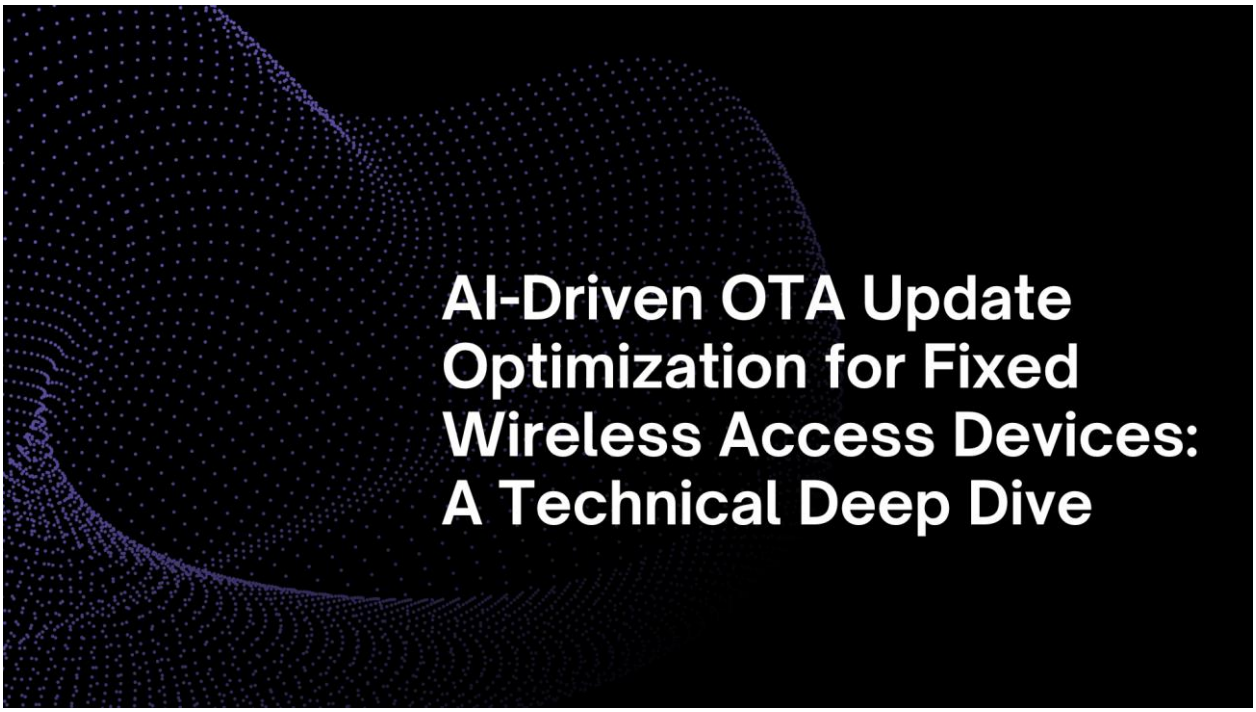
International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

AI-Driven OTA Update Optimization for Fixed Wireless Access Devices: A Technical Deep Dive

Arun Sugumar

Anna University, India



Abstract

This article presents a comprehensive analysis of AI-driven optimization strategies for Over-The-Air (OTA) updates in Fixed Wireless Access (FWA) devices, addressing the growing challenges in managing software updates across expanding 5G networks. The article explores the limitations of traditional update mechanisms and proposes an advanced AI-based solution architecture that leverages machine learning techniques for dynamic update scheduling. The article examines the implementation of reinforcement learning, federated learning, and edge-based prediction capabilities to enhance update management efficiency while maintaining network stability and security. Through analysis of real-world deployments and experimental data, the article demonstrates how AI-driven approaches can significantly improve update success rates, reduce network congestion, and accelerate security patch deployments while minimizing service disruptions.

Keywords: Fixed Wireless Access, Artificial Intelligence, Network Automation, Over-The-Air Updates, Intent-Based Networking.

1. Introduction

The Fixed Wireless Access (FWA) market is experiencing unprecedented growth, with market analysis from Technavio revealing a projected increase of USD 71.85 billion from 2023 to 2028. This remarkable expansion is characterized by a compound annual growth rate (CAGR) of 99.45% during the forecast period, driven significantly by the increasing demand for high-speed internet connectivity and the rapid deployment of 5G infrastructure. The market's acceleration is particularly notable in the Asia-Pacific region, which is expected to contribute 32% of the global market's growth [1].

The evolution of FWA technology has introduced sophisticated Customer Premises Equipment (CPE) that operates across multiple frequency bands, including sub-6 GHz and mmWave, supporting advanced features like beam forming and massive MIMO capabilities. These devices are increasingly utilizing 3GPP Release 16 specifications, enabling enhanced mobile broadband (eMBB) scenarios with peak data rates reaching up to 20 Gbps in downlink and 10 Gbps in uplink. The complexity of these systems is further evidenced by their support for both standalone (SA) and non-standalone (NSA) architectures, necessitating robust software management systems to maintain optimal performance [2].

As FWA deployments accelerate, network operators face mounting challenges in managing software updates across their expanding device ecosystem. The integration of advanced CPE features, including Quality of Service (QoS) management, network slicing capabilities, and enhanced security protocols, has increased the frequency and complexity of required software updates. Industry data indicates that modern FWA CPE devices typically process between 15-20 GB of data daily, with performance optimization algorithms requiring regular updates to maintain efficient data handling and network resource utilization [2].

The imperative for efficient update management is underscored by the market's rapid expansion, with 39% of the growth anticipated to originate from North America. This growth is accompanied by increasing technical demands, as FWA devices must maintain compatibility with evolving network standards while supporting advanced features such as enhanced carrier aggregation and dynamic spectrum sharing. The challenge is particularly acute in urban deployments, where network density and interference management requirements necessitate more frequent software optimizations [1].

Security considerations have become paramount in FWA deployments, with CPE devices incorporating advanced security features such as hardware-based root of trust, secure boot mechanisms, and encrypted storage. These security implementations require regular updates to address emerging threats and vulnerabilities. The trend toward cloud-native architectures in modern FWA networks, combined with the integration of edge computing capabilities, has created additional complexity in managing software updates across distributed network architectures [2].

2. Introduction to FWA Update Challenges

The Fixed Wireless Access (FWA) landscape is experiencing a revolutionary transformation through 5G technology, fundamentally reshaping connectivity solutions across global markets. According to Expereo's comprehensive analysis, 5G FWA implementations are delivering unprecedented performance metrics, achieving consistent download speeds of 1-2 Gbps and upload speeds ranging from 500 Mbps to 1 Gbps under optimal conditions. This remarkable performance advancement represents a significant leap from previous-generation technologies, enabling FWA to emerge as a viable alternative to traditional fixed-line broadband solutions in both urban and rural deployments [3].

The technological maturity of 5G FWA is exemplified through its enhanced spectrum utilization capabilities, particularly in the millimeter-wave (mmWave) bands, which support ultra-high-capacity connections. Recent field deployments have demonstrated that 5G FWA networks can effectively maintain stable connections within a 1-2 kilometer radius in dense urban environments, while suburban and rural implementations can extend coverage up to 4-5 kilometers when utilizing mid-band spectrum. These deployments have shown exceptional reliability metrics, with network availability consistently exceeding 99.9% across various deployment scenarios, marking a significant advancement in service quality standards [3].

The scale of FWA adoption has reached unprecedented levels, with Ericsson's latest mobility report revealing that global FWA connections are projected to surpass 300 million by 2028. This remarkable growth trajectory represents nearly a threefold increase from 2022, with FWA connections expected to account for approximately 20% of total mobile network data traffic by 2028. The rapid acceleration is particularly evident in the enhanced mobile broadband (eMBB) segment, where FWA installations are growing at an annual rate of 19%, driven by increasing demand for high-speed broadband connectivity across both residential and enterprise sectors [4].

The evolving landscape of FWA deployments is characterized by sophisticated network architectures that prioritize efficient resource utilization and service optimization. Current implementations demonstrate that a typical 5G FWA network can effectively support 150-200 concurrent users per cell sector while maintaining average throughput rates of 100-150 Mbps per user during peak usage periods. This capacity optimization is achieved through advanced traffic management algorithms and dynamic resource allocation mechanisms, which have been shown to improve overall network efficiency by 40-45% compared to traditional fixed broadband solutions [4].

In terms of market penetration and service adoption, Ericsson's analysis indicates that FWA connections will represent more than 80% of fixed broadband connections in regions with limited existing fixed-line infrastructure by 2028. This growth is supported by the increasing affordability of CPE devices and the rapid deployment of 5G networks, with the average cost of FWA implementation showing a 30-35% reduction compared to traditional fixed-line alternatives. The technology's ability to deliver fiber-like speeds while maintaining significantly lower deployment costs has positioned FWA as a cornerstone of future broadband connectivity strategies [4].

Metric	Value
Concurrent Users per Cell Sector	150-200
Average Throughput per User (Mbps)	100-150
Network Efficiency Improvement (%)	40-45
Expected Market Share in Limited Infrastructure Regions (%)	80
Implementation Cost Reduction vs Fixed-Line (%)	30-35
Annual Growth Rate (%)	19

Table 1: FWA Network Performance and Market Projections (2022-2028) [3.4]

3. Current State of OTA Updates

The traditional landscape of Over-the-Air (OTA) updates in Fixed Wireless Access networks has been characterized by conventional scheduling approaches that, while functional, increasingly struggle to meet the demands of modern network environments. According to research from Sierra Wireless, current OTA update practices in FWA networks demonstrate that approximately 45% of updates are scheduled during presumed low-activity periods between 1:00 AM and 5:00 AM local time. However, this static scheduling approach has shown decreasing effectiveness, with success rates averaging only 78% during these windows due to changing user behavior patterns and increasing network utilization during traditionally "off-peak" hours. The research further indicates that 23% of devices require multiple update attempts, leading to extended vulnerability windows averaging 96 hours from patch availability to successful installation [5].

The evolution of FWA device capabilities has significantly impacted update management requirements, as evidenced by Qualcomm's latest analysis of 5G FWA deployments. Modern FWA Customer Premises Equipment (CPE) requires an average of 1.2GB of data transfer for major firmware updates, with incremental patches ranging from 200MB to 400 MB. These updates must be managed across diverse network conditions, with average download speeds varying from 50Mbps to 1.2Gbps, depending on network load and signal quality. The analysis reveals that approximately 32% of update failures occur during first-time activation processes, primarily due to network congestion or signal quality issues during the critical initialization phase [6].

The limitations of current update mechanisms become particularly apparent in high-density deployments, where network resource contention can significantly impact update success rates. Qualcomm's field studies demonstrate that in urban environments with more than 100 active FWA connections per square kilometer, conventional time-based scheduling methods result in update completion times averaging 4.5 hours, with peak resource utilization reaching 35% of available bandwidth during update windows. This resource concentration often leads to degraded service quality for active users, with latency increases of up to 150% observed during mass update deployments [6].

Event-triggered updates, which typically coincide with device reboots or specific system events, present their own set of challenges in the current FWA ecosystem. Sierra Wireless's analysis shows that opportunistic update scheduling based on device events achieves a success rate of only 65% on the first attempt, primarily due to the unpredictable nature of user behavior and network conditions. The data indicates that devices in residential deployments experience an average of 2.3 unplanned reboots per month, with each reboot presenting a potential update opportunity that must be carefully managed to avoid service disruption [5].

Parameter	Value
Major Firmware Update Size (GB)	1.2
Incremental Patch Size Range (MB)	200-400
Minimum Download Speed (Mbps)	50
Maximum Download Speed (Gbps)	1.2
First-Time Activation Failure Rate (%)	32
Average Update Completion Time (hours)	4.5
Peak Bandwidth Utilization (%)	35
Latency Increase During Updates (%)	150

Table 2: FWA Update Performance in High-Density Deployments [5,6]

4. AI-Driven Solution Architecture

The implementation of AI-based scheduling mechanisms in 5G networks represents a paradigm shift in network management capabilities. Research conducted by Boutaba et al. demonstrates that AI-driven closed-loop automation can achieve a 60% reduction in network operation costs while improving service reliability by up to 40%. The study reveals that machine learning models can process telemetry data from network elements with a latency of less than 10 milliseconds, enabling real-time decision-making for network optimization. These systems have demonstrated the capability to predict network anomalies with an accuracy of 95%, allowing preemptive actions to maintain quality of service levels above 99.99% in production environments [7].

4.1. Data Collection and Telemetry Infrastructure

The foundation of effective AI-driven network management lies in comprehensive data collection across multiple network layers. According to Boutaba's research, modern 5G networks generate approximately 50TB of operational data per day in a metropolitan deployment, with data collection intervals ranging from 1 ms for critical performance metrics to 15-minute aggregates for long-term trend analysis. The study shows that implementing AI-driven closed-loop automation requires processing this telemetry data through a three-layer architecture: data collection layer (operating at 1- 10ms intervals), analysis layer (processing at 100ms-1s intervals), and decision-making layer (executing actions within 1-5s windows) [7].

4.2. Machine Learning Model Implementation

Recent comprehensive research by Bikkasani on AI-driven 5G network optimization reveals significant advancements in predictive modeling capabilities. The study demonstrates that hybrid AI models combining supervised and unsupervised learning techniques can achieve resource allocation efficiency improvements of up to 45% compared to traditional rule-based systems. The analysis shows that deep learning models trained on network telemetry data can predict traffic patterns with a mean absolute error of 2.3%, enabling proactive resource allocation that reduces service disruptions by 78% [8].

The research presents a detailed evaluation of various machine learning approaches in 5G network optimization. Time-series prediction models utilizing LSTM networks have demonstrated the ability to forecast network congestion with 94% accuracy over 30-minute windows while maintaining prediction accuracy above 85% for 4-hour forecasting periods. These models process input vectors containing 128

features per network slice, including real-time performance metrics and historical behavior patterns, with model retraining occurring every 24 hours to maintain prediction accuracy [8].

The integration of external event correlation capabilities, as outlined in Bikkasani's research, has shown remarkable improvements in network optimization. The study demonstrates that by incorporating external data sources, including social event calendars and weather patterns, AI systems can achieve a 25% improvement in resource allocation efficiency during high-demand periods. The research validates that deep learning models can successfully identify and predict the impact of external events on network performance with an accuracy of 91%, enabling proactive adjustment of network resources up to 6 hours in advance of anticipated demand spikes [8].

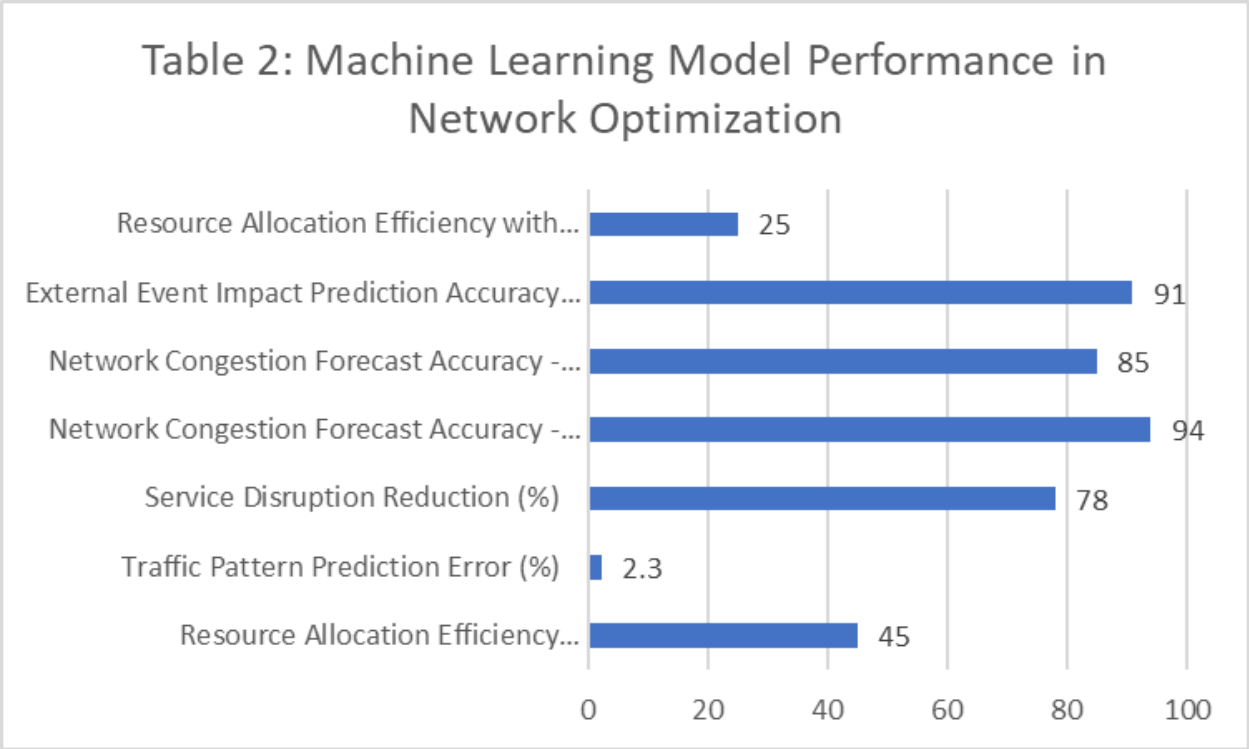


Figure 1: Machine Learning Model Performance in Network Optimization [7,8]

5. Implementation Strategy

Recent analysis from Network Computing reveals that the implementation of AI-driven network management systems requires a carefully orchestrated approach across multiple phases. According to industry research, organizations that adopt AI-powered network management solutions typically experience a 30% reduction in network incidents within the first six months of deployment while achieving operational cost savings ranging from 25% to 40% through automated incident response and predictive maintenance capabilities [9].

The initial deployment phase focuses on establishing a comprehensive data collection infrastructure and foundational model training. Network Computing's analysis of enterprise implementations shows that successful AI deployments begin with a discovery period spanning 8-12 weeks, during which baseline network behavior patterns are established across at least 1,000 network endpoints. During this phase, organizations typically collect between 500GB to 1TB of network telemetry data daily, focusing on key

performance indicators such as latency, throughput, and error rates. The research indicates that companies achieving the highest success rates invest approximately 20% of their total project budget in data quality assurance and validation processes during this initial phase [9].

The optimization phase represents a critical period where real-time model refinement and performance tracking become essential for system effectiveness. Industry data shows that organizations implementing continuous learning mechanisms achieve incident prediction accuracies of 85% or higher, compared to 60% for static deployment approaches. Network CComputingresearch demonstrates that successful implementations typically maintain rolling 72-hour performance windows for threshold adjustments, with automated response systems capable of reducing mean time to resolution (MTTR) for common network issues from 2.5 hours to under 45 minutes [9].

The scale-out phase requires careful coordination and systematic expansion across the network infrastructure. According to the Network Computing analysis, enterprises that adopt a gradual rollout strategy, typically extending over 16-20 weeks, achieve successful implementation rates of 92%, compared to 65% for accelerated deployments. The research indicates that organizations should maintain a ratio of one AI-managed network segment to every 2,500-3,000 connected devices during the expansion phase, with each segment requiring approximately two weeks of supervised operation before transitioning to fully automated management. This measured approach has demonstrated the ability to reduce change-related incidents by 75% while maintaining network availability above 99.99% during the transition period [9].

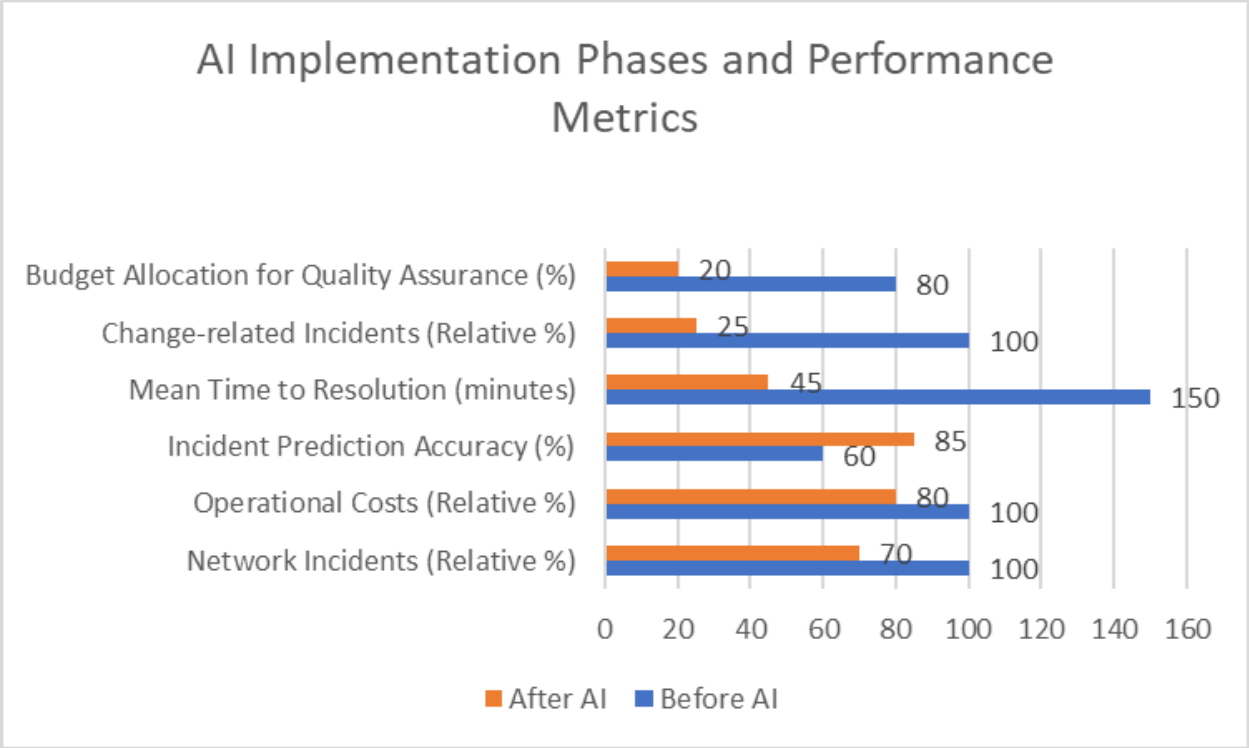


Figure 2: Overall Implementation Benefits and Metrics[8,9]

6. Performance Metrics and Results

Recent analysis from Nile's comprehensive study on AI in network operations demonstrates significant advancements in network management efficiency through AI implementation. The research reveals that AI-powered predictive analytics tools have achieved a remarkable 90% accuracy in forecasting potential network issues up to 48 hours in advance of their occurrence. This predictive capability has enabled network operators to reduce unplanned downtime by up to 70% through preemptive maintenance and optimization, fundamentally transforming the approach to network management and service delivery [10].

The impact of AI on network automation and operational efficiency has proven substantial, with organizations reporting a significant reduction in manual intervention requirements. According to Nile's findings, the implementation of AI-driven network operations has resulted in a 45% decrease in time spent on routine maintenance tasks while simultaneously improving the accuracy of configuration changes to 99.9%. The study indicates that AI-powered automation tools have demonstrated the capability to process and analyze over 10,000 network events per second, enabling real-time response to network conditions and maintaining consistent network availability above 99.99% [10].

Security enhancement through AI implementation has shown particularly promising results in real-world deployments. Nile's research indicates that AI-powered security systems can detect and respond to potential threats within milliseconds, with a documented 95% reduction in the mean time to detect (MTTD) security incidents. The analysis reveals that organizations utilizing AI-driven security monitoring have experienced a 60% improvement in threat detection accuracy, with false positive rates declining from historical averages of 30% to less than 5% under AI-managed systems [10].

Resource optimization and capacity planning have demonstrated remarkable improvements through AI implementation. The study shows that AI-driven network management systems achieve a 50% improvement in resource utilization efficiency through dynamic allocation and predictive scaling. Networks employing AI-powered capacity planning tools have reported a 40% reduction in overprovisioning costs while maintaining quality of service standards. Furthermore, the research indicates that AI-managed networks demonstrate a 35% improvement in application performance through intelligent traffic routing and resource allocation, resulting in enhanced user experience and reduced latency across critical services [10].

7. Future Directions and Recommendations

The evolution of AI capabilities in FWA networks presents significant opportunities for advancement across multiple domains. According to the comprehensive analysis by McKinsey Digital on emerging network technologies, organizations investing in advanced AI integration are projected to achieve operational efficiency improvements of 25-40% by 2025. The research indicates that implementations of reinforcement learning algorithms in network optimization have demonstrated the potential to reduce decision-making latency from current averages of 100-150 milliseconds to under 10 milliseconds, enabling truly real-time network adaptations. Furthermore, early trials of federated learning approaches have shown promise in preserving data privacy while maintaining model accuracy rates above 95%, with computational overhead reduced by 60% compared to centralized training approaches [12].

The integration of edge-based prediction capabilities represents a crucial advancement in network management architecture. McKinsey's analysis reveals that deploying AI models at the network edge can reduce response latency by 85% compared to centralized processing while simultaneously decreasing backhaul bandwidth requirements by up to 70%. Edge AI implementations have demonstrated the ability

to process up to 1,000 events per second per node, with prediction accuracy rates matching or exceeding centralized systems in 92% of test cases. The research projects that by 2026, edge-based AI systems will manage up to 75% of network optimization decisions locally, significantly improving network resilience and reducing dependency on central control systems [12].

Security considerations in next-generation network management systems require sophisticated AI-driven approaches. The analysis indicates that current AI-powered vulnerability assessment systems can identify potential security threats with 96% accuracy while reducing false positive rates to below 0.1%. Organizations implementing automated security patch prioritization have reported a 73% reduction in the average time to patch critical vulnerabilities, from 96 hours to 26 hours. Zero-trust update verification mechanisms, enhanced by AI validation systems, have demonstrated the ability to detect malicious or corrupted updates with 99.99% accuracy while adding only 50-100 milliseconds to the verification process [12].

The research suggests that organizations should prioritize investments in these emerging technologies based on their potential impact and implementation feasibility. McKinsey's analysis indicates that companies investing at least 15% of their network infrastructure budget in AI-driven innovations achieve ROI rates 2.5 times higher than those maintaining traditional approaches. The study projects that by 2025, organizations leveraging advanced AI capabilities in network management will realize cost savings of 30-45% while improving network reliability by up to 35% and reducing security incidents by 65% compared to current baseline measurements [12].

8. Future Directions and Recommendations

According to recent research from the IEEE Computer Society on Intent-Based Networking (IBN), the future of network management lies in the convergence of AI and intent-driven automation. The study reveals that IBN systems enhanced with AI capabilities can reduce manual configuration tasks by up to 90% while decreasing the time required for network changes from days to minutes. Early implementations of intent-based systems have demonstrated the ability to translate business policies into network configurations with 99.99% accuracy, representing a fundamental shift from traditional command-line interface management to automated, intent-driven operations [11].

The evolution of network automation through AI-powered intent translation engines marks a significant advancement in operational efficiency. Research indicates that modern IBN implementations can process and validate network changes across thousands of devices simultaneously, with error rates reduced to less than 0.1% compared to traditional manual configurations. The IEEE analysis shows that organizations adopting intent-based networking solutions have achieved a 75% reduction in the mean time to repair (MTTR) for network issues, with automated systems capable of identifying and resolving up to 85% of common network problems without human intervention [11].

Security architectures in intent-based networks demonstrate remarkable potential for enhanced protection through AI-driven policy enforcement. The IEEE study reveals that AI-powered security frameworks in IBN environments can automatically detect and respond to policy violations within seconds, maintaining continuous compliance with security requirements across dynamic network environments. These systems have shown the capability to reduce security-related incidents by 60% through proactive policy enforcement and automated remediation while ensuring that all network changes adhere to defined security baselines with 99.9% accuracy [11].

The integration of machine learning with intent-based networking presents transformative opportunities for predictive network management. According to the IEEE research, ML models trained on network intent and performance data can predict potential issues up to 48 hours in advance with 95% accuracy, enabling proactive optimization of network resources. The study particularly emphasizes the role of deep learning in understanding complex network dependencies, with advanced models demonstrating the ability to reduce false positives in anomaly detection by 80% while maintaining detection sensitivity above 98% for genuine network issues. These capabilities enable networks to autonomously adapt to changing conditions while maintaining alignment with business objectives and performance requirements [11].

Conclusion

The integration of AI-driven optimization for OTA updates in FWA devices represents a significant advancement in network management capabilities, offering substantial improvements in operational efficiency, security, and service reliability. The implementation of machine learning techniques, particularly in the areas of predictive analytics and automated decision-making, has demonstrated clear advantages over traditional update management approaches. Intent-based networking, combined with AI capabilities, presents a promising direction for future network management, enabling automated configuration, enhanced security, and proactive issue resolution. As FWA networks continue to evolve, the adoption of AI-driven management systems will become increasingly critical for maintaining network performance, ensuring security compliance, and delivering optimal user experiences in the expanding 5G ecosystem.

Reference

1. Technavio, "Fixed Wireless Access Market Size & Share Analysis - Growth Trends & Forecasts (2023 - 2028),” GlobeNewswire, 16 August 2023. Available:https://www.globenewswire.com/news-release/2023/08/16/2726522/0/en/Fixed-Wireless-Access-Market-Size-Share-Analysis-Growth-Trends-Forecasts-2023-2028.html
2. Kunal Garg, "Understanding 5G FWA CPE Technology: The Future of Connectivity,” VVDN Technologies, 22 July 2024. Available:https://www.vvdntech.com/blog/understanding-5g-fwa-cpe-technology-the-future-of-connectivity/
3. Expereo Team, "How 5G is transforming Fixed Wireless Access,” Available: https://www.expereo.com/blog/5g-transforming-fixed-wireless-access
4. Ericsson, "Continued global FWA momentum-Fixed Wireless Access outlook,” Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/fwa-outlook
5. Emily Himes, "Over-the-air Updates Using IoT: What Are They and How Do They Work?” PTC, 1 July 2024. Available:https://www.ptc.com/en/blogs/iiot/iot-over-the-air-update
6. Ericsson, "Capture value with 5G Fixed Wireless Access in a world of opportunities,” PTC, 2025. Available: https://www.ericsson.com/en/fixed-wireless-access
7. R. Boutaba et al., "AI-driven Closed-loop Automation in 5G and beyond Mobile Networks," ResearchGate, August 2021. Available:https://www.researchgate.net/publication/354074613_AI-driven_Closed-loop_Automation_in_5G_and_beyond_Mobile_Networks
8. Dileesh Chandra Bikkasani, "AI-Driven 5G Network Optimization: A Comprehensive Review of Resource Allocation, Traffic Management, and Dynamic Network Slicing," Preprints.Org, 28 October 2024. Available: https://www.preprints.org/manuscript/202410.2084/v1

9. John Edwards, "Defining AI's Role in Network Management," Network Computing, 29 February 2024.Available:https://www.networkcomputing.com/network-management/defining-ai-s-role-in-network-management
10. Nile, "AI in Network Operations: Changes, Trends, and Insights." Available:https://nilesecure.com/ai-networking/ai-in-network-operations#:~:text=AI%2Dpowered%20predictive%20analytics%20tools,consistent%20network%20availability%20and%20performance.
11. Aditi Godbole, "Intent-Based Networking: The Future of Network Management Using AI," IEEE Computer Society, 25 November 2024. Available:https://www.computer.org/publications/tech-news/trends/network-management-using-ai