世界无线局域网应用发展联盟
WLAN Application Alliance

Network Innovation and Development Alliance
全球固定网络创新联盟

Research Report on AI-Powered New Connectivity Enabling the New Industrial Revolution

May 2025

Media: contact@waa-alliance.org
Work: contact@waa-alliance.org
Business: contact@waa-alliance.org
Web: www.waa-alliance.org

Contributing Organizations

China Mobile (Hangzhou) Information Technology Co., Ltd.
Huawei Technologies Co., Ltd.
ZTE Corporation
New H3C Technologies Co., Ltd.
Ruijie Networks Co., Ltd.
Hisilicon（Shanghai） Technologies Co., Ltd.
Shenzhen Longsailing Semiconductor Co., Ltd.
FiberHome Telecommunication Technologies Co.,Ltd.
TÜV Rheinland
Shenzhen iTest Technology Co., Ltd.
Realsil Microelectronics (Suzhou) Co., Ltd.

Key Authors (in no particular order)

Zhao Hangbin, Su Chang, Xie Xudong, Xu Yihua, Li Feng, Li Yanchun, Xu Fan, Liao Qian, Wu Xuming, Zhao Wangsheng, Shi Huaping, Yue Huawei, Zhang Bo, Wang Bo, Zhang Yaodong, Cheng Yongchun, Shi Zhenzhong, Zhu Lihua, Zhong Ke, Xu Fangxin, Huang Qisheng, Wang Zisheng, Qian Yurong, Liu Kailin, Sun Xuhong, Cao Bo, Li Yun, Chen Jinhua, Du Bo, Li Wei, Wang Ge, Liu Daxing, Yao Xiaofeng, Fan Zhichao, Chen Zhixiong, Zhang Tao

Contents

Introduction

The AI-powered Wireless Local Area Network (AI WLAN) provides key fundamental wireless network support for the new industrial revolution by deeply integrating artificial intelligence (AI) with wireless network technology. Through intelligent, adaptive, and highly efficient data transmission, AI WLAN significantly enhances social productivity, driving improvements in the efficiency of production and daily life. With the continuous evolution of WLAN technology, its applications have become increasingly pervasive across work, production, and daily activities, leading to an exponential growth in the volume of data carried by wireless networks. Traditional static management approaches can no longer meet the demands for high density, low latency, and high reliability. The introduction of AI enables WLAN networks to perceive environmental changes in real time, predict interference, and dynamically optimize resource allocation, substantially improving user experience. As the "capillaries" in the information network, the AI-powered intelligence of WLAN serves as the key support for the new industrial revolution.

This research report explores the key AI technologies applied in WLAN networks and demonstrates, through real-world Application Scenarios, how AI addresses core challenges such as interference suppression, channel fluctuation, and load balancing.

Chapter 1

The Core Value of AI in WLAN Networks

1.1 Introduction to AI Technology

AI technology enables machines to possess the capabilities of perception, learning, reasoning, and decision-making by simulating human intelligence. Currently, AI technology represented by Machine Learning can dig patterns and optimize decision from big data, providing new solutions for the efficient operation and maintenance, performance optimization, security protection, resource management, and experience enhancement of WLAN networks.

The core paradigms of AI technology include Supervised Learning, Unsupervised Learning, Reinforcement Learning, and Generative Artificial Intelligence. Their corresponding features are as follows:

- Supervised Learning: As the most widely-applied technical branch, supervised learning can accurately predict network changes and optimize decisions through classification and regression prediction models. For example, time-series prediction based on Received Signal Strength Indicator (RSSI) can assist device roaming decisions and optimize handover latency and stability.
- Unsupervised Learning: Through clustering and dimensionality reduction techniques, it can efficiently process high-dimensional, unlabeled data. For instance, it can compress and extract features from Channel State Information (CSI), reducing air-interface transmission overhead and improving spectrum utilization. It can also identify interference and attacks in the wireless environment, enhancing security and privacy protection.
- Reinforcement Learning: Aiming at dynamic optimization, reinforcement learning realizes policy iteration through the interaction between the agent and the environment. It is applied in fields such as Radio Resource Management (RRM) and network parameter adaptive adjustment, significantly enhancing the network's resilience in complex environments.
- Generative Artificial Intelligence: The generative technology represented by Large Language Models (LLMs) can reconstruct the Human-Computer Interaction (HCI) mode. For example, network configuration can be automated through natural language instructions, lowing the barrier in operation and maintenance.

AI technology can also be classified into traditional machine learning and neural networks based Deep Learning. In practice, these two types of methods can complement each other. Traditional machine learning is generally based on clear statistical inference, with good stability, interpretability and low computing power requirements. However, its model capabilities are often limited, and it is good at handling simple data with a small amount of data. Deep learning has a powerful non-linear fitting ability, is good at handling complex problems with a large amount of data, but has high requirements for training data and computing power resources.

1.2 The Value of AI to WLAN

From relatively traditional methods such as K-means, decision trees, and probabilistic graphical models to emerging methods based on neural networks such as DRL, Transformer, and LLMs, they can help WLAN solve various problems. The following table lists some exploration directions of AI technology applications in WLAN.

Table1 Application Fields of AI Technology in WLAN

AI Technology	Application Fields
K-means	Pattern recognition, such as wireless environment classification
Decision Tree	Pattern recognition, such as device health analysis, coverage hole identification
Probabilistic Graphical Models	Closed-loop problem solving, problem root-cause analysis
Mutual Information	Wireless channel feature engineering
ARIMA	Anomaly Detection in Time-Series Data of WLAN Device Logs
Logistic Regression	WLAN device health analysis
Bayesian inference	Pattern recognition, such as abnormal terminal analysis; WLAN AP location planning
LSTM、RNN	Channel environment prediction, MCS prediction
DRL	Wireless parameter tuning, such as RRM parameter adjustment
Transformer	Predictive analysis, natural language processing
KNN	Generation of dynamic thresholds for wireless algorithms
CNN	Interference detection, wireless path loss simulation
LLM	Accurately understanding human intentions, providing auxiliary analysis and strategy support for planning, operation and maintenance, security, and problem location

These AI technologies (such as machine learning, deep learning, and reinforcement learning) bring the following core advantages to WLAN networks:

- Real-time Optimization: Dynamically adjust channels, power, and modulation methods to cope with environmental changes.
- Interference Suppression: Identify and avoid co-channel/adjacent-channel interference, and improve the Signal-to-Noise Ratio (SNR).
- Performance Prediction, Analysis, and Processing: Predict network congestion risks through historical data analysis and take preventive measures.
- Reliability Enhancement: Anomaly detection and reliability improvement.
- New Business: WLAN perception.

Chapter 2

Key Technologies and Application Scenarios: How AI Optimizes the WLAN Experience

In the field of network communication, the intelligent features of AI provide new ideas for optimizing the WLAN experience. By learning network status in real-time, predicting user behavior, and dynamically adjusting resource allocation, AI can significantly improve the WLAN's throughput, stability, and coverage efficiency.

AI can enable the all-around improvement of WLAN from the physical layer to the MAC layer. At the physical layer, AI can enhance the ability to cope with channel interference and changes in wireless signal propagation conditions, reduce the overhead of CSI information transmission, and improve the accuracy of transmission parameter selection. At the MAC layer, AI can improve the estimation of service requirements, predict dynamic traffic, achieve dynamic link management, optimized channel access, and resource scheduling, and save WLAN network energy consumption. At the network control and management level, AI can simplify the operation and maintenance of network.

2.1 Interference Detection and Suppression

In densely-deployed scenarios, a large number of WLAN devices share the 2.4GHz/5GHz frequency bands, resulting in severe interference in the WLAN spectrum.

AI Solutions:

- Interference Detection:
  - Interference Type and Degree Judgment: Collect a large amount of signal data in normal and interference states as a training set, and use algorithms such as Support Vector Machine (SVM) and decision trees for training to establish an interference detection model. This model can be used to judge in real-time whether there is interference and the type and degree of the interference.

- Interference Hybrid Small Model: Use multi-dimensional information such as the number of time-domain interferences & signal strength, frequency-domain interference bandwidth & Signal-to-Interference plus Noise Ratio (SINR), and spatial-domain interference direction & user MAC address as input parameters. Adopt a Convolutional Neural Network (CNN) to automatically extract the features of WLAN signals and establish a systematic interference model.

- Performance Enhancement:
  - Deep Learning-based Channel Allocation: Use deep learning algorithms such as Deep Neural Network (DNN) or Convolutional Neural Network (CNN) to automatically learn and optimize the WLAN channel allocation according to multi-dimensional information such as interference conditions, user distribution, and service requirements in the environment. The model can take the environmental features as input and output the optimal channel allocation scheme to avoid interference and improve spectrum utilization.
  - Reinforcement Learning-based Power Control: Through the reinforcement learning algorithm, let the agent (such as an access point AP) interact with the environment to learn. The agent takes different transmit power actions according to the current interference state and its own power control strategy, and then obtains reward feedback from the environment (such as improved communication quality or reduced interference). Through continuous learning and optimization, the agent can find the optimal power control strategy to reduce interference to other devices while ensuring communication quality.
  - Artificial Intelligence-assisted Beamforming: Use artificial intelligence algorithms such as linear regression and ridge regression in machine learning to calculate the optimal beamforming vector according to information such as the location of interference sources in the environment and signal propagation characteristics. By adjusting the transmission beam direction of the antenna, the signal is concentrated towards the target user, while suppressing the signal transmission in the direction of the interference source, thereby improving the signal quality and reducing interference.

Application Scenarios: In densely-populated office or home environments with a large number of WLAN device interferences, AI WLAN can improve the user network experience rate.

2.2 Link Adaptation

Wireless channels are severely affected by multipath effects, human body/obstacle occlusion, hand-held terminal shaking, and movement, resulting in significant fluctuations in signal quality.

AI Solutions:

- Channel Sensing:
  - Channel Fluctuation Classification: Take the collected channel features such as signal strength, phase, and Doppler shift as input, train the SVM model, and divide the channel state into categories such as stable, slightly fluctuating, and severely fluctuating. Find the optimal classification hyperplane in the high-dimensional space to achieve accurate classification of channel fluctuations.
  - Channel Fluctuation Judgment: Construct an ensemble learning model composed of multiple decision trees using multiple channel features such as signal strength, signal-to-noise ratio, and multipath delay spread. This can improve the accuracy and stability of channel fluctuation detection, effectively handle the interaction between features, and reduce the overfitting.

- Performance Enhancement:
  - Deep Learning-based MCS Prediction: Use deep learning models (such as CNN, LSTM, etc.) to predict the channel state and automatically select the appropriate Modulation and Coding Scheme according to the prediction results.
  - Reinforcement Learning-based MCS Selection: Through the reinforcement learning algorithm, the agent continuously adjusts the MCS according to the current channel state and transmission effect. After each decision, the agent obtains rewards based on indicators such as the bit error rate and throughput fed back by the receiving end. Through continuous learning and optimization, it finds the optimal MCS adjustment strategy to improve the air-interface performance.
  - AI-based Beam Tracking: In the case of user terminal movement, use AI technology to achieve real-time beam tracking. By learning and predicting the terminal movement trajectory and quickly responding to channel changes, the agent can timely adjust the MIMO beamforming direction and shape of the WLAN signal to ensure that the beam always accurately points to the moving terminal, reducing signal interruption and quality degradation.

Application Scenarios: In office or home environments where people walk/pass through or hand-held mobile phones shake/move, AI WLAN can improve the user network experience rate.

2.3 Physical Layer Parameter Adaptation

The requirements of real-time services are difficult to be met in a dynamic and complex wireless environment. Traditional WLAN rate selection algorithms (such as the Minstrel algorithm based on signal strength, historical packet loss rate or the detection of a small amount of data) are difficult to cope with the dynamic and complex wireless environment, and have the following problems:

- Sub - optimal Rate Selection: Static thresholds cannot adapt to real-time changes such as channel interference, and it is easy to select too high (frequent retransmission) or too low (wasting bandwidth) rates.
- High Latency and Jitter: Frequent rate switching or retransmission will increase the end-to-end latency.
- Inefficient Selection Process: Random sampling and the expansion of the optional physical layer transmission parameter set brought about by the evolution of standard protocols lead to large optimization space faced by the classical algorithm, and further lead to a slow convergence in the parameter selection, affecting real-time services such as video calls , VR and etc.

AI Solutions:

- Environment Sensing and Feature Extraction:
  - Model rate selection as a Markov Decision Process (MDP), and train the agent through algorithms such as Q-learning and DQN:
    - State: Current channel quality, device load, interference level, etc.
    - Action: Select the best combination of transmission parameters, including MCS, GI, Nss, etc.
    - Reward: Maximize throughput, minimize latency or packet loss rate.
  - Through continuous interactive learning, the agent can quickly adapt to the dynamic environment. Relevant progress includes the DARA (Data-driven Algorithm for Rate Adaptation) algorithm.

- Federated Learning (FL) Cooperative Optimization:
  - In multi - AP scenarios, jointly train a global model through FL, share local experience (such as channel occupancy patterns), and avoid the local optimization problem of single-node decision-making.
  - The AI-driven rate selection algorithm significantly enhances the reliability, efficiency, and adaptability of WLAN through a closed loop of real-time perception, decision-making, and optimization, and it is especially suitable for business scenarios with dynamic, dense, or high real-time service requirements.

Application Scenarios: Home network intelligent anti-interference. Smart home devices are intermittently interfered by microwave ovens and Bluetooth. The online learning model can identify the interference cycle, actively reduce the data transmission speed during the interference peak period, and increase the data transmission speed during the stable period, improving the average throughput.

2.4 Multi-Link Load Balancing and Resource Allocation

In the scenario of multi-link operation (MLO) of an access point (AP), traditional WLAN relies on RSSI (Received Signal Strength Indicator) to associate users with the links, and there exists the problem of unbalanced load. In addition, the simultaneous access of multiple devices in the network, along with the fluctuations in bandwidth requirements and traffic patterns, pose further challenges to multi-link management. The introduction of AI technology can help optimize the load balancing and resource allocation of multiple links, thus enhancing the user experience.

AI Solutions:

- Cluster Analysis of User Requirements:
  - Group users according to service types (video, IoT, games) through classifiers or clustering algorithms (such as K-means) and assign different priorities. The AI system can monitor the network usage of each link in real-time, including bandwidth requirements, traffic types, and device workloads. Through machine learning and deep learning, AI can identify the traffic patterns of different devices and applications and predict future bandwidth requirements, thereby optimizing traffic allocation. For example, AI can automatically recognize different types of traffic such as video streams, games, web browsing, and downloads, and map high - priority applications (such as video calls or high-definition video streams) to specified links and allocate more bandwidth through TTLM.

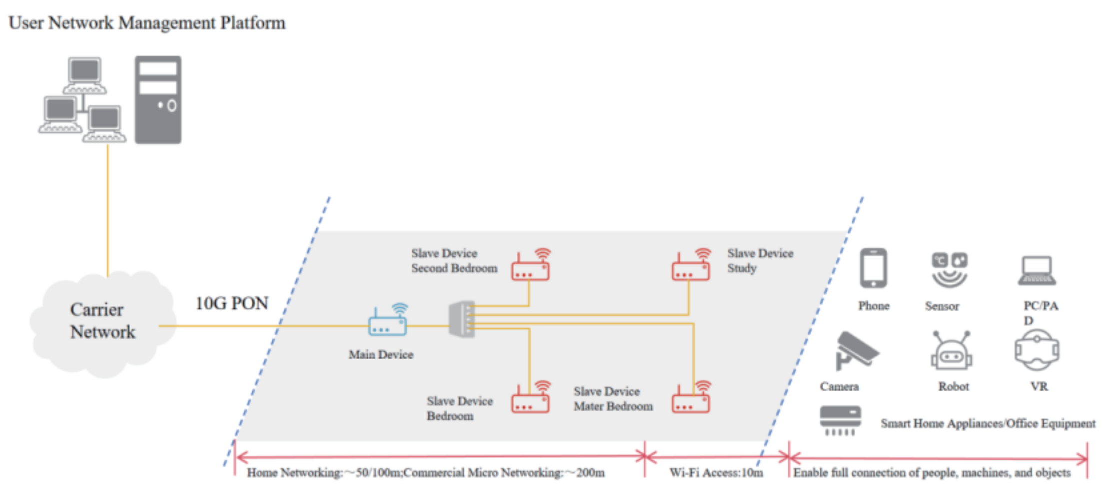
- Dynamic Bandwidth Allocation:
  - Deep learning models can predict traffic peaks and allocate resources for each link in advance. AI-based network multi-link load balancing and resource allocation technology can ensure that each device obtains appropriate resources according to its priority and requirements by dynamically adjusting the bandwidth and traffic in the home network. AI can adaptively adjust the routing according to the link quality and current network conditions to avoid network congestion and ensure the smooth operation of various devices and applications of users.
  - When multiple links in the network (such as 2.4GHz and 5GHz WLAN links, or even wired connections) are in use, the AI system can intelligently allocate traffic to each link according to the real-time network conditions and the loads of different links, thereby avoiding over-congestion of a certain link.

Application scenarios:

Predict traffic through AI and perform load balancing to ensure that the latency in each room is less than 10 milliseconds, or guarantee the throughput and latency of rooms with key businesses. In a home network, multiple smart devices are simultaneously connected to the network. Suppose there is a smart TV playing a 4K video stream at home, a phone with high-definition video call, another phone with an online game, and other devices are browsing the web or downloading files. Traditional network link management methods may not be able to effectively handle the situation of multiple devices and large traffic volumes. As a result, certain applications (such as video calls or games) will be affected by other background traffic, causing stuttering or latency. By introducing an AI-based multi-link load balancing system, AI can monitor the bandwidth usage of each device in real time, and automatically select links and adjust the traffic transmission according to priorities. For example, when the smart TV is playing high-quality video, AI can map its services to low-load links and allocate more bandwidth to it, ensuring smooth video playback, while reducing the impact on other low-bandwidth applications (such as web browsing) and enhancing the overall user experience.

2.5 AI-based Service Scheduling Optimization Strategy

The WLAN network experience is a systematic issue, and the scheduling strategies of the entire link will be reflected in the WLAN experience to some extent. Currently, WLAN networking forms include single WLAN devices, FTTR, AC+AP, etc. The FTTR networking is shown in the following figure. The connection between the central office and the main gateway, as well as between the main and slave devices, generally adopts PON networking. However, the existing DBA scheduling strategy of the PON network is usually determined by the ONU's application first and then the OLT's allocation, without considering the wireless air-interface factors. This makes the scheduling strategy lack global consideration, unable to predict the network state of the entire link and develop targeted progressive optimization strategies, resulting in a waste of link resources and further deterioration of the network.



Note: FTTR equipment comprises master devices and slave devices.

Figure 1 Typical FTTR Networking Diagram

AI Solutions:

- AI-Predicted Optical and Air-Interface Joint Service Guarantee Strategy:
  - Full-Link Traffic Collection and Prediction: Collect the traffic size and service model information of the full link in historical time periods as training set data, and use algorithms such as SVM, decision trees, and convolutional neural networks for training to establish an optical-wireless service traffic prediction model, which can predict the traffic and service scale on optical and wireless links in real-time within a certain period in the future.
  - User Quantity Prediction: The change in the number of users not only affects air-interface scheduling but also tests the rationality of the scheduling algorithm for allocating optical link resources. Adopt algorithms such as linear regression and decision trees to analyze the historical number of users and obtain the predicted value of the future number of users. Provide the prediction and trend of the number of users under each ONU covering the entire network.
  - Air-Interface Quality Prediction and Optical Path Strategy Compensation: Each node of the network link affects the user experience of terminal users, but not all nodes of the network link can be optimized, or some nodes have been optimized to the best. A linkage mechanism between the WLAN environment and optical link scheduling can be established. When the air-interface environment of a certain ONU or area deteriorates to a certain extent, the user experience can be compensated by optimizing the optical link scheduling. Collect the connection quality of the user network everywhere and obtain the user network quality model for future time through algorithms such as time-series analysis. When it is predicted that the air-interface quality is about to deteriorate (such as retransmission, packet loss, high latency, etc.), a more aggressive network guarantee target is selected in the optical link scheduling strategy.
  - Prediction-based Optical and Air-Interface Joint Scheduling: Compared with the traditional "application-first-then-allocation" algorithm, according to the service model and network target predicted by AI, combined with the prediction of the wireless air-interface quality, actively guide the dynamic adjustment of optical link scheduling in advance. Through the prediction and dynamic adjustment strategy, reduce the packet loss and latency of the entire network link, and to a certain extent, reduce the bandwidth waste caused by fixed allocation, making the strategy more in line with the actual traffic requirements of the current network and enhancing the network-using experience of wireless users.

Application Scenarios: In scenarios with periodic tidal changes in traffic, the number of terminals, and services (such as small restaurants, shops, etc.), as well as in cases where the uplink is long (such as FTTR, long-distance all-optical networking, etc.), the optimization of the uplink scheduling strategy based on AI can improve the adaptability of the network to actual services; through proactive and advanced optimization, enhance the timeliness of network optimization; during peak periods, increase the throughput of the entire network.

2.6 WLAN Sensing

WLAN sensing is mainly used for motion detection, target recognition, positioning and tracking, etc. Common motion detections include human intrusion detection, fall detection, etc. There are also more precise motion detections, including the detection of biological quantities such as breathing and heartbeat, as well as gesture recognition. WLAN sensing can achieve personnel identification and authentication based on CSI (Channel State Information) fingerprints. In addition, positioning and tracking have always been hot research topics in WLAN sensing, and the application scope covers indoor and outdoor navigation. High accuracy is a very challenging goal in these applications. Taking the application of detecting the presence of people as an example, dynamic environments (such as the rotation of fans, the fluttering of curtains, the movement of pets, etc.), wireless interference, and movements in adjacent rooms or other floors can all lead to misjudgments in intrusion/presence detection. In addition, the CSI when a human body is motionless is similar to the CSI characteristics of an empty room under Gaussian noise, which makes it easy to misjudge the presence of people as the absence of people. AI-based methods have unique advantages in aspects such as human behavior detection, human health detection, construction of virtual digital spaces, and security detection.

AI Solutions:

- Personnel Activity Sensing:

Based on sensing algorithms such as LSTM and Transformer, the dynamic RSSI, CSI, and other wireless signals within a certain period can be subjected to feature extraction and classification, so as to more accurately describe the environmental features and determine whether events such as personnel entry and falls occur.

  - Dataset Construction: During the wireless communication process, synchronously collect and label data such as CSI in different human activity scenarios, including scenarios such as no one, someone walking, standing, lying down, etc., to form a model training dataset.
  - Model Training: Based on high-quality training data and AI algorithm models such as LSTM, the static and dynamic features of the wireless channel can be combined to more accurately and reliably sense different types of personnel activities. Through the efficient collection of data in different scenarios and time periods during the communication process and the fine-tuning of model training, the scene adaptability and accuracy of AI sensing can be continuously improved. Use a convolutional neural network (CNN) to extract the feature differences between the dynamic environment, wireless device interference, and the effects of human movement on CSI and radar waveforms. Pre-process the waveform data into a two-dimensional matrix and input it into the model. Extract key features through convolution and pooling operations. Through sample training, the CNN can accurately identify waveform differences, filter out noise signals, and improve the sensing accuracy.

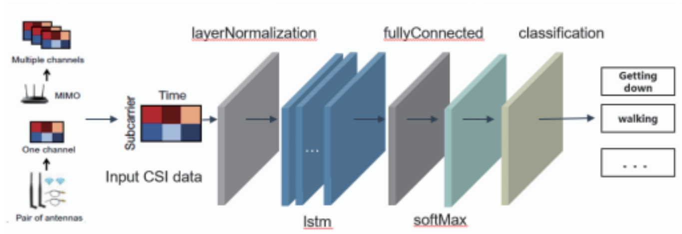


Figure 2 Human Behavior Analysis Based on LSTM

- Respiratory Signal Sensing

The rise and fall of the chest during breathing will cause fluctuations in the CSI (Channel State Information) signal. The Recurrent Neural Network (GRU, Gated Recurrent Unit) is used to process the sensed waveform to capture the subtle periodic changes generated by the rise and fall of the chest caused by human breathing. Through the simplified structure of the update gate and reset gate, the GRU can effectively retain the time-dependent features related to breathing, while reducing computational complexity and resource consumption. It can filter out random noise or aperiodic interference and focus on the respiratory features in weak signals, thus identifying whether there is a stationary human body in the space.

- Optimization of Transmit Power and Sensing Sensitivity

Based on the DDPG (Deep Deterministic Policy Gradient) reinforcement learning algorithm, use information such as the current space size, neighbor AP information, the number of misjudgments (from user feedback), the thickness of building obstacle materials, and the user communication experience as state inputs, and set the transmit power adjustment ratio and sensing sensitivity adjustment coefficient as action outputs. Construct a reward mechanism with a reduction in the number of misjudgments as a positive reward and a decline in the user communication experience, an increase in misjudgments, or insufficient signal coverage as a negative reward. Pre - process the state data and input it into the model. Through sample training, the DDPG model can accurately identify the optimal action in different environmental states, adaptively adjust the transmit power or sensing sensitivity, effectively filter out interference caused by signal penetration, reduce the error detection rate, and enhance the accuracy of the WLAN sensing system while ensuring the basic communication experience of users.

- Multi-device Fusion Sensing

Although AI+CSI can achieve motion sensing, it may not be able to achieve precise sensing in scenarios such as building occlusion and complex electromagnetic environments. In the FTTR architecture, the slave gateway is connected to the master gateway through a branch optical fiber. The master gateway is equipped with a service module, an optical fiber sensing module, a wireless signal sensing module, and an AI fusion algorithm module. The service module realizes communication services. The optical fiber sensing module collects the intensity and phase of the optical fiber scattered light signal through the branch optical fiber. The wireless signal sensing module collects CSI signals. The AI fusion algorithm module fuses and analyzes the features of the optical fiber scattered signals and CSI signals to realize the perception of the external environment. The module configuration of the master gateway can also be deployed on the slave gateway according to resource conditions.

Application Scenarios:

- AI algorithms enable WLAN sensing technology, upgrading it from a communication network to an intelligent environmental sensing platform, promoting the implementation of sensing-related applications in home, campus, and hotel scenarios, and endowing WLAN networks with new value.
- In a test environment with external device and personnel interference, the LSTM-based model can achieve the detection of personnel entry with an accuracy of over 93%. In hotel, home, and other environments, it can enhance security protection and health monitoring capabilities while better protecting user privacy.

2.7 Green Energy Saving

In traditional WLAN networks, energy - saving measures mainly rely on static configurations, such as regularly turning off some modules and reducing link rates. This approach lacks flexibility, has difficulty adapting dynamically to network load changes, often sacrifices performance or QoS, and has limited energy - saving effects.

AI Solutions:

- Business Environment Sensing:
  - Business Characteristics: The AI module can identify business types and load intensities by conducting real-time analysis of packet content, traffic features, and port activity. It uses deep-learning models (such as CNN, RNN, etc.) or traditional machine-learning models (such as decision trees, SVM, etc.) to achieve this. Based on the prediction of the next - period business and the requirements of the Service-Level Agreement (SLA), it designs energy-saving strategies for wireless transmission according to business characteristics.
  - Network Environment: The AI system can perceive the device operating environment, user activity status, and physical network status through a trained model, and then dynamically adjust the device's energy-saving mode. It also analyzes user behavior patterns and historical traffic data. During periods of predicted low traffic, it can proactively put the access point (AP) or link into a low - power mode or hibernation mode in advance, improving energy efficiency.

- Dynamic Energy - Saving Measures:
  - On - Demand Energy Provision: Depending on the business type, it decides whether to enable a high - speed transmission channel or maintain low-power operation. Without affecting the QoS requirements of the business, it adaptively adjusts the transmit power, selects the optimal Modulation and Coding Scheme (MCS), optimizes scheduling resources to enhance link quality and spectrum utilization, and maximizes energy efficiency.
  - Intelligent Sensing and Adaptive Adjustment: By leveraging AI technology, the router (in the FTTR architecture, it can be the master gateway) collects device data to construct a device - state database. This database encompasses information such as the number of connected devices, WLAN traffic data, LAN port traffic, network-environment changes (signal-interference level, channel quality), device-state data, and device power consumption. The AI model deployed in the router (in the FTTR architecture, on the master gateway) analyzes the data in the device-state database, generates energy-saving control instructions, and sends them to the corresponding devices. By continuously learning the daily usage patterns of the router by users,

the AI model can accurately analyze the device's usage frequency and business requirements in different time periods. When it detects that the device is in a low-activity state, it automatically adjusts the router to a low-power mode. When there is a high-activity demand, it automatically wakes up the device or restores its normal performance, effectively avoiding unnecessary energy consumption. In a campus scenario, the FTTR master gateway, as the data hub, continuously collects the PON-port traffic data, WLAN traffic data, and LAN-port traffic of each slave gateway. It also monitors the device status, power consumption, and alarm information of the routers, constructing a device-state database covering the entire campus. Considering the large number of network nodes and high data - transmission volume in the campus network, the master gateway is equipped with a lightweight AI model with feature dimensionality reduction and model compression. This ensures efficient processing of massive data while reducing its own operating energy consumption. The AI model conducts in-depth analysis of the database and formulates energy-saving strategies for the complex scenarios of multiple regions and multiple services in the campus.

- Intelligent Prediction and Intelligent Wake-up: With the help of AI's analysis of the router's historical usage data, it can predict the future usage requirements of users. For example, based on the user's daily fixed Internet-surfing habits, it wakes up the router from standby mode and restores its normal working state in advance before the user's commonly used Internet-surfing time period, ensuring that the network can respond quickly when the user uses it. When the user has not used the router for a long time, it automatically enters a deep standby state, further reducing energy consumption.
- Device-to-Device Cooperative Energy Saving: AI enables the router to interact with other smart devices in the home (such as smart lights, air conditioners, TVs, and home appliances connected to smart sockets). When no one is at home, through information interaction with smart door locks or other sensors, it can detect that all family members are out. Then it automatically turns off or reduces the WLAN Over-the-Air Transmit Power, leaving only the lowest - power mode for IoT devices. In the campus AC+AP structure or the FTTR architecture, the master gateway analyzes the data in the device-state database through the AI model, generates energy-saving control instructions, and sends them to the slave gateways. It can turn off some slave gateways and high-energy-consuming devices (such as NAS and smart TVs in standby mode). At night during low-traffic periods, according to the reduction in network usage, it automatically makes the WLAN AP enter a low-power mode, reduces the signal strength or turns off the 5GHz frequency band. The master gateway also reduces the CPU frequency, decreases background processing tasks, and enters the night-time energy-saving mode through timed scheduling, automatically restoring the normal mode in the morning. In low-load application scenarios, it adjusts the router's power consumption according to the traffic threshold. For example, it reduces the WLAN Over-the-Air transmission power, makes some WLAN APs hibernate, and only maintains the basic connection of the master gateway. At the same time, it only activates low-power IoT-exclusive channels (such as SparkLink, Thread, Zigbee, BLE).

Application Scenarios:After the introduction of AI, WLAN network devices can achieve more refined and intelligent ECO energy-saving control through business sensing and environmental sensing, achieving energy-saving beyond the traditional ECO mode.

2.8 Simplifying Network Operation

Wireless network sites, especially enterprise - level network sites, involve a large number of network devices, and network indicators and configurable parameters are numerous and complex. Wireless networks themselves have invisible and uncertain characteristics. IT personnel using traditional operation and maintenance methods need cumbersome problem-reproduction and problem-location processes, which makes problem-analysis difficult. The configuration of network devices and human-machine interaction also require professional knowledge, posing high requirements for users, and the usability is relatively poor.

AI Solutions:

- Network Fault Sensing/Analysis/Troubleshooting:
  - Terminal-Experience Classification: Real - time monitor key wireless - network indicators of terminals, such as the packet-loss rate and RSSI. Through classification models (such as random forest, neural network, etc.), classify the usage experience of terminals to achieve rapid fault reporting and automatic fault reporting for unmanned devices.
  - Fault Root - Cause Analysis: Collect network data for faults and, in combination with classification models and other algorithms, explore the root causes behind the faults to help operation and maintenance personnel quickly troubleshoot.

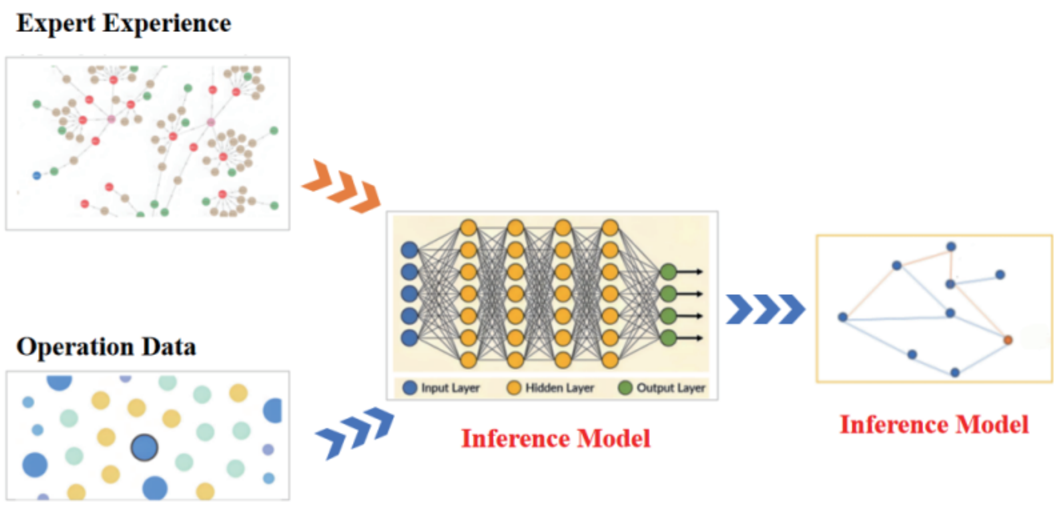


Figure 3 Root - Cause Analysis of Wireless Problems Based on AI Troubleshooting

- Voice Control: Embed AI algorithms in network devices such as gateways and routers to enable voice - input control functions. This allows for easy adjustment of WLAN configuration parameters. Then, control instructions are sent to terminal devices connected to the WLAN network through the WLAN channel, achieving the goal of controlling network devices via voice commands. This interaction method is more natural and convenient, improving the user experience.

- Machine Learning: AI can optimize the performance and security of network devices through machine-learning algorithms. For example, AI can predict future traffic trends based on historical data and adjust the configuration of network devices in advance. Or, it can detect potential security threats through anomaly-behavior detection algorithms and take timely measures.
- Virtual Assistant: AI can implement the virtual-assistant function, providing personalized services and support to people. For example, when a problem occurs, the virtual assistant can help people quickly locate the problem and provide solutions. When technical support is needed, the virtual assistant can offer remote assistance.
- Predictive Maintenance: AI can predict potential problems of network devices through predictive-maintenance algorithms and take preventive measures in advance. This maintenance method can greatly reduce the failure rate and downtime.
- Intelligent Recommendation: AI can recommend relevant services or products based on users' historical behaviors and preferences. This recommendation method can help people better meet their needs and improve the user experience.
- Personalized Settings: AI can set relevant parameters or configurations according to users' personalized requirements. This setting method enables network devices to better adapt to different scenarios and needs, improving the user experience.

In summary, AI technology can greatly simplify the operation of network devices by humans and improve performance in terms of operation and maintenance efficiency, user experience, and security.

Application Scenarios:
In a factory network environment, when an AGV cart experiences wireless-network failures due to interference, the operation and maintenance platform can identify and locate the fault point. Operation and maintenance personnel can quickly locate the root cause of the fault through AI troubleshooting, restoring production efficiency. In a densely-populated office or home environment, users can use AI-enabled voice and gestures to achieve more user-friendly network-device configuration.

2.9 CSI Data Compression and Feedback

In WLAN, CSI (Channel State Information) feedback is crucial for the optimization of MIMO (Multiple-Input Multiple-Output) technology. For Single-User MIMO, CSI can help reduce the interference between multiple spatial data streams of a single user when using a linear receiver. For Multi-User MIMO, CSI can help reduce the interference of data streams between users. However, traditional compression methods (such as codebook quantization) face a trade-off between high feedback overhead and accuracy loss, especially in scenarios of large-scale antenna communication (such as 8x8 MIMO communication and joint beamforming communication), where the efficiency is low.

AI Solutions:

- Compression Feedback Based on Traditional Machine-Learning Algorithms:
  - Rely on unsupervised learning and use clustering algorithms (such as the K-means algorithm) to further compress the compression results provided by the protocol.
- End-to-End Compression Model:
  - Rely on supervised learning and use lightweight autoencoders. At the terminal (STA), compress the CSI matrix (such as the complex-valued H matrix), and decode and restore it at the AP side. The autoencoder can adaptively extract sparse features, compress high-dimensional CSI data into low-dimensional representation data, and then reconstruct it at the receiving end. This can effectively remove redundant information, achieve a higher compression ratio, and reduce the reconstruction error. CSI data may have specific structures, such as spatio-temporal-frequency correlations. Deep-learning methods (such as multi-dimensional convolutional neural networks and recurrent neural networks) can be used to effectively capture these complex non-linear relationships, accurately model and compress CSI, and effectively reduce feedback overhead while enhancing reconstruction accuracy.
    - Encoder: Deployed at the STA, it maps high-dimensional CSI to low-dimensional hidden vectors.
    - Decoder: Deployed at the AP, it reconstructs the CSI through multiple neural-network layers.

- Dynamic Adaptation Mechanism:
  - Based on meta - learning and transfer-learning, train the model to make it adaptable to multiple scenarios (indoor/outdoor, static/moving), reducing performance fluctuations caused by environmental changes. In different scenarios (such as indoor, outdoor, different frequency bands), the distribution of CSI data may vary greatly. Using transfer-learning technology, the model learned in one environment can be adapted to another environment, providing an efficient cross-environment adaptation solution for CSI compression and improving the effectiveness and flexibility of compression.

- Joint Optimization Design:
  - Jointly train CSI compression with downstream tasks (such as beamforming/joint beamforming and resource allocation) to maximize system throughput rather than simply minimizing the reconstruction error.

Application Scenarios:
In large-scale MIMO scenarios, AI-based compression can significantly reduce the CSI feedback volume while keeping the beam-forming gain error within a reasonable range. In high-frequency bands (6 GHz/mmWave), it can achieve fast channel tracking. By compressing time-varying CSI through CNN or temporal convolutional networks (TCN), millisecond-level channel-state updates can be realized.

2.10 Channel Access

8In a WLAN network, the basic access method at the Medium Access Control (MAC) layer used by stations (STA) is the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism known as the Distributed Coordination Function (DCF).

Before sending data, an STA must listen to the wireless medium. If the medium is detected to be continuously idle for the minimum specified duration (such as the Distributed Inter-Frame Space, DIFS), the STA is allowed to send a frame after waiting for an additional random back-off period. The generation range of the random back-off counter is limited by the contention window. The initial range is set from 0 to the minimum contention window size (CWmin). If the medium is occupied, the contention window will be doubled successively until it reaches the maximum contention window size (CWmax), further reducing the collision probability.

Over the past decade or so, there has been an inevitable development trend in the WLAN field towards continuously increasing throughput and strictly controlling latency. To achieve these goals, an efficient channel-access protocol is of great importance. Relevant analyses show that the current competition-based access protocols face significant performance degradation in densely - deployed scenarios. The binary exponential back-off mechanism can cause short-term unfairness, that is, continuous transmission failures will lead to further degradation of latency performance. Therefore, a more efficient channel-access scheme is needed to increase throughput while reducing latency and jitter.

Furthermore, in high-density, multi-service-concurrent Wi-Fi networks, different services (XR/AR, 4K video conferencing, industrial IoT, ordinary web browsing, etc.) have extremely different requirements for latency, jitter, and throughput. The traditional fixed or static-tuned Enhanced Distributed Channel Access (EDCA) mechanism has difficulty simultaneously meeting the three goals of low-latency guarantee, maximum throughput, and fair access, and often faces the following pain points:

- Service Congestion: VO/VI services are blocked by BE/BK queues, resulting in choppy video and intermittent audio.
- Insufficient Adaptability: When the number of users or channel quality changes suddenly, the channel - access parameters cannot be adjusted in a timely manner, causing performance fluctuations.
- Lack of Fairness: Individual traffic over-occupies the air interface, impairing the overall experience.

AI Solutions:

Single - Device:

- High-Priority Service Sensing and Access Protection:
  - The device side uses AI technologies such as XGBoost, DNN, CNN, and LSTM+attention mechanism to distinguish between business categories such as XR, voice, AR, IoT, and BE and their real-time QoS requirements by combining traffic time-series, QoS tags, and packet-size features.
  - The device side maps the data flow into the corresponding queue according to the AI-recognition results. Furthermore, the STA side can also decide whether to initiate P-EDCA competition to obtain higher-priority access based on AI recommendations.
- Traffic-Intensity Prediction and Intelligent Scheduling Mapping:
  - Extract important traffic-related statistical features such as the average frame length and burstiness of the device on the current channel over a period of time as features. Use DNN and other lightweight deep-neural networks to estimate the load intensity of each business queue in the short term based on the feature - extraction results, guiding access-category mapping and spatio-temporal resource allocation.
  - The device on the station side dynamically selects the main-channel access or non-main-channel access based on the traffic-prediction results provided by DNN and other lightweight neural networks to obtain the best service experience.

Multiple Devices:

- Multi-AP Channel Access:
  - Centralized Channel-Access Decision-Making: Each AP uploads necessary statistical parameters such as traffic, packet-error rate, and duty cycle to the edge controller. The edge controller uses artificial-intelligence algorithms such as genetic algorithms, simulated annealing algorithms, and DQN algorithms to make decisions and issue the optimal channel-access strategy configuration for the entire network, achieving collision avoidance and load balancing among multiple APs.
  - Incorporate available means such as non-main-channel access, dynamic OFDMA RU allocation, and dynamic EDCA parameter configuration into the decision-making set for joint consideration. Through reinforcement learning, construct an optimal model for specified goals (such as maximum capacity, minimum latency, etc.) from end to end.

Application Scenarios: In a high-density deployment environment, AI-based channel access can identify diverse services in real-time and dynamically adjust parameters to ensure the smooth operation of all services.

Chapter 3

AI and the Future Air-Interface Coordination of WLAN

Currently, the IEEE 802.11 working group has established two standard task groups, TGbn and TGbq, which are aimed at WLANs in the Sub-7GHz and millimeter-wave frequency bands respectively. Although these two standards are still in the initial stage of development, the proposals of the standard task groups offer a glimpse into the prototype of future WLANs. Future WLANs are expected to include a series of new features such as multi-AP air-interface coordination and millimeter-wave beam management. These advanced features involve more complex multi-AP management and beam management, expanding the optimization and management space to new dimensions. AI is of great value in addressing the issues in these new dimensions.

3.1 Coordinated Beam Forming

The Coordinated Beam Forming (CoBF) technology enables multiple Access Points (APs) to work together to precisely direct signals, reducing the impact of APs on non-target Stations (STAs) during concurrent spatial reuse and thus enhancing the overall network performance. The application of CoBF technology mainly involves two steps: obtaining Channel State Information (CSI) and adjusting and coordinating the transmitted signals based on the CSI.

In IEEE 802.11bn, only two APs are supported for CoBF. The schematic diagram is as follows:

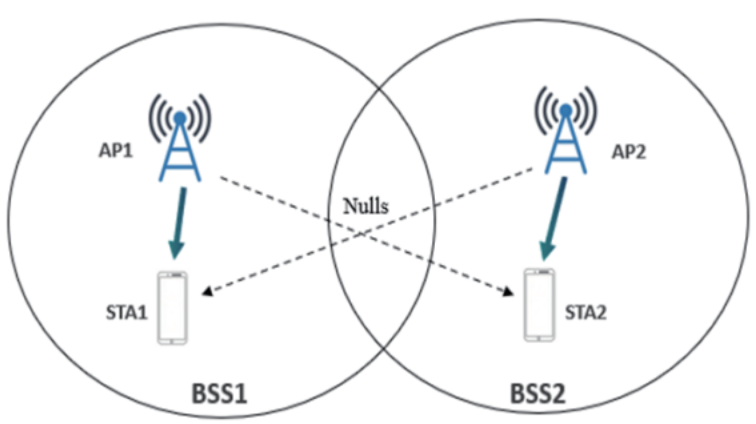


Figure 4 Schematic Diagram of CoBF in IEEE 802.11bn

Different from the situation in Chapter 2.9 where the Access Point (AP) only obtains the Channel State Information (CSI) of the Stations (STAs) within its Basic Service Set (BSS), in the Coordinated Beamforming (CoBF) mechanism, the AP also needs the CSI of the STAs in other BSSs. The following figure shows a candidate process for obtaining CSI in CoBF, called sequential NDP based sounding. In this process, AP1 not only acquires the CSI of associated STA1 but also that of STA2 associated with AP2. Similarly, AP2 obtains the CSI of both STA1 and STA2.

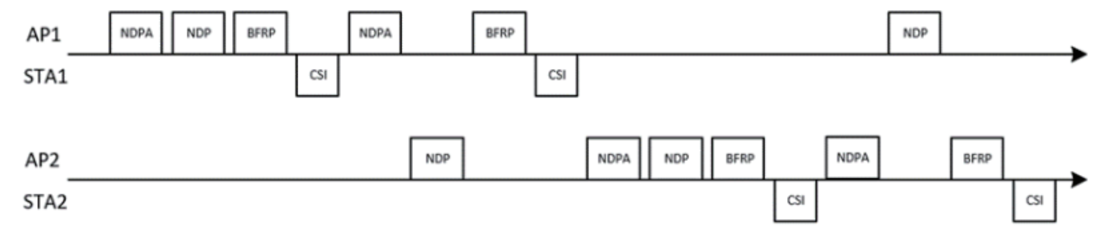


Figure 5 Sequential NDP Based Sounding for CoBF in IEEE 802.11bn

Taking AP1 as an example, after obtaining the CSI information of STA1 and STA2, AP1 processes the CSI. When transmitting signals, it can not only enhance the signal for STA1 but also reduce the interference to STA2. This process is reciprocal for the users of AP1 and AP2, enabling both sides to achieve a larger channel capacity.

From the working principle of CoBF described above, it can be seen that CSI is crucial for beam forming in CoBF. This is because it is necessary to adjust the direction and shape of the beam according to the real-time channel state and use Nulling (i.e., zero-forcing or beam nulling, which means actively suppressing signals in a specific direction to form a zero point of signal strength, thereby suppressing interference or noise) to reduce or weaken the impact of APs on non-target STAs during concurrent spatial reuse. The accuracy of Nulling directly determines the final performance of CoBF, and the change in CSI caused by channel changes will weaken Nulling, thus seriously affecting the performance of CoBF. AI can be used to predict channel changes and optimize CSI feedback.

AI Solutions:

- AI technology can compress CSI, effectively reducing the air-interface overhead required for CSI feedback. For details, refer to 3.9.
- In view of the time-varying channel characteristics, Recurrent Neural Networks (RNNs) are used to predict the future CSI state, improving the real-time performance and robustness of the system. For example, time-domain CSI prediction can efficiently model historical channel data and dynamically adjust the feedback frequency, which can effectively reduce redundant information transmission and improve the prediction accuracy.
- Combining online learning with incremental inference and dynamic parameter update strategies, the model parameters are updated in real-time and dynamically to maintain high-precision CSI prediction and adapt to rapid channel changes. For example, the Long Short-Term Memory (LSTM) model can gradually infer the channel state under multi-modal inputs (such as location, image, vision, sensor data, etc.) and adjust the beam direction in real-time.
- Generative Adversarial Networks (GANs) are used to generate high-fidelity CSI features. Especially in low-bit quantization scenarios, GANs can generate samples highly similar to the real data distribution, effectively retaining the spatio-temporal correlation and non-linear features of CSI. Through adversarial training to enhance the data, the prediction robustness and the generalization ability of the prediction model can be effectively improved.
- Deep learning is used to efficiently process high-dimensional CSI data. For example, Convolutional Neural Networks (CNNs) and Transformers, combined with the terminal location, can capture, extract, and predict the

temporal and spatial characteristics of CSI, and adaptively predict and generate the optimal precoding matrix or parameters. For instance, when adjusting the transmitted signals according to the CSI, AI can be used to adaptively predict, integrate, and adjust the CSI of the target terminal and other terminals, enabling the transmitted signals to be in-phase superimposed at the target terminal to enhance the signal strength. At the same time, for non-target areas and idle terminals, the AP will reduce or stop signal transmission to avoid energy waste and interference.

Application Scenarios:

In scenarios that require multi-AP coordinated coverage, such as dynamic environments (e.g., indoor factories or hospitals with moving obstacles causing sudden channel changes) and high-density terminal environments (e.g., stadiums, conference centers), using AI for CSI prediction can improve the prediction accuracy and robustness, reduce the computational and pilot overhead, adapt to complex environments, and have better real-time performance, effectively enhancing the overall performance of CoBF.

3.2 Coordinated Spatial Reuse

The 802.11bn standard introduces a new feature of Coordinated Spatial Reuse (Co-SR). The goal of this technology is to achieve concurrent transmission by coordinating and controlling the transmission power of multiple APs, thereby making more efficient use of channel resources. A Co-SR transmission is initiated by an AP that has obtained a Transmission Opportunity (TXOP), which becomes the Sharing AP. The AP selected by the Sharing AP for coordinated concurrent transmission is called the Shared AP.

The performance of Co-SR technology highly depends on high-precision interference management algorithms, which involve the selection of Shared APs and power regulation. Traditional heuristic methods are less efficient in complex topologies and dynamic environments and sometimes have difficulty achieving optimal selection, which may lead to mutual interference or resource waste.

AI Solutions

- Intelligent Identification of Sharing Relationships:
  - Based on deep learning, collect the geographical locations, path losses, historical interference information, channel state information, etc., of each AP and its associated STAs, and effectively identify multiple groups of APs with parallel transmission capabilities and their target STAs , and divide them into a sharing group.
  - Use federated learning to share interference model parameters among multiple APs, which can improve the accuracy of global decision-making.

- Power Control Optimization:
  - DRL models power control as a Markov Decision Process (MDP). The state space includes states such as the signal interference level between AP and STA, buffer size, traffic priority, target STA, and estimated SINR. The action space is the power level. The agent optimizes the power setting through interaction with the environment, minimizing the interference between Shared APs and maximizing the total throughput .
  - Incrementally train the model through online learning to adapt to newly added APs or dynamic environments, which reduces the recalibration overhead of traditional methods.

- CoSR Scheduling Optimization:
  - Through online training, the errors in measurement data such as power and received signal strength can be effectively reduced. Thus, when performing CoSR scheduling, more accurate parameters can be selected.
  - By detecting multi-dimensional information such as the number of retransmissions and packet error patterns, it is possible to effectively distinguish packet error scenarios caused by random backoff collisions, hidden nodes, etc., avoiding incorrect inputs to CoSR scheduling due to such abnormal situations and improving the accuracy of scheduling parameters.

Application Scenarios:

In multi-AP scenarios, using AI to dynamically select Shared APs and optimize power can enhance the system capacity and avoid latency jitter caused by interference for real-time services like VR.

3.3 TXOP Sharing

The 802.11bn standard introduces a new feature of Coordinated Time-Division Multiple Access (Co-TDMA), which allows an AP (Sharing AP) to share a part of the TXOP it has obtained with another AP (Shared AP) belonging to a certain AP set for transmitting one or more PPDUs, thereby enhancing the access opportunities of the Shared AP. In a scenario with densely deployed multi-APs, the rational selection of shared APs and optimal time-slot allocation require comprehensive awareness of network load, traffic priorities, and link quality, along with real-time decision-making. Traditional approaches often suffer from inefficiency or unfair resource distribution.

AI Solutions:

- Dynamic Time Slot Allocation:
  - Use the LSTM model to model sequences such as historical time slot utilization and service arrival rate, and combine the attention mechanism to enhance the ability to capture bursty service patterns, predicting future shared resource requirements in advance .
  - Adopt Deep Reinforcement Learning (DRL): Through the competition - cooperation mechanism, learn the optimal time slot request strategy to maximize the overall network utility function, such as weighted throughput and fairness.
- QoS-Aware Scheduling:
  - Adopt hierarchical reinforcement learning. The high-level layer decides the service priority (for example, video has priority over web pages), and the low-level layer fine-tunes the time slot position and length to meet the delay sensitivity of different services .
  - Adopt an offline pre-training and online deployment scheduling strategy. Through large-scale training in the offline stage, the model can adapt to multiple network environments. In the online stage, there is no need for search and iteration, and the decision-making efficiency is high.

Application Scenarios:

In a high-concurrency IoT environment, such as during the morning rush hour in a smart building, some APs face severe congestion. After introducing the AI-optimized Co-TDMA mechanism, the system can dynamically share spatio-temporal resources with low-load and high-priority Shared APs, ensuring low-latency access for critical services (such as security cameras and voice communication).

3.4 Seamless Roaming

In 802.11bn, a new context sharing mechanism is introduced, which enables the STA to minimize the delay and data interruption during the roaming process without dissociating from the associated state(State 4). Although the standard provides a basic mechanism framework, in practical industry applications, due to the significant differences in the roaming behaviors of STAs from different manufacturers and models, the fixed roaming guidance strategies on the AP side have inconsistent effects on different terminals and cannot effectively guarantee the roaming experience.

- User Profiling
  - Network Profiling: Use DRL (Q-Learning) to learn the roaming strategies of different networks. Different networks have different characteristics, and using a fixed guidance strategy cannot most efficiently guide terminal roaming. Use an AI model to abstract the effective roaming guidance strategies corresponding to various networks into different network roaming profiles. For example, an AI model can learn the roaming guidance baseline threshold under a specific network topology.
  - Terminal Profiling: Use neural network algorithms (LSTM-CNN algorithms) to learn the roaming strategies of different terminals and provide different guidance strategies for terminals with different roaming performances. Use an AI model to learn and analyze terminal behaviors and abstract different terminal roaming profiles, with each terminal roaming profile corresponding to a set of personalized roaming guidance strategies.
  - Network portraits and terminal portraits can be drawn based on AI. For each terminal, a corresponding network portrait can be created, and through the coordination between the network and the terminal, the best roaming effect can be achieved.

- Predicting the Target AP:
  - Both APs and STAs will conduct target screening in the list of roaming candidate APs. However, when there are many APs in the network and they are at different distances from the STA, the screening value of each candidate AP is different. By predicting the probability of the STA switching to each AP, the candidate efficiency can be improved, the number of measurements can be reduced, and the roaming decision can be completed more quickly. Deep neural networks (DNNs) can be utilized to predict the likelihood of an STA roaming to each candidate AP. Decision-making factors include inter-AP signal strength, historical STA roaming patterns, and AP load status. This approach aims to maximize the STA's average Modulation and Coding Scheme (MCS), Received Signal Strength Indicator (RSSI), etc., while minimizing the number of roaming handovers.

Effect: In a densely networked environment, for the next-hop roaming of the STA, an optimal target AP list and the best roaming threshold are provided without manual configuration. Using roaming profiles to guide terminals can reduce terminal roaming problems.

3.5 Integrated Millimeter-Wave Beam Management and Tracking

AI can enable the beam management and high-low frequency coordination of the 11bq integrated millimeter-wave system.

Millimeter-wave (mmWave) communication relies on high-gain narrow beams to achieve high-speed transmission, but beam management faces problems such as high beam search overhead, difficult dynamic channel adaptation, and low high-low frequency coordination efficiency. AI (deep learning, reinforcement learning, federated learning, etc.) can significantly optimize the beam management process and make intelligent predictions of link quality and coordinate high-low frequency links based on the channel state.

Problems:
Traditional beam search relies on exhaustive scanning (such as the 11ad SSW beam scan), resulting in high initial access and handover latency, being difficult to adapt to mobile scenarios, and causing high overhead problems for fast handovers between high and low frequency links.
The low-frequency (Sub-7GHz) band has wide coverage but low capacity, while millimeter-wave beams have high gain, low interference, and high capacity but are prone to interruption. Traditional dual-connection handover strategies are static and cannot be optimized in real-time.
Millimeter-wave beams are narrow (3°-15°) and easily blocked. Traditional fixed beam widths cannot adapt to dynamic environments (such as changes in pedestrian flow and vehicle flow).

AI Solutions:
- AI-Assisted Beam Search Trigger Prediction:

Based on terminal-side sensing-assisted beam prediction: For vehicle terminals, on - board radar can be used; for mobile phones, camera data can be used to obtain 3D radar point clouds for environmental perception and beam optimization. For WLAN APs, digital twins can be used to simulate the wireless environment of a local area, and AI can predict channel changes (such as rain attenuation, movement of obstacles) within the next 100ms and adjust beam parameters in advance. For WLAN mobile STAs, motion sensors and AI trajectory prediction can be used, and through a multi-modal fusion model (CNN+RNN), the time of occlusion can be predicted in advance, and the beam can be adjusted accordingly.

- AI-Driven Beam Range Prediction (Beam Coverage Prediction):
  - Beam preselection based on historical data: Use LSTM/Transformer to analyze the user's movement trajectory and historical channel states (such as CSI, AoA/AoD), predict the optimal beam direction, and reduce the scanning range. Based on the estimated user location, speed, and learned environmental features (such as the distribution of building walls), Top-K candidate beam screening can be performed to reduce the scanning time.
  - AI dynamic beam width adjustment: Adopt Reinforcement Learning (RL) to dynamically adjust the beam width according to channel quality (such as SNR, RSSI) and user distribution (for example, wide beams for initial search and narrow beams for data transmission).

- AI Establishing Cross-Band Channel Characteristic Mapping:
  - Use low-frequency channel information (such as CSI) to assist millimeter-wave beam training, and model the spatial correlation between high and low frequency channels through Graph Neural Networks (GNNs) to reduce the millimeter-wave beam search time.

Application Scenarios:

- In stadiums, AI adjusts the beam coverage based on the real-time pedestrian flow heat map to ensure no blind spots in high-density user areas.
- In WLAN millimeter-wave communication for unmanned aerial vehicles, AI predicts the attitude deviation caused by wind and dynamically adjusts the beam direction.
- In AI-powered high-low frequency coordinated beam management, the Sub-6GHz link provides coarse-grained beam indication, and the millimeter-wave link can quickly complete fine alignment.

3.6 Co-RTWT

Collaborative R-TWT (Co-RTWT) enables the Access Point (AP) to coordinate its R-TWT scheduling with the Overlapping Basic Service Set (OBSS) AP, and/or extend protection to the R-TWT scheduling of the OBSS AP. Co-RTWT requires coordinating the R-TWT scheduling of multiple APs to reduce channel conflicts. In complex environments, how to optimize the scheduling to improve the multi-AP collaborative efficiency, reduce latency, enhance reliability, and save power consumption is an urgent problem to be solved. Specifically, the 11bn Co-RTWT mechanism faces the following technical challenges in high-density, high-dynamic, and multi-service scenarios:

- Dynamic channel resource competition
  - Multi-terminal collaborative wake-up is prone to cause channel contention. Traditional scheduling relies on manual rules and cannot adapt to bursty traffic and multi-link interference in real time.
- Ultra-dense device conflict suppression
  - When the density of IoT nodes increases (e.g., > 1000 nodes/m²), the implicit topological correlation between terminals may lead to the overlap of wake-up windows, significantly increasing the probability of conflicts.
- Trade-off between multi-link energy efficiency and throughput
  - Multi-Link Operation (MLO) requires synchronous scheduling of wake-up times across frequency bands. The power consumption of terminals fluctuates due to frequent link switching, and it is difficult to optimize the global energy efficiency.
- Quality of Service (QoS) guarantee in mobility scenarios
  - The randomness of the trajectory of mobile terminals leads to rapid changes in the network topology, and fixed scheduling rules cannot guarantee low-latency and highly reliable connections.

AI solutions:

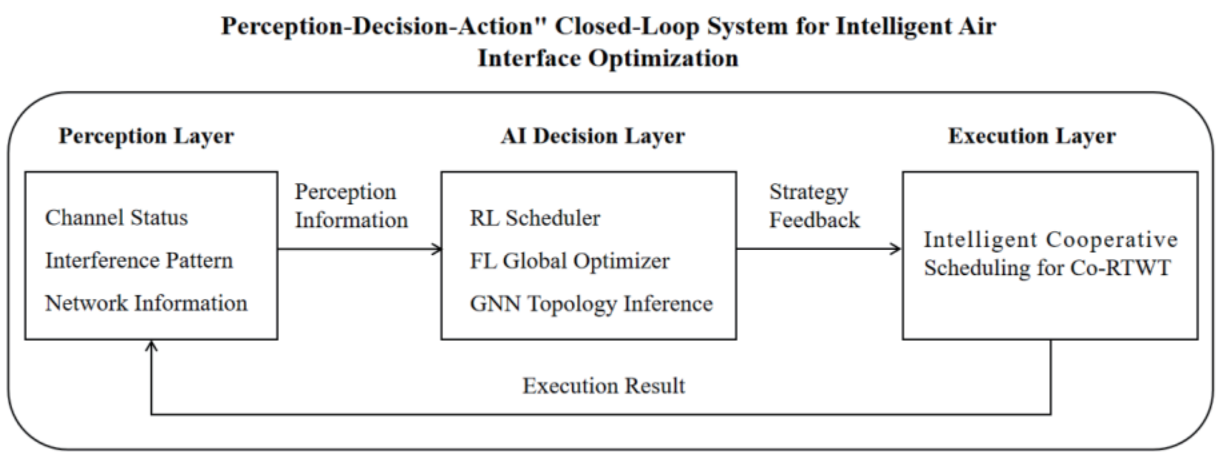


Figure 6 Perception-Decision-Action" Closed-Loop System for Intelligent Air Interface Optimization

- Reinforcement learning-driven collaborative wake-up
  - Input: Multi-dimensional environmental states (channel load matrix, device service priority tags, interference spectrum characteristics).
  - Model: Build an intelligent agent based on the Proximal Policy Optimization (PPO) algorithm, output the spatio-temporal distribution strategy of the TWT window, and achieve non-overlapping allocation of wake-up times.
  - Advantage: Break through the static constraints of the rule engine and dynamically avoid channel competition caused by bursty traffic.

- Global energy efficiency management through federated learning
  - Architecture: Based on the horizontal federated learning framework, the AP and terminals jointly train a lightweight energy efficiency model, and the terminals upload encrypted gradient parameters
  - Strategy: In the MLO scenario, allocate appropriate wake-up frequency band priorities to terminals (e.g., low-power links respond to control signaling with priority), and balance the global energy consumption.

- Conflict prediction using Graph Neural Networks (GNN)
  - Modeling: Abstract the terminal network into a heterogeneous graph, where node attributes include terminal location, service cycle, and the number of historical conflicts, and edge weights represent physical distance and interference coupling strength.
  - Inference: Identify terminal groups with high conflict risks through the Graph Attention Network (GAT) model, and guide the AP to allocate Co-RTWT parameters differentially.

Application scenarios:

The deep integration of AI technology and Co-RTWT (Collaborative Responsive-Traffic Indication Message Window Time) marks a paradigmatic shift in wireless networks from being "rule-driven" to "intelligence-embedded." Through technologies such as reinforcement learning, graph networks, and federated learning, dynamic resource optimization, conflict probability suppression, and cross-link energy efficiency balancing can be achieved in high-density scenarios, laying a low-power and highly reliable wireless connection foundation for scenarios such as Industry 4.0, the metaverse, and smart cities.

3.7 New Channel Access Mechanisms

The 802.11bn standard provides new channel access mechanisms, including non-primary channel access and high-priority EDCA (P-EDCA) access:

- Non-Primary Channel Access: The Non-Primary Channel Access (NPCA) mechanism provided by 11bn aims to address the problem of low utilization of secondary channels when the primary channel is congested. This technology is designed to increase throughput and reduce access latency by dynamically using non-primary channels to optimize spectrum utilization.

The working mechanism is as follows: when the primary channel is busy, the device switches to an idle non-primary channel to continue backoff and transmission; when the primary channel is idle, the device switches back to the primary channel.

- P-EDCA Access: The Priority EDCA (P-EDCA) provided by 11bn is an enhanced version of the EDCA mechanism. It is designed to reduce the tail of the access delay distribution of low-latency AC_VO traffic (other Application Scenarios are to be determined). Ultra-High Reliability (UHR) Stations (STAs) using P-EDCA will balance the impact on stations that do not use P-EDCA through specific rules.

The working mechanism is that when a station has high-priority traffic, it can rely on the P-EDCA mechanism to send a Defer Signal with a shorter arbitration interval to initiate P-EDCA contention, thereby obtaining the right to use the wireless channel more preferentially.

AI Solutions:

- Channel Access Decision:
  - Use deep reinforcement learning (DRL), including algorithms such as DQN and its derivative DDPG, to obtain the optimal channel access parameter combination. The reward function integrates "key service latency, total throughput, and fairness index" to ensure that different services benefit simultaneously.
  - Introduce self-supervised policy smoothing during the DRL tuning process. Through adaptive thresholds, it avoids air interface oscillations caused by frequent jumps.

Chapter 4

Challenges and Directions

4.1 Challenges from Computing Power, Compatibility, and Generalization

4.1.1 Computing Power Limitations of Edge AI

In recent years, with the continuous development of artificial intelligence (AI) technology, more and more WLAN performance optimization functions can be realized with the assistance of AI algorithms. This has led to an increasing demand for local computing power resources on the network edge side, especially in wireless routers. Traditional embedded processors are generally limited in single-precision floating-point operation capabilities, memory bandwidth, and parallel computing efficiency, making it difficult to support the real-time operation requirements of complex models in practical scenarios. To effectively break through the computing power bottleneck and promote the wide application of edge intelligence in WLAN scenarios, it is urgent to conduct systematic optimizations at multiple levels, including computing architecture design, model optimization strategies, and software-hardware collaboration mechanisms.

- Model Compression: In view of the resource-constrained characteristics of edge devices, the system comprehensively introduces AI model compression technology in the model design and deployment stages to reduce the model's storage overhead and inference computing load. The main methods include low-bit quantization and extremely low-bit quantization (such as 1-bit weight networks), which significantly compress the model size while maintaining accuracy as much as possible; knowledge distillation, which migrates knowledge to construct smaller and more efficient inference models; and sparsification and pruning mechanisms, which further reduce redundant computing paths and parameter storage requirements. Through these compression strategies, after shrinking the model, it becomes possible to deploy it on edge devices, especially router devices, which can significantly reduce the average time consumption and power consumption of a single inference and improve the overall system operation efficiency and deployability.
- End-Cloud Collaboration: To balance inference performance, system energy efficiency, and service flexibility, the system constructs an end-cloud collaboration mechanism based on a hierarchical intelligent architecture. This mechanism preferentially executes core tasks with low latency and high frequency locally, while migrating computationally intensive or cross-domain integration complex tasks to the cloud for processing, forming a multi-level computing power collaboration model of local fast response+cloud deep computing.

- Network-Wide Optimization: In the process of the in-depth integration of AI and WLAN and the evolution towards intelligence, the application of AI computing power is no longer limited to local inference of single-point devices but is gradually expanding to distributed intelligent inference capabilities with network-wide collaboration. By deploying lightweight computing power units with perception and inference capabilities on multiple wireless routers and terminal nodes, the system can continuously perceive the network operation status and make data-driven dynamic decisions. Relying on distributed AI models to conduct real-time analysis of multi-dimensional data such as traffic distribution, user access behavior, and interference characteristics in the WLAN network, the system can dynamically optimize key parameters such as channel allocation, transmit power control, and user association guidance on the control plane, achieving intelligent resource orchestration and on-demand scheduling. Under the actual conditions of limited spectrum resources and increasing access density, such intelligent mechanisms have significantly improved the overall network access efficiency, service quality, and environmental adaptability, promoting the evolution of WLAN networks towards a self-optimizing and self-adaptive intelligent architecture.

In actual deployments, the system can dynamically evaluate the resource requirements of various tasks based on real-time indicators such as device computing power status, power consumption thresholds, network bandwidth, and latency, and make intelligent scheduling and offloading decisions. Through this collaborative approach, the system can ensure service quality while achieving the optimal allocation of inference tasks and continuously improving the overall energy efficiency ratio.

Compared with traditional wireless algorithms, AI technology has higher requirements for storage capacity and computing power. Therefore, when designing the deployment architecture, it is necessary to fully consider the communication, computing, and storage overhead it brings. Currently, WLAN devices are mainly divided into APs, Controllers, and cloud servers. An AP is an edge device responsible for radio signal transceiver. Air interface data such as CSI is directly obtained by the AP. Deploying underlying algorithms on the AP can save communication overhead in the network. However, APs are cost-sensitive and generally have limited computing and storage capabilities, making them unsuitable for deploying large models. The Controller is the information and control center of the WLAN. Centralized optimization algorithms across APs are most efficiently deployed on the Controller. Cloud servers currently collect network information and provide advanced functions for user interaction. They have strong computing and storage capabilities and can be flexibly expanded. However, they have high communication costs and uncontrollable latency and are suitable for deploying algorithms with non-real-time algorithms such as AI assistants. As AI technology penetrates deeper into WLAN, the deployment of algorithms will become more and more refined. It may be the future development direction to hierarchically split the data flow or model of a function and let the AP, Controller, and cloud handle the parts they are good at respectively.

4.1.2 Compatibility between New AI-Supported Devices and Old Devices without AI Support

- WLAN wireless networks are gradually introducing intelligent adjustment mechanisms based on artificial intelligence to achieve more efficient spectrum management, interference avoidance, network load balancing, and quality of service assurance. Such mechanisms usually rely on AI-enhanced WLAN routing devices with local inference capabilities and may extend or enhance the current WLAN protocol stack (such as the 802.11 series), thereby introducing intelligent decision-making capabilities in the control plane and data plane. In this context, there is a significant differentiation in device capabilities at the device level: traditional WLAN terminal devices cannot directly participate in AI inference or policy collaboration due to the lack of AI collaboration interfaces and computing power support; while new AI terminals can have deeper state intercommunication and collaborative optimization with AI routers. The resulting problems of protocol compatibility and performance heterogeneity have become one of the key technical challenges for the implementation of AI in WLAN scenarios.

- Nevertheless, at the system level, AI-enhanced WLAN routers can indirectly optimize the network performance of traditional devices through perception and scheduling mechanisms. For example, through real-time learning and dynamic scheduling of the network-wide traffic status, AI routers can provide traditional terminals with better access channels, more stable bandwidth guarantees, and less interference, thereby achieving passive performance improvement without changing the terminal hardware.

- In contrast, new AI-sensing terminal devices (supporting AI signaling collaboration or having local collaboration mechanisms) can achieve finer-granularity network state collaboration and resource allocation when collaborating with AI routers, further enhancing connection stability, load balancing efficiency, and QoS assurance capabilities, and obtaining a significantly better performance improvement space than traditional devices.

In summary, for traditional devices, AI can enhance the overall network performance through perception and intelligent scheduling mechanisms. For new-type devices, AI can drive the WLAN network architecture towards a more intelligent, adaptive, and efficient evolution.

4.1.3 Generalization of Sensing Capabilities in Complex Scenarios

As can be seen from Section 2.6, AI WLAN sensing technology has achieved a large number of research results. However, due to the diversity of application scenarios and the complexity of WLAN device deployment environments, most of the current WLAN sensing algorithms are dedicated customized solutions. To meet actual needs, WLAN devices usually need to integrate more than ten sensing algorithms, usually significant development and operation and maintenance burdens. To address this issue, we turn our attention to large model technology, hoping to simplify the algorithm design and system integration process with its unified modeling and generalization capabilities.

Looking back at the development history of natural language processing, there were also multiple technical routes and scattered applications in its early stage. The rise of large models has broken this situation. Through a unified architecture and large-scale pre-training, complex algorithm design problems have been transformed into computing power-driven problems, greatly improving the technical level. Currently, the data foundation of AI WLAN sensing is gradually being improved. In the future, it is expected to build a general WLAN sensing pre-training model based on large model methods. WLAN sensing also has the potential to become an important part of the multi-modal capabilities of large models, further promoting the development of WLAN sensing technology and accelerating the implementation and explosion of the integrated communication and sensing industry.

4.2 Future Development Prospects

Short-Term Capabilities: Realize the configuration and optimization of intelligent capabilities.Users can achieve parameter configuration and optimization with a one-click minimalist operation. The air interface of the 802.11be system can be deeply integrated with AI to achieve sub-millisecond-level scheduling and service guarantee.

Medium-Term Capabilities: Intelligent capabilities will complete expert-level parameter configuration and optimization, conduct long-term on-site observation, carry out performance enhancement and fault handling by time period and business. As the WLAN technology evolves towards the new air interfaces of IEEE 802.11bn/bq, intelligent capabilities will be combined with the features of the new air interfaces to further enhance the network performance.In the future, as WLAN moves towards millimeter waves, AI will play a significant role in both millimeter-wave beam management and high/low-frequency link management in WLAN 802.11bq.

Long-Term Capabilities: Network devices can adaptively share intelligent experience, conduct joint learning and optimization, predict and prevent faults in advance, conduct AI-customized negotiations on the custom parts of the protocol, and generate optimal protocol behaviors and rules for local scenarios among intelligent devices, such as channel competition rules and collision resolution rules. AI-powered spectrum sharing can achieve efficient spectrum coordination and transmission collaboration within WLAN and between WLAN and other wireless systems.

Conclusion

The new industrial revolution has put forward many new requirements for access network technologies. For example, the volume of data processing has increased dramatically, and there are strict requirements for real-time performance and reliability. It also requires support for the access of a large number of terminals and the collaboration of multiple access technologies. Among them, WLAN represented by technologies such as WiFi, SparkLink, Bluetooth and Zigbee plays a key role. For instance, WiFi undertakes 70% of the access traffic and 31% of IoT connections, and Bluetooth also undertakes 29% of IoT connections. Therefore, improving the performance of WLAN such as bandwidth, latency, reliability, and the number of concurrent terminals is of great significance for the new industrial revolution.

AI enables WLAN to shift from "passive response" to "active optimization", solving the problem that traditional algorithms cannot meet the complex and time-varying wireless environment and wireless services. It significantly improves reliability, efficiency, and user experience, meets the requirements of new-quality productivity for bandwidth, latency, and stability, and becomes a key data path for the new industrial revolution. The gradual deepening of the application of AI technology in communication networks will transform traditional network elements from functional units to Intelligent Agent. Through the collaboration among Intelligent Agent, traditional communication networks will be transformed into intelligent networks.

This report summarizes the current technical directions of AI applications in WLAN networks. These technical directions will be iteratively updated over time. For these technical directions, it is recommended to implement them gradually through technical standards, testing and certification, etc., so as to effectively promote the digital transformation of the industry and the development of the new industrial revolution.

Abbreviation Table

Abbreviation	Full Name
AGV	Automated Guided Vehicle
ARIMA	Autoregressive Integrated Moving Average Model
Co-SR	Coordinated Spatial Reuse
Co-BF	Coordinated Beamforming
Co-RTWT	Cooperative Restricted Target Wake-up Time
Co-TDMA	Coordinated Time-Division Multiple Access
CSI	Channel State Information
CSMA	Carrier Sense Multiple Access
DCF	Distributed Coordination Function
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
DQN	Deep Q-Network
DARA	Data-Driven Rate Adaptation algorithm
DDPG	Deep Deterministic Policy Gradient
EDCA	Enhanced Distributed Channel Access
FTTR	Fiber to the Room
GAN	Generative Adversarial Network
GAT	Graph Attention Network
GNN	Graph Neural Network
IoT	Internet of Things

Abbreviation	Full Name
LLM	Large Language Model
LSTM	Long Short-Term Memory
MIMO	Multiple-Input Multiple-Output
MDP	Markov Decision Process
MLO	Multi-Link Operation
MCS	Modulation and Coding Scheme
OLT	Optical Line Terminal
ONU	Optical Network Unit
PON	Passive Optical Network
PPO	Proximal Policy Optimization
P-EDCA	Priority-Enhanced Distributed Channel Access
QoS	Quality of Service
RNN	Recurrent Neural Network
RRM	Radio Resource Management
RSSI	Received Signal Strength Indicator
KNN	K-Nearest Neighbor algorithm
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TCN	Temporal Convolutional Network
TWT	Target Wake-up Time
UHR	Ultra-High Reliability
XGBoost	Extreme Gradient Boosting

References

[1] Lazaridou, Angeliki, and Marco Baroni. "Emergent multi-agent communication in the deep learning era." arXiv preprint arXiv:2006.02419 (2020).

[2] Zou, Hang, Qiyang Zhao, Lina Bariah, Mehdi Bennis, and Merouane Debbah. "Wireless multi-agent generative AI: From connected intelligence to collective intelligence." arXiv preprint arXiv:2307.02757 (2023).

[3] ZS. Ji, Q. Wang, S. Wu, J. Tian, X. Li and W. Wang, "Deep learning based user grouping for FD-MIMO systems exploiting statistical channel state information," in China Communications, vol. 18, no. 7, pp. 183-196, July 2021.

[4] W. Zhou, T. Zhu, D. Ye, W. Ren and K. -K. R. Choo, "A Concurrent Federated Reinforcement Learning for IoT Resources Allocation With Local Differential Privacy," in IEEE Internet of Things Journal, vol. 11, no. 4, pp. 6537-6550, 15 Feb.15, 2024.

[5] H. Xiang, J. Peng, Z. Gao, L. Li and Y. Yang, "Multi-Agent Power and Resource Allocation for D2D Communications: A Deep Reinforcement Learning Approach," 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall), London, United Kingdom, 2022, pp. 1-5.

[6] F. Lu, F. Yuan, Y. Li, X. Song, C. Liu and F. Liu, "5G VoNR Traffic Real-time Prediction Method based on LSTM," 2023 5th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qiangdao, China, 2023, pp. 915-919.

[7] Q. Cui, Z. Zhang, Y. Shi, W. Ni, M. Zeng and M. Zhou, "Dynamic Multichannel Access Based on Deep Reinforcement Learning in Distributed Wireless Networks," in IEEE Systems Journal, vol. 16, no. 4, pp. 5831-5834, Dec. 2022.

[8] R. Huang, M. Guo, C. Gu, S. He, J. Chen and M. Sun, "Toward Scalable and Efficient Hierarchical Deep Reinforcement Learning for 5G RAN Slicing," in IEEE Transactions on Green Communications and Networking, vol. 7, no. 4, pp. 2153-2162, Dec. 2023.

[9] https://www.ieee802.org/11/Reports/tgbn_update.htm

[10] https://www.ieee802.org/11/Reports/tgbn_update.htm

[11] Xia, Dong, Jonathan Hart, and Qiang Fu. "Evaluation of the Minstrel rate adaptation algorithm in IEEE 802.11 g WLANs." In 2013 IEEE International Conference on Communications (ICC), pp. 2223-2228. IEEE, 2013.

[12] Sammour, Ibrahim, and Gerard Chalhoub. "Evaluation of rate adaptation algorithms in IEEE 802.11 networks." Electronics 9, no. 9 (2020): 1436.

[13] Mortaheb, Matin, Mohammad A. Amir Khojastepour, Srimat T. Chakradhar, and Sennur Ulukus. "Deep Learning-Based Real-Time Rate Control for Live Streaming on Wireless Networks." In 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN), pp. 263-267. IEEE, 2024.

[14] Queirós, Rúben, Eduardo Nuno Almeida, Helder Fontes, José Ruela, and Rui Campos. "WLAN rate adaptation using a simple deep reinforcement learning approach." In 2022 IEEE Symposium on Computers and Communications (ISCC), pp. 1-3. IEEE, 2022.

[15] F. Qi, J. Guo, Y. Cui, X. Li, C. Wen and S. Jin, "Deep Learning-Based CSI Feedback in WLAN Systems," 2024 5th International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), Wuhan, China, 2024, pp. 43-48, doi: 10.1109/ISCEIC63613.2024.10810137.

[16] 802.11-23/0275r2, Improved AIML Enabled Index Based Beamforming CSI Feedback Schemes

[17] 802.11-23/0290r2, Study on AI CSI Compression

[18] S. Szott, K. Kosek-Szott, P. Gawłoicz, J. T. Gómez, B. Bellalta, A. Zubow, F. Dressler, "WLANMeets ML: A Survey on Improving IEEE 802.11 Performance with Machine Learning," IEEE Communication Surveys & Tutorials, Vol.24, Issue 3, Juen 2022

[19] Z. Guo, Z. Chen, P. Liu, J. Luo, X. Yang and X. Sun, "Multi-agent reinforcement learning-based distributed channel access for next generation wireless networks", IEEE Journal on Selected Areas in Communications, Vol. 40, Issue 5, May 2022

[20] Y. Ding, S. Liew and T. Wang, "Non-Uniform Time-Step Deep Q-Network for Carrier-Sense Multiple Access in Heterogeneous Wireless Networks", IEEE Transactions on Mobile Computing, Vol. 20, Issue 9, Sept 2021

[21] Ali Z A, Abduljabbar Z H, Tahir H A, et al. eXtreme gradient boosting algorithm with machine learning: A review [J]. Academic Journal of Nawroz University, 2023, 12(2): 320-334.