frontiers | Frontiers in The Internet of Things

TYPE Review
PUBLISHED 04 September 2025
DOI 10.3389/friot.2025.1658273

OPEN ACCESS

EDITED BY
Kuo-Hui Yeh,
National Yang Ming Chiao Tung University, Taiwan

REVIEWED BY
Rolando Herrero,
Northeastern University, United States
Ali Hassan,
HITEC University, Pakistan

*CORRESPONDENCE
Fabian Chukwudi Ogenyi,
ogenyi@kiu.ac.ug

RECEIVED 02 July 2025
ACCEPTED 13 August 2025
PUBLISHED 04 September 2025

CITATION
Ogenyi FC, Ugwu CN and Ugwu OP-C (2025) Securing the future: AI-driven cybersecurity in the age of autonomous IoT.
Front. Internet Things 4:1658273.
doi: 10.3389/friot.2025.1658273

Securing the future: AI-driven cybersecurity in the age of autonomous IoT

Fabian Chukwudi Ogenyi[1]*, Chinyere Nneoma Ugwu[1] and Okechukwu Paul-Chima Ugwu[2]

[1]Department of Electrical, Telecommunication and Computer Engineering, Kampala International University, Kampala, Uganda, [2]Department of Publication and Extension, Kampala International University, Kampala, Uganda

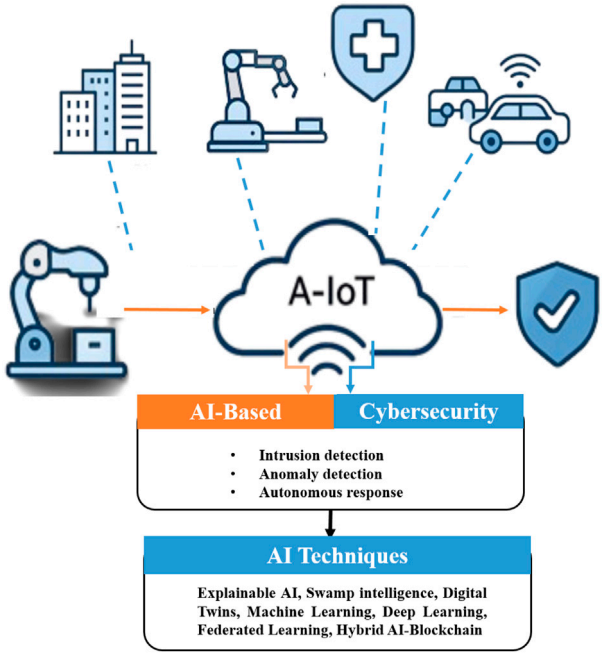
The Autonomous Internet of Things (A-IoT) represents a major advancement in interconnected systems, enabling self-governing smart devices to operate collaboratively across domains such as smart cities, industrial automation, healthcare, and autonomous vehicles. However, the complexity, scale, and heterogeneity of A-IoT environments introduce severe cybersecurity challenges, including expanded attack surfaces, real-time data processing demands, sophisticated adversarial threats, and privacy risks. Traditional security measures are not always adequate to address these emerging threats, and this is why intelligent adaptive defence systems are required. This narrative review offers an extensive and systematic presentation of AI-based cybersecurity strategies that are specific to the peculiarities of A-IoT ecosystems. It examines fundamental methods, including machine learning, deep learning, federated learning, and swarm intelligence, as well as the latest paradigms, such as explainable AI, generative adversarial networks, and digital twins. The approaches are discussed within the scope of the most important security tasks, such as intrusion detection, anomaly detection, malware analysis, secure authentication, and autonomous threat response. The review also locates crucial issues related to data quality, model interpretability, adversarial vulnerabilities and ethical limitations of the application of AI in security-critical applications. Moreover, it describes future research directions using hybrid AI-blockchain frameworks, self-healing autonomous agents, and trust-aware AI systems.

KEYWORDS

autonomous internet of things (A-IoT), AI-driven cybersecurity, intrusion detection systems, federated learning, explainable artificial intelligence (XAI), and autonomous self-healing security

1 Introduction

The rapid development of the Internet of Things (IoT) has ushered in a new era of transformation, with billions of devices, including wearables and sensors, industrial equipment, and so on, being linked to execute automated, data-driven activities (Alaba, 2024). The Autonomous Internet of Things (A-IoT) is the current stage of the IoT ecosystem's growth, which is growing increasingly autonomous, intelligent, and decentralized (Vermesan et al., 2022). A-IoT combines standard IoT with machine intelligence to allow systems to observe, assess, and respond with minimal human interaction. Applications such as autonomous automobiles, smart manufacturing,



GRAPHICAL ABSTRACT
AI-driven cybersecurity in autonomous iot ecosystems.

remote healthcare, and intelligent energy grids demonstrate A-IoT's expanding importance in commercial and critical infrastructure (Chataut et al., 2023). Nonetheless, such high levels of automation and interconnection create a dynamic and complex cyber threat environment. Because of their numerous components, limited resources, and real-time operations, A-IoT setups are prone to complex security concerns, in contrast to traditional systems (Goudarzi et al., 2022). These systems are typically located at the network's edge, where it is more difficult to control everything from a single location, increasing the risk of sophisticated cyberattacks such as adversarial machine learning, data poisoning, spoofing, botnet propagation, and 0-day vulnerabilities (Zhukabayeva et al., 2025). As A-IoT devices gain more autonomous control over safety-critical operations, cybersecurity becomes a technological and operational need (Kabir et al., 2022).

To address these problems, there has been a boom in the use of Artificial Intelligence (AI) in cybersecurity systems. AI may enhance security systems by learning from data, detecting anomalous activity, and adapting to new sorts of threats, allowing it to surpass the limitations of traditional rule-based and signature-based systems (Ahmad et al., 2024). Machine learning can predict possible attacks, deep learning may reveal hidden patterns in network behaviour, and reinforcement learning can enable autonomous threat response strategies (Sewak et al., 2023). Furthermore, the rise of federated learning, comprehensible AI, and self-healing systems suggests that AI has greater potential for securing decentralised and privacy-sensitive A-IoT systems (Ding et al., 2023). Recent reviews on IoT security and design strategies, such as those by Hassan et al. (2024) and Hassan et al. (2025), provide valuable frameworks for presenting insights into security features, antenna architectures, and AI-enabled protection mechanisms.

This analysis presents a critical and extensive synthesis of current achievements in AI-based cybersecurity that are relevant to the autonomous IoT ecosystem. It especially examines techniques for building and applying AI methodology to secure A-IoT systems against a wide range of cyber threats, focusing not only on underlying technology but also on emerging paradigms. It also exposes basic problems in present practice and suggests future research directions to increase scalability, interpretability, and resilience in hostile environments. The review is grounded in literature sourced from IEEE Xplore, SpringerLink, Scopus, and the ACM Digital Library, covering studies published between 2020 and 2025, selected based on relevance to AI-enabled cybersecurity in A-IoT contexts. The remainder of this work is structured as follows: Section 2 addresses the development and technological environment of A-IoT systems, as well as their important characteristics and cyber-risk profiles. Section 3 of the paper outlines the cybersecurity problems that autonomous IoT infrastructures provide. Section 4 discusses the ideas and taxonomies of AI in cybersecurity, while Section 5 delves into specific AI-driven techniques and defence mechanisms in A-IoT environments. Section 6 covers new and emerging AI approaches, whereas Section 7 includes real-world instances. Section 8 outlines evaluation measures, whereas Section 9 identifies existing gaps and unsolved difficulties. Section 10 discusses possible future study directions. Section 11 concludes with major facts and perspectives.

2 Evolution of IoT and autonomous systems

The IoT has undergone a rapid transformation from basic sensor-based networks to complex, intelligent, and autonomous ecosystems,

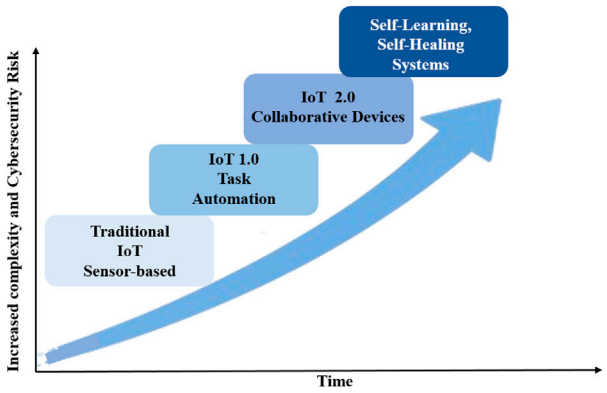


FIGURE 1
Evolution of IoT systems toward autonomy.

as illustrated in Figure 1. This evolution can be categorized into three generations (Choudhary, 2024). The first generation of IoT was characterized by passive and static data collection. Devices in this phase, including sensors and actuators, functioned primarily as data acquisition tools connected to centralized monitoring and control units (Abduljawwad et al., 2023). These systems lacked cognitive capabilities and real-time flexibility, limiting their use to basic tasks such as environmental sensing and inventory management (Kabir et al., 2022). Decision-making was entirely human-driven, and system behavior remained rigid and predictable.

The second generation, often termed Smart IoT, introduced context awareness and basic intelligence. With the integration of cloud platforms, edge computing, and more capable embedded systems, IoT devices gained the ability to process data locally and make limited decisions based on environmental inputs (Bablu and Rashid, 2025; Khriji et al., 2022). This era saw the rise of smart homes, connected vehicles, and industrial automation systems, where devices could dynamically respond to certain events. However, intelligence and coordination were still heavily dependent on centralized infrastructure (Chataut et al., 2023).

The emergence of the third generation, known as the Autonomous IoT (A-IoT), marks a paradigm shift in the design and function of connected systems. A-IoT systems are not only intelligent but also self-governing, adaptive, and capable of autonomous decision-making (Sewak et al., 2023; Valsalan et al., 2024). The convergence of IoT with artificial intelligence, distributed edge computing, and high-speed communication technologies like 5G and 6G enables this shift. Designed to learn from data, make predictions, and execute real-time actions, A-IoT devices operate with minimal human input. For example, in precision agriculture, autonomous drones can analyze crop health and initiate spraying without human intervention. Similarly, self-driving vehicles can communicate with roadside infrastructure to adapt to changing traffic conditions (Hossain M. S. et al., 2025).

Key characteristics of A-IoT systems include contextual intelligence, decentralized decision-making, continuous learning, and multi-agent collaboration (Vermesan et al., 2022; Anjosi et al., 2023). These capabilities enable real-time responsiveness in critical applications such as intelligent transport systems, robotic surgery, industrial automation, smart grids, and emergency response. However, the same features that empower autonomy also introduce new system requirements, particularly the need for dependable, scalable, and secure performance in dynamic, heterogeneous environments (Santoso and Surya, 2024).

As A-IoT systems grow in complexity and autonomy, they also become increasingly vulnerable to advanced cybersecurity threats (Zhukabayeva et al., 2025). Unlike traditional IoT systems with predefined communication flows and perimeter-based defenses, A-IoT networks are highly adaptive, open-ended, and often *ad hoc* in structure. Mechanisms that enable autonomy, such as self-learning models, continuous interdevice communication, and collaborative behaviour, also expose these systems to sophisticated attack surfaces (Roy, 2023). For instance, by manipulating training data or injecting malicious inputs during inference, adversarial machine learning can corrupt AI models. In addition, threats such as firmware tampering, data spoofing, botnet propagation, and insider compromise are amplified by the decentralized nature and heterogeneity of A-IoT devices (Asadi et al., 2024).

The lack of centralized control and a unified security policy in the majority of A-IoT deployments is another critical issue because it makes real-time threat detection and coordinated mitigation activities challenging. The traditional cybersecurity paradigm is not sufficient in this new era, and thus there is a need to develop an AI-based, context-aware, and autonomous cybersecurity paradigm that can evolve along with the systems that it defends (Tallam, 2025). Machine learning, encryption, edge analytics, and adaptive system design are becoming not only a requirement to enable autonomy but also to protect it. Finally, A-IoT presents not only a technological revolution but also a new cybersecurity frontier, which needs interdisciplinary innovation to ensure the dependability of operations in A-IoT cybersecurity (Alfahaid et al., 2025).

3 Cybersecurity challenges in autonomous IoT (A-IoT) ecosystems

The increased usage of A-IoT systems across society's most vital sectors, such as healthcare, transportation, manufacturing, and energy, the complexity and scale of the cybersecurity environment have grown tremendously (Kabir et al., 2022). Such systems differ significantly from traditional information infrastructures because to their distributed intelligence, autonomy, mobility, and real-time operation, which together provide a novel and complex set of security issues (Narayanan et al., 2022). The heterogeneity and scalability of the ecosystem is one of the most pressing issues in A-IoT cybersecurity. A-IoT environments consist of a large number of devices with varying hardware architectures, communication protocols, software stacks, and functional functions (Bouzidi et al., 2022). This variability complicates not just the application of standard security principles, but also the authentication of devices, firmware integrity, and secure data transit throughout the system (Catuogno and Galdi, 2023). Furthermore, there might be billions of autonomous nodes scattered throughout the world, and typical centralised security systems' latency, bandwidth, and control overhead are untenable, necessitating decentralised, adaptive, and

scalable defence systems. Aside from heterogeneity, resource restrictions and real-time processing present another significant obstacle to A-IoT system security. Most edge devices, such as sensors, actuators, and embedded controllers, have limited computational, memory, and energy resources (Cardoso et al., 2023). These limits typically prevent the use of traditional encryption techniques and machine learning models, which might be computationally intensive or need regular communication with cloud resources. Furthermore, A-IoT is real-time, which means that choices must be made autonomously within milliseconds, necessitating lightweight and latency-aware security techniques (Mondal et al., 2024). Intrusion detection systems (IDS) and anomaly detection models must be rapid and have a low false positive rate to avoid interfering with time-sensitive operations such as autonomous driving, remote surgery, or industrial automation. One of the research challenges is developing real-time, low-overhead, distributed AI-based security solutions (Arulmurugan et al., 2024).

Another critical concern in A-IoT cybersecurity is the risk of data leakage and integrity breaches. Autonomous devices are used to continually collect, process, and communicate sensitive data, such as personal health information, user behaviour, geolocation, and environmental measurements, without the user's knowledge or consent (Alam, 2024). A-IoT is decentralised and frequently mobile, increasing the likelihood of data interception, alteration, or exfiltration during transmission or storage (Asadi et al., 2024). In the case of edge device learning combined with federated learning or swarm intelligence, the confidentiality and integrity of raw data and model updates are crucial (Lazaros et al., 2024). Unless A-IoT systems have robust encryption, data routing, and tamper-resistant storage, they might become a source of significant privacy breaches and disinformation. This difficulty is exacerbated by the fact that in constrained edge settings, traditional Public Key Infrastructure (PKI) and blockchain-based solutions may not be feasible due to resource limits (Ni et al., 2024).

New and highly intelligent risks to A-IoT systems are also developing as a result of AI-powered cyber threat development. Artificial intelligence attacks, such as deepfake command injection, synthetic data poisoning, and generative adversarial examples, can modify sensor inputs, confuse AI classifiers, and circumvent traditional IDS systems (Ghiurău and Popescu, 2024). Enemies might use AI model flaws in A-IoT devices to launch adversarial assaults that silently change inputs, resulting in misclassification or faulty decision-making, which can cause bodily harm in safety-sensitive circumstances (Camerota, 2025). Such attack vectors are especial concerning for A-IoT systems that rely heavily on perception and decision-making, such as self-driving cars or intelligent surveillance drones. What makes the situation worse is that such assaults are difficult to detect and can appear innocent to human users and traditional signature-based security measures (Asiri et al., 2023). Defending AI itself, through adversarial training, robust learning, and explainability, has become a critical frontier in A-IoT cybersecurity.

To further complicate the security picture, the number of 0-day vulnerabilities in decentralised A-IoT systems is increasing (Chataut et al., 2023; Hossain S. et al., 2025). They are security flaws that have yet to be found and can be exploited by attackers until developers learn about them or provide patches. Zero-day vulnerabilities can propagate fast and remain unpatched for lengthy periods of time in decentralised systems where firmware and software updates are not centrally managed or devices are not always connected (Zengeni and fadli Zolkipli, 2024). This offers up opportunities for large-scale botnet formation, backdoor implants, and firmware takeovers (Chen et al., 2024). The most major issue is that most A-IoT devices lack safe over-the-air update methods, and the range of manufacturers and platforms makes it difficult to distribute fixes on time as shown in Table 1.

Furthermore, peer authentication and consensus, which are widely utilised in decentralised trust models, are subject to Sybil, spoofing, and insider attacks (AlMarshoud et al., 2024). The protection of A-IoT ecosystems requires a paradigm shift, with static, reactive security techniques being replaced by dynamic, proactive, and AI-enhanced approaches that can manage heterogeneity, resource constraints, privacy issues, and ever-changing threats (Commey et al., 2024). These difficulties highlight the significance of doing multidisciplinary research in artificial intelligence, embedded systems, cryptography, and real-time systems engineering to create the next-generation of robust A-IoT security systems (Allioui and Mourdi, 2023).

4 Role of artificial intelligence in cybersecurity

As the complexity and dynamism of cyber threats in A-IoT settings have grown, the shortcomings of traditional cybersecurity procedures have become increasingly obvious (Tariq et al., 2023). Traditional defence solutions, such as signature-based intrusion detection, rule-based access control, and periodic patching, are struggling to keep up with the latest cyberattacks, which are adaptable, polymorphic, and stealthy. These outdated techniques are often based on pre-defined threat signatures or set rules, and thus are ineffective against new or 0-day threats, particularly in decentralised, resource-constrained, and high-velocity A-IoT ecosystems (Sadhu et al., 2022). Furthermore, they lack situational awareness and on-the-fly flexibility to respond to challenges that occur when the cyber and physical worlds intersect, such as spoofing sensor readings or adversarial manipulation of AI models (Guesmi et al., 2023). AI has revolutionized cybersecurity by allowing for data-driven, autonomous, and adaptive threat identification and response. Machine Learning (ML) is one of the most popular AI applications because it allows you to learn from data patterns and generate predictions without using explicit programming (Taye, 2023). Supervised learning is commonly used in cybersecurity contexts to perform tasks such as malware classification and intrusion detection, where labelled datasets of known threats exist (Allioui and Mourdi, 2023). Unsupervised learning, in turn, is important for anomaly detection, which is the identification of unexpected behaviour or departures from prior patterns that might indicate an ongoing assault (Usmani et al., 2022). Reinforcement learning (RL) (Shehzadi, 2024), which may be used to tune automated firewalls or manage adaptive honeypots, is less mature in deployment but promise for autonomous defence systems that can learn optimal techniques by trial and error in dynamic threat situations. In addition to classical ML, Deep Learning (DL) has the potential to model complicated

TABLE 1 Key cybersecurity challenges in A-IoT ecosystems.

Challenge	Description	Example scenario	Implications	References
Expanded Attack Surface	High number and diversity of autonomous devices increase the potential entry points for attackers.	Smart cities with massive sensor and edge networks.	Increases susceptibility to lateral movement, DDoS attacks, and firmware hijacks.	Tariq et al. (2023), Sadhu et al. (2022), Kabir et al. (2022)
Heterogeneity and Scalability	Devices vary in architecture, OS, protocols, and roles; system must support billions of devices.	Cross-vendor autonomous logistics networks.	Limits interoperability; complicates authentication, updates, and data protection.	Catuogno and Galdi (2023), Bouzidi et al. (2022)
Resource Constraints	Limited computation, memory, and battery make traditional security and AI methods unsuitable.	Remote battery-powered environmental sensors.	Hinders the use of heavy encryption or deep learning; needs lightweight security.	Cardoso et al. (2023), Hudda and Haribabu (2025), Tayyab et al. (2023)
Real-Time Processing Requirements	Autonomous decisions must be made within strict timeframes to ensure safety and continuity.	Millisecond reactions in autonomous vehicles or robots.	Delayed detection can result in physical harm or operational failures.	Reddy et al. (2024), Shehzadi (2024), Mondal et al. (2024)
Data Privacy and Integrity	Continuous sensing and data exchange raise privacy and integrity concerns during storage or transmission.	Federated learning for personalized medical diagnostics.	Violations of regulations like GDPR; risk of surveillance and tampering.	Chang et al. (2023), Asadi et al., 2024, Alam (2024), Rao and Deebak (2023)
Adversarial Threats	Attackers can manipulate inputs to deceive AI-based security systems through synthetic or poisoned data.	Deepfake commands in voice-controlled industrial systems.	AI models make incorrect or unsafe decisions; lowers system trust.	Shayea et al. (2025), Ghiurău and Popescu (2024), Camerota (2025)
Zero-Day Vulnerabilities	Unknown flaws in firmware/software are exploited before patches are available or applied.	Botnet propagation in IoT-enabled smart factories.	Hard to detect or patch across decentralized devices; long-term persistence.	Chataut et al. (2023), Zengeni and fadli Zolkipli (2024), Hossain S. et al. (2025), Chen et al. (2024)
Authentication and Trust Issues	Peer devices and updates are vulnerable to spoofing, Sybil, and insider attacks in decentralized networks.	Drone-to-drone authentication in a surveillance swarm.	Trust breakdown leads to false decisions, data leaks, or control hijacking.	AlMarshoud et al. (2024), Commey et al. (2024)

high-dimensional data as shown in Figure 2 and Table 2. Convolutional and recurrent deep neural networks have been utilized for network traffic analysis, encrypted malware detection, and behavioural profiling of IoT nodes. DL models can learn hierarchical features from raw inputs, allowing them to be very accurate even in noisy or encrypted contexts (Tayyab et al., 2023). They require significant amounts of labelled data and computer power, which may be challenging to deliver in A-IoT devices with limited storage and processing capacity. Federated Learning (FL) presents a novel paradigm that preserves privacy (Chang et al., 2023). In FL, several A-IoT devices collaborate to train a global model, providing only local model updates rather than raw data. The strategy is more privacy and data locality friendly, as well as communication overhead efficient, making it ideal for sensitive applications such as smart healthcare or industrial IoT systems (Rao and Deebak, 2023).

Swarm Intelligence (SI) is another emerging AI paradigm in cybersecurity that is inspired by the collective activities of biological species like ants and birds. SI may be used to create distributed, cooperative, and adaptive security solutions in A-IoT ecosystems, in which autonomous agents communicate intelligence and respond to threats in a coordinated manner. A swarm-based intrusion detection system, for example, allows devices to communicate anomaly scores and coordinate responses in a decentralised fashion, making the system more robust to local failures or assaults (Reddy et al., 2024). Swarm-based defence tactics are especially useful in highly mobile or

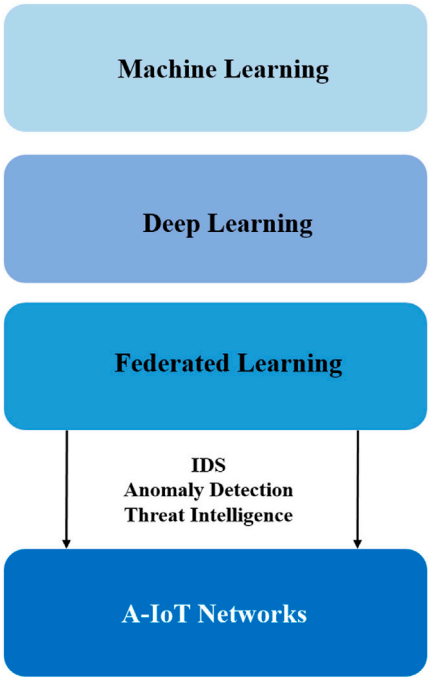


FIGURE 2
AI-powered cybersecurity architecture for A-IoT ecosystems.

TABLE 2 Comparison of traditional vs. AI-driven security methods in A-IoT.

Criteria	Traditional security	AI-driven security	References
Adaptability	Rule-based; static configurations often require manual updates.	Continuously adapts to evolving threats *via* learning from new data.	Tariq et al. (2023), Sadhu et al. (2022), Shayea et al. (2025)
Detection Accuracy	Often limited to known threats and signatures; prone to false negatives.	High accuracy in identifying known and unknown threats through data-driven modeling.	Allioui and Mourdi, 2023, Usmani et al. (2022), Tayyab et al. (2023)
Latency	May introduce delays due to rigid processing and lack of parallelism.	Can offer real-time threat detection and fast mitigation, especially at the edge.	Muppalaneni et al. (2024), Islam et al. (2024)
Scalability	Difficult to scale in highly distributed A-IoT environments.	Easily scalable *via* cloud-edge integration and distributed learning (e.g., federated).	Chang et al. (2023), Rao and Deebak, 2023, Hudda and Haribabu, 2025
Real-Time Response	Reactive; slow to respond to emerging attacks.	Proactive and autonomous response mechanisms using reinforcement and online learning.	Shehzadi (2024), Reddy et al. (2024), Seo et al. (2023)

mission-critical A-IoT systems, such as autonomous drone swarms, battlefield networks, or disaster recovery systems (Seo et al., 2023). AI offers significant benefits in cybersecurity. To begin, AI will enable real-time and autonomous threat identification that can be tailored to changing assault patterns without the continual participation of humans (Muppalaneni et al., 2024). Second, it enhances scalability and generalization, allowing models to be utilised on a variety of devices and protocols (Islam et al., 2024). Third, AI models can improve threat intelligence correlation by combining data from diverse sources to identify multi-vector threats (Alhakami, 2024). Furthermore, privacy-sensitive approaches such as federated learning can help with data protection rules, which is critical in sensitive industries like as healthcare and banking.

However, AI-based security is not without flaws and risks. One of the significant challenges is the vulnerability of AI models to adversarial assaults, which use malicious inputs to trick the system (Shayea et al., 2025). An attacker can utilize model biases to avoid or misclassify training data, or he can modify it. Furthermore, the majority of AI models, particularly deep learning systems, are neither interpretable nor explainable, and human operators cannot trust or confirm their judgements (Şahin et al., 2025). Data dependence is also an issue: in order to train effective models, high-quality, labelled datasets are typically required, which might be a hurdle in the event of emerging threats when labelled data is not yet accessible. Finally, the computational overhead and energy needs of AI models are realistic constraints to implementing AI models on lightweight A-IoT devices (Hudda and Haribabu, 2025).

Although artificial intelligence can provide an impressive toolkit for improving cybersecurity in autonomous systems, its implementation should be approached holistically, with automation balanced with robustness, performance with interpretability, and intelligence with ethical and privacy protection (Singh et al., 2025). The hybrid AI framework, which blends numerous learning paradigms with domain knowledge, ensures both technological efficacy and operational dependability in A-IoT cybersecurity (Alfahaid et al., 2025).

5 AI-driven cybersecurity techniques for A-IoT

The A-IoT ecosystem is being extended to include smart cities, self-driving cars, healthcare, and industrial automation, which necessitates the need to have more context-aware, intelligent, and scalable cybersecurity (Alhakami, 2024). The scale, speed, and sophistication of current cyber threats are proving to be more than what can be handled by traditional rule-based defense mechanisms. Artificial intelligence (AI) is a paradigm shift, and it allows cybersecurity systems to shift into proactive and autonomous defense instead of reactive positions (Mohamed, 2025). This part focuses on the key AI-based cybersecurity methods that are specific to A-IoT systems and their scientific basis, practical implementation, and drawbacks.

5.1 AI-enhanced intrusion detection and prevention systems (IDS/IPS)

The most critical A-IoT security applications of AI are intrusion detection and prevention. Machine learning (ML) and deep learning (DL) allow AI-based IDS/IPS to examine large volumes of real-time traffic and device behaviour to detect known and novel attack patterns (Albulayhi, 2022). To the extent that they can identify behavioural anomalies that indicate 0-day threats, AI-enhanced IDS systems can go beyond signature-based detection (Shaik and Shaik, 2024). Adaptive IPS systems surpass traditional methods by implementing real-time prevention measures, such as blocking malicious traffic or isolating compromised nodes. A-IoT is highly decentralized, and lightweight IDS models can also be trained at the edge, using federated learning to guarantee low latency and on-site threat detection (Kanzouai et al., 2025). Although such systems are very responsive and scalable, it is still difficult to adjust the detection thresholds to reduce the number of false positives and ensure model accuracy in heterogeneous environments.

5.2 Anomaly and threat detection using unsupervised learning

Unsupervised learning techniques, such as k-means clustering, autoencoders, and one-class Support Vector Machines (SVMs), are increasingly used to identify anomalous behavior in A-IoT networks without relying on labeled datasets (Kaliyaperumal et al., 2024). Such techniques are especially useful in identifying new or evasive threats in situations where normal behavior is situational. For instance, a drone deviating unexpectedly from its flight path or

altering its communication protocol may indicate a cyber-physical attack (Pavithra et al., 2023). AI models trained on contextual and temporal data can detect such anomalies in real time and initiate pre-emptive countermeasures. More advanced implementations introduce graph neural networks (GNNs) to be used to model the inter-device interactions, thus offering system-wide visibility of complex A-IoT infrastructures (Sha et al., 2025). Unsupervised techniques can, however, be less precise and need close calibration to achieve the trade-off between sensitivity and specificity.

5.3 Malware classification and behavioral analysis

AI has changed the malware detection process, especially for obfuscated or polymorphic malware that cannot be detected using traditional static analysis (Chandran et al., 2025). Malware variants can be classified with convolutional neural networks (CNNs) and recurrent neural networks (RNNs) by processing opcode sequences, API call patterns and run-time behavior (Almaleh et al., 2023). Such models are further enhanced with behavioral monitoring tools that evaluate the interaction of a process with the system resources in order to detect fileless or stealth attacks. Identification is possible with the predictive capabilities of AI, but so is prediction of the probable behavior of new malware strains, which is critical in the rapidly changing threat landscape of A-IoT (Jeffrey et al., 2023). Although they are accurate, DL-based malware classifiers are computationally demanding and need large labeled data which is not always possible in all A-IoT devices.

5.4 Intelligent authentication and access control

In A-IoT ecosystems, traditional authentication mechanisms, such as pre-shared keys and static credentials, are ill-suited for diverse, large-scale deployments (Hossain et al., 2024). With AI, dynamic, behavior-based authentication is possible based on biometric profiling, device fingerprinting, and ongoing user or device activity monitoring. Methods such as decision trees and reinforcement learning adjust access privileges on a real-time basis depending on the level of risk (Usmani et al., 2022). These systems enhance security by identifying insider threats or unauthorized access and are still usable. But these methods also present issues of user privacy and data protection in circumstances where behavioral data is centrally gathered or insufficiently anonymised.

5.5 AI-based encryption and lightweight cryptography

Resource constraints in A-IoT environments necessitate encryption techniques that are both secure and efficient. An increasing trend in the application of AI has been the use of AI to develop lightweight cryptographic algorithms that can dynamically switch key sizes, cypher strength, and mode of operation based on the degree of threat and available resources (Zafir et al., 2024). These security mechanisms allow tradeoffs between security and latency and power. AI also supports intelligent key management, enabling automatic key rotation, breach detection, and secure key distribution across mesh or *ad hoc* networks (Pothumarti et al., 2021). While promising, AI-based encryption still requires further validation against advanced cryptographic attacks and standardization for cross-vendor compatibility as shown in Table 3.

5.6 Autonomous response systems and swarm intelligence

Autonomous response is one of the most sophisticated ways of utilizing AI in cybersecurity. Such systems powered by reinforcement learning and swarm intelligence have the ability to take immediate actions without human intervention. In a particular case, it could become isolated when identifying a compromise, reorganize the network association, or deploying decoy services to confuse the attackers (Reddy et al., 2024). Swarm-based solutions add robustness to the system as nodes are able to behave in a defensive manner through coordination. To facilitate transparency and trust, explainable AI (XAI) frameworks are becoming increasingly embedded, which enables human stakeholders to have an idea of the reasoning behind the automated decisions (Mahto, 2025). Although these systems are self-regulated, they have to be controlled to prevent unintentional disturbances or self-reinforcing mistakes.

Cybersecurity approaches powered by AI provide disruptive abilities throughout the A-IoT security stack, including threat identification and categorization, authentication, and dynamic response (Mohamed, 2025). They are critical in contemporary defense systems due to their ability to learn and adapt to emerging threats, their real-time processing, and their scalability. Nevertheless, they should be deployed in harmony with resource-efficiency, ethical compliance, false-positive management, and resistance to adversarial manipulation (Albulayhi, 2022; Shaik and Shaik, 2024). There is no one single method of AI that is always best. To take the example of malware analysis, deep learning is very effective in that task, but it can be computationally infeasible at the edge, whereas federated learning allows decentralization without sacrificing privacy but has a harder time converging models on heterogeneous data. As A-IoT networks are increasingly autonomous and mission-critical, hybrid and explainable AI methods, in combination with advanced evaluation frameworks, will become necessary in order to realize resilient and trustworthy cybersecurity.

6 Emerging trends and novel AI approaches

As the danger scenario for A-IoT ecosystems grows more dynamic and sophisticated, classic AI methodologies, while foundational, are being supplemented by a new generation of advanced, explainable, and adaptable artificial intelligence paradigms (Zafir et al., 2024). These innovative ideas are not

TABLE 3 Mapping of AI techniques to A-IoT domains and cybersecurity functions.

AI technique	A-IoT Domain(s)	Cybersecurity function	Advantages	Limitations	References
Machine Learning	Smart homes, wearables, industrial IoT	Anomaly detection, malware classification	High accuracy with sufficient data	Vulnerable to data drift, adversarial input	Albulayhi (2022), Shaik and Shaik (2024)
Deep Learning	Video surveillance, smart cities	Intrusion detection, pattern recognition	Automatically extracts complex features	Requires high computation, black-box nature	Chandran et al. (2025), Jeffrey et al. (2023)
Federated Learning	Healthcare, finance, smart grid	Privacy-preserving training across devices	No raw data sharing, decentralized learning	Communication overhead, non-IID data challenges	Kanzouai et al. (2025), Tayyab et al. (2023)
Swarm Intelligence	Environmental monitoring, UAVs, logistics	Distributed threat detection, routing defense	Decentralized and adaptive	Sensitive to noisy environments	Reddy et al. (2024), Mahto (2025)
Explainable AI (XAI)	Autonomous vehicles, critical infrastructure	Transparent decision-making in detection	Improves trust and compliance	Trade-off between explainability and accuracy	Mahto (2025), Usmani et al. (2022)
Generative Adversarial Networks (GANs)	Intrusion simulation, data augmentation	Attack scenario modeling, synthetic data generation	Enhances model robustness via simulated threats	Training instability, potential misuse	Sha et al. (2025), Chandran et al. (2025)
Digital Twins	Smart manufacturing, predictive maintenance	Simulated threat response, system-level testing	Risk-free testing of cybersecurity measures	Requires accurate modeling and real-time syncing	Pavithra et al. (2023), Alfahaid et al. (2025)

only changing the way cyber defences are deployed, but they are also solving long-standing issues with transparency, flexibility, and scalability. This section looks at cutting-edge AI approaches that are helping to build a more robust and intelligent cybersecurity framework for A-IoT (Mba, 2025). Explainable Artificial Intelligence (XAI) is a significant invention that is gaining popularity in cybersecurity. Unlike traditional AI models, which frequently function as "black boxes," XAI provides interpretability and openness in decision-making processes (Chinnaraju, 2025). Explainability is critical in A-IoT cybersecurity for verifying security warnings, explaining automated mitigation measures, and adhering to regulatory frameworks like General Data Protection Regulation (GDPR) and National Institute of Standards and Technology (NIST). For example, XAI-integrated intrusion detection systems can identify abnormalities and explain which parameters (e.g., packet frequency, device location, and protocol behaviour) influenced the detection decision (Javed et al., 2023). This transparency builds confidence, makes human-machine collaboration easier, and allows security analysts to better understand, audit, and modify AI models over time (Van Hoang, 2023). Generative AI, particularly Generative Adversarial Networks (GANs) and diffusion models, is proving to be an effective tool for threat simulation and defensive strategy development. GANs may be used to model realistic adversarial attack patterns that resemble 0-day vulnerabilities or polymorphic malware, allowing defensive systems to be educated in a more diversified and realistic threat environment (Peppes et al., 2023). This proactive exposure greatly improves model generalisation and robustness. Furthermore, generative models may be used to synthesise attack data in situations when real-world datasets are limited, allowing for the creation of more effective threat classifiers and automatic red-teaming frameworks for penetration testing in A-IoT systems (Ali and Ghanem, 2025).

Transfer Learning and Meta-Learning are gaining traction as threat vectors continue to evolve, particularly those influencing previously undetected device behaviours or settings (Sha et al., 2025). Transfer learning allows pre-trained models to be rapidly fine-tuned on tiny, domain-specific datasets, resulting in greatly reduced training time and resource needs for edge-based security applications. This is especially useful in A-IoT systems, when labelled data is scarce or scattered. Meta-learning, often known as "learning to learn," goes one step further by creating models that can swiftly adapt to new sorts of assaults while exposing as little data as possible (Fadhilla et al., 2022). These skills are crucial for dealing with rapidly changing malware, adaptive adversaries, and dynamic device behaviours in diverse A-IoT contexts.

Digital Twins, a unique cybersecurity paradigm, are being used to bridge the physical and digital domains of A-IoT (Salim et al., 2024). A digital twin is a real-time virtual counterpart of a physical object or system that enables predictive analytics, anomaly detection, and cyber-physical simulations. When combined with AI, digital twins can mimic the effects of hypothetical cyber assaults on A-IoT infrastructures like self-driving cars or smart manufacturing facilities. This not only improves threat prediction, but it also enables scenario-based training, resilience testing, and proactive risk management (Hossain et al., 2024). The dual-loop interaction of the physical world and its digital doppelganger enables cyber defenders to dynamically monitor system integrity and optimize security postures (Singh et al., 2025).

Furthermore, Federated and Distributed Learning approaches are revolutionising the deployment of edge AI. Traditional AI training necessitates centralised data aggregation, which creates privacy concerns and scalability challenges. Federated Learning (FL) tackles this issue by training models locally on IoT devices and selectively sharing model updates, protecting data privacy and lowering connection cost (Shayea et al., 2025). In A-IoT systems, this decentralised intelligence enables scalable, privacy-preserving, real-time threat detection over a wide network of edge devices. Distributed AI models, when reinforced with blockchain or consensus methods, can enable trustless collaboration across

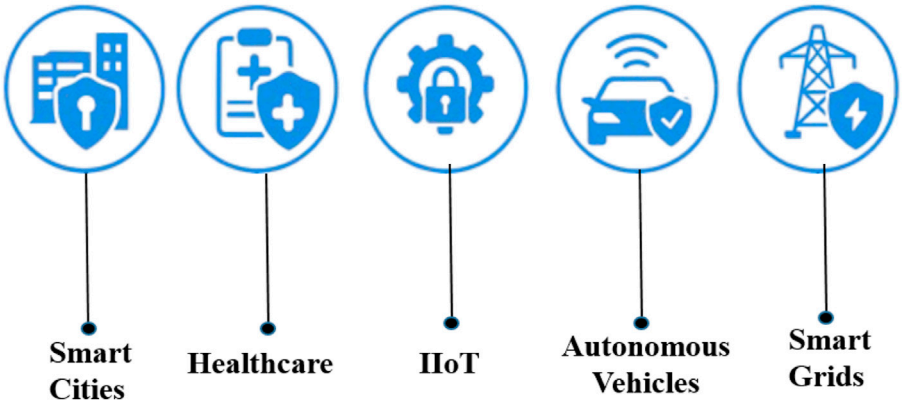


FIGURE 3
Use case mapping of AI-driven cybersecurity across domains.

devices in hostile contexts, reducing the chance of single-point failures and improving system resilience against distributed denial-of-service (DDoS) assaults and insider threats (Albshaier et al., 2024).

In conclusion, the integration of innovative AI technologies such as XAI, generative models, transfer/meta-learning, digital twins, and federated intelligence represents a paradigm change in A-IoT cybersecurity (Chinnaraju, 2025). These solutions not only address the increasing sophistication of cyber threats, but they also take into account the operational limits and ethical issues that autonomous, resource-constrained IoT systems provide. Together, they open the way for the development of transparent, adaptable, and scalable cybersecurity systems that can learn, evolve, and defend in real time, ushering in a new age of intelligent security for the autonomous digital frontier (Oliha et al., 2024).

7 Case studies and real-world applications

Cybersecurity and AI are not just theoretical concepts in A-IoT ecosystems but are actively being used to shape critical infrastructure in such areas as smart cities, autonomous mobility, healthcare, industry, and energy as shown in Figure 3. Such practical applications show the power of autonomous systems and at the same time reveal substantial cybersecurity risks (Vermesan et al., 2022).

The section presents the leading examples of case studies in which AI-based cybersecurity methods are implemented to secure A-IoT systems against emerging cyber threats (Allioui and Mourdi, 2023). Smart cities A-IoT technologies are part of intelligent traffic systems, surveillance networks, connected infrastructure, and environmental monitoring. Nonetheless, the complexity and interconnectedness make these systems have many attack vectors (Kanellopoulos et al., 2023). As an example, video surveillance systems based on AI and able to recognize faces and analyze behavior can be targeted by adversarial input attacks, which can interfere with identity verification procedures (Albshaier et al., 2024; Vermesan et al., 2022).

Cities are implementing federated learning-based intrusion detection systems that can process data in a more localized manner at edge nodes to avoid centralizing sensitive data, ensuring privacy and scalability (Hamid and Bawany, 2024). At the same time, researchers are investigating swarm intelligence approaches to conduct distributed anomaly detection across geographically dispersed nodes within urban infrastructure, enabling a coordinated response to threats in real-time across large-scale environments (Chinnaraju, 2025).

Another vital use of A-IoT is autonomous vehicles (AVs) and drones, particularly in mission- and adversarial-critical settings. These systems are based on AI navigation, object recognition, and decision-making but are susceptible to cyber-physical attacks, including Global Positioning System (GPS) spoofing, Light Detection and Ranging (LIDAR) manipulation, and adversarial attacks on the AI model through malicious road signs (Pavithra et al., 2023).

A prominent real-life application can be seen in an AI-based security system in automobiles, which uses anomaly detection algorithms to detect Controller Area Network (CAN) bus operations and command injections that are not authorized. Deep Reinforcement Learning (DRL) is applied to optimize flight routes in threat scenarios and adapt to the communication protocols depending on the aerial environment in drone ecosystems (Sarikaya and Bahtiyar, 2024). Due to the continued development of AVs towards full autonomy, digital twin simulation and threat intelligence platforms are becoming more critical in their ability to simulate and mitigate multimodal and complex cyber threats (Allioui and Mourdi, 2023).

In the sphere of Industrial IoT (IIoT), AI-empowered cybersecurity is a critical factor in the security of automated production lines, robotics, and supply chain networks. Most IIoT infrastructures have yet to transition to modern systems with robust in-built security, which means they are vulnerable to ransomware, insider attacks, and 0-day attacks (Fadhilla et al., 2022). To counter this, organizations are implementing AI-enhanced Security Information and Event Management (SIEM) systems that use machine learning to identify anomalous patterns of behavior on Industrial Control Systems (ICS) and Supervisory Control and Data Acquisition (SCADA) systems.

Remarkably, generative AI has been applied to modeling Advanced Persistent Threats (APTs) so that security teams can predictively model and test industrial security measures beforehand (Zhuwankinyu et al., 2024). Moreover, edge AI is being implemented on factory floors to monitor in real-time and automatically mitigate threats, which means that the latency of the attack detection and response is lower by far.

The A-IoT devices used in healthcare, like AI-equipped wearables, pose a significant cybersecurity threat because of the sensitivity of personal data that they process and due to the constant connection to the network. Such devices can track vital signs and send the information to cloud-based diagnostic systems, which makes them the most viable targets of data manipulation and privacy violation (Putra et al., 2024). To illustrate, manipulation of AI models may lead to incorrect diagnoses or treatment delays. Medical practitioners are adopting federated learning systems to overcome these risks by training models in collaboration across hospitals and preserving patient privacy. Another related trend is the use of explainable AI (XAI), which aims to increase transparency in diagnostic algorithms and promote trust and responsibility in clinical decision-making (Hamid and Bawany, 2024; Pavithra et al., 2023).

Energy and smart grid systems represent another critical A-IoT application space, and the national security implications are enormous. Such systems use AI to predict energy loads, identify faults, and automatically control energy distribution but are also becoming the subject of attack by adversaries aiming to deny energy continuity or tamper with usage data. The use of AI-based anomaly detection is common for tracking consumption trends from millions of smart meters, and the application of distributed AI with blockchain technology ensures tamper-resistance and data provenance (Jayavarma et al., 2025). European and North American case studies demonstrate how deep learning is being successfully used to identify and counter cyber threats in wind farms, solar installations, and power substations. Furthermore, responsive AI-based demand-response algorithms can enhance cyberattack resilience and maintain grid stability (Deshpande, 2024).

To sum up, the days when cybersecurity by AI was a speculative topic are over, and it is a realistic requirement in various industries. The variety of real-world applications of federated learning, digital twins, DRL, XAI, and generative AI shows how essential context-aware, scalable, and interpreted cybersecurity solutions are. The cases highlight the need to urgently develop adaptive frameworks that can address the dynamic threat environments in autonomous, decentralized, and data-intensive environments.

8 Evaluation metrics and benchmarking

In order to guarantee the efficiency of the AI-based cybersecurity approaches specialized in A-IoT environments, which are resource-limited and face dynamic and complicated threat landscapes, a rigorous, multidimensional analysis framework has to be used (Mahto, 2025). This section talks about the fundamental measures and benchmarking parameters that define the performance, efficiency, resiliency and deployability of AI-based security solutions in A-IoT environments.

8.1 Detection performance and classification accuracy

The foundation of any cybersecurity evaluation lies in the assessment of detection capabilities. Metrics such as accuracy, precision, recall, and F1-score are crucial in evaluating how effectively an AI model distinguishes malicious activities from benign ones (Sharma et al., 2024). While detection accuracy offers a general sense of correctness across predictions, it can be misleading when datasets are imbalanced, as is typical in cybersecurity, because it may overestimate model effectiveness when benign instances dominate (Hamid and Bawany, 2024). Therefore, precision and recall become more critical: precision measures the proportion of true positives among all detected positives, indicating the model's reliability in reducing false alarms, while recall assesses the proportion of actual malicious activities that the model successfully detects, representing its completeness (Bold et al., 2022). The F1-score harmonizes these two, balancing the need for accuracy and completeness, and is especially valuable in skewed datasets, as summarized in Table 4.

8.2 Latency and real-time responsiveness

Latency is also important, considering that A-IoT systems are in real-time when it comes to detecting and mitigating threats. The detection accuracy of AI-based defences should not be the only criterion of their effectiveness, but also their response time, especially when they have to operate under real-time conditions where such delays may result in catastrophic outcomes (Sharma et al., 2024). The overall detection latency (data acquisition, feature extraction, model inference, and mitigation response) ought to be measured accurately (Jayavarma et al., 2025). Low-latency and low-computational-overhead systems are more appropriate to be deployed in edge environments, where there is limited resource availability. In addition, benchmarking must involve the simulated real- or near-real-time attack scenarios to maintain operational continuity and flexibility (Allioui and Mourdi, 2023).

8.3 Resource and energy efficiency

Resource usage is another crucial assessment criterion, particularly given that a large portion of A-IoT devices runs on batteries and has limited computational capabilities. Energy efficiency metrics measure power used in training and inference processes, which directly determines the longevity of the device and the sustainability of the whole system (Bai et al., 2024). At the same time, the computational efficiency is quantified regarding the memory consumption, processor load, and bandwidth consumption. These constraints should be benchmarked against lightweight AI models, which are trained using model pruning, quantization, or edge-friendly architectures. The energy/resource efficiency/detection performance trade-off needs to be thoroughly examined to achieve a balance between them in applied cases (Desislavov et al., 2023; Zhukabayeva et al., 2025).

TABLE 4 Evaluation metrics for AI-Driven cybersecurity in A-IoT.

Metric	Definition	Relevance	Use case example	References
Accuracy	Percentage of correct predictions among total samples	Basic measure of detection performance	Intrusion detection system classification	Sharma et al. (2024), Hamid and Bawany (2024)
Precision/Recall/ F1-Score	Precision: ratio of true positives to all predicted positives. Recall: ratio of true positives to all actual positives. F1: harmonic mean of the two.	Useful in class-imbalanced datasets to minimize false positives or false negatives	Malware detection in smart grids	Bold et al. (2022), Sharma et al. (2024)
False Positive Rate (FPR)	Proportion of benign actions incorrectly labeled as malicious	High FPR leads to alert fatigue, reduces trust in system	IDS in smart homes	Sharma et al. (2024)
Latency	Time between attack onset and system response	Critical for real-time threat detection in mission-critical applications	Vehicle-to-everything (V2X) communication	Allioui and Mourdi (2023), Putra et al. (2024)
Energy Efficiency	Power consumed per prediction or detection cycle	Vital for battery-operated and resource-constrained A-IoT devices	Wearable health monitors	Zhukabayeva et al. (2025), Bai et al. (2024)
Model Interpretability	Degree to which decision-making logic can be understood	Important for debugging, transparency, and regulatory compliance	Explainable AI in smart factories	Desislavov et al. (2023), Akash (2025)
Robustness	Resilience of AI models to adversarial inputs and concept drift	Validates reliability under attack or changing conditions	GAN-generated spoofing in security cameras	Sharma et al. (2024), Allioui and Mourdi (2023), Shyaa et al. (2024)
Scalability	Ability to maintain performance with increasing devices or data	Key for deployment across large-scale and heterogeneous A-IoT environments	Smart city cybersecurity framework	Allioui and Mourdi (2023), Hazra et al. (2021)
Deployment Readiness	Readiness for integration with current IoT protocols and regulatory standards	Determines real-world applicability of AI-based cybersecurity frameworks	Federated model deployment across healthcare IoT	Akash, 2025; Hazra et al. (2021)

8.4 Robustness against adversarial attacks

With the growing number of threat actors taking advantage of the weaknesses of the AI systems, the resistance to adversarial attacks has become one of the essential metrics. This means testing the robustness of AI models in the presence of well-designed malicious inputs that aim to deceive or detect (Sharma et al., 2024). Robustness testing involves the creation of adversarial examples by following different attack strategies, e.g., evasion or poisoning attacks, and quantifying the loss in detection performance. A suitable model must achieve high accuracy and recall when faced with adversarial perturbations, particularly in security-focused A-IoT infrastructures (Allioui and Mourdi, 2023). The evaluation must also take concept drift into account, the fact that changing patterns of attack require constant learning and model flexibility (Shyaa et al., 2024).

8.5 Scalability and deployment readiness

Finally, we cannot ignore the metrics of operational scalability and deployment readiness in AI-based defences. Scalability is the degree to which AI models can be used under conditions of increased connected devices, the volume of data streams, or network complexity (Allioui and Mourdi, 2023). Alternatively, deployment readiness assesses the simplicity of deployment with available IoT infrastructures, compatibility with existing protocols, and adherence to security and regulatory requirements (Hazra et al., 2021). This involves the capacity of the system to accommodate decentralized training methods like federated learning and the simplicity of updating models after deployment. Not only is a solution of high deployment readiness theoretically sound, but it is also scalable in practice (Akash, 2025).

In summary, the overall benchmarking scheme of assessing AI-based cybersecurity in A-IoT systems should combine the traditional measures of detection with the more recent factors like latency, energy efficiency, robustness, and scalability. This type of approach is holistic and makes the proposed solutions not only correct and secure but also efficient, resilient, and deployable in real-world autonomous IoT systems.

9 Challenges, gaps, and open issues

Despite the development of AI-based cybersecurity solutions for A-IoT ecosystems, there remain many challenges and unaddressed problems in securely and efficiently deploying them (Singh et al., 2024). Addressing these gaps is crucial for building strong, reliable, and scalable security models that meet the specific needs of autonomous, distributed IoT networks.

One of the basic issues is the quality of data, its availability, and labeling of the data on which AI models are trained. More specifically, supervised and deep learning methods are based on large amounts of quality-labelled data that depict both benign and malicious actions (Hossain S. et al., 2025). Nevertheless, the heterogeneity of data sources, the existence of proprietary communication protocols, and the stringent privacy policies in A-IoT settings are major impediments to the complete data gathering process, which results in biased, incomplete, and even fragmented datasets (Qudus, 2025). In addition, labelled attack data,

particularly 0-day or very sophisticated attacks, are limited, which hinders model training and validation (Hirsi et al., 2025). Such data constraints highlight the necessity of creating powerful unsupervised or semi-supervised learning algorithms and the creation of synthetic datasets, such as through the use of generative adversarial networks (GANs) (Peppes et al., 2023).

Tightly connected with data issues is the problem of AI model explainability and reliability. Most existing AI systems, especially deep neural network-based systems, are opaque black boxes that make it unintelligible to human experts how they are making decisions. Such a lack of transparency creates substantial adoption barriers in cybersecurity, where automated decisions may have operational and safety-relevant consequences (Hassija et al., 2024). Explainable AI (XAI) approaches provide potential solutions to explain model predictions and improve user confidence, but these approaches are still in their early stages in real-time A-IoT security applications, and they tend to have extreme computational costs. Therefore, future research aims to come up with reliable AI systems that are both complex and interpretable and formulate uniform metrics to measure explainability, which is crucial to its universal acceptance (Chander et al., 2025).

The resulting threat of adversarial machine learning compounds the security picture. Malicious actors are becoming increasingly skilled at identifying weaknesses in AI models and creating adversarial inputs that can negatively impact the performance of classifiers or reduce the effectiveness of detectors, potentially nullifying cybersecurity protections. In addition, poisoning attacks, which introduce contamination into training data, are a long-term threat to model integrity (Tian et al., 2022). To combat these new threats, there is a dire need to come up with a strong method of hardening the models, continuous monitoring, and dynamic retraining. Nevertheless, these approaches have not been studied or tested at scale with large amounts of heterogeneous A-IoT. Furthermore, there is a major deficiency in the creation of systematic structures to track and counter adversarial threats without compromising the speed and accuracy of detection (McCarthy et al., 2022).

The second urgent problem is the trade-off between security and performance of the system, particularly in A-IoT devices with limited resources. Increased security is generally associated with an increased computational burden, memory consumption, and communication overhead, which may have a negative impact on battery life, latency, and the experience (Bai et al., 2024). On the other hand, efficiency could make systems susceptible to attacks or limit the complexity of the threat detection algorithms (Shyaa et al., 2024). The process of finding the optimal balance between them requires security models that can dynamically adapt the level of protection to the level of threat, the state of devices, and the priorities of the work (Mallick and Nath, 2024). Nevertheless, in-depth frameworks that combine such trade-offs into a variety of use cases are limited, which points out a divide between theory and practice (Qudus, 2025).

In addition to technical issues, legal and ethical and standardization challenges are major impediments to the mainstreaming of AI-powered cybersecurity in A-IoT systems. The legal frameworks governing data privacy, data security responsibility, and cross-border data transfer vary significantly across different jurisdictions, and this makes the compliance of globally distributed IoT devices challenging (Allioui and Mourdi, 2023). Other ethical issues, including algorithm bias, responsibility in AI decision-making, and user permission, are also impediments to responsible AI in critical infrastructure. Additionally, there are no universal AI model validation, cybersecurity, and interoperability standards among heterogeneous IoT devices, which do not allow building coherent defense strategies (Qudus, 2025). Collaboration between policymakers, industry stakeholders, and academia is necessary to create comprehensive regulations and standards that promote innovation and protect societal values (Agrawalla and Banerjee, 2025).

In summary, while AI-driven cybersecurity holds transformative potential for A-IoT ecosystems, it confronts significant challenges related to data management, model explainability, adversarial robustness, performance-security trade-offs, and governance. Addressing these open issues will require multidisciplinary research, cross-sector cooperation, and ethically grounded innovation to realize autonomous IoT networks that are secure, resilient, and socially responsible.

10 Future directions and research opportunities

The ever-changing nature of A-IoT ecosystems, combined with the ongoing sophistication of cyber threats, necessitates future-oriented research and new frameworks to advance AI-based cybersecurity capabilities, close existing gaps, and make autonomous networks truly resilient (Kumar et al., 2025). This section discusses the important future research paths and prospective research possibilities that will revolutionise the security paradigm of A-IoT systems, with an emphasis on the use of emerging technologies, new AI techniques, and adaptive frameworks.

One of the most promising directions is the development of hybrid AI-blockchain security systems that can leverage the synergistic potential of artificial intelligence and decentralised ledger technologies to address the underlying issues of trust, data integrity, and secure device authentication in A-IoT networks (Bhumichai et al., 2024). Blockchain's immutable, distributed, and decentralized nature can serve as a solid foundation for secure data sharing and provenance, lowering the risk of data tampering and unauthorised access, while AI can be used to improve anomaly detection and adaptive threat response through intelligent analytics (McCarthy et al., 2022). Future work should focus on developing lightweight, scalable blockchain implementations that fit within the resource constraints of IoT devices and are readily integrated with AI-based intrusion detection and trust management systems. It will be critical to investigate consensus techniques that are optimised for real-time security operations and evaluate their impact on latency and energy usage (Al-awamy et al., 2025). The next disruptive frontier is the development of self-healing security systems capable of monitoring, analysing, and resolving cyber threats in real time without requiring human interaction (Allioui and Mourdi, 2023). These systems would employ high-level AI algorithms to continually monitor network health, detect vulnerabilities or assaults, and automatically take defensive or recovery actions, therefore decreasing downtime and

operating risks in mission-critical A-IoT applications (Johnphill et al., 2023). The study should focus on reinforcement learning and meta-learning approaches that allow these systems to dynamically adapt to the changing threat landscape and system configurations. The challenge is to strike a balance between autonomy and control, so that self-healing operations do not unintentionally interfere with lawful activities, and compliance and confidence may be gained through openness and auditability (Tyagi and Seranmadevi, 2024). Another potential research area is the development of AI-powered cyber threat intelligence systems, which use big data, natural language processing, and predictive analytics to collect, process, and exchange actionable threat intelligence in diverse A-IoT contexts (Fuentes-Peñailillo et al., 2024). Early warning systems may be installed on these platforms to correlate data from a range of sources, including device logs, network traffic, and external threat feeds, enabling for proactive defence actions (Aminu et al., 2024). In the future, it is critical to focus on the use of federated learning to retain privacy and security in collaborative intelligence sharing, as well as the development of real-time inference models that can scale with the rising amount and speed of threat data. Explaining and human-in-the-loop strategies will be researched to improve the use and credibility of such platforms among security operators (Azeri et al., 2024).

The nature of risks to A-IoT systems necessitates study into cross-domain AI adaptability, such as multi-modal threat detection systems, which incorporate data from several sources, such as network signals, sensor readings, audio-visual inputs, and user behaviour analytics (Fuentes-Peñailillo et al., 2024). This complete technique enables deeper contextual awareness and improved detection, particularly against modern multi-vector threats that evade unimodal systems (Alhakami, 2024). Research questions include how to create unified feature representations, scalable fusion systems, and adaptive learning algorithms that can transfer knowledge across domains and modalities. This type of cross-domain information is required for total situational awareness and strong security postures in autonomous systems that operate in complex and dynamic contexts (Zou et al., 2025).

Finally, the ideal of fully autonomous cybersecurity agents capable of operating autonomously on scattered A-IoT networks summarises future research objectives (Tyagi and Seranmadevi, 2024). These bots would use advanced AI characteristics like continuous learning, reasoning, decision-making, and teamwork to automatically detect, forecast, and neutralise emerging cyber threats on a large scale. This ambitious goal necessitates advancements in multi-agent systems, trust management, ethical AI, and durable real-time communication protocols (Chaffer et al., 2024). The study must address concerns such as coordination among autonomous agents, dispute resolution, policy enforcement across the system, and robustness against targeted attacks on the agents themselves (Huang et al., 2025). Furthermore, when these agents assume critical security tasks, ethical considerations must be addressed and aligned with human supervision systems to assure control and responsibility (Vaseashta, 2022).

To summarize, the next-generation of AI-based cybersecurity in Autonomous IoT ecosystems will rely on synergistic technologies and adaptive intelligence paradigms that go beyond detection and provide proactive, self-sustaining defence and intelligence capabilities. The mix of multidisciplinary effort, spanning AI, blockchain, network security, and systems engineering, and real-world validation through real-world A-IoT deployments will be critical in shaping the next-generation of robust, autonomous cyber defence systems capable of protecting the increasingly linked globe.

11 Conclusion

This review has illuminated the critical convergence of artificial intelligence and cybersecurity within the rapidly growing A-IoT ecosystem, underscoring both its transformative potential and the complex challenges involved in securing such large-scale, heterogeneous, and dynamic networks. We explored the evolution of IoT towards autonomy, demonstrating how increased device intelligence and interconnectivity significantly expand the attack surface and intensify security concerns, thereby necessitating novel, adaptive defense mechanisms. Our analysis revealed the limitations of traditional security paradigms in addressing the scale, diversity, and real-time demands of A-IoT systems and established AI-driven cybersecurity as a pivotal paradigm shift that enables proactive, context-aware, and self-adaptive protection.

In a comprehensive study of AI methods such as machine learning, deep learning, federated learning, and swarm intelligence, we were able to isolate their distinctive advantages in intrusion detection, anomaly recognition, malware classification, and orchestration of autonomous responses, and also note enduring challenges of explainability, limited data availability, and susceptibility to adversarial attacks. We also noted emerging innovations of explainable AI, generative adversarial models to simulate threats, transfer learning, and digital twins, which have the potential to improve the transparency, robustness, and simulation quality of cyber defense. The practical use of AI-powered security solutions is evident in smart cities, autonomous cars, industrial IoT, healthcare, and smart grids, and this is an indication of the wide applicability and practicality of AI-enhanced security solutions.

Important gaps still exist in data quality, model reliability, adversarial robustness, and ethical and regulatory frameworks, highlighting the necessity for multidisciplinary collaboration and responsible development of AI. In perspective, some of the potentially fruitful research avenues can be found in hybrid AI-blockchain systems, self-healing security systems, AI-based threat intelligence systems, cross-domain multi-modal detection, and fully autonomous cybersecurity agents. These are likely to propel the next-generation of innovation in securing autonomous IoT ecosystems.

In summary, the future of AI-powered cybersecurity for autonomous IoT depends on the seamless integration of adaptive intelligence with decentralized trust architectures, resulting in autonomous, resilient, and transparent defenses capable of safeguarding increasingly complex and mission-critical cyber-physical systems. To attain such a vision, long-term research, standardization, and ethical governance will be needed to make sure that the integration of AI and IoT yields secure, trustworthy, and sustainable autonomous networks that will support the digital future.

Author contributions

FO: Writing – original draft, Conceptualization, Writing – review and editing, Methodology. CU: Writing – review and editing, Writing – original draft, Conceptualization. O-CU: Supervision, Writing – review and editing, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Abduljawwad, M., Khaleel, M., Ogedengbe, T. S., and Abraheem, S. (2023). Sensors for daily utilization. *Int. J. Electr. Eng. Sustain.* 106–119. Available online at: https://ijees.org/index.php/ijees/article/view/53

Agrawalla, B., and Banerjee, S. (2025). Academia and industry: an essential alliance for a sustainable future. *Innovation Chem. Mater. Sustain.* 2 (1), 1–4. doi:10.63654/icms.2025.02001

Ahmad, H., Gulzar, M. M., Aziz, S., Habib, S., and Ahmed, I. (2024). AI-based anomaly identification techniques for vehicles communication protocol systems: comprehensive investigation, research opportunities and challenges. *Internet Things* 27, 101245. doi:10.1016/j.iot.2024.101245

Akash, T. R. (2025). Sustainability in business security: leveraging analytics for cyber risk mitigation. *J. Bus. Manag. Stud.* 7 (1), 126–139. doi:10.32996/jbms.2023.5.5.24

Al-awamy, A. A., Al-shaibany, N., Sikora, A., and Welte, D. (2025). Hybrid consensus mechanisms in blockchain: a comprehensive review. *Int. J. Intelligent Syst.* 2025 (1), 5821997. doi:10.1155/int/5821997

Alaba, F. A. (2024). "The evolution of the IoT," in *Internet of things: a case study in Africa* (Cham: Springer Nature Switzerland), 1–18.

Alam, T. (2024). Data privacy and security in autonomous connected vehicles in smart city environment. *Big Data Cognitive Comput.* 8 (9), 95. doi:10.3390/bdcc8090095

Albshaier, L., Budokhi, A., and Ahmed, A. (2024). A review of security issues when integrating iot with cloud computing and blockchain. *IEEE Access* 12, 109560–109595. doi:10.1109/access.2024.3435845

Albulayhi, K. (2022). An adaptive deep-ensemble anomaly-based intrusion detection system-of-systems for the internet-of-things. (Doctoral dissertation, Moscow, ID, USA: University of Idaho).

Alfahaid, A., Alalwany, E., Almars, A. M., Alharbi, F., Atlam, E., and Mahgoub, I. (2025). Machine learning-based security solutions for IoT networks: a comprehensive survey. *Sensors* 25 (11), 3341. doi:10.3390/s25113341

Alhakami, W. (2024). Evaluating modern intrusion detection methods in the face of gen V multi-vector attacks with fuzzy AHP-TOPSIS. *PLos One* 19 (5), e0302559. doi:10.1371/journal.pone.0302559

Ali, A., and Ghanem, M. C. (2025). Beyond detection: large language models and next-generation cybersecurity. *SHIFRA* 2025, 81–97. doi:10.70470/shifra/2025/005

Allioui, H., and Mourdi, Y. (2023). Exploring the full potentials of IoT for better financial growth and stability: a comprehensive survey. *Sensors* 23 (19), 8015. doi:10.3390/s23198015

Almaleh, A., Almushabb, R., and Ogran, R. (2023). Malware API calls detection using hybrid logistic regression and RNN model. *Appl. Sci.* 13 (9), 5439. doi:10.3390/app13095439

AlMarshoud, M., Sabir Kiraz, M., and H. Al-Bayatti, A. (2024). Security, privacy, and decentralized trust management in VANETs: a review of current research and future directions. *ACM Comput. Surv.* 56 (10), 1–39. doi:10.1145/3656166

Aminu, M., Akinsanya, A., Dako, D. A., and Oyedokun, O. (2024). Enhancing cyber threat detection through real-time threat intelligence and adaptive defense mechanisms. *Int. J. Comput. Appl. Technol. Res.* 13 (8), 11–27. doi:10.7753/IJCATR1308.1002

Anjos, J. C. D., Matteussi, K. J., Orlandi, F. C., Barbosa, J. L., Silva, J. S., Bittencourt, L. F., et al. (2023). A survey on collaborative learning for intelligent autonomous systems. *ACM Comput. Surv.* 56 (4), 1–37. doi:10.1145/3625544

Arulmurugan, L., Thakur, S., Dayana, R., Thenappan, S., Nagesh, B., and Sri, R. K. (2024). "Advancing security: exploring AI-driven data encryption solutions for wireless sensor networks," In 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). Piscataway, NJ, USA: IEEE. 1–6.

Asadi, M., Jamali, M. A. J., Heidari, A., and Navimipour, N. J. (2024). Botnets unveiled: a comprehensive survey on evolving threats and defense strategies. *Trans. Emerg. Telecommun. Technol.* 35 (11), e5056. doi:10.1002/ett.5056

Asiri, M., Saxena, N., Gjomemo, R., and Burnap, P. (2023). Understanding indicators of compromise against cyber-attacks in industrial control systems: a security perspective. *ACM Trans. Cyber-Physical Syst.* 7 (2), 1–33. doi:10.1145/3587255

Azeri, N., Hioual, O., and Hioual, O. (2024). A distributed intelligence framework for enhancing resilience and data privacy in dynamic cyber-physical systems. *Clust. Comput.* 27 (5), 6289–6304. doi:10.1007/s10586-024-04349-y

Bablu, T. A., and Rashid, M. T. (2025). Edge computing and its impact on real-time data processing for IoT-driven applications. *J. Adv. Comput. Syst.* 5 (1), 26–43. doi:10.69987/JACS.2025.50103

Bai, G., Chai, Z., Ling, C., Wang, S., Lu, J., Zhang, N., et al. (2024). Beyond efficiency: a systematic survey of resource-efficient large language models. *arXiv Prepr. arXiv: 2401.00625*. doi:10.48550/arXiv.2401.00625

Bhumichai, D., Smiliotopoulos, C., Benton, R., Kambourakis, G., and Damopoulos, D. (2024). The convergence of artificial intelligence and blockchain: the state of play and the road ahead. *Information* 15 (5), 268. doi:10.3390/info15050268

Bold, R., Al-Khateeb, H., and Ersotelos, N. (2022). Reducing false negatives in ransomware detection: a critical evaluation of machine learning algorithms. *Appl. Sci.* 12 (24), 12941. doi:10.3390/app122412941

Bouzidi, M., Gupta, N., Cheikh, F. A., Shalaginov, A., and Derawi, M. (2022). A novel architectural framework on IoT ecosystem, security aspects and mechanisms: a comprehensive survey. *IEEE Access* 10, 101362–101384. doi:10.1109/access.2022.3207472

Camerota, C. (2025). Machine learning application to IoT and IIoT security and reliability.

Cardoso, P., Moura, J., and Marinheiro, R. N. (2023). Elastic provisioning of network and computing resources at the edge for IoT services. *Sensors* 23 (5), 2762. doi:10.3390/s23052762

Catuogno, L., and Galdi, C. (2023). Secure firmware update: challenges and solutions. *Cryptography* 7 (2), 30. doi:10.3390/cryptography7020030

Chaffer, T. J., Goldston, J., Okusanya, B., and A I, G. D. (2024). On the ETHOS of AI agents: an ethical technology and holistic oversight system. *arXiv Prepr. arXiv: 2412.17114*. doi:10.48550/arXiv.2412.17114

Chander, B., John, C., Warrier, L., and Gopalakrishnan, K. (2025). Toward trustworthy artificial intelligence (TAI) in the context of explainability and robustness. *ACM Comput. Surv.* 57 (6), 1–49. doi:10.1145/3675392

Chandran, S., Syam, S. R., Sankaran, S., Pandey, T., and Achuthan, K. (2025). From static to ai-driven detection: a comprehensive review of obfuscated malware techniques. *IEEE Access* 13, 74335–74358. doi:10.1109/access.2025.3550781

Chang, Y., Zhang, K., Gong, J., and Qian, H. (2023). Privacy-preserving federated learning *via* functional encryption, revisited. *IEEE Trans. Inf. Forensics Secur.* 18, 1855–1869. doi:10.1109/tifs.2023.3255171

Chataut, R., Phoummalayvane, A., and Akl, R. (2023). Unleashing the power of IoT: a comprehensive review of IoT applications and future prospects in healthcare, agriculture, smart homes, smart cities, and industry 4.0. *Sensors* 23 (16), 7194. doi:10.3390/s23167194

Chen, S., Guang, Y., and Han, Q. (2024). "A review of backdoor control techniques for embedded devices," in *Proceeding of the 2024 6th international conference on information technology and computer communications*, 1–7.

Chinnaraju, A. (2025). Explainable AI (XAI) for trustworthy and transparent decision-making: a theoretical framework for AI interpretability. *World J. Adv. Eng. Technol. Sci.* 14 (3), 170–207. doi:10.30574/wjaets.2025.14.3.0106

Choudhary, A. (2024). Internet of things: a comprehensive overview, architectures, applications, simulation tools, challenges and future directions. *Discov. Internet Things* 4 (1), 31. doi:10.1007/s43926-024-00084-3

Commey, D., Mai, B., Hounsinou, S. G., and Crosby, G. V. (2024). Securing blockchain-based IoT systems: a review. *IEEE Access* 12, 98856–98881. doi:10.1109/access.2024.3428490

Deshpande, V. (2024). Smart grids integration with AI-Powered demand response. *Res. J. Comput. Syst. Eng.* 5 (1), 45–58. Available online at: https://technicaljournals.org/RJCSE/index.php/journal/article/view/94

Desislavov, R., Martínez-Plumed, F., and Hernández-Orallo, J. (2023). Trends in AI inference energy consumption: beyond the performance-vs-parameter laws of deep learning. *Sustain. Comput. Inf. Syst.* 38, 100857. doi:10.1016/j.suscom.2023.100857

Ding, S., Ward, H., Cucurachi, S., and Tukker, A. (2023). Revealing the hidden potentials of internet of things (IoT)-An integrated approach using agent-based modelling and system dynamics to assess sustainable supply chain performance. *J. Clean. Prod.* 421, 138558. doi:10.1016/j.jclepro.2023.138558

Fadhilla, C. A., Alfikri, M. D., and Kaliski, R. (2022). Lightweight meta-learning BotNet attack detection. *IEEE Internet Things J.* 10 (10), 8455–8466. doi:10.1109/jiot.2022.3229463

Fuentes-Peñailillo, F., Gutter, K., Vega, R., and Silva, G. C. (2024). New generation sustainable technologies for soilless vegetable production. *Horticulturae* 10 (1), 49. doi:10.3390/horticulturae10010049

Ghiurău, D., and Popescu, D. E. (2024). Distinguishing reality from AI: approaches for detecting synthetic content. *Computers* 14 (1), 1. doi:10.3390/computers14010001

Goudarzi, A., Ghayoor, F., Waseem, M., Fahad, S., and Traore, I. (2022). A survey on IoT-enabled smart grids: emerging, applications, challenges, and outlook. *Energies* 15 (19), 6984. doi:10.3390/en15196984

Guesmi, A., Hanif, M. A., Ouni, B., and Shafique, M. (2023). Physical adversarial attacks for camera-based smart systems: current trends, categorization, applications, research challenges, and future outlook. *IEEE Access* 11, 109617–109668. doi:10.1109/access.2023.3321118

Hamid, S., and Bawany, N. Z. (2024). "Federated learning for enhanced intrusion detection in smart city environments," in *2024 18th international conference on open source systems and technologies (ICOSST)* (IEEE), 1–6.

Hassan, A., Nizam-Uddin, N., Quddus, A., Hassan, S. R., Rehman, A. U., and Bharany, S. (2024). Navigating IoT security: insights into architecture, key security features, attacks, current challenges and AI-driven solutions shaping the future of connectivity. *Comput. Mater. Continua* 81 (3), 3499–3559. doi:10.32604/cmc.2024.057877

Hassan, A., Nizam-uddin, N., Quddus, A., Hassan, S. R., Bharany, S., Rehman, A. U., et al. (2025). Designs strategies and performance of IoT antennas: a comprehensive review. *Discov. Comput.* 28 (1), 39. doi:10.1007/s10791-025-09536-y

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., et al. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cogn. Comput.* 16 (1), 45–74. doi:10.1007/s12559-023-10179-8

Hazra, A., Adhikari, M., Amgoth, T., and Srirama, S. N. (2021). A comprehensive survey on interoperability for IIoT: taxonomy, standards, and future directions. *ACM Comput. Surv. (CSUR)* 55 (1), 1–35. doi:10.1145/3485130

Hirsi, A., Alhartomi, M. A., Audah, L., Salh, A., bin Mad Sahar, N., Ahmed, S., et al. (2025). Comprehensive analysis of DDoS anomaly detection in software-defined networks. *IEEE Access* 13, 23013–23071. doi:10.1109/access.2025.3535943

Hossain, M., Kayas, G., Hasan, R., Skjellum, A., Noor, S., and Islam, S. R. (2024). A holistic analysis of internet of things (IoT) security: principles, practices, and new perspectives. *Future Internet* 16 (2), 40. doi:10.3390/fi16020040

Hossain, M. S., Rahman, M., Rahman, A., Kabir, M. M., Mridha, M. F., Huang, J., et al. (2025). Automatic navigation and self-driving technology in agricultural machinery: a state-of-the-art systematic review. *IEEE Access* 13, 94370–94401. doi:10.1109/access.2025.3573324

Hossain, S., Senouci, S. M., Brik, B., and Boualouache, A. (2025). A privacy-preserving self-supervised Learning-based intrusion detection system for 5G-V2X networks. *Ad Hoc Netw.* 166, 103674. doi:10.1016/j.adhoc.2024.103674

Huang, J., Huang, K., Jackson, K., and Hughes, C. (2025). "AI agent safety and security considerations," in *Agentic AI* (Cham: Springer), 369–407.

Hudda, S., and Haribabu, K. (2025). A review on WSN based resource constrained smart IoT systems. *Discov. Internet Things* 5 (1), 56–46. doi:10.1007/s43926-025-00152-2

Islam, M. S., Rahman, M. A., Bin Ameedeen, M. A., Ajra, H., Ismail, Z. B., and Zain, J. M. (2024). Blockchain-enabled cybersecurity provision for scalable heterogeneous network: a comprehensive survey. *CMES-Computer Model. Eng. Sci.* 138 (1), 43–123. doi:10.32604/cmes.2023.028687

Javed, A. R., Ahmed, W., Pandya, S., Maddikunta, P. K. R., Alazab, M., and Gadekallu, T. R. (2023). A survey of explainable artificial intelligence for smart cities. *Electronics* 12 (4), 1020. doi:10.3390/electronics12041020

Jayavarma, A., Parakkat Kesava Panikker, P., and Nair, M. G. (2025). Revolutionizing the energy sector: exploring diversified blockchain platforms for a sustainable future. *Front. Blockchain* 8, 1544770. doi:10.3389/fbloc.2025.1544770

Jeffrey, N., Tan, Q., and Villar, J. R. (2023). A review of anomaly detection strategies to detect threats to cyber-physical systems. *Electronics* 12 (15), 3283. doi:10.3390/electronics12153283

Johnphill, O., Sadiq, A. S., Al-Obeidat, F., Al-Khateeb, H., Taheir, M. A., Kaiwartya, O., et al. (2023). Self-healing in cyber–physical systems using machine learning: a critical analysis of theories and tools. *Future Internet* 15 (7), 244. doi:10.3390/fi15070244

Kabir, S., Gope, P., and Mohanty, S. P. (2022). A security-enabled safety assurance framework for IoT-based smart homes. *IEEE Trans. Industry Appl.* 59 (1), 6–14. doi:10.1109/tia.2022.3176257

Kaliyaperumal, P., Periyasamy, S., Thirumalaisamy, M., Balusamy, B., and Benedetto, F. (2024). A novel hybrid unsupervised learning approach for enhanced cybersecurity in the IoT. *Future Internet* 16 (7), 253. doi:10.3390/fi16070253

Kanellopoulos, D., Sharma, V. K., Panagiotakopoulos, T., and Kameas, A. (2023). Networking architectures and protocols for IoT applications in smart cities: recent developments and perspectives. *Electronics* 12 (11), 2490. doi:10.3390/electronics12112490

Kanzouai, C., Bouarourou, S., Zannou, A., Boulaalam, A., and Nfaoui, E. H. (2025). Enhancing IoT scalability and interoperability through ontology alignment and FedProx. *Future Internet* 17 (4), 140. doi:10.3390/fi17040140

Khriji, S., Benbelgacem, Y., Chéour, R., Houssaini, D. E., and Kanoun, O. (2022). Design and implementation of a cloud-based event-driven architecture for real-time data processing in wireless sensor networks. *J. Supercomput.* 78 (3), 3374–3401. doi:10.1007/s11227-021-03955-6

Kumar, A., Masud, M., Alsharif, M. H., Gaur, N., and Nanthaamornphong, A. (2025). Integrating 6G technology in smart hospitals: challenges and opportunities for enhanced healthcare services. *Front. Med.* 12, 1534551. doi:10.3389/fmed.2025.1534551

Lazaros, K., Koumadorakis, D. E., Vrahatis, A. G., and Kotsiantis, S. (2024). Federated learning: navigating the landscape of collaborative intelligence. *Electronics* 13 (23), 4744. doi:10.3390/electronics13234744

Mahto, M. K. (2025). "Explainable artificial intelligence: fundamentals, approaches, challenges, XAI evaluation, and validation," in *Explainable artificial intelligence for autonomous vehicles* Boca Raton, FL, United States: CRC Press (Taylor & Francis Group), 25–49.

Mallick, M. A. I., and Nath, R. (2024). Navigating the cyber security landscape: a comprehensive review of cyber-attacks, emerging trends, and recent developments. *World Sci. News* 190 (1), 1–69.

Mba, J. U. (2025). Advancing maritime operations sustainable practices and enhanced safety protocols for global shipping. *World J. Adv. Res. Rev.* 25 (1), 152–173. doi:10.30574/wjarr.2025.25.1.0028

McCarthy, A., Ghadafi, E., Andriotis, P., and Legg, P. (2022). Functionality-preserving adversarial machine learning for robust classification in cybersecurity and intrusion detection domains: a survey. *J. Cybersecurity Priv.* 2 (1), 154–190. doi:10.3390/jcp2010010

Mohamed, N. (2025). Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms. *Knowl. Inf. Syst.* 67, 6969–7055. doi:10.1007/s10115-025-02429-y

Mondal, M. K., Mandal, R., and Biswas, U. (2024). Big IoT data analytics in fog computing. *Fog Comput. Intelligent Cloud IoT Syst.*, 279–307. doi:10.1002/9781394175345.ch12

Muppalaneni, R., Inaganti, A. C., and Ravichandran, N. (2024). AI-driven threat intelligence: enhancing cyber defense with machine learning. *J. Comput. Innovations Appl.* 2 (1), 1–11.

Narayanan, A., Korium, M. S., Melgarejo, D. C., Hussain, H. M., De Sena, A. S., Silva, P. E. G., et al. (2022). Collective intelligence using 5G: concepts, applications, and challenges in sociotechnical environments. *IEEE Access* 10, 70394–70417. doi:10.1109/access.2022.3184035

Ni, J., Fang, G., Zhao, Y., Ren, J., Chen, L., and Ren, Y. (2024). Distributed group key management based on blockchain. *Electronics* 13 (11), 2216. doi:10.3390/electronics13112216

Oliha, J. S., Biu, P. W., and Obi, O. C. (2024). Securing the smart city: a review of cybersecurity challenges and strategies. *Open Access Res. J. Multidiscip. Stud.* 7 (1), 94–101.

Pavithra, R., Kaliappan, V. K., and Rajendar, S. (2023). Security algorithm for intelligent transport system in cyber-physical systems perceptive: attacks, vulnerabilities, and countermeasures. *SN Comput. Sci.* 4 (5), 544. doi:10.1007/s42979-023-01897-9

Peppes, N., Alexakis, T., Adamopoulou, E., and Demestichas, K. (2023). The effectiveness of zero-day attacks data samples generated *via* GANs on deep learning classifiers. *Sensors* 23 (2), 900. doi:10.3390/s23020900

Pothumarti, R., Jain, K., and Krishnan, P. (2021). A lightweight authentication scheme for 5G mobile communications: a dynamic key approach. *J. Ambient Intell. Humaniz. Comput.*, 1–19. doi:10.1007/s12652-020-02857-4

Putra, K. T., Arrayyan, A. Z., Hayati, N., Damarjati, C., Bakar, A., Chen, H. C., et al. (2024). A review on the application of internet of medical things in wearable personal health monitoring: a cloud-edge artificial intelligence approach. *IEEE Access* 12, 21437–21452. doi:10.1109/access.2024.3358827

Qudus, L. (2025). Resilient systems: building secure cyber-physical infrastructure for critical industries against emerging threats. *Int. J. Res. Publ. Rev.* 6 (1), 3330–3346. doi:10.55248/gengpi.6.0125.0514

Rao, P. M., and Deebak, B. D. (2023). Security and privacy issues in smart cities/industries: technologies, applications, and challenges. *J. Ambient Intell. Humaniz. Comput.* 14 (8), 10517–10553. doi:10.1007/s12652-022-03707-1

Reddy, D. K. K., Nayak, J., Behera, H. S., Shanmuganathan, V., Viriyasitavat, W., and Dhiman, G. (2024). A systematic literature review on swarm intelligence based intrusion detection system: past, present and future. *Archives Comput. Methods Eng.* 31 (5), 2717–2784. doi:10.1007/s11831-023-10059-2

Roy, S. (2023). Empowering intrusion detection in IoT networks with efficient machine learning strategies. Doctoral dissertation. North Dakota State University.

Sadhu, P. K., Yanambaka, V. P., and Abdelgawad, A. (2022). Internet of things: security and solutions survey. *Sensors* 22 (19), 7433. doi:10.3390/s22197433

Şahin, E., Arslan, N. N., and Özdemir, D. (2025). Unlocking the black box: an in-depth review on interpretability, explainability, and reliability in deep learning. *Neural Comput. Appl.* 37 (2), 859–965. doi:10.1007/s00521-024-10437-2

Salim, M. M., Camacho, D., and Park, J. H. (2024). Digital twin and federated learning enabled cyberthreat detection system for IoT networks. *Future Gener. Comput. Syst.* 161, 701–713. doi:10.1016/j.future.2024.07.017

Santoso, A., and Surya, Y. (2024). Maximizing decision efficiency with edge-based AI systems: advanced strategies for real-time processing, scalability, and autonomous intelligence in distributed environments. *Q. J. Emerg. Technol. Innovations* 9 (2), 104–132.

Sarıkaya, B. S., and Bahtiyar, Ş. (2024). A survey on security of UAV and deep reinforcement learning. *Ad Hoc Netw.* 164, 103642. doi:10.1016/j.adhoc.2024.103642

Seo, S., Lee, J., Kim, B., Lee, W., and Kim, D. (2023). MF2-DMTD: a formalism and game-based reasoning framework for optimized drone-type moving target defense. *Comput. Mater. Continua* 77 (2), 2595–2628. doi:10.32604/cmc.2023.042668

Sewak, M., Sahay, S. K., and Rathore, H. (2023). Deep reinforcement learning in the advanced cybersecurity threat detection and protection. *Inf. Syst. Front.* 25 (2), 589–611. doi:10.1007/s10796-022-10333-x

Sha, Z., Layton, A., Heydari, B., and Van Bossuyt, D. L. (2025). Special issue: networks and graphs for engineering systems and design. *J. Comput. Inf. Sci. Eng.* 25, 060301. doi:10.1115/1.4068457

Shaik, A. S., and Shaik, A. (2024). "AI enhanced cyber security methods for anomaly detection," in *International conference on machine intelligence, tools, and applications* (Cham: Springer Nature Switzerland), 348–359.

Sharma, A., Kumar, V. G., and Poojari, A. (2024). Prioritize threat alerts based on false positives qualifiers provided by multiple AI models using evolutionary computation and reinforcement learning. *J. Institution Eng. (India) Ser. B*, 1305–1322. doi:10.1007/s40031-024-01175-z

Shayea, G. G., Zabil, M. H. M., Habeeb, M. A., Khaleel, Y. L., and Albahri, A. S. (2025). Strategies for protection against adversarial attacks in AI models: an in-depth review. *J. Intelligent Syst.* 34 (1), 20240277. doi:10.1515/jisys-2024-0277

Shehzadi, T. (2024). Reinforcement learning-based autonomous systems for cyber threat detection and response. *East. Eur. J. Multidiscip. Res.* 1 (1), 123–137.

Shyaa, M. A., Ibrahim, N. F., Zainol, Z., Abdullah, R., Anbar, M., and Alzubaidi, L. (2024). Evolving cybersecurity frontiers: a comprehensive survey on concept drift and feature dynamics aware machine and deep learning in intrusion detection systems. *Eng. Appl. Artif. Intell.* 137, 109143. doi:10.1016/j.engappai.2024.109143

Singh, N., Buyya, R., and Kim, H. (2024). Securing cloud-based internet of things: challenges and mitigations. *Sensors* 25 (1), 79. doi:10.3390/s25010079

Singh, T., Kumar, S., Singh, S. K., Gupta, B. B., Wu, J., and Castiglione, A. (2025). "Enhancing autonomous system security with AI and secure computation technologies," in *AI developments for industrial robotics and intelligent drones* (Hershey, PA, USA: IGI Global Scientific Publishing), 159–186.

Tallam, K. (2025). Transforming cyber defense: harnessing agentic and frontier AI for proactive, ethical threat intelligence. *arXiv preprint arXiv:2503.00164*.

Tariq, U., Ahmed, I., Bashir, A. K., and Shaukat, K. (2023). A critical cybersecurity analysis and future research directions for the internet of things: a comprehensive review. *Sensors* 23 (8), 4117. doi:10.3390/s23084117

Taye, M. M. (2023). Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers* 12 (5), 91. doi:10.3390/computers12050091

Tayyab, M., Marjani, M., Jhanjhi, N. Z., Hashem, I. A. T., Usmani, R. S. A., and Qamar, F. (2023). A comprehensive review on deep learning algorithms: security and privacy issues. *Comput. Secur.* 131, 103297. doi:10.1016/j.cose.2023.103297

Tian, Z., Cui, L., Liang, J., and Yu, S. (2022). A comprehensive survey on poisoning attacks and countermeasures in machine learning. *ACM Comput. Surv.* 55 (8), 1–35. doi:10.1145/3551636

Tyagi, A. K., and Seranmadevi, R. (2024). Blockchain for enhancing security and privacy in the smart healthcare. *Digital Twin Blockchain Smart Cities*, 343–370. doi:10.1002/9781394303564.ch16

Usmani, U. A., Happonen, A., and Watada, J. (2022). "A review of unsupervised machine learning frameworks for anomaly detection in industrial applications," in *Science and information conference* (Cham: Springer International Publishing), 158–189.

Valsalan, P., Hasan, N. U., Baig, I., Zghaibeh, M., Farooq, U., and Suhail, S. (2024). Unleashing the potential: the joint of 5G and 6G technologies in enabling advanced IoT communication and sensing systems: a comprehensive review and future prospects. *J. Commun.* 19 (11), 523–535. doi:10.12720/jcm.19.11.523-535

Van Hoang, N. (2023). Human expertise and machine learning in collaborative intelligence frameworks for robust cybersecurity solutions. *J. Appl. Cybersecurity Anal. Intell. Decision-Making Syst.* 13 (12), 1–12. Available online at: http://sciencespress.com/index.php/JACAIDMS/article/view/2023-12-04

Vaseashta, A. (2022). "Applying resilience to hybrid threats in infrastructure, digital, and social domains using multisectoral, multidisciplinary, and whole-of-government approach," in *Building cyber resilience against hybrid threats* (Amsterdam, Netherlands: IOS Press), 42–59.

Vermesan, O., Bröring, A., Tragos, E., Serrano, M., Bacciu, D., Chessa, S., et al. (2022). "Internet of robotic things–converging sensing/actuating, hyperconnectivity, artificial intelligence and IoT platforms," in *Cognitive hyperconnected digital transformation* (Aalborg, Denmark: River Publishers), 97–155.

Zafir, E. I., Akter, A., Islam, M. N., Hasib, S. A., Islam, T., Sarker, S. K., et al. (2024). Enhancing security of internet of robotic things: a review of recent trends, practices, and recommendations with encryption and blockchain techniques. *Internet Things* 28, 101357. doi:10.1016/j.iot.2024.101357

Zengeni, I. P., and fadli Zolkipli, M. (2024). Zero-day exploits and vulnerability management. *Borneo Int. J. eISSN* 2636-9826 (7), 26–33. Available online at: https://majmuah.com/journal/index.php/bij/article/view/648

Zhukabayeva, T., Zholshiyeva, L., Karabayev, N., Khan, S., and Alnazzawi, N. (2025). Cybersecurity solutions for industrial internet of things–edge computing integration: challenges, threats, and future directions. *Sensors* 25 (1), 213. doi:10.3390/s25010213

Zhuwankinyu, E. K., Moyo, T. M., and Mupa, M. N. (2024). Leveraging generative AI for an ethical and adaptive cybersecurity framework in enterprise environments. *IRE Journals* 8 (6), 654–675.

Zou, X., Yan, Y., Hao, X., Hu, Y., Wen, H., Liu, E., et al. (2025). Deep learning for cross-domain data fusion in urban computing: taxonomy, advances, and outlook. *Inf. Fusion* 113, 102606. doi:10.1016/j.inffus.2024.102606