

Spectrum Utilization Analytics and Forecasting for Open RAN Deployment

Sriker Reddy Palla
Dish Network LLC, USA

ARTICLE INFO

Received: 01 Nov 2025

Revised: 26 Dec 2025

Accepted: 06 Jan 2026

ABSTRACT

Radio frequency spectrum is an invaluable resource that calls for smart management solutions with the advent of the fifth generation wireless networks and the use of the Open RAN approach. Spectrum use analytics is the use of advanced data science methods for the interpretation of real-time data and historical data extracted from distributed radios on the network. Autoregressive integrated models and deep learning models provide the means for the prediction of capacity needs and congestion maps along the timeline and geography. Synergistic methods with RAN Intelligent Controller provide the means for closed-loop automation and the use of smart applications for automatic allocation according to demands for the use of the spectrum resource. The use of reinforcement learning provides models for the autonomous allocation process along the timeline with the interpretation of the environment and the use of rewards for the learning process. Implementation methods in this context involve the use of computing infrastructure and the integration with the existing legacy systems for the management of networks with testing for the continuity of the service. The use of smart spectrum management solutions provides benefits such as increased efficiency in the spectrum resource use, reduced latency in the networks, and increased quality of experience.

Keywords: Spectrum Utilization Analytics, Open RAN Architecture, Time-Series Forecasting, RAN Intelligent Controller, Reinforcement Learning

1. Rarity of Spectrum & Allocation Issues in Contemporary Wireless Communications

The radio frequency spectrum is an inherently limited natural resource in wireless communication networks, with the portion of the electromagnetic spectrum suitable for mobile broadband use limited by the combined effect of propagation properties and global regulatory regulation. The advent of the fifth-generation cellular network of unprecedented data speeds and lower latency performance requirements than all previous generations of wireless cellular networks has increased this demand for wireless spectrum. The introduction of network slicing into this situation raises more complexities with regard to spectrum resource management, with this technology now requiring spectrum isolation between several virtual networks, which require different levels of service and resource allocation objectives to be mutually fulfilled within the limited resource allocation of the radio frequency spectrum [1].

While network slicing represents an architectural evolution in wireless communications, it also poses resource management issues in terms of multidimensional resource allocation that go well beyond basic assignment of frequency bands. This is because network slice service map management is required by network service providers when assigning infrastructure resources and radio network resources in the radio access network and core network infrastructure, while respecting end-to-end service level agreements concerning various service categories, including massive enhanced mobile

broadband communications, massive machine-type communications, and ultra-reliable and low-latency communications. Mobility management is more complex when network slicing is involved because user equipments change cells and hence network slice reconfiguration and spectrum reassignment are necessary in order to provide end-to-end service continuity. Radio resource management then needs to consider interference coordination, admission control, and handovers among other issues [1].

Strategies for spectrum allocation need to take into account the diverse nature of licensed frequency spectra. This ranges from low-band frequency below one gigahertz, which has range but lacks capacity; medium-band frequency between one and six gigahertz, which has medium range and medium capacity; and millimeter wave frequency above twenty-four gigahertz, which has huge bandwidth and very limited range. The technical issue of mitigating the interference complexity increases with the rising need to support macrocellular, microcellular, picocellular, and femtocellular communications within overlapping geographical regions and differing transmission powers and antenna size and type. Frameworks for dynamic spectrum sharing between fourth and fifth generations of communications technology further complicate the issue and need real-time spectrum allocation to retain compatibility while allowing for a smooth transition to next-generation communication technology.

2. Spectrum Analytics Framework Based on Data and Telemetry

For the purpose of comprehensive spectrum resource utilization monitoring, performance measurement data needs to be continuously collected from radio access network elements dispersed across the cell network, involving measurement instrumentation of physical layer indicators of channel state, medium access control layers of resource allocation statistics, and radio resource management decision records. Contemporary cellular base station technology provides fine-grained monitoring of network activity via standardized management interfaces to facilitate centralized analytical platforms to compile performance measurement data of synchronized measurement intervals across a network of thousands of cell sites. The magnitude of measurement intervals of collected network activity will influence the accuracy of analytical models to be developed, varying from millisecond intervals of resource block allocations to minute intervals of traffic volume contributions [3].

Data pipeline designs follow hierarchical processing topologies, with edge filtering and initial aggregations conducted at the edge RU sites to lower backhaul capacity-utilization factors ahead of the distribution of purified data to analytics analysis systems via the sanctioned data stream. Principles of network functions virtualization facilitate the elastic scaling of the data telemetry monitoring software components as lightweight software containers running atop general-purpose computing servers with arbitrarily elastic processing capacity scaling demand functions, independent of initial purpose-built hardware designs with limited capacity for downstream expansion via software updates alone [1]. Time-series storage systems optimized for rapid data ingestion velocities store the full historical record of measurements with user-specified retention periods to enable the support of operational monitoring analysis views as well as modeling analysis in the batch processing mode. Microservices-oriented architecture decouples data ingestion, processing, and analysis phases to ensure fault tolerance with separate scaling of compute components depending upon the nature of computational workload [4].

Important new key performance indicators are those that focus on spectral occupancy information such as the proportion of available logical resource blocks actually reused in active communication transmissions, signal quality information describing the reference signal received power and signal-to-interference plus noise ratio values in defined coverage regions, and data traffic information describing data volume throughputs on defined frequency bands and time windows. More sophisticated feature engineering methods relate new spectral usage information with additional metadata information such as geographical coordinates, time information describing hour and day of

the week, weather information, and so on, describing events such as sporting events or public assemblies with high abnormal usage patterns. Machine learning processing pipelines are used in data preprocessing to normalize data values coming from various hardware vendors, handle missing values resulting from data gathering failures, and identify points with abnormal values pointing towards hardware malfunctions or improper configuration affecting data analysis models [4].

| Measurement Layer | Telemetry Type | Collection Interval | Primary Purpose | Processing Location |
|-------------------|--------------------------------|----------------------|-----------------------------|----------------------|
| Physical Layer | Channel State Indicators | Millisecond-scale | Signal Quality Assessment | Edge Radio Units |
| MAC Layer | Resource Allocation Statistics | Sub-second intervals | Spectrum Occupancy Tracking | Edge Aggregation |
| RRM Layer | Decision Logs | Second-scale | Policy Validation | Centralized Platform |
| Network Layer | Traffic Volume Aggregations | Minute-scale | Capacity Planning | Analytics Platform |
| Application Layer | Quality of Experience Metrics | Periodic sampling | Service Level Monitoring | Cloud Analytics |

Table 1: Telemetry Data Collection Parameters and Performance Metrics in Spectrum Analytics Framework [3, 4]

3. Time Series Forecasting Methodologies for Spectrum Demand Prediction

Statistical forecasting methods offer a set of basic solutions to the problem of predicting future patterns of use, autoregressive integrated moving averages being a set of models that yield interpretable models of temporal relationships within stationary or trend-stationary time series data. These traditional statistical models break down observed data into systematic parts consisting of deterministic trend, periodic seasonality, and stochastic error terms describing unpredictable changes. Identification algorithms examine the autocorrelation and partial autocorrelation functions to decide on the order of autoregressive and moving-average terms, and differencing operations render non-stationary data stationary when showing long-term trends of growth or decline. Estimation of parameters via maximum likelihood optimization delivers point and interval predictions of the uncertainty surrounding future predictions for use demand [7].

Designs inspired by recurrent neural networks defy some of the restrictions posed by linear statistical models when dealing with complex nonlinear dynamics and high-dimensional input space inherent in large-scale cellular networks. Long short-term memory neural networks are designed using special gating modules such as input gates for controlling information flows to the cells, forget gates for eliminating irrelevant information from previous cells along a particular temporal trajectory, and output gates for controlling activation values broadcasted into downstream layers of a neural network design. These design features make it possible for learning to proceed on a large temporal window by mitigating vanishing or exploding gradients inherited by recurrent architectures during backpropagation through time learning algorithms. Convolutional neural networks apply spatial filtering techniques for deriving a hierarchical representation of spectra measurement matrices through both time and space [7].

Training techniques for deep predictive models involve the use of supervised learning. For these models, the input variables and target values involve the sequence of spectrum usage in the past, while the target involves the values of the predicted forecasts. Optimization techniques involve the use of stochastic gradient descent with momentum and adaptive learning rate techniques, including the use of the Adam algorithm, aiming at the objective of minimizing the loss function that involves the mean

squared error or the mean absolute percentage error. Regularization techniques involve the use of techniques such as dropout, where the model randomly shuts down neurons during training, and the use of early stopping, where the optimization technique is paused once the performance of the model ceases to improve on the validation data. Techniques that involve the averaging of the predictions of multiple models that were trained separately involve the use of techniques that entail the use of weighted averages and the use of stack techniques, and these techniques have greater predictive accuracy, as explained by an author [8].

| Measurement Layer | Telemetry Type | Collection Interval | Primary Purpose | Processing Location |
|-------------------|--------------------------------|----------------------|-----------------------------|----------------------|
| Physical Layer | Channel State Indicators | Millisecond-scale | Signal Quality Assessment | Edge Radio Units |
| MAC Layer | Resource Allocation Statistics | Sub-second intervals | Spectrum Occupancy Tracking | Edge Aggregation |
| RRM Layer | Decision Logs | Second-scale | Policy Validation | Centralized Platform |
| Network Layer | Traffic Volume Aggregations | Minute-scale | Capacity Planning | Analytics Platform |
| Application Layer | Quality of Experience Metrics | Periodic sampling | Service Level Monitoring | Cloud Analytics |

Table 2: Comparative Analysis of Time-Series Forecasting Methodologies for Spectrum Demand Prediction [7, 8]

4. Integration of Spectrum Intelligence with RAN Intelligent Controller Architecture

The disaggregated RAN architecture defines a hierarchical control plane topology with a separation of intelligence functions from infrastructure resources using standardized open interfaces. The RAN intelligent controller framework includes near-real-time controllers with a timescale from ten milliseconds to one second for supporting closed-loop optimized control of RR management functions, in addition to non-real-time controllers for handling policy decision-making and model training processes with a timescale above one second. This timescale separation enables a suitable timescale for operation of spectrum analytics engines, where applications for demand forecasting and capacity planning are normally non-real-time intelligent apps with a decision-making timescale above one second [5].

Within the RIC framework, intelligent applications that reside in the RIC ecosystem can retrieve the state of the network through standardized northbound interfaces specified in open radio access network alliance documentation. The applications run optimization procedures that consider trade-offs for multiple objectives, which range from the maximization of spectral efficiency to the distribution of loads in adjacent cells and comprehensive quality of experience fairness for different subscribers that have different service agreements. Artificial intelligence techniques used in intelligent applications can analyze real-time network data for anomalous usage patterns, foretell potential congestion points, and suggest preventive measures such as carrier aggregation or inter-cell handover optimization [5].

In closed-loop automation chains, the entire life cycle from monitoring and predictions, decision execution, and validation is managed, and the southbound transmission of configuration commands via E2 interfaces sends configuration updates from the RIC platforms to the radio resource management modules in the distributed or central units. E2 interface specifications also address parameter updates, policy compliance, and decision delegation between RIC platforms and the underlying radio access network components. Feedback routines assess measured values after executing actions and match predictions, allowing automated improvements of optimization strategies with reinforcement learning, where successful actions are reinforced and suboptimal

behavior is discouraged by reduced performance of the network. Security frameworks that safeguard control plane communications use mutual authentication, verification, and encryption to prevent malicious users from intercepting and altering spectrum allocation strategies [6].

| RIC Component | Operational Timescale | Primary Functions | Interface Type | Control Loop Nature | Typical Applications |
|--------------------|-----------------------------|--|---------------------------|-----------------------|-----------------------------|
| Near-Real-Time RIC | 10 milliseconds to 1 second | Radio Resource Management Optimization | E2 Interface (Southbound) | Closed-Loop Reactive | Dynamic Carrier Aggregation |
| Non-Real-Time RIC | Above 1 second | Policy Management and Model Training | A1 Interface (Northbound) | Closed-Loop Proactive | Demand Forecasting |
| xApp Platform | Sub-second execution | Real-time Analytics Processing | Service Model APIs | Event-Driven | Anomaly Detection |
| rApp Platform | Multi-second intervals | Long-term Optimization | REST APIs | Policy-Based | Capacity Planning |
| RIC Data Lake | Continuous ingestion | Historical Data Storage | Data Collection APIs | Passive Monitoring | Model Training Datasets |

Table 3: RAN Intelligent Controller Architecture Components and Operational Timescales [5, 6]

5. Adaptive Spectrum Allocation Using Reinforcement Learning Methods

Reinforcement learning paradigms find applications in modeling spectrum management as a sequence of decision-making steps accomplished by autonomous entities perceiving network conditions, choosing resource allocation actions, and receiving reward signals defining measures of policy efficiency based on predefined operator targets. The MDP model describes network evolution based on state transition probabilities to reflect causal linkages between existing allocation actions and future network behaviors regarding statistical distributions related to spectrum usage, interference, and service qualities. Network state description models relevant network characteristics like intra-cell ratios of utilized/spectral resources, buffer sizes, inter-cell RSRP reporting between adjacent cells, and mobility parameters describing UE trace behaviors within network coverage regions [9].

Actions spaces represent the set of allocation choices fed into learning agents, which include frequency band mappings to target cells, transmission power control, modulation/coding schemes, and Beamforming parameters. Continuous action spaces enable precise control over parameters during resource allocation, whereas discrete action spaces facilitate learning by allowing allocation choices to be made between pre-defined configuration settings. Reward functions represent translation of high-level business goals into signals that influence optimization of learning policies, involving maximum throughput optimization, latency penalties, and fairness requirements disallowing starvation in low-priority traffic classes. Negative rewards corresponding to high overheads or termination of service trigger stability in allocation policies with potential adaptations according to network variations [9]. The policies are optimized for parameterized policies by deep neural networks to map the observed state to the probability distributions for actions, with actor-critic networks breaking down the learning process into distinct sub-problems for the value function and policies. The value function network predicts the weighted cumulative rewards for each state to provide baselines for variance reduction in the estimated gradient with Monte Carlo trajectory sampling. Experience buffers are used to store interaction traces as ordered tuples of states, actions, rewards, and next states to re-use for learning and to reduce instability in gradient-based optimizations by absorbing temporal correlations in trajectories. Multi-agent reinforcement learning

approaches extend to coordinated decisions for multiple cells/network spectra with centralized learning and decentralized control paradigms by training cooperative policies for offline learning stages and executing independently for runtime tasks [10].

| RL Framework Element | Component Type | Information Encoded | Representation Format | Optimization Objective | Learning Mechanism |
|-----------------------------|-----------------------|--|------------------------------|-------------------------------|---------------------------|
| State Space | Network Observations | Spectrum Occupancy, Queue Depths, RSRP | Multidimensional Vectors | Environment Characterization | Passive Sensing |
| Action Space | Allocation Decisions | Frequency Assignments, Power Levels | Discrete or Continuous | Resource Distribution | Agent Selection |
| Reward Function | Performance Signals | Throughput Gains, Latency Penalties | Scalar Values | Policy Effectiveness | Objective Translation |
| Policy Network | Decision Mapping | State-to-Action Probability | Deep Neural Network | Action Selection Optimization | Gradient Ascent |
| Value Network | Reward Estimation | Expected Cumulative Returns | Function Approximation | Baseline Prediction | Temporal Difference |
| Experience Replay | Trajectory Storage | State-Action-Reward-NextState Tuples | Memory Buffer | Sample Efficiency | Batch Training |

Table 4: Reinforcement Learning Components for Adaptive Spectrum Allocation [9, 10]

6. Operational Benefits and Implementation Factors for Intelligent Spectrum Management

Predictive spectrum analytics provides quantifiable gains in terms of efficiency of network resource utilization by matching resource allocation capabilities and expected demand behavior patterns, minimizing occurrences and durations of congestion events, causing user experience dissatisfaction. Proactive resource allocation schemes migrate resources to involved cells when traffic starts increasing before quality of service is affected, maintaining system throughput at levels better than those promised in service-level agreements. By training algorithms on past experiences, machine learning algorithms detect repeat daily, weekly, and yearly demand behavior patterns to facilitate early resource migration when ahead of reactive schemes that only respond after congestion occurs. Latency improvement is obtained from resource allocation schedules when delay-sensitive traffic is scheduled when spectrum is limited, and other traffic is postponed when spectrum availability is high [11].

The implementation methods of the spectrum intelligence systems involve computational infrastructure planning, from edge computing to cloud platforms for inference and training of the models, respectively. Graphics processing units can speed up the training of artificial networks, with parallel processing systems handling the high-dimensional space of states and large batches to increase the accuracy of the gradient estimates. The model deployment methods, however, consider the latency of the networks during inference, with model quantization reducing the precision from thirty-two-bit floating point to eight-bit integers for faster computations on the edge devices with negligible errors. The containerized systems enable easy migrations to different hardware platforms, with the orchestration systems considering the model versioning, release, and rollback operations [11].

Integration with the existing management frameworks for the broader network requires the formulation of middleware pieces to map the analytics-driven suggestions in vendor-agnostic configuration command formats required by the older element management frameworks. Standardization of interfaces for intent-based networks seeks to abstract the equipment-specificities through declarative policy definitions, making it easier for analytics frameworks to articulate the required outcome specifications without detailing the execution specifications. Validation and test methodologies are required to check the adherence to seamless service delivery by automated changes to the resource allocations, as well as within regulatory boundaries regarding the use of the allocated spectrum resources, coupled with maintaining interoperabilities between borders via standardized signal interfaces. Roll-out models starting from the 'shadow mode' operation phase, where the suggestions are computed but not acted upon, facilitate empirical verification regarding the correctness and safety of the policies before the activation of the closed-loop automation process in the production networks [12].

Conclusion

Smart spectrum management via analytics and forecasting is a basic need for large-scale Open RAN deployments facing a scarcity of frequencies and mounting capacity demands. The convergence of time-series forecasting approaches with a disaggregated RAN architecture assists in predictive resource allocation, foreseeing congestion points ahead of time, thereby preventing service quality degradation. RIC platforms support standardized hooks for integrating machine learning models engaged in perpetual analysis of time-series telemetry and providing optimization guidance targeted at spectral efficiency, latency needs, and quality of experience goals. Reinforcement learning algorithms move ahead of rule-based allocation approaches by learning adaptive policies through direct interaction with their environment, allowing automated spectrum resource allocation in dynamically varying network conditions, typical of dense-urban scenarios. To be successfully adopted, it is essential to be mindful of computing infrastructure support, interfacing with legacy support, and testing mechanisms guaranteeing automated actions preserve continuity of services and do not violate regulatory parameters. With evolving wireless systems embracing increasingly complex networking for a wide range of application domains, intelligence in spectrum resource allocation is deemed a pivotal facilitator of sustainable resource utilization and networking.

References

- [1] Haijun Zhang et al., "Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges," ResearchGate, 2017. [Online]. Available: https://www.researchgate.net/publication/313611684_Network_Slicing_Based_5G_and_Future_Mobile_Networks_Mobility_Resource_Management_and_Challenges
- [2] Francesco Restuccia and Tommaso Melodia, "Deep Learning at the Physical Layer: System Challenges and Applications to 5G and Beyond," arXiv:2004.10113v2, 2020. [Online]. Available: <https://arxiv.org/pdf/2004.10113>
- [3] Oriol Sallent et al., "On Radio Access Network Slicing from a Radio Resource Management Perspective," IEEE Wireless Communications, Volume 24, Issue 5, 2017. [Online]. Available: <https://ieeexplore.ieee.org/document/7891795>
- [4] Xin Li et al., "Network Slicing for 5G: Challenges and Opportunities," IEEE Internet Computing, Volume 21, Issue 5, 2017. [Online]. Available: <https://ieeexplore.ieee.org/document/8039298>
- [5] Michele Polese et al., "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/367369842_Understanding_O-RAN_Architecture_Interfaces_Algorithms_Security_and_Research_Challenges
- [6] Leonardo Bonati et al., "Intelligence and Learning in O-RAN for Data-driven NextG Cellular Networks," arXiv:2012.01263v2, 2021. [Online]. Available: <https://arxiv.org/pdf/2012.01263>

- [7] Nei Kato et al., "The Deep Learning Vision for Heterogeneous Network Traffic Control: Proposal, Challenges, and Future Perspective," ResearchGate, 2016. [Online]. Available: https://www.researchgate.net/publication/311782493_The_Deep_Learning_Vision_for_Heterogeneous_Network_Traffic_Control_Proposal_Challenges_and_Future_Perspective
- [8] Chaoyun Zhang et al., "Deep Learning in Mobile and Wireless Networking: A Survey," ResearchGate, 2018. [Online]. Available: https://www.researchgate.net/publication/323722699_Deep_Learning_in_Mobile_and_Wireless_Networking_A_Survey
- [9] Zhiyuan Xu et al., "A Deep Reinforcement Learning based Framework for Power-Efficient Resource Allocation in Cloud RANs,". [Online]. Available: <https://raas.syr.edu/wp-content/uploads/2017/12/ICC17.pdf>
- [10] Cong Luong Nguyen et al., "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," ResearchGate, 2018. [Online]. Available: https://www.researchgate.net/publication/328380592_Applications_of_Deep_Reinforcement_Learning_in_Communications_and_Networking_A_Survey
- [11] Yaohua Sun et al., "Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues," arXiv:1809.08707v2, 2019. [Online]. Available: <https://arxiv.org/pdf/1809.08707>
- [12] Fengxiao Tang et al., "Future Intelligent and Secure Vehicular Network Toward 6G: Machine-Learning Approaches," ResearchGate, 2019. [Online]. Available: https://www.researchgate.net/publication/337803870_Future_Intelligent_and_Secure_Vehicular_Network_Toward_6G_Machine-Learning_Approaches