A Comprehensive Review of Artificial Intelligence of Things (AIoT): Towards Energy-Efficient and Secure Node Architectures

Dr. Rishi Kumar Sharma[1]., Mr. Raghu Nandan Singh Hada[2]., Md. Shahbaz Khan[3]

[1]Associate professor, CSE, Jagannath University, Jaipur, Rajasthan

[2]Assistant Professor, CSE, Jagannath University, Jaipur, Rajasthan

[3]Diploma Jagannath University, Jaipur Rajasthan

DOI: https://doi.org/10.51244/IJRSI.2026.1306000334

Received: 27 June 2025; Accepted: 02 July 2026; Published: 09 July 2026

ABSTRACT

The proliferation of Internet of Things (IoT) devices across healthcare, smart cities, industrial automation, transportation, and agriculture has ushered in an era of unprecedented connectivity and data generation. Despite these advances, conventional IoT architectures remain fundamentally constrained by limited battery capacity, restricted computational resources, and growing exposure to sophisticated cyber-threats. Artificial Intelligence of Things (AIoT)—the convergence of artificial intelligence (AI) with IoT infrastructure—has emerged as a transformative paradigm capable of addressing these systemic limitations. This paper presents a thorough and structured review of AIoT with a dual emphasis on energy optimization and node-level security. We systematically examine the AIoT layered architecture, state-of-the-art machine learning and deep learning approaches for energy management, AI-driven intrusion detection systems, blockchain-based trust frameworks, lightweight cryptographic primitives, and federated learning mechanisms. Furthermore, the interplay between energy efficiency and security overhead is critically analyzed. Key open challenges, including scalability, real-time attack detection, privacy preservation, and standardization gaps, are identified, and a forward-looking research agenda is proposed. This review aims to serve as a consolidated reference for researchers and practitioners seeking to develop secure, energy-aware, and intelligent AIoT ecosystems.

Index Terms—AIoT, Internet of Things, Energy Optimization, Node Security, Edge Computing, Machine Learning, Deep Learning, Intrusion Detection Systems.

INTRODUCTION

The Internet of Things (IoT) has fundamentally reshaped the way digital systems interact with the physical world. With projections indicating over 75 billion connected devices by the end of the decade [1], IoT has penetrated virtually every domain of human activity—from wearable health monitors and autonomous vehicles to precision agriculture and smart power grids. Billions of sensors and actuators now generate continuous streams of data, enabling real-time monitoring, automation, and decision-making. At the core of this technological revolution lies the capacity for billions of heterogeneous nodes to sense, communicate, and act upon real-world data without direct human intervention.

However, the wide-scale deployment of IoT introduces a dichotomy of critical challenges. On the one hand, the resource-constrained nature of IoT nodes—typically operating on limited battery reserves and with restricted computational capabilities—makes energy management a persistent concern that directly impacts network lifetime, reliability, and operational cost [2]. On the other hand, the open and distributed architecture of IoT networks makes them attractive targets for adversarial activities, including distributed denial-of-service (DDoS) attacks, node impersonation, data tampering, and malware propagation [3]. Traditional IoT systems often treat energy saving and security as separate, non-interacting objectives. Yet the two are fundamentally intertwined: every security mechanism consumes precious energy, while a successful attack can rapidly drain a node's battery through denial-of-service floods or forced retransmissions.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

The integration of Artificial Intelligence (AI) with IoT infrastructures, colloquially termed Artificial Intelligence of Things (AIoT), offers a promising avenue to holistically address these challenges. By embedding intelligent analytics, adaptive learning, and autonomous decision-making into the IoT stack, AIoT enables systems to dynamically optimize resource utilization while simultaneously strengthening resilience against evolving cyber threats [4]. Techniques such as reinforcement learning (RL), deep neural networks, federated learning, and edge intelligence are increasingly being deployed to create self-managing, self-healing AIoT ecosystems.

This review paper makes the following contributions: (i) it provides a structured examination of the AIoT architecture with emphasis on energy and security dimensions; (ii) it systematically surveys AI-based energy optimization techniques including routing, scheduling, harvesting, and edge offloading; (iii) it reviews intelligent node security mechanisms such as ML-based intrusion detection, blockchain integration, and lightweight cryptography; (iv) it analyzes the interdependencies and trade-offs between energy efficiency and security overhead; and (v) it identifies open research gaps and delineates a future research agenda. The remainder of this paper is organized as follows. Section II presents the AIoT architectural framework. Sections III and IV cover energy optimization techniques. Section V addresses node security challenges. Section VI reviews AI-based security mechanisms. Section VII analyzes the energy–security trade-off. Section VIII highlights application domains. Sections IX and X present challenges and future directions, respectively. Section XI concludes the paper.

AIoT Architecture And Framework

AIoT represents the seamless convergence of AI-driven intelligence with the pervasive sensing and communication capabilities of IoT. The result is a multi-layered architecture capable of adaptive, context-aware, and autonomous operations across diverse deployment environments [4], [5]. This model extends the Open Systems IoT Reference Model (OSiRM) by incorporating AI-specific components at each tier.

A. Five-Layer AIoT Architecture

AIoT systems are decomposed into five functional layers, each performing distinct roles in data acquisition, processing, and service delivery. AI is vertically integrated across all layers. At the perception layer, on-device TinyML routines preprocess sensor signals. At the edge, reinforcement learning agents adjust resource allocation in real time. In the cloud, large-scale model training and federated learning aggregation take place.

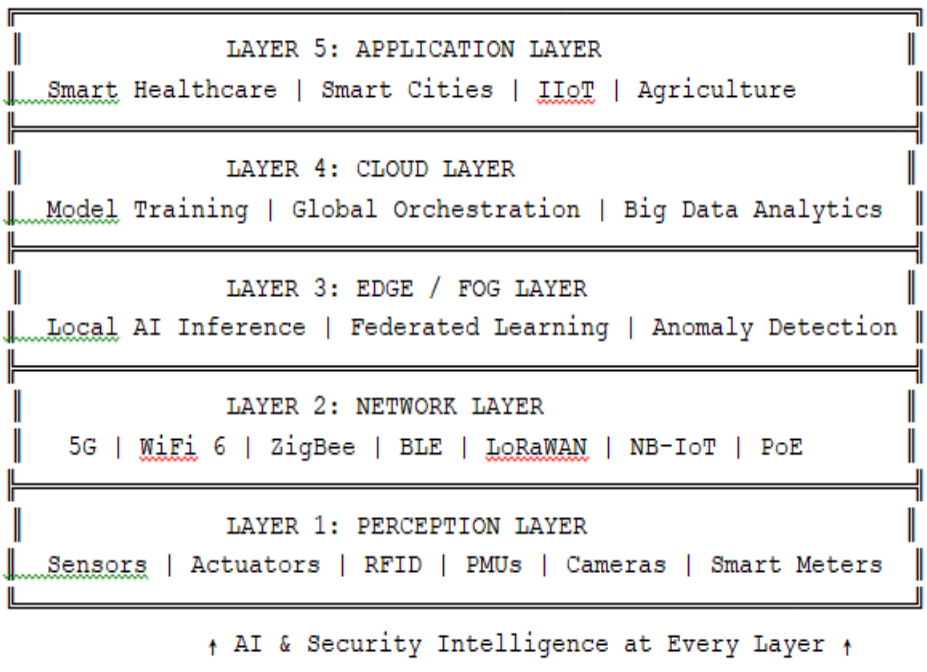

Fig. 1. Five-Layer AIoT Architecture showing vertical AI integration across all tiers .

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

The Perception Layer constitutes the physical interface between the AIoT system and the real world. It encompasses heterogeneous sensing devices—including temperature, pressure, optical, and inertial sensors, RFID readers, cameras, and actuators—that continuously collect environmental data. In Electric Power and Energy Systems (EPES), this layer includes phasor measurement units (PMUs) and smart inverters [2]; in smart buildings, it encompasses humidity, occupancy, and light sensors [3]. AI models embedded at this layer enable on-device preprocessing, local anomaly flagging, and sensor fusion.

The Network Layer is responsible for reliable and efficient transmission of raw and preprocessed data across the AIoT ecosystem. It leverages a range of wireless and wired communication technologies, including IEEE 802.15.4 (ZigBee), Bluetooth Low Energy (BLE), Wi-Fi 6, LoRaWAN, NB-IoT, 5G cellular networks, and Power over Ethernet (PoE). 5G, with its native support for massive machine-type communication (mMTC), is expected to become the backbone for many AIoT applications [1]. AI-optimized routing protocols at this layer reduce transmission energy and improve end-to-end latency.

The Edge/Fog Layer performs intermediate data processing in close proximity to IoT end nodes. By executing AI inference workloads—such as classification, regression, and anomaly detection—at the network edge, this layer dramatically reduces communication overhead, response latency, and cloud dependency. Edge gateways and fog nodes can run compressed deep learning models for real-time anomaly detection or adaptive lighting control [3], [6]. Federated learning frameworks operating at this layer enable collaborative model training without centralizing sensitive raw data.

The Cloud Layer provides large-scale computational resources and persistent storage for complex AI model training, historical data analytics, and global coordination. Cloud-based analytics enable long-term trend analysis, policy generation, and the training of sophisticated models that can later be distilled for edge deployment [2], [3].

The Application Layer delivers domain-specific intelligent services to end users, spanning smart healthcare, intelligent transportation, industrial automation, smart agriculture, and urban infrastructure management. Examples include dynamic stochastic energy management systems (DSEMS) for power grids [2], smart home energy monitoring [3], and city-wide traffic management. AI integration across all layers facilitates closed-loop optimization and real-time adaptability [5].

Energy Consumption In AIoT: Sources And Impact

Energy consumption remains one of the most pervasive and consequential limitations of large-scale AIoT deployments. The majority of IoT nodes are battery-powered and are frequently deployed in remote or inaccessible environments where battery replacement is impractical or cost-prohibitive. Understanding the principal sources of energy expenditure is a prerequisite for designing effective optimization strategies.

Empirical studies have consistently demonstrated that wireless communication—encompassing packet transmission, reception, and channel negotiation—accounts for the dominant share of energy consumption in IoT nodes, often exceeding local computation by an order of magnitude [2], [7]. Studies confirm that transmitting one bit can cost several orders of magnitude more energy than executing a local computation. Beyond communication, continuous sensing operations contribute to baseline energy drain, particularly for high-frequency or high-resolution sensors. Other notable sources include idle listening, wherein a node continuously monitors the channel even in the absence of incoming traffic, and security protocol execution, which imposes cryptographic and authentication overheads.

The deployment of AI capabilities at the edge introduces an additional energy dimension. While edge inference eliminates the energy cost of round-trip cloud communication, the local execution of neural network models on resource-constrained hardware demands careful architectural consideration. Model compression, quantization, and hardware-aware neural architecture search (NAS) are emerging as important tools for managing this trade-off.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

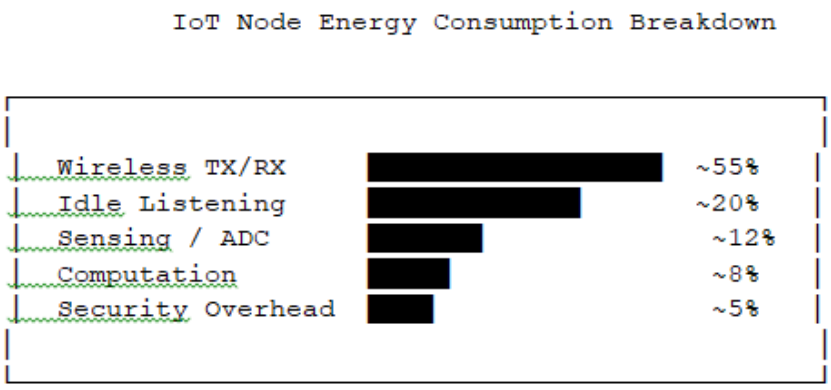

Fig. 2. Approximate energy consumption breakdown for a typical IoT node. Wireless communication dominates, motivating AI-driven optimization strategies.

AI-Based Energy Optimization Techniques

A. Machine Learning for Resource Management

Machine learning has proven highly effective in optimizing resource allocation, communication scheduling, and routing in AIoT networks. Reinforcement learning (RL), in particular, has gained prominence for its ability to enable agents to learn optimal policies through interaction with dynamic environments without requiring explicit prior knowledge of system models [2]. Deep Q-Networks (DQNs) and Proximal Policy Optimization (PPO) algorithms have been applied to adaptively control transmission power levels, dynamically select relay nodes, and intelligently switch devices between active and sleep states. Q-learning approaches applied to cluster-based wireless sensor networks have demonstrated network lifetime improvements of 20–40% over conventional fixed-policy protocols [5]. Support Vector Machines (SVMs) and decision trees have been employed for workload prediction and proactive energy state management.

B. Deep Learning-Based Predictive Optimization

Deep learning architectures extend the capabilities of classical ML by capturing complex temporal and spatial dependencies in network traffic and energy data. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks excel at modeling time-series data, making them highly suitable for predicting device activity patterns, traffic load distributions, and energy consumption trajectories [2]. Convolutional Neural Networks (CNNs) have been applied to spatial analysis of sensor field data to identify regions of high activity, enabling event-driven sensing that suspends data collection during periods of low information content. Attention-based Transformer architectures are increasingly being explored for multi-step energy forecasting in heterogeneous AIoT environments.

C. AI-Enhanced Clustering and Routing

Clustering is one of the most effective topology management strategies for prolonging network lifetime [1]. Nodes are organized into groups, each led by a cluster head (CH) that aggregates data and forwards it towards the sink. Classical protocols such as LEACH pioneered randomized CH election based on a probabilistic threshold, while HEED improved load balancing by selecting CHs according to residual energy and intra-cluster communication cost [8], [9]. DWEHC supports multi-hop intra-cluster communication, and BEEM introduces coverage-sensitive longevity [10], [11].

AI-driven adaptive clustering algorithms dynamically select cluster heads based on residual energy, node centrality, and traffic load, extending network lifetime substantially [2]. Fuzzy logic systems (e.g., FL-LEACH) combine multiple parameters—residual energy, node density, distance to base station—to make soft clustering decisions [12]. Bio-inspired metaheuristic algorithms—including Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Genetic Algorithms (GA)—have been widely applied to routing

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

optimization in large-scale AIoT networks, efficiently exploring high-dimensional solution spaces to identify near-optimal routing topologies.

D. Edge and Fog Computing

Offloading computational tasks from energy-constrained IoT nodes to edge servers constitutes one of the most impactful strategies for network-wide energy conservation. AI-enabled task schedulers dynamically partition workloads between local execution and edge offloading based on real-time estimates of channel quality, queue length, and residual battery energy [6]. In smart buildings, local gateways process HVAC and lighting sensor streams, executing AI-based predictive models that adjust setpoints without consulting the cloud [3]. Multi-access Edge Computing (MEC) frameworks augmented with deep reinforcement learning have demonstrated significant reductions in end-to-end task latency and device-level energy consumption.

In 5G-IoT scenarios, devices can dynamically choose among multiple radio access technologies (RATs)—LTE, Wi-Fi, ZigBee—based on energy cost, required data rate, and security level [1]. AI-powered network selection modules weigh these factors in real time, switching to the most energy-efficient interface that still meets QoS and security constraints.

E. Energy Harvesting with AI Control

Energy harvesting technologies that convert ambient energy—including solar radiation, thermal gradients, mechanical vibration, and radio frequency signals—into usable electrical power offer a pathway to perpetually operating AIoT nodes. The intermittent and stochastic nature of harvested energy necessitates intelligent management strategies that balance energy intake, storage, and consumption. AI-based energy management controllers, trained on historical harvesting profiles and workload demands, have demonstrated superior performance over rule-based approaches in sustaining system operation during low-harvesting periods. RL-based dynamic duty cycling adjusts sensor sleep intervals based on predicted traffic and harvested energy, achieving near-optimal energy neutrality [5]. In EPES, solar harvesting with a power density of 24 mW/cm² has been demonstrated for grid-monitoring sensor nodes [2].

Table I Comparison of AI-Based Energy Optimization Techniques in AIoT

Technique	**Primary Application**	**Key Algorithms**	**Energy Saving**	**Overhead**
Reinforcement Learning	Dynamic routing & power control	DQN, PPO, Q-learning	20–40%	Moderate
Deep Learning (CNN/RNN/LSTM)	Traffic prediction, scheduling	LSTM, CNN, Transformer	High	High
Federated Learning	Distributed model training	FedAvg, FedProx	Moderate	Low
Swarm Metaheuristics (PSO/ACO/GA)	Cluster-head selection, routing	PSO, ACO, GA	Moderate	Low
Edge/Fog Computing (MEC)	Local task offloading	Deep RL, DNN	Very High	Very Low
Energy Harvesting + AI	Adaptive harvesting scheduling	RL, LSTM forecasting	Perpetual op.	Very Low
Fuzzy Logic Clustering	CH selection in WSNs	FL-LEACH, FLSEP	15–30%	Low

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Node Security In AIoT: Threat Landscape

The security posture of AIoT systems is inherently constrained by the limited computational and communication capabilities of IoT nodes, which preclude the direct deployment of conventional cryptographic and intrusion-detection mechanisms designed for resource-rich platforms [3]. The result is a broad and expanding attack surface that adversaries can exploit through a range of vectors.

A. Principal Security Threats

Distributed Denial-of-Service (DDoS) attacks remain the most prevalent and disruptive threat in AIoT environments. By flooding network nodes with spurious traffic, attackers exhaust both communication bandwidth and node battery reserves, rendering services unavailable. Botnet infrastructures, increasingly composed of compromised IoT devices, amplify the scale and sophistication of such attacks.

Node impersonation and Sybil attacks involve adversaries masquerading as legitimate network participants to inject falsified sensor readings, subvert routing decisions, or gain unauthorized access to protected resources. Physical node capture attacks, feasible in unattended deployments, can expose cryptographic keys and facilitate large-scale network compromise. Eavesdropping on unencrypted wireless channels enables passive data exfiltration, while replay attacks exploit the absence of timestamp validation to reuse captured authentication tokens [3], [7].

Malware injection, facilitated by the absence of secure boot and firmware integrity verification on many commodity IoT platforms, enables persistent adversarial control. Data tampering at the perception layer—whether by modifying sensor outputs or intercepting actuator commands—can have catastrophic consequences in safety-critical applications such as medical devices and industrial control systems. In power systems, GPS spoofing attacks targeting PMU time synchronization and Man-in-the-Middle (MITM) attacks are additional critical threats [2].

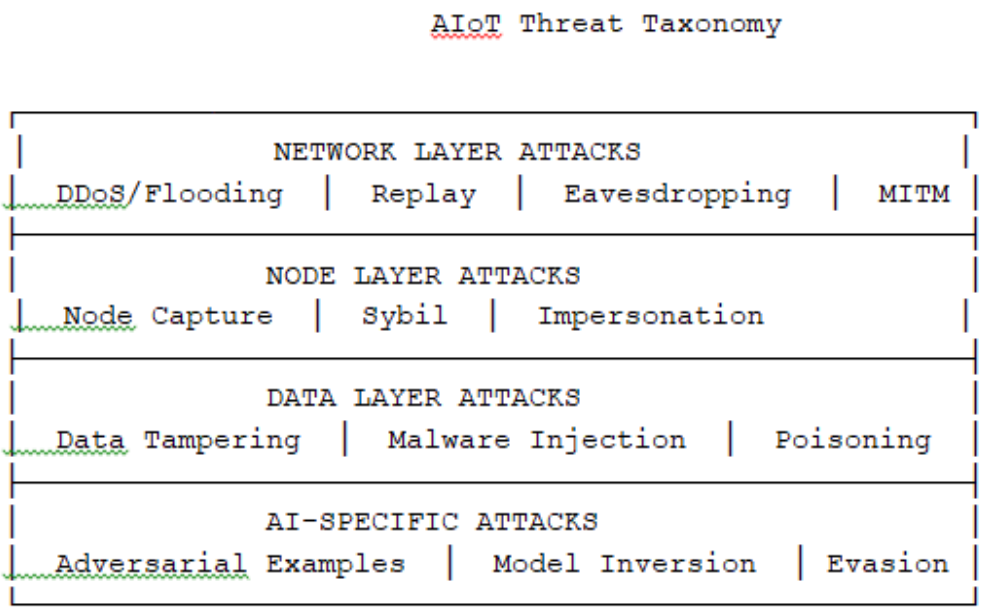

Fig. 3. AIoT threat taxonomy spanning network, node, data, and AI-specific attack vectors.

AI-Based Security Mechanisms

A. Intrusion Detection Systems (IDS)

AI-enabled Intrusion Detection Systems (IDS) have emerged as the primary defense mechanism against network-layer attacks in AIoT environments. By training classifiers on labeled traffic datasets, these systems learn to discriminate between benign and malicious network behavior with accuracy levels that significantly

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

surpass traditional signature-based detection approaches, which are ineffective against zero-day exploits [1], [3]. Artificial Neural Networks (ANNs) and Random Forests have demonstrated high detection rates against DDoS and port-scanning attacks, while LSTM networks excel at detecting temporally coordinated attacks that exploit short-term behavioral patterns. Deep Belief Networks (DBNs) and stacked autoencoders are particularly effective for anomaly detection in settings where labeled attack data is scarce, by learning compressed representations of normal behavior and flagging significant deviations.

Lightweight versions of these models, obtained through pruning and quantization, can be deployed directly on edge gateways or even on sensor motes, enabling real-time mitigation without overwhelming the battery [5]. Federated IDS architectures, wherein local models are trained on each node's private traffic data and periodically aggregated at an edge server, address both privacy and data siloization concerns [5].

B. Blockchain for Decentralized Trust

Blockchain technology provides a decentralized, tamper-evident ledger that is well-suited to address the trust and authentication challenges inherent in distributed AIoT systems. By encoding device identities, access policies, and transaction histories in immutable smart contracts, blockchain eliminates the single point of failure associated with centralized authentication servers [6]. In power systems, a private blockchain can guarantee the integrity of PMU data, making post-incident forensics reliable [2].

Lightweight blockchain variants, including Directed Acyclic Graph (DAG) structures such as IOTA Tangle and proof-of-authority consensus mechanisms, have been developed to reduce the computational and energy overheads that render conventional proof-of-work blockchains unsuitable for resource-constrained IoT deployments. Smart contracts automate access control and can revoke compromised node credentials without human intervention. Blockchain-secured firmware update mechanisms ensure that over-the-air (OTA) updates can be cryptographically verified by recipient nodes before installation, mitigating malware injection risks.

C. Lightweight Cryptography

Securing data confidentiality and message integrity on IoT nodes demands cryptographic algorithms that deliver robust security guarantees within strict resource budgets. The NIST Lightweight Cryptography Standardization process, concluded in 2023 with the selection of the ASCON cipher suite, reflects the growing recognition of this need [7]. Standards such as PRESENT, SPECK, and SIMON offer strong encryption with a fraction of the energy required by AES. Block cipher PRESENT (64-bit block, 80/128-bit key) offers hardware-optimized implementations requiring as few as 1,000 gate equivalents, making it viable for deeply embedded platforms. Elliptic Curve Cryptography (ECC) provides asymmetric key exchange and digital signature capabilities with key lengths an order of magnitude shorter than RSA at equivalent security levels.

AI can further optimize the use of these ciphers by adaptively selecting the cryptographic suite according to the current battery level, data sensitivity, and perceived threat, thus minimizing energy waste [2], [7].

D. Federated Learning for Privacy-Preserving Security

Federated learning (FL) has emerged as a powerful paradigm for training AI-based security models across distributed AIoT nodes without requiring the centralization of sensitive raw data [5]. In the FL framework, each participating node trains a local model on its private dataset and transmits only the resulting model gradients or parameter updates to a central aggregator, which synthesizes a global model without accessing any individual node's data. This approach significantly reduces communication bandwidth, preserves data privacy, and can be combined with differential privacy for formal guarantees.

This architecture is particularly well-suited to healthcare and industrial AIoT deployments where regulatory frameworks such as GDPR and HIPAA impose strict constraints on data sharing. Differential privacy mechanisms can be applied to gradient submissions to provide formal privacy guarantees against model inversion attacks. Secure aggregation protocols utilizing homomorphic encryption further ensure that individual gradient contributions remain confidential even from the aggregating server.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Table II Summary of AI-Based Node Security Mechanisms in AIoT Systems

Mechanism	Threats Addressed	Algorithms / Tools	Energy Overhead	Maturity
ML-based IDS	DDoS, Intrusion, Port Scan	Random Forest, LSTM, ANN, DBN	High	High
Blockchain	Auth., Tamper-proofing, OTA	Smart Contracts, IOTA Tangle	Medium	Medium
Lightweight Crypto	Data confidentiality, Integrity	PRESENT, SPECK, ASCON, ECC	Low	High
Federated Learning	Privacy-preserving training	FedAvg, Differential Privacy	Low	Medium
Anomaly Detection	Botnet, Sybil, Zero-day	Autoencoder, DBN, VAE	Medium	Medium
Adaptive Crypto Selection	Resource-constrained security	AI-based policy engine	Very Low	Emerging

The Energy–Security Trade-Off In AIoT

A fundamental and often overlooked tension in AIoT system design is the inherent trade-off between security robustness and energy efficiency. Security mechanisms—whether cryptographic, authentication-based, or AI-driven—consume computational cycles and communication bandwidth that directly translate into increased energy expenditure. Conversely, weakened or absent security exposes nodes to energy-exhaustion attacks that can degrade network lifetime more severely than the cost of the protective measures themselves [6].

DDoS attacks represent the most direct example of security failures cascading into energy crises. Flood-based attacks saturate the communication interfaces of victim nodes, forcing continuous reception and processing of spurious packets until battery reserves are depleted. Sybil attacks that subvert cluster-head elections increase the transmission distances of legitimate nodes, inflating per-packet energy costs. Replay attacks that force repeated re-authentication cycles impose cumulative cryptographic overhead on constrained nodes.

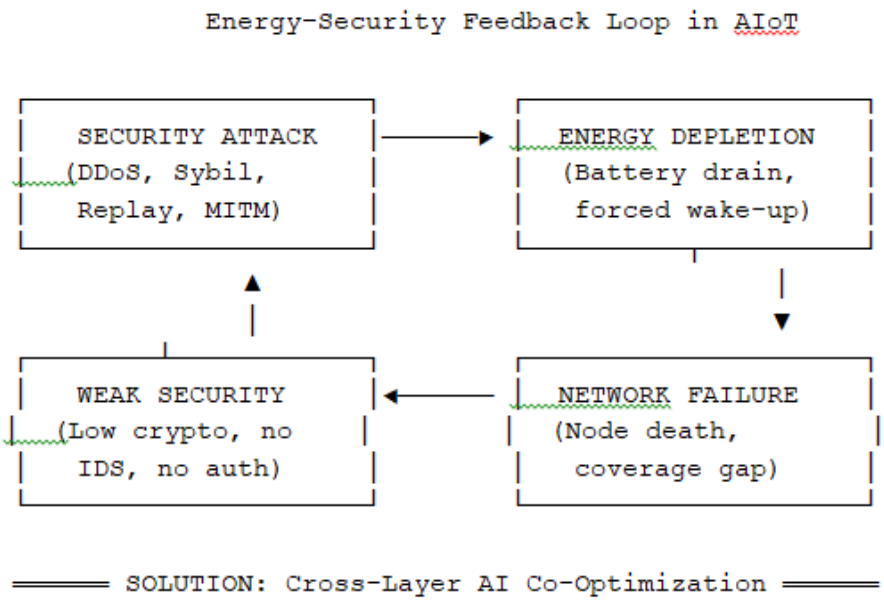

Fig. 4. Energy–security feedback loop in AIoT systems. Security breaches cascade into energy crises and vice versa, motivating cross-layer co-optimization [2], [4].

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Cross-layer optimization frameworks that co-design energy management and security policies have shown considerable promise in resolving this tension. Such frameworks leverage shared contextual information—including residual energy levels, channel conditions, and threat intelligence—to dynamically adjust the aggressiveness of security checks and the granularity of energy conservation measures. For instance, a node under suspected attack may temporarily elevate authentication frequency while disabling non-essential sensing tasks, maintaining security without catastrophic energy expenditure. Energy-aware blockchain consensus mechanisms that adaptively scale proof-of-work difficulty based on network-wide energy availability represent another compelling realization of this paradigm.

Table III Qualitative Comparison of AIoT Techniques: Energy–Security Balance

Technique	Energy Saving	Security Enhancement	AI Involvement	Maturity
AI Clustering (RL/FL)	High	Moderate	High	Mature
Edge Computing	Very High	Moderate	High	Mature
Energy Harvesting + AI	Very High	Low	Medium	Developing
ML-based IDS	Low overhead	Very High	Very High	Mature
Blockchain (Lightweight)	Medium cost	Very High	Medium	Developing
Lightweight Cryptography	Low overhead	High	Medium	Mature
Federated Learning IDS	Low overhead	Very High	Very High	Emerging
Cross-layer Co-optimization	High	High	Very High	Emerging

Applications Of Secure And Energy-Efficient AIoT

A. Smart Healthcare

Smart healthcare represents one of the most consequential and demanding application domains for AIoT. Wearable biosensors monitoring cardiac rhythms, blood glucose levels, and neurological activity generate continuous high-volume data streams that must be transmitted reliably with strict latency and accuracy constraints. AI-driven duty cycling and adaptive compression algorithms extend the operational lifetimes of implantable and wearable devices to clinically viable durations. End-to-end encryption of physiological data streams, blockchain-secured medical records, and federated learning-based diagnostic models that train across hospital networks without centralizing patient data collectively define the state of the art in healthcare AIoT security.

B. Smart Cities and Urban Infrastructure

Smart city deployments integrate AIoT across traffic management, public safety surveillance, environmental monitoring, and utility grid management. AI-driven adaptive street lighting systems reduce energy consumption by up to 70% compared to conventional fixed-schedule operation by adjusting illumination levels based on pedestrian and vehicular traffic density. AI-driven clustering of small cells and IoT gateways balances load and saves energy, while network selection algorithms switch between 5G, Wi-Fi, and LPWAN based on application QoS requirements and energy budgets [1]. AI-powered Security Operations Centers (SOCs) integrating real-time threat intelligence, automated incident response, and cross-system anomaly correlation are increasingly essential.

C. Electric Power and Energy Systems (EPES)

Industrial AIoT deployments in manufacturing, energy production, and logistics leverage predictive maintenance, quality assurance, and autonomous process control to improve operational efficiency and reduce

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

unplanned downtime. Intelligent electronic devices equipped with AI-IDS protect against GPS spoofing and DoS attacks, while reinforcement learning dynamically adjusts power flows and line ratings, reducing losses and improving resilience [2]. Time-series anomaly detection models deployed at the edge enable millisecond-scale identification of mechanical failures before they progress to catastrophic events. PoE-based LED lighting and HVAC systems in smart buildings leverage AI-optimized scheduling to cut energy costs by up to 30%, with edge-based IDS monitoring for unauthorized access [3].

D. Precision Agriculture

Agricultural AIoT systems deploy sensor networks across large outdoor environments to monitor soil moisture, ambient temperature, crop health, and livestock vital signs. The geographic expanse of agricultural deployments and the absence of reliable power infrastructure place extreme demands on node energy autonomy. AI-optimized solar-powered sensing schedules that concentrate data collection during peak sunlight and defer energy-intensive transmission to favorable propagation conditions have demonstrated multi-year autonomous operation in field trials.

Open Challenges And Research Gaps

Despite the remarkable progress surveyed in this paper, AIoT systems continue to face a set of fundamental challenges that constrain practical deployment at scale.

Resource Heterogeneity and Scalability: The enormous diversity in hardware capabilities, communication protocols, and energy profiles across AIoT deployments complicates the design of universally applicable optimization and security frameworks. Managing authentication, key distribution, and model aggregation at the scale of billions of nodes remains an unsolved problem. Many AI-enhanced clustering algorithms assume homogeneous sensor networks, whereas real-world AIoT deployments involve highly heterogeneous devices [1]–[3].

Real-Time Threat Detection: The latency constraints of many AIoT applications require that security decisions be made in milliseconds, whereas state-of-the-art deep learning IDS models often require inference times that are incompatible with these requirements on resource-constrained hardware. Developing lightweight yet highly accurate models capable of real-time inference on IoT-class processors is an active and urgent research priority.

Adversarial Robustness of AI Models: AIoT systems that rely on AI for both energy management and security are susceptible to adversarial attacks specifically targeting the AI components. Adversarial examples crafted to evade ML-based IDS, data poisoning attacks targeting federated learning aggregation, and model inversion attacks against edge inference models represent emerging threat vectors that require dedicated countermeasures.

Lack of Benchmarks and Standardization: There are no standard datasets or testbeds that simultaneously measure energy consumption and security posture in AIoT settings. The AIoT landscape is fragmented by a proliferation of incompatible communication protocols, data formats, and security standards [3]. Coordinated standardization efforts by bodies such as IEEE, ISO/IEC, and IETF are critical to resolving this fragmentation.

Privacy and Regulatory Compliance: The aggregation of sensor data from wearable devices, smart homes, and connected vehicles raises significant privacy concerns. Ensuring compliance with evolving regulatory frameworks such as GDPR and HIPAA while preserving the utility of AI analytics is a complex sociotechnical challenge. The energy–security trade-off—finding the Pareto-optimal operating point in real time—remains an open problem.

Future Research Directions

Based on the analysis of current limitations and emerging technological developments, the following research directions are identified as high-priority avenues for advancing secure and energy-efficient AIoT systems.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VI June 2026

Joint Energy–Security Optimization Frameworks: Cross-layer controllers that use multi-objective RL to dynamically allocate energy budgets among communication, computation, and security tasks. Such frameworks leverage shared contextual information—residual energy, channel conditions, and threat intelligence—to simultaneously minimize energy consumption and maximize security resilience.

TinyML and Lightweight Adaptive AI Models: Developing neural architecture search (NAS) methodologies tailored to IoT hardware constraints, yielding binary neural networks and ultra-lightweight deep learning models that run intrusion detection directly on microcontrollers with negligible battery impact. These models maintain high task performance within strict memory, latency, and energy budgets.

Post-Quantum Cryptography for AIoT: Integrating NIST-standardized post-quantum cryptographic primitives into IoT communication stacks to ensure long-term security against quantum computing threats. Quantum-resistant lightweight cryptographic schemes tailored to IoT's severe resource constraints remain an open design challenge.

Explainable AI (XAI) for Security Analytics: Designing interpretable AI models for intrusion detection that provide actionable explanations for their decisions, facilitating human-in-the-loop security operations and regulatory compliance. Human-interpretable explanations for IDS decisions improve trust and facilitate compliance auditing.

AI-Native 6G IoT: Exploiting the ultra-reliable low-latency communication (URLLC) and integrated sensing and communication (ISAC) capabilities of 6G networks to support energy-efficient and secure AIoT at unprecedented scale [1]. 6G's native AI capabilities can offload training and inference to the network infrastructure, achieving energy-neutral intelligent security.

Digital Twin Integration: Leveraging digital twin models of physical AIoT deployments to enable predictive maintenance, proactive security threat modeling, and energy consumption forecasting without disturbing live systems. Creating virtual replicas of IoT deployments to simulate attack scenarios and energy profiles before real-world hardening.

Quantum-Secured Federated Learning: Combining quantum key distribution (QKD) with federated learning frameworks to provide information-theoretic security guarantees for distributed model training in high-security AIoT applications. Green AIoT Architectures that treat carbon footprint and energy consumption as first-class optimization objectives alongside accuracy and latency are also a pressing priority.

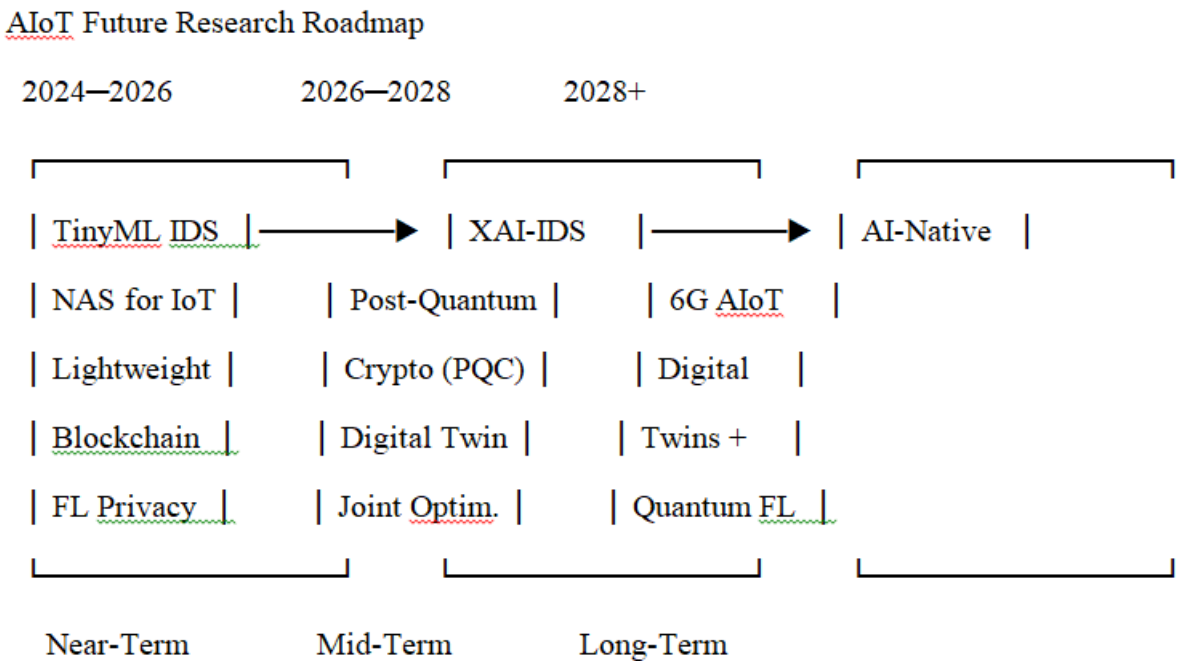

Fig. 5. AIoT research roadmap showing near-term, mid-term, and long-term priority directions.

CONCLUSION

This paper has presented a comprehensive and structured review of Artificial Intelligence of Things (AIoT) with a primary focus on energy optimization and node security—two dimensions that are individually critical and collectively interdependent in determining the viability of real-world AIoT deployments. We examined the five-layer AIoT architecture and systematically surveyed AI-driven techniques for energy management, including reinforcement learning-based routing, deep learning-based predictive scheduling, edge computing offloading, energy harvesting control, and bio-inspired clustering across wireless sensor networks, power grids, and smart buildings.

On the security front, we reviewed AI-powered intrusion detection systems, blockchain-based decentralized trust frameworks, lightweight cryptographic standards including ASCON and ECC, and federated learning-based privacy-preserving security architectures. A critical analysis of the energy–security feedback loop highlighted both the risks of underinvesting in security and the computational costs of robust protection, motivating cross-layer co-optimization frameworks. Key application domains—smart healthcare, smart cities, electric power systems, and precision agriculture—were examined to illustrate the practical stakes of this research agenda.

Open challenges including scalability, adversarial robustness, real-time detection, privacy compliance, absence of unified benchmarks, and standardization deficits were systematically identified. A forward-looking research agenda was proposed encompassing TinyML, post-quantum cryptography, explainable security analytics, 6G-enabled architectures, digital twin integration, and quantum-secured federated learning. As AIoT systems assume increasing responsibility for safety-critical and privacy-sensitive functions across society, the development of intelligent, energy-efficient, and inherently secure architectures transitions from a research aspiration to an engineering imperative.

REFERENCES

1. W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in Proc. 33rd Hawaii Int. Conf. Syst. Sci., 2000.
2. O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," IEEE Trans. Mobile Comput., vol. 3, no. 4, pp. 366-379, 2004.
3. P. Ding, J. Holliday, and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks," in Proc. DCOSS, 2005, pp. 322-339.
4. L. Xu, G. O'Hare, and R. Collier, "A balanced energy-efficient multi-hop clustering scheme for wireless sensor networks," in Proc. 7th IFIP WMNC, 2014.
5. P. Nayak and A. Devulapalli, "A fuzzy logic-based clustering algorithm for WSNs to extend the network lifetime," IEEE Sensors J., vol. 16, no. 1, pp. 137-144, Jan. 2016
6. L. Xu, R. Collier, and G. M. P. O'Hare, "A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios," IEEE Internet Things J., vol. 6, no. 1, pp. 1-22, Jul. 2017.
7. D. Minoli, K. Sohraby, and B. Occhiogrosso, "IoT considerations, requirements, and architectures for smart buildings: Energy optimization and next-generation building management systems," IEEE Internet Things J., vol. 4, no. 1, pp. 269-283, Feb. 2017.
8. D. M. Mendez, I. Papapanagiotou, and B. Yang, "Internet of Things: Survey on security and privacy," arXiv:1707.01879, 2017.
9. G. Bedi, G. K. Venayagamoorthy, R. Singh, R. Brooks, and K.-C. Wang, "Review of Internet of Things (IoT) in electric power and energy systems," IEEE Internet Things J., vol. 5, no. 2, pp. 847-870, Apr. 2018.
10. "Probabilistic Security Interface for IOT Devices in the Cloud Network", IJRTE, International , 2019.
11. J. Zhang and D. Tao, "Empowering Things with Intelligence: A survey of progress, challenges, and opportunities in AIoT," arXiv:2011.08612, 2020.
12. Identity Based Remote Device Authentication system for Enhancing the Security among Remote Machines in IoT or Cloud Robotics Network", Neuro Quantology, International ,2022.

13. M. Pooyandeh and I. Sohn, "Edge network optimization based on AI techniques: A survey," Electronics, vol. 10, no. 22, p. 2830, 2021. doi: 10.3390/electronics10222830.
14. Y. Ali, H. U. Khan, and M. Khalid, "Engineering the advances of artificial neural networks for IoT security: A systematic review," J. Big Data, vol. 10, no. 128, 2023. doi: 10.1186/s40537-023-00813-3.
15. NIST, "Lightweight Cryptography Standardization: Final selection of ASCON," National Institute of Standards and Technology, Gaithersburg, MD, USA, 2023.
16. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015. doi: 10.1038/nature14539.
17. R. Mustafa et al., "A cross-layer secure and energy-efficient framework for IoT: A comprehensive survey," Sensors, vol. 24, no. 22, p. 7209, 2024. doi: 10.3390/s24227209.
18. Ghaffari et al., "Securing Internet of Things using machine and deep learning methods: A survey," Cluster Computing, vol. 27, pp. 9065-9089, 2024. doi: 10.1007/s10586-024-04321-4.
19. S. I. Siam et al., "Artificial Intelligence of Things: A survey," ACM Trans. Sensor Networks, vol. 20, no. 2, pp. 1-58, 2024.
20. O. Aouedi et al., "A survey on intelligent Internet of Things: Applications, security, privacy, and future directions," arXiv:2406.03820, 2024.
21. S. K. Sahu and K. Mazumdar, "Exploring security threats and solutions techniques for IoT," Frontiers Artif. Intell., vol. 7, 2024. doi: 10.3389/frai.2024.1380608.
22. B. Isong and K. Moeti, "A comprehensive review of energy optimization techniques in the Internet of Things," J. Inf. Syst. Inform., vol. 7, no. 1, pp. 1-35, 2025.