Deep Learning-Based Intrusion Detection Framework for Securing IoT-Enabled Smart Homes

M.Sangeetha[1], Ben Sujin[2]
[1,2]Lecturer, Computer Engineering Department, University of Technology and Applied Sciences,
Nizwa, Sultanate of Oman.

Article Info

Article history:

Received Dec 25, 2024
Revised Jan 22, 2025
Accepted Feb 21, 2025

Keywords:

Smart homes
Internet of Things (IoT)
Intrusion Detection System (IDS)
Deep Learning
Cybersecurity
CNN
RNN
Edge AI

ABSTRACT

The explosion of Internet of Things (IoT) technologies has allowed smart home devices to cooperate and offer effortless access to many benefits. With everyone more connected these days, criminals have a larger chance to attack your smart home, since it makes them more likely to get into network breaches, be targeted by malware, face denial-of-service attacks and have their data stolen. Most existing IDS systems which depend on fixed rules or known signatures, have difficulty dealing with the fast changes and limited resources typical in IoT networks. To overcome these problems, deep learning introduced an automated way for detecting intrusion attempts that greatly boosted both the accuracy and flexibility of the tools. This study gives a detailed look at recent deep learning IDS frameworks built for IoT smart homes. It sorts existing methods into three categories: those for centralized architecture, distributed architecture and those depending on both host and network deployment; the list of deep learning models included are Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory networks (LSTM), Auto encoders and Generative Adversarial Networks (GANs). The study also investigates the following challenges in IoT: shortage of datasets, difficulty in detecting events at the right moment, a lack of tools to interpret how IoT algorithms function and the small amount of processing power that the devices have. Besides, it assesses frequently encountered datasets, the main performance measures used and the various methods of optimization employed for practical uses. This paper hopes to offer a starting point for those who want to build efficient, strong and flexible deep learning systems for detecting attacks in smart home environments.

Corresponding Author:

M.Sangeetha,
Computer Engineering Department,
University of Technology and Applied Sciences, Sultanate of Oman.
Email: sangeetha.mani@utas.edu.om

1. INTRODUCTION

When smart homes were first developed, the main goal for IoT was to improve ease of use and save resources through automated lighting, controlling the temperature, watching for security threats and managing household appliances [1]. Unfortunately, because IoT devices are usually resource-poor, heterogeneous and unsecure, these improvements created an attack surface on a broad scale. Originally, catching smart home attacks depended mainly on intrusion detection entailing standard rules and clearly known threats. Although they managed to detect common attacks, they couldn’t recognize new or unknown risks on IoT networks [2,3].

Whole research teams are shifting more toward machine learning (ML) and deep learning (DL) in recent times since they help create effective IDS systems for immediate threat detection and handling the changes found in networks. It is because of Convolutional Neural Networks (CNNs), Recurrent Neural

Networks (RNNs) and Long Short-Term Memory (LSTM) networks that, by automatically finding complex features of traffic information, the system can detect difficult and delicate attack behaviors [4]. Also, BoT-IoT and TON_IoT datasets have made it easy to develop and evaluate DL systems for smart home settings. The current solutions are able to notice attempts at device spoofing, command injection, data exfiltration and botnet activities accurately and rarely falsely detect such attacks [6].

Looking at the future, IDS solutions based on DL in smart homes will likely adapt to be more scattered, transparent and compact. Because of edge computing and federated learning, intrusion detection processing can now be spread to devices which makes privacy better and cuts down latency [7]. Alongside that, using XAI methods will make it easier for users to trust the decisions the model takes for security reasons. Dealing with problems related to unbalanced datasets, robustness under attacks and resource limitations is essential for future IDS frameworks. As a result, combining cybersecurity, deep learning, embedded systems and human-computer interaction in future research will be key to designing resilient and brain-like security for upcoming smart homes [8]. Figure 1 represents the IDS evolution.

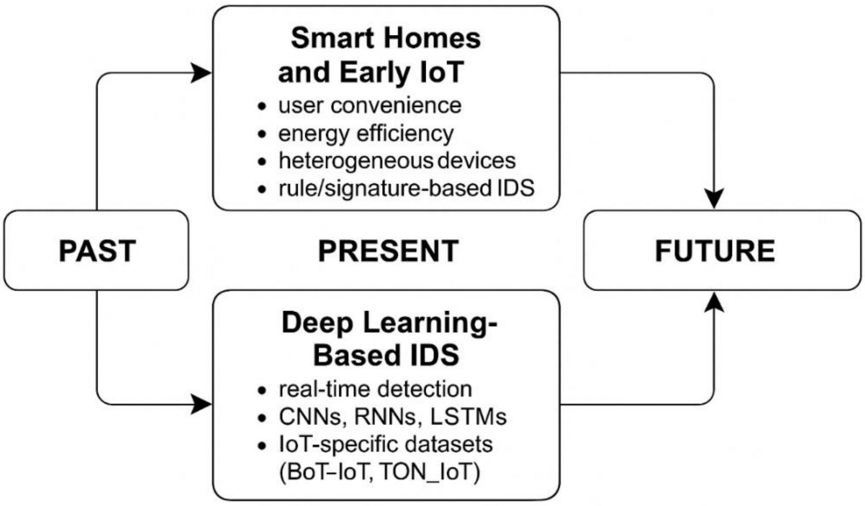

Figure 1. Evolution of IDS in Smart Homes – Past, Present, and Future Approaches

2. BACKGROUND AND MOTIVATION

2.1. Smart Homes and IoT Security Landscape

With the rise of smart homes, regular homes are now connected systems made up of smart thermostats, smart lighting, security cameras, voice control assistants and intelligent appliances. To provide automation, comfort and better energy use, home devices talk to each other through gateways and various wireless systems. Yet, the same traits that attract people to smart homes—being everywhere, accessible remotely and including several brands—also put them at high risk for cyberattacks [9]. Today’s consumer IoT devices are often developed so lightly in terms of protection that some lack important features such as robust encryption, safe firmware updates and reliable authentication. Because of these vulnerabilities, attackers can use weak passwords, default settings, unpatched software or unsafe ways to send information to break into systems, hijack connected devices or attack a network through DDoS. Besides, because smart homes often deal with private audio, video, health and behavioral data, any privacy breach carries serious safety threats for those living there. Thus, strong security systems are important to defend smart homes from upcoming cyber challenges and maintain the confidence and long-lasting success of Internet of Things systems. Figure 2 depicts the Key Components and Security Vulnerabilities in the Smart Home and IoT Security Landscape.

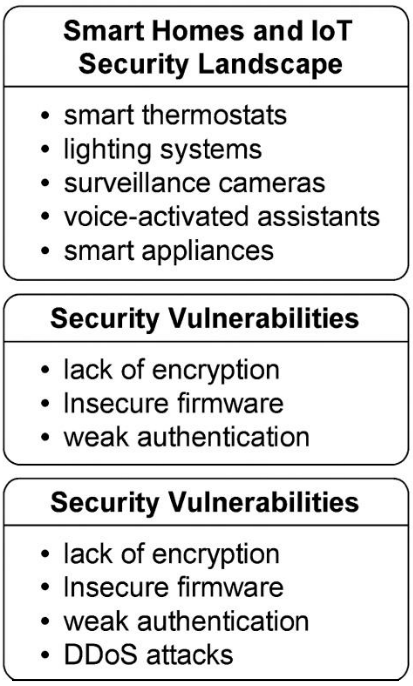

Figure 2. Key Components and Security Vulnerabilities in the Smart Home and IoT Security Landscape

2.2. Limitations of Traditional IDS

For years, IDS systems based on signatures and rules have helped detect attacks that networks already recognize. Such systems depend on rules, shortcuts or libraries of malware to spot anything suspicious or unintended. Such systems are practical in networks that stay fixed and unchanging, but not in the modern, unpredictable and strict conditions of smart home networks [10]. These tools are normally unable to spot unnoticed threat behaviors, as they do not have the ability to learn from unknown attacks. Besides, smart homes use many different devices to create high volumes of data and makes it harder for typical IDSs to adapt and scale. Additionally, too many false positives cut down on the system’s ability to detect threats, while also bothering system administrators with multiple unnecessary alerts. Because of this, rule-based systems normally need to be handled and updated manually, becoming difficult whenever new threats in IoT emerge [11]. Because of these issues, experts have started looking for smarter, flexible and scalable approaches to instantly spot threats within smart home systems. Key limitations of traditional IDS is portrayed in Figure 3.

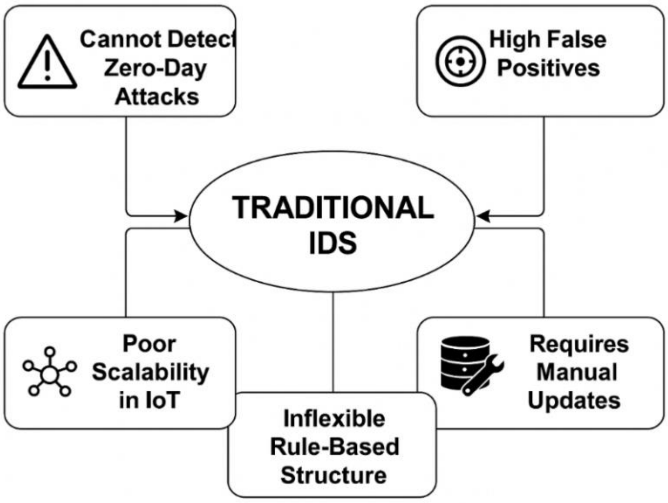

Figure 3. Key Limitations of Traditional IDS in Smart Home Environments

2.3. Role of Deep Learning in IDS

In the context of IoT-enabled smart homes, deep learning has been chosen for intrusion detection because it can automatically learn from complex data without the need to handpick important features as shown in Figure 4. Unlike regular machine learning systems, CNNs, RNNs, LSMT and Autoencoders in deep learning can examine raw network traffic without the need for manual feature extraction to discover secret attack patterns and changes over time. CNNs do extremely well at noticing various levels or hierarchies, in groups of packets, whereas RNNs and LSTMs are ideal for discovering ongoing or continuing patterns or hazards [12]. They have been shown to work better at detection, remembering and remaining consistent in rough conditions, mainly in huge and uneven IoT datasets. Furthermore, because deep learning frameworks identify changes in regular activity, they are useful for anomaly-based IDS in smart homes. So, adding DL techniques to IDS designs shows great potential for forming smart systems that are intelligent, can adjust to changes and are very efficient for IoT-based smart homes.

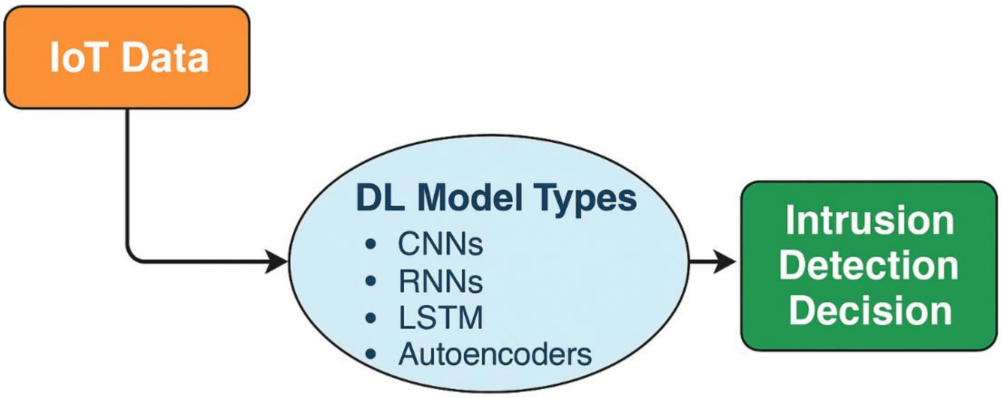

Figure 4. Deep Learning Workflow for Smart Home Intrusion Detection

3. DEEP LEARNING MODELS FOR IDS IN SMART HOMES

Table 1 gives the comparative analysis of deep learning models for intrusion detection in smart home environments.

Table 1. Comparative Analysis of Deep Learning Models for Intrusion Detection in Smart Home Environments

DL Model	Application in IDS	Strengths	Limitations
CNN	Spatial feature extraction from network traffic data	High accuracy, fast inference	Needs large datasets
RNN/LSTM	Temporal sequence modeling for attack detection	Captures sequential behavior	Prone to vanishing gradients
Autoencoders	Anomaly detection using reconstruction error	Unsupervised learning, low overhead	Poor classification performance
GANs	Synthetic data generation for IDS training	Tackles data imbalance	Training instability

4. TAXONOMY OF DL-BASED IDS ARCHITECTURES

4.1. Based on Deployment

The way IDS are used in a smart home environment helps determine their ability to detect threats, how much they consume and if they can scale. IDS are divided into three major types depending on if they are installed on a single host (HIDS), on a network (NIDS) or use both approaches together (Hybrid IDS) as shown in Figure 5.

On each IoT device, a HIDS is in place to look at system logs, watch for changes in application functions, inspect file statuses and spot processes that are not allowed. Because of this approach, it is easier to spot threats like those resulting from firmware tampering, gaining elevated privileges or software targeted at only specific devices. Yet, putting HIDS on smart home devices is not easy because most Internet of

Things components lack the right amount of processor, memory and energy. With such restrictions, deep learning models cannot be very complex and can result in delayed responses in detection.

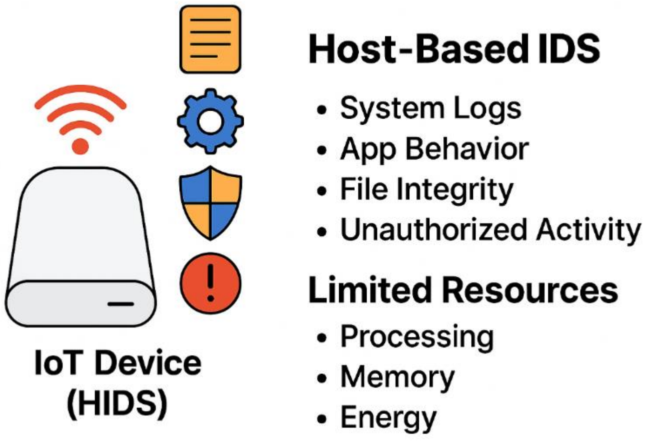

Figure 5. Functional Scope and Resource Constraints of Host-Based IDS in IoT Devices

A Network-Based IDS is located at the gateway or router in a smart home and reviews both incoming and outgoing traffic between devices inside the network and on the internet. By analyzing the way packets are exchanged, communication happens and types of used protocols, NIDS can find out about external threats including distributed denial-of-service attacks, port scans and sudden unusual traffic surges, which is depicted in Figure 6. NIDS are equipped for deep learning because they pool all processing power in the same place. However, they are less likely to catch attacks if they remain hidden within a single device or are hidden in encrypted chats, not in network traffic.

Figure 6. Centralized Network-Based IDS Monitoring Smart Home Devices and Detecting External Threats

In Figure 7, local security is handled by HIDS and NIDS is used to check network flows. With this method, threats inside and outside an organization are detected easier, the overall sense of awareness increases and the detection accuracy rises from events seen in multiple layers. In this way, a hybrid system might spy out malware by tracking abnormal changes in a device and changes in its network activity. Hybrid IDS systems are very powerful, but they add extra challenges for synchronizing, merging data and keeping the system up to date. Still, moving the computational load equally across variable devices and guaranteeing that the system remains responsive is an important challenge in smart home environments.

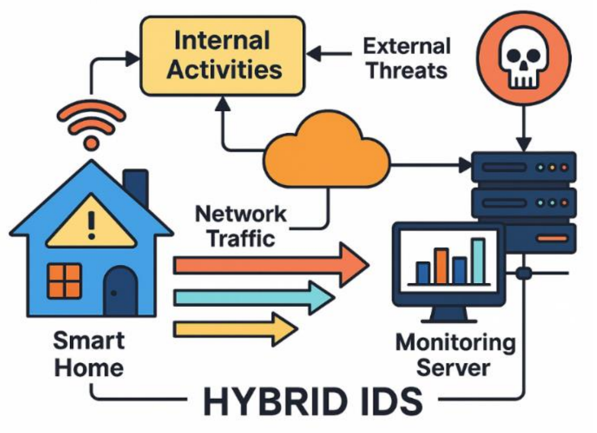

Figure 7. Hybrid IDS Architecture for Smart Homes Combining Internal and Network Threat Detection

4.2. Based on Learning Type

Deep learning IDS in smart homes can adapt and be smart because of the learning paradigm they were built on. IDS methods are divided into supervised, unsupervised and reinforcement learning branches, which is given in Table 2. Each with pros and cons that change based on what data is available for the application.

In Supervised Learning, datasets are arranged so that every input (e.g., network traffic record) has an output label indicating whether it's safe (e.g., benign) or malicious. Among the supervised methods, CNNs, LSTMs and MLPs have been successful in detecting network threats. CNNs work well with spatial data such as those in matrixes or headers and LSTMs are the best choice for modeling patterns in sequences, for example, in the order of events or headers. Even though supervised learning gives great results in lots of cases, its usefulness relies on having plenty of good-quality, evenly divided, labeled datasets which are usually missing in smart homes. Identifying examples manually is slow and training models on one set of attacks may make it more difficult to find and classify new or newly changed attacks.

These techniques help, as smart home data may be neither complete nor properly labeled. They target anomalies by studying how normal behavior looks and highlighting behavior that deviates as a probable intrusion. In IDS, common unsupervised approaches are based on autoencoders, clustering (for example, k-means) and Principal Component Analysis (PCA). In fact, autoencoders are good at restoring regular inputs and estimating how much different the inputs are from them to flag outliers. Its primary benefit is that it can identify attacks that haven't been seen before. The sensitivity to common things people do on their devices could result in an increased number of false-positives.

Reinforcement Learning is considered a fresh perspective for IDS research, especially helpful in smart home situations where things can change quickly. An RL agent learns to do things such as flag or ignore various traffic, depending on if it rewards the agent with praise or not. As a result of trial and error, RL-based IDS is able to update its detection policy with experience. The application of DRL, the combination of RL and neural networks, is being tested for automating intrusion responses and dynamic anomaly detection. But smart home applications are not widely explored, especially since it's hard to develop accurate reward rules, results are only seen over a long period and you need powerful computers to converge. Still, RL has the potential to produce smart IDS systems that act independently when facing unknown threats.

Table 2. Comparison of Learning Paradigms for DL-Based IDS in Smart Homes

Learning Type	Description	Common DL Models	Strengths	Limitations
Supervised Learning	Learns from labeled datasets (benign vs. malicious traffic)	CNN, LSTM, MLP	- High detection accuracy - Suitable for known attack types	- Requires large labeled datasets - Poor at detecting novel/zero-day attacks

Unsupervised Learning	Learns normal patterns; detects deviations as anomalies	Autoencoders, Clustering (e.g., k-means), PCA	- No need for labeled data - Effective for unknown/zero-day threats	- High false positive rate - Sensitive to behavior variations
Reinforcement Learning	Learns optimal detection policy through reward-based feedback from environment	Deep Q-Networks (DQN), Policy Gradient, DRL	- Adaptive over time - Capable of self-improvement in dynamic conditions	- Requires well-defined reward functions - High computational cost

5. DATASETS FOR SMART HOME IDS RESEARCH

Establishing and testing deep learning-based IDS systems for smart homes is mainly possible due to datasets that correctly model network behavior, events happening in the system and types of attacks. IDS models are trained, validated and judged based on such datasets. Many different datasets have emerged over the years and they are used with varying characteristics, benefits and shortcomings when building smart environments. More research studies in the IDS field use the NSL-KDD dataset which is a refined version of the original KDD'99 dataset, than any other. Though its structure is well-balanced and avoids duplication, it is now thought to be old-fashioned for new IoT applications because it does not cover many new attack methods or types of smart device traffic. It improves on earlier datasets by containing a wider range of real attacks, for example fuzzing, analysis and backdoors, plus well-defined features suitable for detecting network intrusions. Tested commonly in standard network environments, it isn't often suited to IoT applications. Traffic data from the CICIDS2017 and CICIDS2018 sets, produced by the Canadian Institute for Cybersecurity, is more realistic and recent and it shows traffic generated by both safe and malicious flows in pretend enterprise backgrounds. They enjoy great popularity among those who perform deep learning research because of their detailed labeling, extensive variety of protocols and mention of modern attacks such as botnets and brute-force attacks. Table 3 represents the public datasets comparison.

Since smart home networks and IoT ecosystems are different from other types of networks, new datasets, TON_IoT and BoT-IoT, are now available. This dataset was made to demonstrate smart environments, recording telemetry, operating system logs and traffic from different IoT devices and services. Both regular and abnormal activity happen in data from multiple sources (sensors and logs) which is important when designing IDS with multimodal deep learning. At the same time, the BoT-IoT dataset created by the University of New South Wales includes threats like distributed denial-of-service (DDoS), denial-of-service (DoS), keylogging and data exfiltration. Having over 70 million records, it's designed to imitate actual attacks on IoT networks, making it perfect for training large deep neural networks. In spite of their significance, it remains difficult for many datasets to handle class imbalance, generate synthetic traffic and lack data representation in encrypted form. Therefore, we must always work to create detailed, easy-to-use and accurate datasets to support developing IDS solutions that work accurately in real smart homes.

Table 3. Comparison of Public Datasets for Smart Home Intrusion Detection Research

Dataset	Developed By	Key Features	Smart Home Relevance	Limitations
NSL-KDD	Univ. of New Brunswick	Deduplicated version of KDD'99; labeled network connections	Low – lacks IoT-specific features	Outdated, limited to traditional IT traffic
UNSW-NB15	Australian Cyber Security Centre	Realistic modern attack vectors; 49 features	Medium – general-purpose network IDS	Less IoT-specific data
CICIDS2017/18	Canadian Institute for Cybersecurity	Benign + malicious traffic; botnets, DDoS, brute-force attacks	Medium – good traffic realism	Enterprise-oriented; limited IoT protocol diversity
BoT-IoT	Univ. of New	70M records; IoT-	High – designed	Synthetic

	South Wales	specific attacks (DDoS, DoS, data exfiltration, keylogging)	for IoT	environment; lacks encrypted data patterns
TON_IoT	Univ. of New South Wales	Telemetry, logs, and traffic from smart devices; supports multimodal analysis	Very High – tailored for smart homes	Still evolving; may require preprocessing

6. CHALLENGES IN DL-BASED SMART HOME IDS

While DL could offer a lot to security in IoT homes, researchers must still deal with several serious challenges before its usefulness is unrestricted. The challenges represented in Figure 8. One of the biggest hurdles is not hidingenough realistic, labeled data that truly shows the variety and shifts happening in smart homes, making development and research difficult. Most of the time, benign traffic is heavily represented in datasets, making it more challenging for DL models to identify rare or new kinds of threats. Besides, resource limitations on most consumer IoT devices normally keep them from running advanced DL architectures that are usually very demanding to deploy and use. DL models can also get criticized for users and security experts being unable to understand or check their conclusions due to being largely unclear about their inner workings. Because DL-based IDS systems are not clear about their processes, they cannot be trusted in places such as smart homes, where needing reliable security is highest. In addition such models are prone to adversarial attacks, in which carefully created inputs cause the system to incorrectly judge potentially dangerous threats as safe. Also, it is a complicated task for IDS solutions, since they must manage a broad range of IoT devices and changing security threats, all without having to be retrained or adjusted by users very frequently. Solving such challenges requires a variety of researchers to work together on designing lightweight models, what AI means, solid training for autonomous AI and selection of suitable learning models for the varying and limited situations in smart homes.

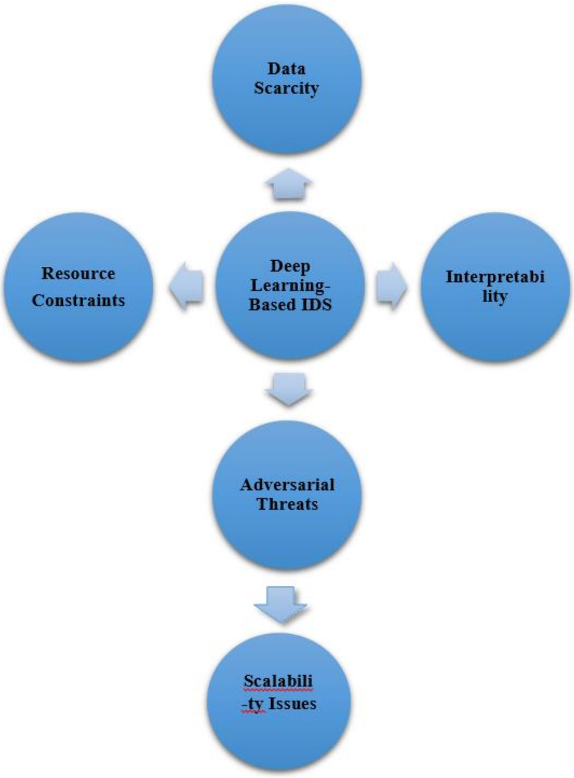

Figure 8. Key Challenges Hindering the Deployment of Deep Learning-Based IDS in Smart Homes

7. FUTURE DIRECTIONS

ID systems for smart homes using deep learning are expected to get more contextual, transparent and efficient as the use of the Internet of Things spreads and gets more complex. One eagerly awaited approach is the creation of context-aware IDS that add features like device activity, the home’s location, what happens over time and who uses the system to the typical security measures. Because these systems recognize the normal operation of a device, they can prevent many false alarms and easily spot soft or camouflaged attacks. Blockchain technology is also now being used with DL-based IDS systems to guarantee that security actions and events are safely stored and checked. Blockchain is valuable for protecting data which makes it well suited to helping investigate and trace attacks on distributed smart home systems. IDS on edge devices is becoming more popular now, thanks to the creation of energy-saving deep learning models that can still work within the power and capability of smart devices. So, security threats are detected easily on mobile devices in a fast and private way with less need for cloud services. At present, the research community is missing key benchmarks designed exclusively for smart homes. Threat intelligence datasets frequently fail to capture details from real-world attacks, provide regular labels or represent new threats. Making suitable, rich and community-accessible benchmarks gives a good basis for realistically testing and judging the achievements of various IDS. All of these areas are set to help smart home IDS become more reliable, wise and useful for actual use.

8. CONCLUSION

Although IoT has made smart homes easier to control, automate and manage real-time, it has also made cybersecurity a greater issue that existing systems cannot fully address. Because today's smart homes work together more closely, intelligent and scalable security is needed to catch any new or complex threat. Deep learning is now considered a key approach because it supports robust feature extraction, detailed study of temporal information and effective anomaly detection in various and changing conditions. This section comprehensively investigated deep learning-based systems for intrusion detection in smart homes, organized the existing approaches according to their deployment methods and learning techniques and listed the most used datasets, indicators of performance and technical issues affecting them. In spite of deep learning's promises, important issues such as having too little data, understanding how models work, limited resources and threats from adversarial attacks are preventing full deployment. In addition, concerns about how well these methods will scale, how they can be used generally and how they work in real time on the edge must be properly considered for many to use them. Going forward, practices such as federated learning, explainable AI, transfer learning and blockchain technology can strongly improve how IDS systems work and adapt. For smart home safety to improve, we need lightweight models that are easily understood, aware of situations and capable of functioning efficiently in small devices with no risk to user privacy. If these efforts are made, the field can advance to develop secure, smart and flexible IDS frameworks that look after the safety and privacy of smart homes while pervasive computing is active.

REFERENCES

[1] A. Abeshu and N. Chilamkurti, "Deep learning: The frontier for distributed attack detection in Fog-to-Things computing," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 169–175, 2018, doi: 10.1109/MCOM.2018.1700418.

[2] M. A. Ferrag, L. Maglaras, A. Derhab, M. Mukherjee, and H. Janicke, "Federated machine learning for cybersecurity and privacy: Concepts, challenges, and future directions," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 466–488, 2020, doi: 10.1109/COMST.2020.3031970.

[3] R. Li, Z. Liu, C. Liu, Z. Cui, Y. Zhou, and Y. Zhang, "A hybrid deep learning-based intrusion detection system for software-defined networks," *Sensors*, vol. 20, no. 22, p. 5827, 2020, doi: 10.3390/s20215827.

[4] A. Alsarhan, K. Salah, and A. Al-Fuqaha, "Intrusion detection for IoT networks using deep learning and blockchain: Challenges and future directions," *Future Generation Computer Systems*, vol. 140, pp. 243–259, 2023, doi: 10.1016/j.future.2022.10.005.

[5] A. Mosenia and N. K. Jha, "A comprehensive study of security of Internet-of-Things," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 4, pp. 586–602, 2017, doi: 10.1109/TETC.2016.2606384.

[6] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, and Y. Elovici, "N-BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, 2018, doi: 10.1109/MPRV.2018.03367731.

[7] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, 2018, doi: 10.1016/j.future.2017.08.043.

[8] M. Husák, M. Čermák, T. Jirsík, and P. Čeleda, "Survey of attack projection, prediction, and forecasting in cyber security," *Computers & Security*, vol. 75, pp. 132–147, 2018, doi: 10.1016/j.cose.2017.12.004.

[9] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021, doi: 10.1016/j.future.2020.10.007.

[10] Z. Zhou, H. Xiong, X. Zhang, H. Yu, and Q. Jin, "A review on multimodal learning for cyber intrusion detection," *IEEE Access*, vol. 10, pp. 125734–125751, 2022, doi: 10.1109/ACCESS.2022.3220075.

[11] R. Rahim, "Effective 60 GHz signal propagation in complex indoor settings," *National Journal of RF Engineering and Wireless Communication,* vol. 1, no. 1, pp. 23–29, 2023, doi: 10.31838/RFMW/01.01.03.

[12] M. Sudhir, K. Maneesha, G. Anudeepthi, T. Anusha, and A. Chandini, "Untangling Pancard by designing optical character reader tool box by correlating alphanumeric character," *International Journal of Communication and Computer Technologies,* vol. 10, no. 1, pp. 7–10, 2022.