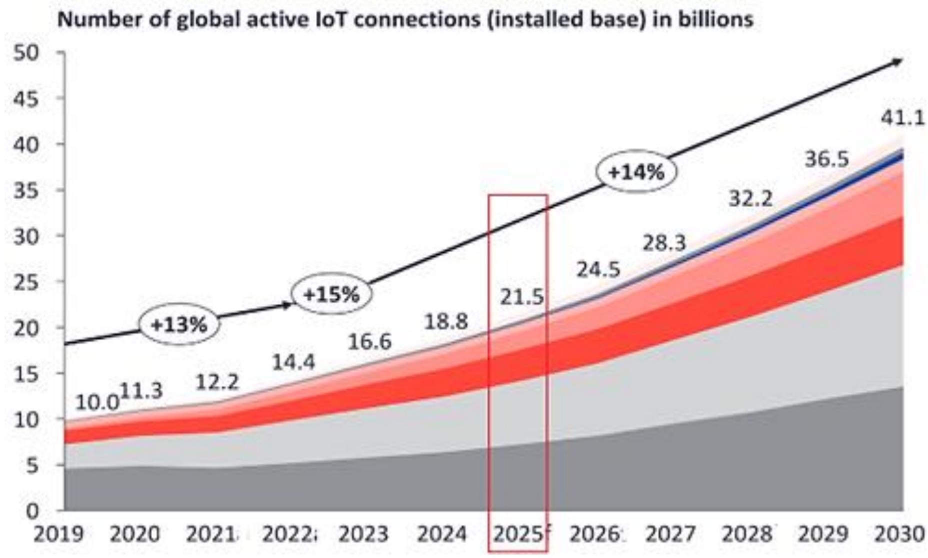

Fig. 3. Worldwide IoT device connectivity and forecast

Through a variety of applications, Internet of Things devices become more accessible when they are connected to the Internet (Zeinab & Elmustafa, 2017). Although the Internet of Things (IoT) makes life more technologically advanced, convenient, and adaptable, it also exposes users' privacy to a variety of risks and attacks. When it comes to the Internet of Things (IoT), security is a major concern because anyone can access it without authorization. Protecting IoT devices requires a variety of different security techniques.

Internet of Things (IoT) devices present significant security vulnerabilities due to their pervasive nature, limited resource availability, and diverse communication protocols. The confidentiality, integrity, and availability of data and services are all threatened by these vulnerabilities, making Internet of Things (IoT) systems susceptible to both passive and active attacks (Shaikh et al., 2018). An illustration of the Internet of Things attacks is shown in Figure 4 (Bout et al, 2022).

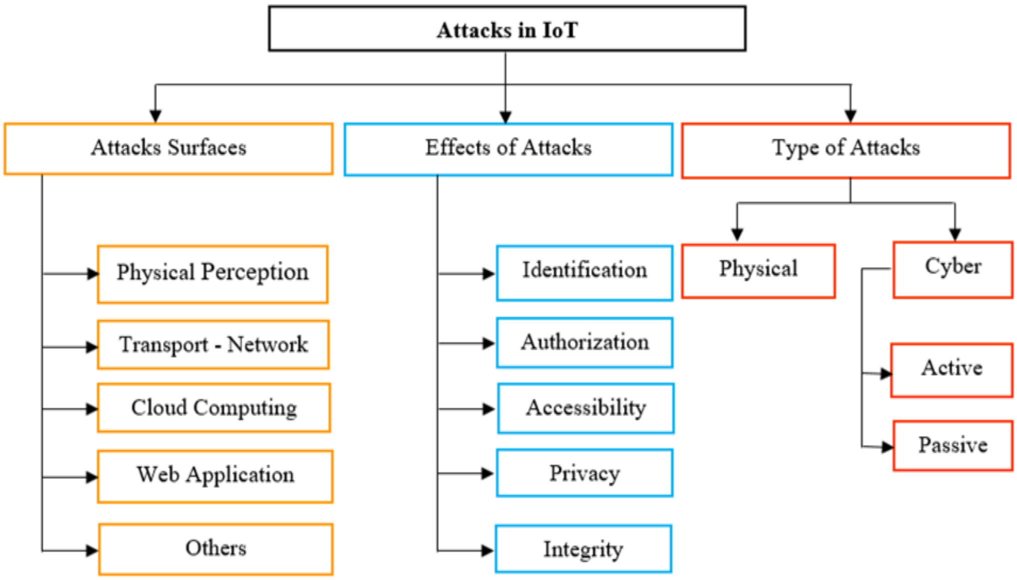

Fig. 4. Attacks of the IoT

To clarify, there are different types of cyber risks relevant to IoT security that one needs to be aware of in the field of cybersecurity today. Basically, there are two categories of threats: active and passive attacks. An active attack is one in which the perpetrators directly compromise your computer systems. They can cause a variety of problems, including file corruption, data theft, and others. A passive attack is one in which the attackers discreetly observe and gather information without the victim's knowledge (Song et al., 2020). A compilation of several types of active and passive attacks, along with their impact in the IoT environment, is described in table 1 (Bout et al., 2022; Song et al., 2020).

Tab. 1. Various active and passive attacks and their effects

Attack type	**Attacks affectation**	**Cybersecurity attacks**
Active	Identification	Sybil attack
		Spoofing attack
	Authorization	Hole attack
		Jamming attack
	Accessibility	DoS attack
		MITM attack
	Confidentiality	Selective forwarding
		Data tampering attack
	Integrity	Malicious inputs attack
		DDoS attack
Passive	Privacy	Eavesdropping attack
		Traffic analysis attack

As mentioned in the introduction, the Internet of Things architecture has three layers, as shown in Figure 1. The vulnerability of each layer to attack represents the main security risk in different Internet of Things scenarios. Table 2 summarizes the architecture of the Internet of Things layers and the associated attacks within each layer (Al-Hazaimeh & Al-Smadi, 2019; Song et al., 2020).

Tab. 2. IoT architecture layers and related attacks

Layer	Layer description	Cybersecurity attacks
Application	The application layer provides many services to users, including smart grid, transportation, city, home, healthcare, and building. It is possible to access and manage the IoT using numerous applications on platforms such as computers, mobile devices, and smart hardware (Swamy et al., 2017).	Buffer overflow, SQL injection, XSS, password attacks, and social engineering attacks
Network	Between the application layer and the perception layer, the network layer transmits information. Various networks, including the Internet, cellular networks, satellite networks, GSM, GPRS, WIFI, 3G, 4G, and so on, make up the network layer (Bello et al., 2017).	Sniffing attack, signal interference attack, data replay attack, data tampering attack, and DDOS attack.
Perception	Using sensors, GPS, RFID, and other hardware devices, the sensing or physical layer collects data from the physical world and converts it to digital form. Typically, nodes in the perception layer are lightweight, have low power, limited processing capacity, limited storage, and are unattended. Typical information security practices are not employed at the perception layer (Ali Khattak et al, 2019).	Skimming attack, eavesdropping attack, spoofing attack, shielding, jamming, killing and cloning attacks.

3. RELATED LITERATURE

Numerous studies have been conducted in the past to examine IoT security and privacy issues. References (Kumar & Patel, 2014) examines security vulnerabilities in various IoT applications, while in (Lin & Bergmann, 2016) evaluates smart home security. In addition, recent studies on potential vulnerabilities within the IoT ecosystem are analyzed in (Borgohain et al., 2015; Lin et al., 2017; Al-Nawashi et al., 2024; 2025). In addition, new investigations into privacy and protection from the perspective of technology and protocols have attracted considerable interest (Al-Hazaimeh et al., 2014; Shaqboua et al., 2022; Yang et al., 2017). All of these studies primarily address security concerns and solutions. Table 3 provides a summary of selected articles in each area that address privacy and protection in an Internet of Things (IoT) context from a technology and protocol perspective.

4. METHODOLOGY

The rapid growth of IoT devices has transformed healthcare, smart homes, and industrial automation. The low computing resources, lack of defined security standards, and complicated network designs of IoT systems make them vulnerable to hackers. To address these vulnerabilities, penetration testing is critical to security assessments of the IoT ecosystem. By simulating real-world attacks, penetration testing identifies vulnerabilities in IoT devices and networks, enabling proactive threat mitigation. In this paper, we propose the use of the Belief-Desire-Intention (BDI) paradigm to improve IoT penetration testing. The BDI framework organizes penetration testing decision making to respond to changing threats and system states. As shown in Figure 5, the proposed BDI block diagram illustrates how belief updates (system knowledge), desires (security goals), and intentions interact. This platform improves vulnerability detection, automates, and intelligently responds to threats. The proposed BDI-based approach improves penetration testing and protects IoT systems from passive and active threats. Addressing IoT security issues with this technology is a major advancement. To clarify the flowchart design in Figure 5, we divided its implementation into six critical components for IoT security: data collection, human knowledge database, machine learning, beliefs, desires, and intentions. Each component is implemented in Python to show how it interacts with the Belief-Desire-Intention (BDI) architecture (see Algorithm 1 - 6 respectively). This method provides a systematic description of how IoT systems can detect anomalies, assess risks, and make informed decisions to improve security. By combining these components, we aim to increase the transparency and effectiveness of IoT security systems, enabling proactive detection and mitigation of possible cyber threats.

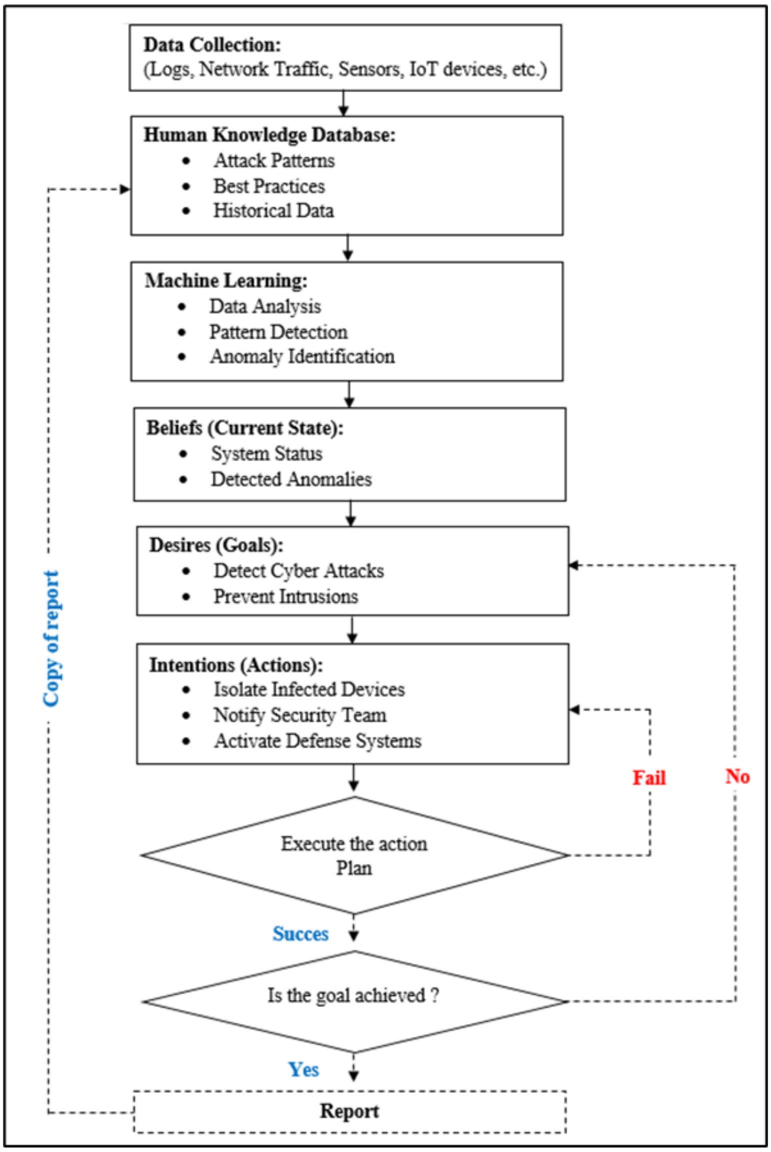

Fig. 5. Flow chart diagram – Proposed BDI – model

More specifically, the human knowledge base consists of a collection of domain-specific inference rules that mimic the thinking of experts in the process of identifying actions that are either abnormal or potentially harmful in Internet of Things environments. The formulation of these rules takes the form of conditional logic statements, such as: "If a device exceeds a predefined packet transmission threshold, classify the activity as indicative of a potential Distributed Denial of Service (DDoS) attack". Anomalous data transfer rates, frequent attempts to gain unauthorized access, communication from unregistered or unknown devices, and behavioral deviations from normal operating profiles are some of the security-related scenarios addressed by the rules introduced. The human knowledge base is not just a data retrieval system, but an essential component for decision making, providing the agent with information that informs its beliefs and initiates intentions consistent with those beliefs. This functionality enhances the system's proactive protection capabilities, which is particularly important for early detection and mitigation of intrusion threats.

Tab. 3. A summary of selected IoT privacy and protection articles in each IoT layer

IoT Layer	Reference	Year	Description
Application	(Xiao et al., 2016)	2023	Researchers established radio channel-based physical layer authentication for wireless networks. Q-learning reinforcement learning and Dyna-Q algorithms determined public data thresholds for the radio channel.
	(Outchakoucht et al., 2017)	2017	The researchers used ML and blockchain. In their solution, smart contracts use limited disclosure information to dynamically make control decisions based on environmental inputs via reinforcement learning.
	(Ni et al., 2009)	2020	The researchers used machine learning to automatically provision new applications based on roles. They also provided answers to two related challenges in this area: adapting to changes in job descriptions and setting constraints.
	(Shokri et al., 2017)	2017	Google and Amazon ML were used for membership inference. The authors tested a shadow training method on various datasets, including patient records from a Texas hospital. We found significant vulnerabilities that allow attackers to infer records.
	(Rouhani et al., 2018)	2023	The authors protected the input and learning parameters. This work used analytical and synthetic methods to use the garbled circuit cryptography method.
	(Mohassel & Zhang, 2017)	2017	A privacy-preserving machine learning protocol was developed in this research. The writers used stochastic gradient descent for logistic regression, linear regression, and the neural network model. A two-server model was enhanced with an offline phase that encrypts datasets prior to using them for ML model training.
Network	(S. Wang et al., 2019)	2019	For IoT systems with more edges, the researchers suggested federated learning. The only variables between learning blocks should be regional and international, as this approach does not transmit raw data. The scientists created a global aggregation frequency control technique to prevent learning loss.
	(X. Wang et al., 2019)	2023	A privacy-preserving architecture using reinforcement learning and deep federation on IoT platform edge devices was proposed.
	(Borthakur et al., 2017)	2020	The researchers proposed clustering behavioral data with low-resource machine learning. This fog node system analyzes health data with privacy using smart wearables and the unsupervised k-means algorithm.
	(Bonawitz et al., 2017)	2021	The authors proposed a paradigm for protecting data aggregation in federated learning with limited resources. The authors used a server to route messages using traffic data. This server simplifies and speeds up the model.
	(Xu et al., 2018)	2018	Before cloud storage, edge computing should aggregate data, according to the authors. They developed a local differential privacy strategy to de-identify data. The study recommends regression analysis for data distribution estimation.
Perception	(Lee & Chung, 2005)	2022	In their article, the authors describe a clustering method that constructs a classification model using aggressive learning neural networks.
	(Rooshenas et al., 2010)	2024	The authors developed a distributed approach for performing Principal Component Analysis (PCA) to allow the base station to access the observations. The recommended technique was developed based on the transmission load of the intermediate nodes.
	(Su et al., 2016)	2023	The differential k-means clustering algorithm has been improved in several ways. They improved an interactive differentially k-means clustering approach with systemized error analysis and developed a non-interactive method.

```python
import numpy as np

# Simulate IoT sensor data collection
def collect_sensor_data():
    """
    Simulates the collection of IoT sensor data.
    Returns a list of sensor readings (normal and anomalous data).
    """
    np.random.seed(42)
    normal_data = np.random.normal(loc=50, scale=5, size=80).tolist()  # Normal sensor readings
    anomalous_data = np.random.uniform(low=90, high=100, size=20).tolist()  # Anomalous readings
    sensor_data = normal_data + anomalous_data  # Combine normal and anomalous data
    print(f"Collected {len(sensor_data)} data points from IoT sensors.")
    return sensor_data
```

Fig. 6. Algorithm 1 Data Collection

```python
class HumanKnowledgeDatabase:
    def __init__(self):
        """
        Represents a human knowledge database with predefined rules or patterns.
        """
        self.knowledge = {
            "normal_range": (40, 60),  # Expected range for normal sensor readings
            "anomaly_threshold": 80    # Threshold above which readings are considered
                anomalous
        }

    def get_normal_range(self):
        """Returns the expected normal range for sensor readings."""
        return self.knowledge["normal_range"]

    def get_anomaly_threshold(self):
        """Returns the threshold for detecting anomalies."""
        return self.knowledge["anomaly_threshold"]
```

Fig. 7. Algorithm 2 Human Knowledge Database

```python
import pandas as pd
from sklearn.ensemble import IsolationForest

def analyze_data_with_ml(sensor_data, knowledge_db):
    """
    Analyzes sensor data using Machine Learning to detect anomalies.
    """
    if len(sensor_data) == 0:
        print("No data available for analysis.")
        return []

    # Convert sensor data to DataFrame for ML processing
    data = pd.DataFrame(sensor_data, columns=["value"])

    # Train an Isolation Forest model for anomaly detection
    model = IsolationForest(contamination=0.05, random_state=42)
    data["anomaly"] = model.fit_predict(data[["value"]])

    # Identify anomalies (-1 indicates an anomaly)
    anomalies = data[data["anomaly"] == -1]
    print(f"Detected {len(anomalies)} anomalies in the data.")
    return anomalies.to_dict(orient="records")
```

Fig. 8. Algorithm 3 Machine Learning

```python
class BeliefSystem:
    def __init__(self):
        """
        Represents the belief system in the BDI framework.
        """
        self.beliefs = {
            "sensor_data": [],           # Store sensor data
            "detected_anomalies": [],    # Store detected anomalies
            "system_status": "normal"    # Current system status
        }

    def update_beliefs(self, new_data):
        """Updates beliefs with new sensor data."""
        self.beliefs["sensor_data"].extend(new_data)
        print(f"Updated beliefs with {len(new_data)} new data points.")

    def update_anomalies(self, anomalies):
        """Updates detected anomalies in beliefs."""
        self.beliefs["detected_anomalies"] = anomalies
        if anomalies:
            self.beliefs["system_status"] = "under_attack"
        else:
            self.beliefs["system_status"] = "normal"
        print(f"System status updated to: {self.beliefs['system_status']}.")
```

Fig. 9. Algorithm 4 Beliefs (Current State)

```python
class DesireSystem:
    def __init__(self):
        """
        Represents the desire system in the BDI framework.
        """
        self.desires = {
            "detect_anomalies": True,  # Goal: Detect anomalies
            "prevent_attacks": True    # Goal: Prevent attacks
        }

    def evaluate_goals(self, beliefs):
        """Evaluates desires based on current beliefs."""
        if self.desires["detect_anomalies"]:
            if beliefs["system_status"] == "under_attack":
                print("Goal: System under attack! Take preventive actions.")
                return True
            else:
                print("Goal: System is normal. No action required.")
                return False
```

Fig. 10. Algorithm 5 Desires (Goals)

```python
class IntentionSystem:
    def __init__(self):
        """
        Represents the intention system in the BDI framework.
        """
        self.intentions = {
            "isolate_device": False,  # Action: Isolate compromised device
            "notify_admin": False     # Action: Notify admin about anomalies
        }

    def update_intentions(self, goal_achieved):
        """Updates intentions based on evaluated goals."""
        if goal_achieved:
            self.intentions["isolate_device"] = True
            self.intentions["notify_admin"] = True
            print("Intentions updated: Isolate device and notify admin.")
        else:
            self.intentions["isolate_device"] = False
            self.intentions["notify_admin"] = False
            print("Intentions updated: No actions required.")

    def execute_intentions(self):
        """Executes actions based on intentions."""
        if self.intentions["isolate_device"]:
            print("Action: Isolating compromised device from the network.")
        if self.intentions["notify_admin"]:
            print("Action: Notifying admin about detected anomalies.")
```

Fig. 11. Algorithm 6 Intentions (Actions)

5. SIMULATION AND EVALUATION

Simulations are used to evaluate the ability of our BDI-based system to solve real-world problems in dynamic and complicated environments. The Belief-Desire-Intention (BDI) framework is essential for creating intelligent systems that make autonomous decisions and adapt. By using beliefs to describe the system's understanding of the environment, desires to set goals, and intentions to act. To evaluate the system's ability to dynamically update its beliefs, prioritize desires, and execute intentions according to the Belief-Desire-Intention (BDI) model, various operational scenarios were simulated. These simulations demonstrate the system's adaptability, scalability, and robustness under different scenarios, as well as its ability to achieve desired outcomes. This study attempts to prove the usability and superiority of the BDI-based system over traditional methods for difficult IoT and cybersecurity activities such as anomaly detection, resource management, and decision making. This study used the IoTID20 dataset (Kang et al., 2019) to evaluate the performance of a BDI-based system in identifying events as normal or anomalous. These labels categorize different threats (e.g., DoS synflooding, UDP flooding, etc.) that IoT networks may suffer from. We present performance metrics to evaluate the predictive capabilities of different models: Belief Accuracy, Goal Attainment Rate, Response Time, False Positive and False Negative Rates, Scalability, Resource Efficiency, F1 Score, Precision, and Recall. These performance measures are defined according to widely accepted standards. Table 4 shows the equations for each metric (Al-Hazaimeh & Al-Smadi, 2023; Al-Hazaimeh et al., 2025; Al-Qasrawi & Al-Hazaimeh, 2013; Nahar et al., 2020; Al-Nawashi et al., 2024; 2025; Shaqboua et al., 2022).

Tab. 4. Equations for metric evaluation

Metric	Equation
Belief Accuracy	$\frac{\text{Number of Correctly Updated Beliefs}}{\text{Total Number of Belifs}} \text{X100}$
Goal achievement rate	$\frac{\text{Number of Achieved Goals}}{\text{Total Number of Goals}} \text{X100}$
Response time	$T_{\text{intenstion execution}} - T_{\text{Belief update}}$
False positive	$\frac{\text{Number of False Positives}}{\text{Total Number of Normal Events}} \text{X100}$
False negative rates	$\frac{\text{Number of False Negative}}{\text{Total Number of Actual Anomalies}} \text{X100}$
Scalability	$\frac{\text{System Performance at Scale N}}{\text{System Performance at Scale 1}}$
Resource efficiency	$\frac{\text{Tasks Completed}}{\text{Resource Consumption}}$
F1-score	$2 * \frac{\text{Precision} * \text{Recall}}{\text{Recall} + \text{Precision}}$
Precision	$\frac{\text{TN} + \text{TP}}{\text{TP} + \text{FP}} \text{X 100}$
Recall	$\frac{\text{TP}}{\text{TP} + \text{FN}} \text{X 100}$

The BDI-based system runs on a laptop equipped with a 12th generation Intel(R) Core(TM) i7-12700 CPU at 2.10 GHz and 16 GB of RAM. Figure 6 illustrates the simulation experiment, showing the three layers of the IoT and the BDI agent. Our model is implemented using Python 3.7 (RLlib package) for experimental design.

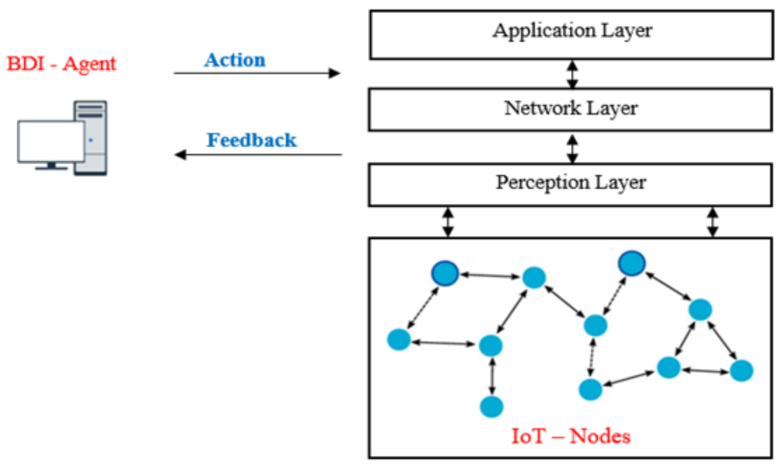

Fig. 12. BDI agent-IoT interaction

Based on the information in table 5, we have pre-defined the information of IoT targets in three layers. These layers include services and associated vulnerabilities. We stored this information in a belief set. To simulate the three structural layers of the Internet of Things (IoT), we needed four agents, one for the application layer, one for the network layer, and eleven nodes for the perception layer. The network layer is responsible for transmitting data between the application and perception layers; however, data can be transmitted between any two layers or any number of nodes. In addition, we use a randomization number to determine the outcome of an attack, making the scenario more unpredictable. The default privilege level of the BDI agent is none, with the initial goal of gaining root power in the application layer or controlling the IoT. We propose to probe and attack three levels of agents using penetration testing for IoT targets.

Tab. 5. Information related to IoT

IoT Structure	Vulnerability	Service
Application	Weak password: SSH:456, and CVE-local CVE-remote	Port, SSH, MySQL, Linux, Nginx, and App
Network	Absence of encryption	WiFi
Perception	Replay attack, and absence of encryption	Light, Lightness ZigBee network for sensor perception

In the process of applying the BDI-based system to the IoTID-20 dataset, the results of the system were considered. The evaluation metrics of the proposed framework are calculated and listed in table 6. As a result, the result is shown in the column chart shown in Figure 7.

Tab. 6. Evaluation metrics-BDI system results

Metric	Obtained result
Belief Accuracy	95%
Goal Achievement Rate	90%
Response Time	200 ms
False Positive Rate	5%
F1-Score	87.4%
Precision	90%
Recall	85%
Note: Response Time refers to the duration required to isolate a compromised device following the detection of an anomaly.	

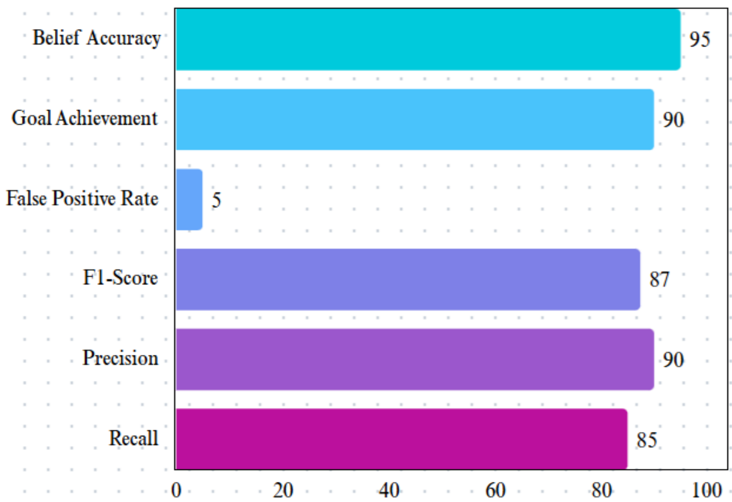

Fig. 13. Metrics column chart

BDI-based technology automates the detection and response to cyberattacks, improving the security of IoT systems. The proposed BDI-based recall rate is 85%. A precision of 90% means considerable accuracy. An F1 score of 87.4% and an accuracy of 95% are demonstrated. Consequently, the BDI-based cyberattack detection and defense system is viable for IoT environments. For IoT cybersecurity, a penetration testing model (BDI) must perform well in six key areas: Compatibility with new devices and protocols; Upgrades and bug fixes are easy with maintainability. Domain coverage, which examines supported devices, attack vectors, and network topologies; consistent performance; usability and trustworthiness, which consider ease of use and vulnerability detection for security professionals. Figure 8 shows the quality dimensions of the BDI model (Shanley & Johnstone, 2015). These dimensions can be used to assess the quality and suitability of the PT (Penetration Testing) model to address the dynamic challenges of IoT security.

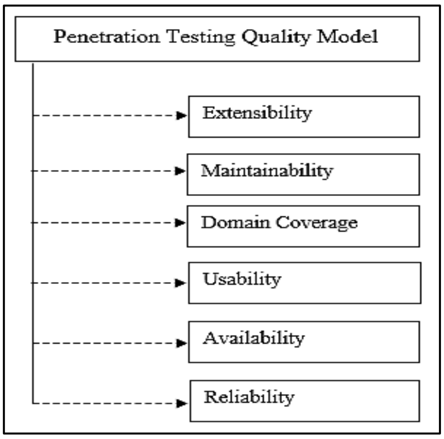

Fig. 14. Dimensions of penetration testing model

According to the metrics mentioned earlier, the proposed penetration testing methodology is of excellent quality in all six aspects. The results are shown in table 7.

Tab. 7. Dimensions evaluation -BDI system quality

Dimensions	Findings
Extensibility	The proposed BDI model includes seven Internet of Things protocols: CoAP, MQTT, Z-Wave, Zigbee, BLE, HTTPS, and HTTP.
Maintainability	On average, bug fixes take 2 hours.
Domain Coverage	The proposed BDI model has 95% coverage of IoT devices such as smart cameras, thermostats, wearables, and industrial and educational sensors.
Usability	The proposed DBI model has a 4.6/5 user satisfaction score, a 15-minute setup time, and an accessible GUI for non-experts to configure.
Availability	The proposed BDI model ensures guaranteed availability with 99.99% uptime, 3-minute fault recovery, and low resource consumption.
Reliability	With 95% detection accuracy and 5% false positives, the proposed BDI model ensures reliable vulnerability identification.

6. DISCUSSION

Intelligent agent-based models hold great promise for securing IoT environments, and the experimental evaluation of the proposed BDI-based framework provides encouraging insights into this potential. When they can reliably detect cybersecurity risks and respond with appropriate actions, their performance accuracy exceeds 95%. These results show that the design of proactive and dynamic protection mechanisms for IoT systems can be achieved by modeling intelligent behavior through desires, intentions, beliefs, and other subjective concepts. The proposed system is distinguished from others by its multi-criteria evaluation method, which outperforms the standard detection accuracy. F1 score, recall, accuracy, response time, and target achievement rate are all included to provide a complete picture of the system's performance. The detection capability, operational efficiency, and decision effectiveness of the system are validated in real-time scenarios through this complete study. The flexibility and independence of the BDI-based technique are clearly superior to those of static rule-based intrusion detection systems or traditional penetration testing methodologies. The proposed architecture uses the cognitive thinking of an agent to understand new circumstances and adapt its behavior accordingly, in contrast to most current methods that rely heavily on passive monitoring or predefined attack signatures. This represents a shift towards security paradigms that are smarter and more aware of attackers. However, even though the results are positive, it is important to recognize some limitations. Due to testing limitations, the current approach may not be sufficient to handle the complexity of actual IoT setups. In addition, the rule set of the human knowledge database could be improved with the help of domain experts or machine learning generated rules, although it is working now.

Increasing the diversity and size of the Internet of Things attack datasets used for evaluation should be the focus of future efforts. In addition, incorporating more advanced feature extraction techniques and exploring hybrid learning approaches have the potential to further improve the adaptability and accuracy of the model.

When it comes to practical applications, addressing these factors will be absolutely necessary to improve the generalizability and robustness of the system. In general, this work contributes to the growing body of research advocating for autonomous and intelligent systems in cybersecurity. Furthermore, it emphasizes the importance of agent-based models in the process of building the next generation of Internet of Things defense solutions.

7. CONCLUSION

This paper first introduces the concepts of IoT and penetration testing. Second, we examined the security of the IoT and outlined its security aspects. Third, we proposed a framework for performing automated penetration testing for the IoT. The experimental results show that the proposed BDI-based system achieved a high level of accuracy, with optimal performance exceeding 95%. A comprehensive set of evaluation criteria, including goal attainment rate, response time, F1 score, precision, recall, and false positive rate, were used to evaluate the effectiveness of the system. The results indicate that the BDI-based system not only excels in detecting cybersecurity attacks, but also has robust decision-making capabilities by selecting appropriate and timely actions to achieve its goals. The results highlight the potential of the proposed system as a reliable and effective solution to address cybersecurity issues, which represents a significant advancement in intelligent threat detection and response mechanisms. Future research should address several challenges, including expanding the IoT attack dataset, extracting additional features, and including a wider variety of IoT attack features to improve the accuracy of the results.

Conflict of interest

The authors confirm that they have no affiliations with or involvement in any organization or entity that has a financial or non-financial interest in the subject matter or materials discussed in this manuscript.

REFERENCES

Abu-Dabaseh, F., & Alshammari, E. (2018). Automated penetration testing: An overview. *4th International Conference on Natural Language Computing (NATL 2018)*. http://dx.doi.org/10.5121/csit.2018.80610

Abu-Ein, A. A., Abuain, W. A., Alhafnawi, M. Q., & Al-Hazaimeh, O. M. (2025). Security enhanced dynamic bandwidth allocation-based reinforcement learning. *WSEAS Transactions on Information Science and Applications*, *22*(1), 21-27. https://doi.org/10.37394/23209.2025.22.3

Al-Hazaimeh, O. M., & Al-Smadi, M. (2019). Automated pedestrian recognition based on deep convolutional neural networks. *International Journal of Machine Learning and Computing*, *9*(5), 662-667. http://dx.doi.org/10.18178/ijmlc.2019.9.5.855

Al-Hazaimeh, O. M., & Al-Smadi, M. A. (2023). Vehicle to vehicle and vehicle to ground communication-speech encryption algorithm. *3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)* (pp. 1-4). IEEE. http://dx.doi.org/10.1109/ICECCME57830.2023.10252814

Al-Hazaimeh, O. M., Abu-Ein, A. A., Al-Nawashi, M. M., & Gharaibeh, N. Y. (2022). Chaotic based multimedia encryption: a survey for network and internet security. *Bulletin of Electrical Engineering and Informatics*, *11*(4), 2151-2159. https://doi.org/10.11591/eei.v11i4.3520

Al-Hazaimeh, O. M., Alhindawi, N., & Otoum, N. A. (2014). A novel video encryption algorithm-based on speaker voice as the public key. *2014 IEEE International Conference on Control Science and Systems Engineering* (pp. 180-184). IEEE. http://dx.doi.org/10.1109/CCSSE.2014.7224533

Al-Hazaimeh, O. M., Al-Smadi, A., Abuain, T., & Abu-Ein, A. A. (2025). End-to-end cybersecurity encryption-video algorithm. *WSEAS Transactions on Computer Research, 13*, 116-123. http://dx.doi.org/10.37394/232018.2025.13.12

Ali Khattak, H,. Ali Shah, M., Khan, S., Ali, I., Imran, M. (2019). Perception layer security in Internet of Things. *Future Generation Computer Systems, 100*, 144-164. https://doi.org/10.1016/j.future.2019.04.038

Al-Nawashi, M. M., Al-Hazaimeh, O. M., & Khazaaleh, M. Kh. (2024). New approach for breast cancer detection based on machine learning techniques. *Applied Computer Science, 20*(1), 1-16. https://doi.org/10.35784/acs-2024-01

Al-Nawashi, M. M., Al-hazaimeh, O. M., Nedal, T. M., Gharaibeh, N., Abu-Ain, W., & Abu-Ain, T. (2025). Deep reinforcement learning-based framework for enhancing cybersecurity. *International Journal of Interactive Mobile Technologies*, *19*(3), 170-190. http://dx.doi.org/10.3991/ijim.v19i03.50727

Al-Qasrawi, I. S., & Al-Hazaimeh, O. M. (2013). A pair-wise key establishment scheme for AD HOC networks. *International Journal of Computer Networks & Communications, 5*(2), 125-136. http://dx.doi.org/10.5121/ijcnc.2013.5210

Al-Sarawi, S., Anbar, M., Abdullah, R., & Al Hawari, A. B. (2020). *Internet of things market analysis forecasts, 2020–2030. 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)* (pp. 449-453). IEEE. https://doi.org/10.1109/WorldS450073.2020.9210375

Bella, G., Biondi, P., Bognanni, S., & Esposito, S. (2023). Petiot: Penetration testing the internet of things. *Internet of Things, 22*, 100707. https://doi.org/10.1016/j.iot.2023.100707

Bello, O., Zeadally, S., & Badra, M. (2017). Network layer inter-operation of device-to-device communication technologies in Internet of Things (IoT). *Ad Hoc Networks*, *57*, 52-62. https://doi.org/10.1016/j.adhoc.2016.06.010

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. *2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17)*. Association for Computing Machinery. https://doi.org/10.1145/3133956.3133982

Borgohain, T., Kumar, U., & Sanyal, S. (2015). Survey of security and privacy issues of internet of things. *ArXiv, abs/1501.02211*. https://doi.org/10.48550/arXiv.1501.02211

Borthakur, D., Dubey, H., Constant, N., Mahler, L., & Mankodiya, K. (2017). Smart fog: Fog computing framework for unsupervised clustering analytics in wearable internet of things. *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (pp. 472-476). IEEE. https://doi.org/10.1109/GlobalSIP.2017.8308687

Bout, E., Loscri, V., & Gallais, A. (2022). Evolution of IoT security: The era of smart attacks. *IEEE Internet of Things Magazine*, *5*(1), 108-113. http://dx.doi.org/10.1109/IOTM.001.2100183

Butun, I., Österberg, P., & Song, H. (2020). Security of the internet of things: Vulnerabilities, attacks, and countermeasures. *IEEE Communications Surveys & Tutorials*, *22*(1), 616-644. https://doi.org/10.1109/COMST.2019.2953364

Cao, K., Ding, H., Wang, B., Lv, L., Tian, J., Wei, Q., & Gong, F. (2022). Enhancing physical-layer security for IoT with nonorthogonal multiple access assisted semi-grant-free transmission. *IEEE internet of things journal*, *9*(24), 24669-24681. https://doi.org/10.1109/JIOT.2022.3193189

Gokhale, P., Bhat, O., & Bhat, S. (2018). Introduction to IOT. *International Advanced Research Journal in Science, Engineering and Technology*, *5*(1), 41-44.

Hu, Z., Beuran, R., & Tan, Y. (2020). Automated penetration testing using deep reinforcement learning. *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 2-10) . IEEE. https://doi.org/10.1109/EuroSPW51379.2020.00010

Kang, H., Ahn, D. H., Lee, G. M., Do Yoo, J., Park, K. H., & Kim, H. K. (2019, September 27). IoT network intrusion dataset. IEEE Dataport. https://dx.doi.org/10.21227/q70p-q449

Kumar, J. S., & Patel, D. R. (2014). A survey on internet of things: Security and privacy issues. *International Journal of Computer Applications*, *90*(11), 20-26. http://dx.doi.org/10.5120/15764-4454

Lee, S., & Chung, T. (2005). Data aggregation for wireless sensor networks using self-organizing map. In T. G. Kim (Ed.), *Artificial Intelligence and Simulation* (Vol. 3397, pp. 508–517). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30583-5_54

Lin, H., & Bergmann, N. W. (2016). IoT privacy and security challenges for smart home environments. *Information*, *7*(3), 44. https://doi.org/10.3390/info7030044

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. *IEEE internet of things journal*, *4*(5), 1125-1142. https://doi.org/10.1109/JIOT.2017.2683200

Mohassel, P., & Zhang, Y. (2017). SecureML: A system for scalable privacy-preserving machine learning. *2017 IEEE Symposium on Security and Privacy (SP)* (pp. 19-38). IEEE. https://doi.org/10.1109/SP.2017.12

Mphale, O., Gorejena, K. N., & Nojila, O. (2024). The future of things: A comprehensive overview of internet of things history, definitions, technologies, architectures, communication and beyond. *Journal of Information Systems and Informatics*, *6*(2), 1263-1286. http://dx.doi.org/10.51519/journalisi.v6i2.738

Nahar, K. M. O., Al-Hazaimeh, O. M., Abu-Ein, A., & Gharaibeh, N. (2020). Phonocardiogram classification based on machine learning with multiple sound features. *Journal of Computer Science*, *16*(11), 1648-1656. https://doi.org/10.3844/jcssp.2020.1648.1656

Ni, Q., Lobo, J., Calo, S., Rohatgi, P., & Bertino, E. (2009). Automating role-based provisioning by learning from examples. *14th ACM symposium on Access control models and technologies (SACMAT '09)* (pp. 75-84). Association for Computing Machinery. http://dx.doi.org/10.1145/1542207.1542222

Outchakoucht, A., Es-Samaali, H., & Leroy, J. P. (2017). Dynamic access control policy based on blockchain and machine learning for the internet of things. *International journal of advanced Computer Science and applications*, *8*(7). http://dx.doi.org/10.14569/IJACSA.2017.080757

Prince, N. U., Al Mamun, M. A., Olajide, A. O., Khan, O. U., Akeem, A. B., & Sani, A. I. (2024). IEEE standards and deep learning techniques for securing internet of things (IoT) devices against cyber attacks. *Journal of Computational Analysis and Applications*, *33*(7), 1270-1289.

Rooshenas, A., Rabiee, H. R., Movaghar, A., & Naderi, M. Y. (2010). Reducing the data transmission in wireless sensor networks using the principal component analysis. *2010 Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing* (pp. 133-138). IEEE. https://doi.org/10.1109/ISSNIP.2010.5706781

Rouhani, B. D., Riazi, M. S., & Koushanfar, F. (2018). Deepsecure: Scalable provably-secure deep learning. *2018 55th ACM/ESDA/IEEE Design Automation Conference (DAC)* (pp. 1-6). IEEE. http://dx.doi.org/10.1109/DAC.2018.8465894

Safitra, M. F., Lubis, M., & Widjajarto, A. (2023). Security vulnerability analysis using penetration testing execution standard (PTES): case study of government's website. *2023 6th International Conference on Electronics, Communications and Control Engineering (ICECC '23)* (pp. 139-145). Association for Computing Machinery. http://dx.doi.org/10.1145/3592307.3592329

Santos, A., Macedo, J., Costa, A., & Nicolau, M. J. (2014). Internet of things and smart objects for M-health monitoring and control. *Procedia Technology*, *16*, 1351-1360. https://doi.org/10.1016/j.protcy.2014.10.152

Shaikh, F., Bou-Harb, E., Neshenko, N., Wright, A. P., & Ghani, N. (2018). Internet of malicious things: Correlating active and passive measurements for inferring and characterizing internet-scale unsolicited iot devices. *IEEE Communications Magazine*, *56*(9), 170-177. http://dx.doi.org/10.1109/MCOM.2018.1700685

Shanley, A., & Johnstone, M. N. (2015). Selection of penetration testing methodologies: A comparison and evaluation. *13th Australian Information Security Management Conference* (pp. 65-72). Cowan University. https://doi.org/10.4225/75/57b69c4ed938d

Shaqboua, R., Tahat, N., Ababneh, O., & Al-Hazaimeh, O. M. (2022). Chaotic map and quadratic residue problems-based hybrid signature scheme. *International Journal for Computers & Their Applications*, *29*(4), 229-235.

Shaukat, K., Alam, T. M., Hameed, I. A., Khan, W. A., Abbas, N., & Luo, S. (2021). A review on security challenges in internet of things (IoT). *2021 26th international conference on automation and computing (ICAC)* (pp. 1-6). IEEE. http://dx.doi.org/10.23919/ICAC50006.2021.9594183

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. *2017 IEEE Symposium on Security and Privacy (SP)* (pp. 3-18). IEEE. https://doi.org/10.1109/SP.2017.41

Su, D., Cao, J., Li, N., Bertino, E., & Jin, H. (2016). Differentially private k-means clustering. *Sixth ACM Conference on Data and Application Security and Privacy (CODASPY '16)* (pp. 26-37). Association for Computing Machinery. https://doi.org/10.1145/2857705.2857708

Swamy, S. N., Jadhav, D., Kulkarni, N. (2017). Security threats in the application layer in IOT applications. *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)* (pp. 477-480). IEEE. http://dx.doi.org/10.1109/I-SMAC.2017.8058395

Tahat, N., Tahat, A. A., Abu-Dalu, M., Albadarneh, R. B., Abdallah, A. E., & Al-Hazaimeh, O. M. (2020). A new RSA public key encryption scheme with chaotic maps. *International Journal of Electrical and Computer Engineering*, *10*(2), 1430-1437. http://doi.org/10.11591/ijece.v10i2.pp1430-1437

Tawalbeh, L., Muheidat, F., Tawalbeh, M., & Quwaider, M. (2020). IoT privacy and security: Challenges and solutions. *Applied Sciences*, *10*(12), 4102. https://doi.org/10.3390/app10124102

Ujjwal, K. C., & Chodorowski, J. (2019). A case study of adding proactivity in indoor social robots using Belief–Desire–Intention (BDI) model. *Biomimetics, 4*(4), 74. https://doi.org/10.3390/biomimetics4040074

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K. (2019). Adaptive federated learning in resource constrained edge computing systems. *IEEE Journal on Selected Areas in Communications*, *37*(6), 1205-1221. http://dx.doi.org/10.1109/JSAC.2019.2904348

Wang, X., Han, Y., Wang, C., Zhao, Q., Chen, X., & Chen, M. (2019). In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Network*, *33*(5), 156-165. https://doi.org/10.1109/MNET.2019.1800286

Xiao, L., Li, Y., Han, G., Liu, G., & Zhuang, W. (2016). PHY-layer spoofing detection with reinforcement learning in wireless networks. *IEEE Transactions on Vehicular Technology*, *65*(12), 10037-10047. https://doi.org/10.1109/TVT.2016.2524258

Xu, C., Ren, J., Zhang, D., & Zhang, Y. (2018). Distilling at the edge: A local differential privacy obfuscation framework for IoT data analytics. *IEEE Communications Magazine*, *56*(8), 20-25. http://dx.doi.org/10.1109/MCOM.2018.1701080

Yalli, J. S., Hasan, M. H., & Badawi, A. A. (2024). Internet of things (IoT): Origin, embedded technologies, smart applications and its growth in the last decade. *IEEE access*, *12*, 91357-91382. http://dx.doi.org/10.1109/ACCESS.2024.3418995

Yang, Y., Wu, L., Yin, G., Li, L., & Zhao, H. (2017). A survey on security and privacy issues in Internet-of-Things. *IEEE Internet of Things Journal*, *4*(5), 1250-1258. https://doi.org/10.1109/JIOT.2017.2694844

Zeinab, K. A. M., & Elmustafa, S. A. A. (2017). Internet of things applications, challenges and related future technologies. *World Scientific News*, *67*(2), 126-148.