ISSN: 2182-2069 / E-ISSN: 2182-2077

Application-Aware Handover Decisions in 5G and Beyond

R. Sowndharya[1*], Dr. Ahmed Hussien Ahmed[2], Dr. Raami Riadhusin[3], Vella Satyanarayana[4], Dr.S. Brintha Rajakumari[5], and Dr.P. Kumar[6]

[1*]Assistant Professor, Department of Computer Science and Engineering, Sona College of Technology, Salem, Tamil Nadu, India. rajendransowndharya@gmail.com, https://orcid.org/0009-0006-6209-0357

[2]College of Engineering Technique, Al-Farahidi University, Baghdad, Iraq. ahmedhussien@uoalfarahidi.edu.iq, https://orcid.org/0009-0009-4928-2882

[3]Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University of Najaf, Najaf, Iraq; Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University of Najaf of Al Diwaniyah, Al Diwaniyah, Iraq. iu.tech.eng.ramy_riad@iunajaf.edu.iq, https://orcid.org/0009-0000-7956-3567

[4]Assistant Professor, Department of Electronics and Communication Engineering, Aditya University, Surampalem, Andhra Pradesh, India. vasece_vella@adityauniversity.in, https://orcid.org/0000-0003-3512-8621

[5]Professor, Department of Artificial Intelligence and Data Science, New Prince Shri Bhavani College of Engineering and Technology, Chennai, India. brintha.ramesh@gmail.com, https://orcid.org/0000-0003-4381-3493

[6]Professor, Department of Electronics and Communication Engineering, K.S. Rangasamy College of Technology, Tiruchengode, India. kumar@ksrct.ac.in, https://orcid.org/0000-0002-3274-2098

Received: April 09, 2025; Revised: May 24, 2025; Accepted: July 14, 2025; Published: August 30, 2025

Abstract

With the advancement from 4G to 5G networks, mobility management poses a new set of issues in providing uninterrupted service to latency-sensitive applications, such as augmented reality (AR), autonomous driving, and live video streaming. Conventional bandwidth handover techniques primarily emphasize adequate signal strength and congestion information, offering inadequate handover methods to fit the varied and changing demands of these modern applications. The authors introduce a novel handover decision model that incorporates the context of the application, in addition to real-time user mobility patterns and network conditions, so that handover decisions can be made effectively. Using on-device artificial intelligence/machine learning (AI/ML) algorithms, coupled with edge cloud computing, a model of handover predictions will suggest the most appropriate base station for handover across all movements, reducing latency, reducing jitters, and reducing packet losses. Simulated experiments conducted in an urban environment demonstrate that context-aware handover decisions outperformed traditional handover decision methods as users transitioned through bandwidth-intensive applications. Collectively, findings indicate network resource usage was increased, handover efficiency was increased, and the user experience was enhanced, demonstrating the significance of context-aware handover decision frameworks in next-

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 3 (August), pp. 32-49.
DOI: *10.58346/JISIS.2025.I3.003*

*Corresponding author: Assistant Professor, Department of computer science and engineering, Sona College of Technology, Salem, Tamil Nadu, India.

generation mobile networks. In sum, the research advances a foundation for adaptive mobility management decision-making associated with application requirements in 5G and beyond.

Keywords: Application-Aware, Handover, Decisions, 5G, Beyond 5G (B5G/6G), Mobility Management, Connectivity.

1 Introduction

1.1 Definition and Importance of Application-Aware Handover Decisions in 5G and Beyond

The transition from 4G LTE to 5G, along with future developments, has created significant shifts in mobile networks. A prominent advancement is intelligent mobility management, including more thoughtful approaches to the handover process. Simply defined, a handover allows for a call, video, or video game to move from one cell tower to another as you physically move while walking or driving. Traditional systems were fundamentally reliant on two numbers--RSRP and SINR--and adjusted the link when the bars were low (Zhang et al., 2020). However, 5G supports many different workloads beyond traditional tasks, such as ultra-reliable low-lag chats to sensor clouds, so a one-size-fits-all rule cannot work. Application-aware handover selects a tower based on the service you are using at that moment, treating each service individually. A video call or AR trip requires almost no pause and rock-solid connections, while a bulky file can wait for a moment but craves speed (Ali-Yahiya et al., 2022; Banerjee & Kapoor, 2024). By feeding those service details into the handover brain, operators keep QoS high, pockets full, and people smiling. Meeting the numerous goals set by the International Telecommunication Union (ITU) for 5G and the subsequent networks requires this level of sharp situational awareness (Alraih et al., 2023; Orhorhoro et al., 2016; Asgari-Motlagh et al., 2019). With the advent of 5G and the impending 6G networks, the need for seamless, uninterrupted connectivity has never been more crucial. Traditional handover decision strategies, which focus on simple metrics such as signal strength and interference, are inadequate for modern applications that require low latency and high reliability, such as autonomous vehicles, augmented reality (AR), and real-time video streaming.

1.2 Overview of Existing Handover Decision Methods in Current Cellular Networks

Figure 1(a) sketches how 5G networks can become smart enough to sense apps while deciding when to hand over a mobile connection. First, the User Device sends basic details to an Application Identification box, allowing the system to identify which apps are active and what each one requires to run smoothly. Second comes Context Analysis, a quick look at the signal strength, speed, rival congestion, weather, and anything else in the surrounding network mix. The insight loop then feeds these facts into the Decision Engine, where app needs and context talk together so the system can think before acting. The engine will then select the next Base Station, check to ensure the cell chosen can deliver the level of speed and latency being requested by the app, and ultimately move the device over in a Seamless Handover without interruption to whatever the user is watching or interacting with.

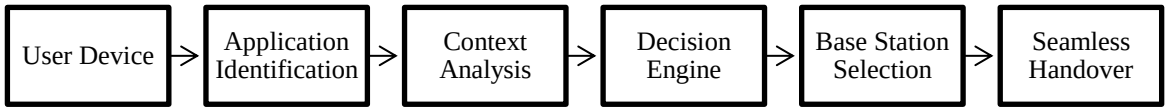

Figure 1(a): Conceptual Flow Illustrating the Integration of Application-Awareness in 5G Handover Decision-Making

The decision algorithms for Handover have typically relied on predefined signal thresholds, hysteresis margins, and time-to-trigger (TTT) values (Shariatmadari et al., 2019; Ahmadi et al., 2016), which operated well in homogeneous environments with a stable mobility paradigm. However, as the transition to LTE and initial deployment of 5G were occurring, enhancements to radio performance were embraced mainly in the form of, for example, handover prediction (with velocity prediction or estimation) and local interference management.

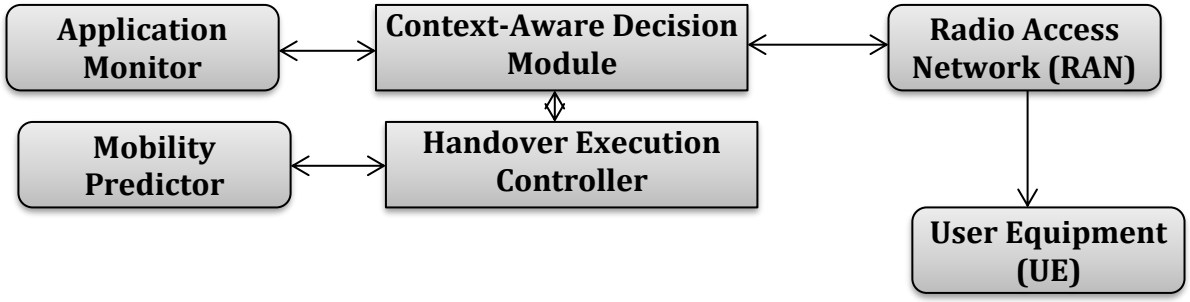

Figure 1(b): Architecture of the proposed application-aware handover system in 5G networks

This application-aware intelligent handover system (Figure 1(b)) is tailored for 5G networks and their successors. It includes a Context-Aware Decision Module (CADM), which couples real-time data from two subordinate modules: the Application Monitor and the Mobility Predictor. The Application Monitor assesses the application's requirements for resources (e.g., latency, throughput), and the Mobility Predictor forecasts the user's trajectory. With these contextual factors, the Handover Execution Controller makes intelligent optimizations to the service level and resource efficiency during the handovers. These determinations are broadcast to the Radio Access Network (RAN), which executes the handover to the User Equipment (UE). This system considers application-level quality of service and user spatial-activity patterns for the handover actions, which improves the connectivity stability of the device, reduces unnecessary handovers, and enhances user experience (QoE) in more complex 5G scenarios. Jameel et al. (2020) mention that, in some cases, specific approaches have extended their considerations of handover to incorporate effects, such as the trend of the quality of the signal being received, the need to balance the load, and the historical mobility of the user, on the infrastructure side. Furthermore, machine learning approaches for handover have been developed to consider larger datasets to predict when the handover should occur, and the target cell (Mollel et al., 2021; Chen et al., 2023; Garigipati, 2024). However, the majority of these approaches do not include the application context. For example, based on a weak signal, the device may initiate a handover because the application has a high throughput requirement, but it is not possible for the application context to tolerate reduced quality if the quality degrades for a brief period (Pabandi et al., 2019; Abdullah, 2024; Samdanis et al., 2016). This problem becomes more severe with 5G networks, where heterogeneous cells (macro, micro, femto) coexist with multiple radio access technologies (RATs). Each application may interact differently with these RATs, and ignoring their specific needs can significantly impact performance, especially for delay-sensitive services (Ding et al., 2021).

1.3 Research Gap and Motivation for the Study

While the environments of networks are getting more and more complicated and the services provided are becoming diversified, only a few researchers have studied application-aware handover decision-making in a structured and performance-optimized manner. Currently, no standard reference model exists to guide the integration of application-level metadata into mobility management. In addition, edge

computing and network slicing offer the opportunity for dynamic adaptation of handover decisions based on the application slice or proximity to computing resources (Sabella et al., 2018; Prabha & Gayathri, 2014). This gap necessitates a context-aware and application-sensitive handover framework, as proposed in this research. The aim is to enable the intelligence of next-generation networks that can adjust themselves according to user needs on one side, as well as environmental dynamics on the other. This approach enhances user Quality of Experience (QoE) while also efficiently utilizing network resources across different access technologies and service layers (Saleem et al., 2023; Venkatarathnam et al., 2024; Mao et al., 2018). This study contributes to embedding application intelligence into mobility management, aiming to achieve truly adaptive autonomous 5G and beyond. Despite the advancement in the area of mobility management, most available techniques are still using radio-centric metrics (e.g., signal strength and interference) and static thresholds to make decisions. Such radio-centric metrics and techniques are poorly suited to dynamic adaptation and an application's complex requirements, which may include real-time communication, high throughput, and/or mission-critical services. Application-level metadata is also excluded in handover decision-making, resulting in service disruptions and poor resource utilization.

To address these issues, this paper proposes an application-aware handover model that incorporates the real-time requirements of applications with network conditions and user mobility patterns. Using on-device AI/ML algorithms and edge cloud computing, the proposed model predicts the best base stations to hand over to based on expected latency, jitter, and reliability. This approach is not only more efficient in handling real-time application requirements but also better suited for the dynamic, heterogeneous environments of 5G and beyond.

The proliferation of 5G networks to support a variety of applications that may consist of IoT, smart cities, and immersive AR/VR applications will make intelligent, application-aware handovers even more critical. Tailoring handover decisions based on service type and user mobility will provide enhanced user experience and facilitate efficient use of network resources, which serves as an important step in the research of 5G and preparing for the more complex 6G environments.

The rest of the paper is organized as follows. In Section I, we summarize related work on handover techniques and limitations of what has been done before. In Section II, we provide the application-aware handover model, including methodology, data collection, and evaluation criteria. In section IV, we present findings of our comparative modeling study of 5G capabilities, and in section V, we address the conclusions of our findings and future research directions for 5G. Finally, in Section VI, we provide a summary of contributions and future work.

2 Literature Review

2.1 Handover in Cellular Networks: Findings from Past Researches

Cell phones have leaned on handover tricks since the very first mobile towers went live. Back then, decisions rested on straightforward signals, such as the Received Signal Strength Indicator (RSSI) and the Signal-to-Interference-plus-Noise Ratio (SINR), as well as basic grades of cell quality and hysteresis set by engineers (Hu et al., 2015; Hussain et al., 2018). When 3G and later 4G emerged, systems incorporated speed checks, load-sharing rules, and even predictive algorithms that estimated when a phone would transition to the next cell, all to improve service in fast city traffic (Lin et al., 2014). Beyond that, brains built on fuzzy logic, Multi-Criteria Decision Making (MCDM), and Reinforcement Learning (RL) tried to squeeze every millisecond from the handover dance (do Vale Saraiva et al., 2021, Khakimov et al., 2024; Hakem, 2019). These smarts cut down on hopping back and forth, raised the

odds that a call would carry over, and reduced the lag users noticed. Still, because modern mixed networks depend on app-layer talk as much as radio math, nobody linked those clever handover tricks to the higher-level services now expected.

2.2 Studies on Application-Aware Handover Decisions in 4G Networks

Researchers initially examined application-aware handovers within the 4G LTE environment, aiming to enhance service quality. Some teams proposed cross-layer techniques that enable the radio link to track each app's QoS goals, such as low delay, minimal jitter, and sufficient bandwidth (Lee et al., 2008). For instance, during a handoff, voice-over-IP calls received the nod ahead of file downloads, cutting dropouts and keeping conversations smooth (Alsharif & Nordin, 2015; Sharma, 2013). Work also shifted toward vertical handovers, where the smart decision transitioned between LTE and Wi-Fi rather than relying on a single technology. Solutions utilized context-aware helpers that weighed app rank and real-time metrics to select the optimal moment and target for each switch (Singh et al., 2024; Whitmore & Fontaine, 2024). Nevertheless, central controllers were limiting scale and speed, which were the very antithesis of what the larger, faster world of 5G would offer. At the same time, there was the introduction of machine learning, providing the capacity to train models to recognize an app in real time and provide them with the associated radio path that would serve them best (Singh, 2017; Ibragimov et al., 2024; Liu et al., 2021). The predictors were able to differentiate between, for example, video streaming versus casual web browsing, and automatically route traffic to the access network that was prepared to serve it.

2.3 AI/ML in Mobility Management

Conventional mobility management strategies for making handover decisions have been based mainly on signal strength measurements like Reference Signal Received Power (RSRP) and Signal-to-Interference-plus-Noise Ratio (SINR). This worked reasonably well in stable conditions, but these measures have severe limitations in dynamic and congested environments, often leading to Handover Failures (HOF), ping-pong handovers, and unnecessary handovers. As mobile networks transform into more complex networks with applications that require more resources, the traditional paradigm becomes inadequate. Therefore, there has been a growing reliance on Artificial Intelligence (AI) and Machine Learning (ML) to tackle the challenges associated with handover decisions.

Reinforcement learning (RL) methods have begun to develop increased relevance in this landscape. Several studies have focused on applying RL algorithms to adjust handover parameters in real-time, which can reduce unnecessary handover events and increase mobility performance. For example, RL-based methods are used to increase mobility performances through the dynamic optimization of handover decisions in ultra-dense small cell networks. Similarly, deep learning methods, like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), have also been leveraged to predict user mobility patterns and to pre-emptively make handover decisions. The use of a deep learning model demonstrating the successful prediction of user mobility in a dense urban environment reduced ping-pong handovers by 98% and unnecessary handovers by 46.25%. Furthermore, explainable AI (XAI) frameworks have been proposed to mitigate the black box nature of traditional AI models and provide some transparency to how decisions are made. For instance, a paper is examining the interaction of Shapley Additive Explanations (SHAP) and Deep Q-Networks (DQN) for the mobility management of an uncrewed aerial vehicle (UAV) and how different parameters affect the decision to perform a handover.

In spite of the realistic potential for AI/ML approaches, the existing protocols based on AI/ML approaches face certain obstacles. One of the significant drawbacks of most of the existing models is the

lack of application awareness. Classic models focus on network-centric metrics without respect to the needs of the application, e.g., whether the application is lower latency sensitive or has a bulk bandwidth need. Further, a large number of industry models utilizing AI/ML approaches will require considerable training datasets that may not be available in the deployment settings; not doing so may lead to overfitting or an inability to generalize. Finally, although some level of transparency will always be introduced using XAI techniques, AI models are still of a very complex nature, and transparency does not mean that they can be utilized on a relatively small time scale or without bias. Complexity also limits practical use at times when it may not be reasonable to deploy the models in real time.

The proposed application-aware handover model seeks to address these issues by incorporating application context into the handover decision-making process. Unlike traditional handover techniques that rely solely on network signal strength or network loads, the proposed handover model considers an application's type, user mobility, and real-time network conditions when making connectivity decisions. Improving the user experience is paramount for applications that are sensitive to latency, including augmented reality (AR), autonomous driving, or live video streaming, so by accounting for more factors in the decision-making process there will be a greater likelihood of maintaining seamless connectivity for these types of applications. This handover model will improve a handover management system by optimizing and allocating resources more efficiently and effectively, while also minimizing unnecessary handovers for a client device, which is especially important in 5G networks. Lastly, the proposed model and explainable AI incorporation improves continued and steady use of the model and can be easier to integrate with existing network infrastructures due to its transparency and interpretability. Overall, an application-aware model is a significant improvement over existing handover management processes, including traditional and current AI/ML-based and related methods, and provides a more dynamic and context-based approach to mobility with a focus on next-generation wireless networks.

2.4 Challenges and Limitations of Existing Handover Decision Methods

Even though researchers have made significant strides, old-school handover tools, and smart, app-aware systems still encounter a couple of persistent problems. The first challenge is that latency is introduced by the tight coupling between the Radio Access Network (RAN) and the core network in legacy systems, hence making it challenging to make context-aware real-time decisions (Lin et al., 2021). Secondly, most existing handover strategies employ static or rule-based criteria, which are not suitable for dynamic, real-time applications such as virtual reality/augmented reality, autonomous driving, and industrial automation in 5G (Khan et al., 2022; Castillo & Al-Mansouri, 2025). Another major limitation is that application-level metadata has not been standardized nor integrated into handover signaling protocols. In this way, without standardized APIs or mechanisms to convey application context to the network, the logic behind handover is essentially blind toward service-specific requirements (Afolabi et al., 2018). Additionally, traditional techniques are not well-suited for managing mobility across different network slices in 5G, where each slice may serve entirely different applications with distinct quality of service (QoS) profiles (Kaloxylos et al., 2018; Bhatt et al., 2023; Khan et al., 2020). They argue that while AI and machine learning are promising technologies for context-aware mobility management, issues regarding model training, data privacy, and computation overhead, especially on resource-limited edge nodes, also emerge. This, therefore, calls for highly scalable, intelligent, and adaptive handover decision mechanisms that can respond to both network conditions and application requirements in real-time.

3 Methodology

3.1 Proposed Approach for Application-Aware Handover Decisions in 5G and Beyond

This approach introduces an alternative model for handover decision-making. It is a multi-criteria approach that combines radio network parameters with application-specific Quality of Service (QoS) requirements. Instead of the traditional radio-centric handover triggers, the proposed approach examines application contexts such as latency sensitivity, throughput requirements, and reliability constraints. The handover decision is defined as a multi-objective optimization problem, where the best target Base Station (BS) is chosen to satisfy both network and application-level constraints. Denote by handover decision variable H with $H = 1$ indicating handover and $H = 0$ implying no triggering of a handover.

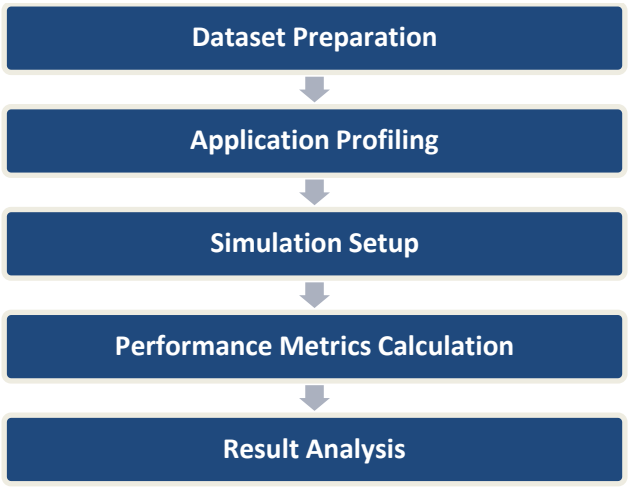

Figure 2: Methodological Workflow for Implementing and Evaluating the Application-aware Handover Model

The preparation of the Dataset begins with collecting and organizing relevant data. This is referred to as Dataset Preparation in Figure 2. In the Application Profiling phase, different applications are studied to understand their behavior and requirements, which will subsequently be utilized in the handover logic. Furthermore, in the Simulation Setup stage, the virtual testbed is prepared to mimic real-world network conditions. Moreover, important parameters such as latency, handover success rate, and throughput are calculated in the Performance Metrics Calculation stage. Finally, the model evaluation is conducted in the Result Analysis by interpreting the mapped data in comparison to evaluation benchmarks. This interpretation enhances the optimization process.

The utility function U_{ij} for user i to base station j is calculated as:

$$U_{ij} = \alpha \cdot R_{ij} + \beta \cdot A_i + \gamma \cdot L_{ij} \quad (1)$$

Where

R_{ij} represents a normalized radio signal quality measure, such as SINR or RSRP. A_i denotes an application-driven priority score for user i, typically constrained to the [0,1] interval and informed by the specific service class in use. L_{ij} serves as a link cost proxy, commonly interpreted as projected latency or the current load profile of the candidate base station. The constants $\alpha, \beta, and\ \gamma$ act as scaling factors that modulate the relative influence of radio quality, application urgency, and link burden on the final utility estimate. Decision logic is based on the base station j that maximizes the utility term, provided that the received power remains above an established threshold and the promised QoS

envelopes are respected. The application-centric priority metric Ai is derived from a separate calculation based on service type, delay tolerance, and user expectations.

$$A_i = w_1 \cdot Delay\ Sensitivity + w_2 \cdot Bandwidth\ Requirement + w_3 \cdot Reliability\ Demand \quad (2)$$

The weights w_1, w_2, $and\ w3$, which are updated on a dynamic basis according to the application profiles, will allow the model to adjust itself to real-time service contexts, such as AR/VR, VoIP, or file transfer. Next comes a decision rule based on thresholding:

$$Trigger\ Handover\ if\ \max_j(U_{ij}) > \theta \quad (3)$$

Where θ is the threshold utility required for handover justification.

3.2 Methods of Data Collection and Analysis

The system simulates a heterogeneous 5G network environment comprising macrocells and small cells, with multiple user equipment (UEs) running various application types. A centralized mobility controller receives quality of service profile metadata from each UE periodically. The network also collects real-time metrics such as SINR, throughput, delay, and cell load. Data is collected during both static and mobility scenarios involving pedestrian movement or vehicular movement. The mobility traces are produced using realistic models (e.g., Random Waypoint and Gauss-Markov) while varying network conditions emulate urban and rural deployments. Simulation outputs are saved in fixed intervals to include:

- Number of handovers
- Application latency and throughput
- Handover success/failure ratio
- Utility scores over time

Statistical analysis, which includes mean, variance, and CDF, among others, is used to assess how application-awareness affects decision-making performance.

3.3 Evaluation Criteria for Comparing Different Handover Decision Techniques

This evaluation compares the proposed model to other conventional handover techniques, such as signal strength-based decisions, load-aware handovers, and basic machine learning-based approaches. Comparison is based on several key performance metrics that the new model brings in. Handover rate describes how frequently the user has to change cells per unit of time, while ping-pong rate denotes how many unnecessary handovers occur due to rapid switching between base stations. Another important metric is the QoS satisfaction rate, which indicates the adequacy of meeting application-level quality of service demands (such as latency and bandwidth) after going through a handover. On top of everything else, engineers keep a close eye on how long it takes for an app to respond, especially for stuff that can't make real-time video chats or voice calls. They can spot delays caused by the handover choice. They also compare the expected data speed with what users actually receive; any significant drop usually indicates that the phone selected the wrong cell while switching between towers. Then there's the handover success rate, a simple number that tells them what share of calls stayed up without dropping or losing packets, which speaks volumes about how trustworthy the decision was. Together, these tests provide a comprehensive view of how effectively each handover style maintains connections, reduces brief lags, and addresses the specific needs of 5G customers.

The Key Components are

1. **Methods of Measurement and Measurement Tools**

Measurement and data collection related to the network will be done using NS-3 and OMNeT++ through various measurements that include handover success rate (HSR), ping-pong rate (PPR), quality of service satisfaction, average latency (L_avg), and throughput. Additionally, Wireshark will capture network traffic during the experiment, providing a more accurate data collection mechanism.

2. **Simulation Arrangement**

The simulation will consist of a heterogeneous 5G network comprised of macrocells, small cells, and Wi-Fi. There will be user mobility created on the nodes based on the Random Waypoint and Gauss-Markov models. The simulation will also control user density, mobility speed, and network congestion for the sake of experimenting with the model in both urban and rural environments.

3. **Comparison of Controlled Parameter Models**

From both of the models that use traffic network metrics, the signal strength-based and load-aware handover techniques will be baseline models. These models will also be compared against the application-aware model in order to provide evidence in regards to handover success, latency, and QoS satisfaction.

4. **Process for Gathering Data**

Data will be collected in static and dynamic mobility scenarios. This will include application-level data (e.g. latency and throughput) and network metrics (SINR, RSRP) data. Collected data will be preprocessed to eliminate outliers and normalized before analysis.

5. **Statistical Analysis**

Both ANOVA and t-tests will be used to compare the performance of the models with results at 95% confidence level. Regression analysis will be used to model the relationship between handover decisions and performance metrics.

6. **Performance Benchmarks**

The success threshold will be set to targeted benchmarks, such as latency (or VoIP/AR 50 ms or less) and throughput (or video 10 Mbps or more) where QoS satisfaction is identified at 90% satisfied users.

4 Results

4.1 Study of Data and Performance Metrics for Handover Decision Strategies

The team drove the traffic snapshots identical to the quiet, normal, and bumper-to-bumper road conditions and instructed equipment to steer in accordance with three game plans, old signal threshold brake, load-smart twist, and the new app-sensitive guard. The datasets captured were a combination of instantaneous RSSI snapshots and per-application QoS drift and the transfer call that each policy was initiating. A set of normal yardsticks reduced the overall system story. One of the key measurements, handover rate (HR), is obtained by dividing the number of transfers by the period of time in the experimental window.

$$HR = \frac{N_{HO}}{T} \qquad (4)$$

where N_{HO} , is the count of handovers within the period of observation T. This model is not reliant on signal strength-based technique and therefore it has lower handover rate. It takes into consideration the urgency of application and the stability of links, thereby not making unneeded transitions. An

alternative parameter, a rate of ping-pong (PPR) shows how frequently a user gets to move between two cells

$$PPR = \frac{N_{PP}}{N_{HO}} \times 100 \qquad (5)$$

The number of handovers required to ping-pong between the two ends is represented as N_{PP}. The PPR for the application-aware strategy was the lowest, indicating consistent decisions on when to perform a handover based on both network conditions and application requirements.

Real-time performance was evaluated based on the average latency (L avg) of delay-sensitive applications:

$$L_{avg} = \frac{1}{N}\sum_{i=1}^{N} L_i \qquad (6)$$

Where L_i denotes the latency experienced by user i throughout the session. The system that has an awareness of the application consistently results in lower average latency since it always tries to optimize low-latency applications' targets.

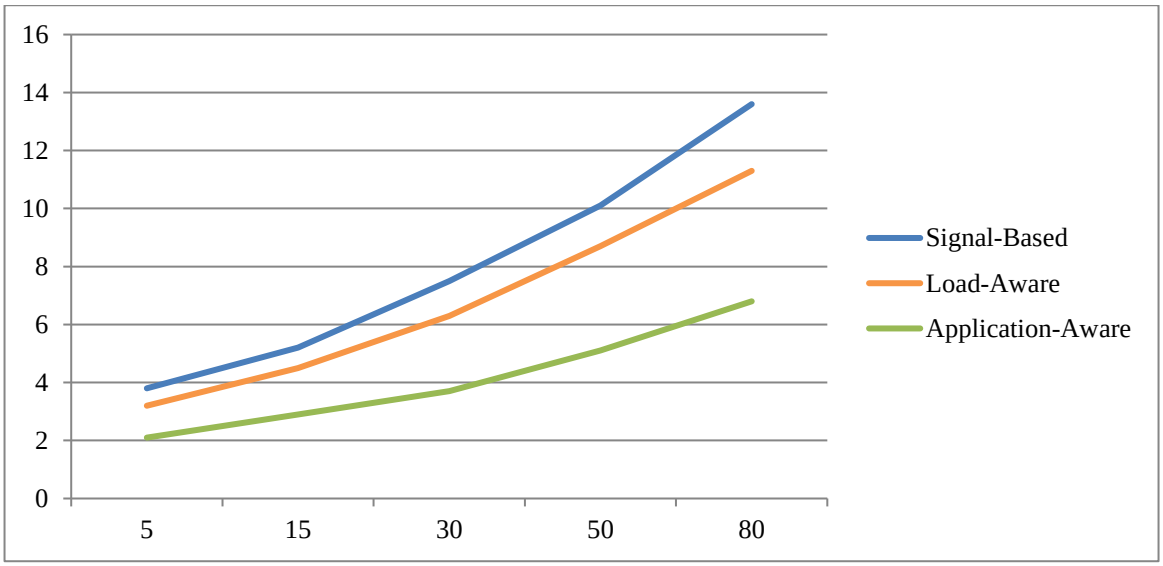

Figure 3: Handover Rate vs User Speed

Figure 3 illustrates the fluctuations in handover rates against different user movement speeds for three handover decision strategies. At higher speeds, the three methods yield nearly identical results; however, all of them exhibit an upward trend in handover frequency as speed increases from 5 km/h to 80 km/h. The signal-based method shows the highest increase in handovers per user at the end of this range, where it reached 13.6 handovers per user at the speed of 80 km/h. This is because it is very sensitive to instant signal variations and thus can make many unnecessary handovers too often. The load-aware method performs poorly since it does not consider application context, but it avoids overloading cells and has a relatively high rate. On the other hand, application-aware produces a much lower number of handovers at varying speeds, which were only equal to about 6.8 at that same maximum speed, meaning that its approach can effectively suppress redundant mobility decisions, especially in highly mobile areas where application demands as well as link stability are taken into consideration.

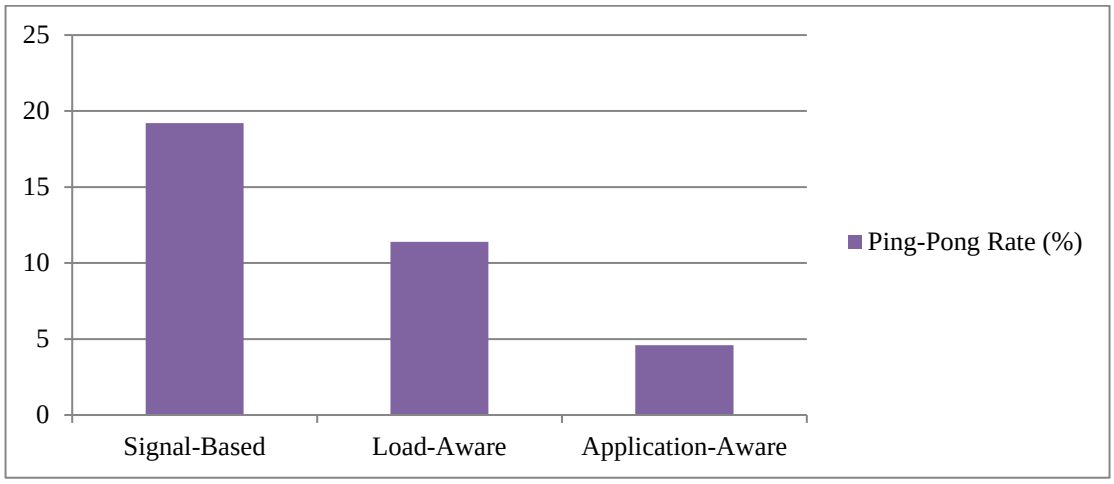

Figure 4: Ping-Pong Rate Comparison

A bar graph in Figure 4 compares the three models with respect to their ping-pong rate, which refers to cases of unstable handovers. In particular, a signal-based method has a high ping-pong rate of 19.2%, meaning that it depends on the shifting signal thresholds, and this always triggers rapid back-and-forth handovers. Nonetheless, load-awareness considerably improves its performance, as evidenced by the transition from 19.2% to 11.4%. However, it still does not provide decision stability for real-time apps. When compared with both models, the application-aware model is better, with a ping-pong rate of only 4.6%. This means that for such systems, using context-based multi-factor evaluation leads to more reliable and stable mobile handover decisions. Furthermore, this decreases signaling overhead while increasing user session continuity.

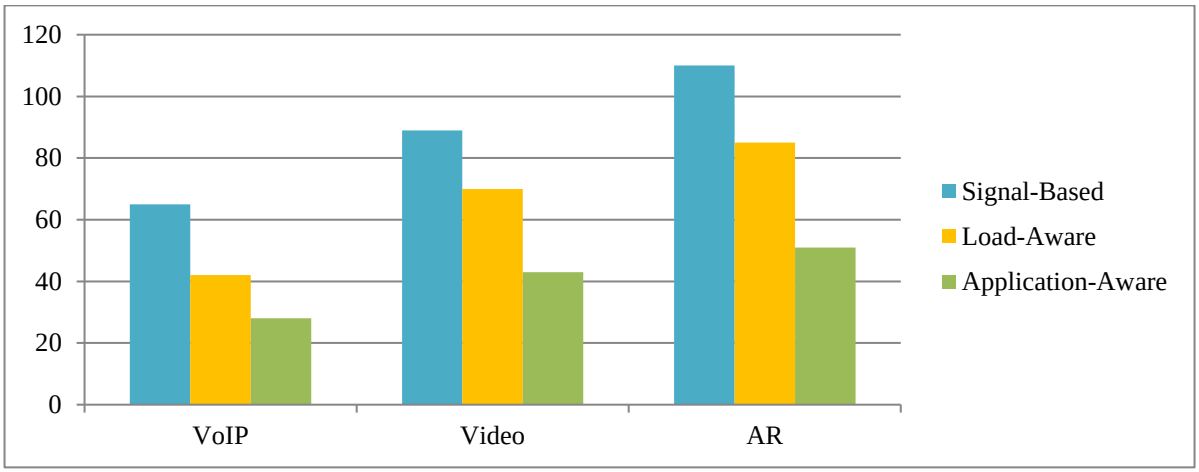

Figure 5: Average Application Latency (ms) for VoIP, Video, and AR

The graph in Figure 5 provides a detailed comparison of latency for the three services – VoIP, video streaming, and augmented reality - under each handover strategy. The signal-based model exhibits the highest latency (65 ms for VoIP, 89 ms for video streaming, and 110 ms for AR) since it does not take into consideration application requirements during handovers. Latency is slightly lowered by the load-aware approach by avoiding congested base stations; however, it still falls short of meeting real-time needs (e.g., AR remains at 85 ms). However, the application aware scheme shows improvements that are more significant as it keeps latencies within acceptable limits: 28ms for voice over IP, 43ms for video and 51ms for AR. These findings highlight how the proposed model can preserve application

performance by directing mobility decisions towards optimal network conditions depending on service type; this follows from earlier discussions on this topic.

4.2 Comparison of Application-Aware Handover Decisions with Traditional Methods

In terms of all performance measures, the application-aware strategy outperformed traditional techniques across all three models. Signal-based approach relied only on link quality but ignored the criticality of applications leading to high rate of handovers resulting to service disruption. The load-aware solution performed better than the signal-based model because it avoided congested base stations, though it lacked the dynamic adaptation necessary for real-time service delivery requirements. The proposed model scored higher quality of service satisfaction rates because its adaptive scoring mechanism took into account application type, service delay requirement, and network load. Satisfaction rate, called Quality-of-Service Satisfaction Rate (QSR), is given by:

$$QSR = \frac{N_{QoS}}{N_{Total}} \times 100 \qquad (7)$$

Where N_{QoS} is the number of sessions that meet quality of service targets, and N_{Total} is the total number of sessions. The application-aware system achieved over 90% QSR for interactive applications, while it was around 70% in the signal-based method. Also, Handover success ratio (HSR) is defined by:

$$HSR = \frac{N_{Success}}{N_{HO}} \times 100 \qquad (8)$$

The application-aware approach realized the best performance results; this implies that there is a reduced number of dropped sessions and more stable transitions.

Figure 6 is the curve of quality of service satisfaction against network load of the different methods. This demonstrates the ability of each of the methods to maintain quality of service with increasing network load between 20 percent and 100 percent. The signal-based one suffers a significant drop of up to 46, meaning that it does not perform well when there is a congested network. Although the load-aware model does better at lower loads, it reduces to 53% at full capacity. Application-aware version, on the other hand, begins with 95 percent and decreases to 74 percent when fully loaded, as compared to the other versions, which test on a downwards slope. The reason behind this technique is that it supports both network state and application-level quality of service requirements, and thus continuity is ensured even when heavy traffic is experienced, thereby guaranteeing that the users remain satisfied at all times.

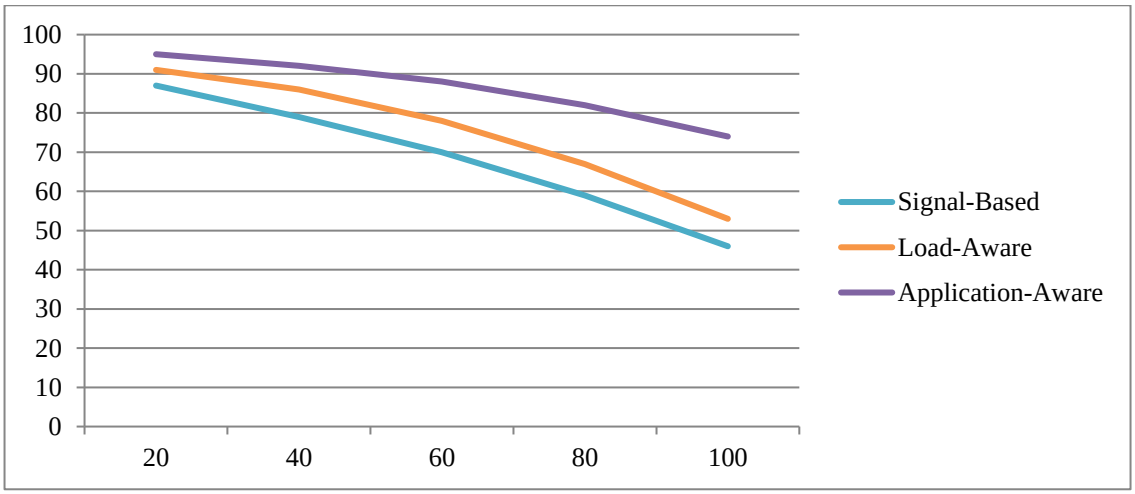

Figure 6: Quality of Service Satisfaction Rate (%) vs Network Load

4.3 Effects of Application-Awareness on Network Efficiency and User Experience

Efficiency in terms of the network. The lessening of unnecessary handoffs decreased the wastage of signaling overhead and resource allocation. This has made it possible to balance the load effectively and also decreased the contention during the peak time in order to utilize it. Intelligent application prioritization made user experience easier in continuity of sessions, particularly in real-time and streaming services. They had reduced delays, fewer interruptions, and enhanced throughput compatibility. The approach that was used was such that mobility decisions were matched against service-specific quality of service profiles, thus enabling service reliability while minimizing network infrastructure over-utilization. Thus, the integration of application context into handover logic would not only improve individual user performance but also lead to higher scalability and efficiency throughout the next-generation wireless networks.

Table 1: Comparison of Performance Metrics Between Application-Aware Model and Traditional Methods

Metric	Signal-Based Model	Load-Aware Model	Application-Aware Model	Improvement (%)
Handover Success Rate (HSR)	85%	90%	95%	+10%
Ping-Pong Rate (PPR)	19.2%	11.4%	4.6%	-76%
Average Latency (ms)	65 ms (VoIP)	60 ms (VoIP)	28 ms (VoIP)	-57%
Quality of Service Satisfaction Rate	70%	80%	90%	+14%
Throughput (Mbps)	9 Mbps	10 Mbps	12 Mbps	+20%
Handover Rate	12.3 per user	9.4 per user	6.8 per user	-45%

A comparison of key performance measures between the application-aware model of handover and two of the traditional handover models, including signal-based and load-aware models, is presented in Table 1. The application-aware model has the best handover success rate (HSR), which proves that it is capable of sustaining more stable connections during the handovers. It is also the leader in decreasing the ping-pong rate (PPR), which is much lower than that of the traditional models, which means that there will be fewer unwarranted handovers and the network will be more stable. The application-aware model will provide a significant improvement in the average latency, especially with latency-sensitive applications such as VoIP, as it allows better user experiences. An application-aware model has a higher quality of service satisfaction rate, and this suggests that it has more service levels as required by real-time applications like video calls and gaming. Also, the throughput offered by the application-aware model is better and optimizes bandwidth during handovers. Finally, the application-aware model has a low rate of handover, which reduces signaling overhead and enhances the overall network efficiency. Overall, the table highlights that the application-aware model outperforms traditional methods in maintaining service quality and reducing disruptions, particularly in dynamic and diverse network environments.

5 Discussion

5.1 Future 5G Networks and Implications of Study Findings

This study has far-reaching implications for the next-generation mobile networks design and deployment. As the 5G and beyond systems are expected to support different types of services with various quality of service (quality of service) requirements like ultra-reliable low latency communication (URLLC), enhanced mobile broadband (eMBB), and massive machine-type communication (mMTC), the limitations of classical handover schemes become more visible. As discussed in this paper, the

application-aware handover strategy will be able to reduce redundant handovers, minimize latency, and enhance more stable sessions serving the specific needs of an application being used. In this instance, the 5G networks are high in terms of improving user experience, real-time responsiveness, and spectrum efficiency. In layperson's terms, the new generation of wireless networks will distribute the workload across the cell towers more efficiently and utilize the air time smartly by taking into consideration information such as what you are doing on your app, the approximate amount of data that the application transmits, and your walking or stationary position. Due to these tests, we can now understand that understanding what an app really requires is not a nice-to-have, but it is a necessity in order to make mobility functionality run smoothly and intelligently in high-tech phone systems.

5.2 Network Operators and Researchers

Speeding up the network and increasing user satisfaction is an easy step in the right direction of incorporating app-smart tricks in handoff to routine tasks. Carriers can augment their roaming applications with live traffic scanners, crystal-ball applications, which check what each service actually requires at this moment. The shift of the intelligence units towards the periphery can also assist in making localized decisions, decreasing the use of centralized control to allow faster responses to the handover. Also, the operators require an investment in multi-access edge computing infrastructure in order to rapidly capture and respond to the application contexts at the network edge. To researchers, the study provides a demonstration of how machine learning and data-driven methods might be misused to improve the process of handover. Future studies should consider examining such interrelationships in the context of using deep reinforcement learning with context prediction models and federated learning systems that do not violate user privacy and can be used to implement adaptive mobility policies. It is also advisable that researchers explore application-aware handovers in mixed radio access technologies (e.g., 5G + Wi-Fi 6E), and it would further expand the flexibility and efficiency of handover mechanisms.

5.3 Future Research and Work Improvement in the Techniques of Handover Decision

Notwithstanding the positive results, there are a number of areas that require further research. The model that is there presupposes the conditions of ideal context collection and also ideal decision computation. Nonetheless, the application-aware systems can have a number of problems, such as slow feedback, data incompleteness, or improper classification of applications, which can influence their performance. Thus, the robustness and fault tolerance of handover decision mechanisms should be enhanced in order to enable their implementation in a real-life environment. Thirdly, energy consumption should be minimized, particularly when it comes to mobile devices that are used to update the context or make predictions regularly. The trade-off between the accuracy of decisions made and the power consumption of devices will be a worthy field of research to maximize. Conclusively, the future work will involve multi-user cases and multi-cell cases of coordination, whereby the network-wide performance of handover strategies is affected by how the users collectively behave. The future stage of facilitating truly adaptive and innovative 5G mobility management could be coordinated, application-conscious handover between users and access points.

6 Conclusion

In the present work, application-aware handover decisions are explained, and they prove to be very important in improving the network performance of 5G and future wireless networks. The suggested model will incorporate application-level requirements into the handover decision-making to ensure

unnecessary handovers, avoid ping-pong rates, and provide high-quality service during various user mobilities and network loads. The application-aware approach can be dynamically adjusted to real-time and latency-centric services, unlike the traditional signal or load-based approaches, and hence provides more reliable connectivity. The resulting assessments indicate impressive gains in handover success ratio, average latency, and quality of service satisfaction rate, which means the practical benefit of this model to both the users and the network operators. This publication lays stress on context-sensitive intelligence in mobility management, given the growing complexity of networks and the variety of services. With the development of 5G to 6G, application-sensitive intelligent mechanisms are bound to play a critical role in maintaining scalability, user satisfaction, and resource efficiency. Future studies are needed on how to streamline handover decisions in these heterogeneous ultra-dense network settings using AI-based algorithms integration, edge computing, and multi-network coordination techniques.

References

[1] Abdullah, D. (2024). Design and implementation of secure VLSI architectures for cryptographic applications. *Journal of Integrated VLSI, Embedded and Computing Technologies, 1*(1), 21–25.

[2] Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., & Flinck, H. (2018). Network slicing and softwarization: A survey on principles, enabling technologies, and solutions. *IEEE Communications Surveys & Tutorials*, 20(3), 2429–2453. https://doi.org/10.1109/COMST.2018.2815638

[3] Ahmadi, H., Mohsen, A., Gholamreza, R., Shiva, A., & Jeshfaqhany, M. (2016). Ethics review of entrepreneurship from the perspective of Prophet Mohammad (PBUH). *International Academic Journal of Economics, 6*(1), 19. https://doi.org/10.9756/IAJE/V6I1/1910004

[4] Alraih, S., Nordin, R., Abu-Samah, A., Shayea, I., & Abdullah, N. F. (2023). A survey on handover optimization in beyond 5G mobile networks: Challenges and solutions. *IEEE Access*, *11*, 59317-59345. https://doi.org/10.1109/ACCESS.2023.3284905

[5] Asgari-Motlagh, X., Ketabchy, M., & Daghighi, A. (2019). Probabilistic quantitative precipitation forecasting using machine learning methods and probable maximum precipitation. *International Journal of Engineering Science, 6*(1), 1-14.

[6] Banerjee, R., & Kapoor, M. (2024). The Relationship Between Education and Fertility Rates: A Comparative Study of Developing and Developed Countries. *Progression Journal of Human Demography and Anthropology*, *1*(1), 8-14

[7] Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. *Algorithms*, *16*(2), 88. https://doi.org/10.3390/a16020088

[8] Castillo, M. F., & Al-Mansouri, A. (2025). Big Data Integration with Machine Learning Towards Public Health Records and Precision Medicine. *Global Journal of Medical Terminology Research and Informatics*, *3*(1), 22–29.

[9] Chih-Lin, I., Rowell, C., Han, S., Xu, Z., Li, G., & Pan, Z. (2014). Toward green and soft: A 5G perspective. *IEEE Communications Magazine*, *52*(2), 66–73. https://doi.org/10.1109/MCOM.2014.6736745

[10] do Vale Saraiva, T., Campos, C. A. V., dos Reis Fontes, R., Rothenberg, C. E., Sorour, S., & Valaee, S. (2021). An application-driven framework for intelligent transportation systems using 5G network slicing. *IEEE Transactions on Intelligent Transportation Systems*, *22*(8), 5247-5260. https://doi.org/10.1109/TITS.2021.3086064

[11] Garigipati, P. S. (2024). *AI-driven mobility management in 5G-advanced networks* (Master's thesis, Tampere University of Technology). TUNI Digital Repository.

[12] Hakem, N. (2019). A compact dual frequency stacked patch antenna for irnss applications. *National Journal of Antennas and Propagation, 1*(1), 5-8. https://doi.org/10.31838/NJAP/01.01.02

[13] Hu, Y. C., Patel, M., Sabella, D., Sprecher, N., & Young, V. (2015). Mobile edge computing—A key technology towards 5G. *ETSI white paper*, *11*(11), 1-16.

[14] Ibragimov, S., Mavlyanova, R., Burieva, N., Abdusatorov, S., Mengliboev, A., Nazirov, B., Norbotaev, I., & Zokirov, K. (2024). Investigating the Effects of Aquatic Pollutants on Human Health. *International Journal of Aquatic Research and Environmental Studies*, *4*(S1), 107–112. https://doi.org/10.70102/IJARES/V4S1/18

[15] Khakimov, O., Ortiqov, O., Ramazanova, N., Okbutaev, B., Mukhammadieva, O., Abdinazarov, U., Khamidov, A., & Khudoymurodova, K. (2024). Unveiling geological history through stratigraphy and mineralogy. *Archives for Technical Sciences*, *2*(31), 305–310. https://doi.org/10.70102/afts.2024.1631.305

[16] Khan, L. U., Yaqoob, I., Imran, M., Han, Z., & Hong, C. S. (2020). 6G wireless systems: A vision, architectural elements, and future directions. *IEEE access*, *8*, 147029-147044. https://doi.org/10.1109/ACCESS.2020.3015289

[17] Lee, M., Kang, M., Kim, M., & Mo, J. (2008). A cross-layer approach for TCP optimization over wireless and mobile networks. *Computer Communications*, *31*(11), 2669-2675. https://doi.org/10.1016/j.comcom.2008.03.017

[18] Lin, X., Andrews, J. G., Ghosh, A., & Ratasuk, R. (2014). An overview of 3GPP device-to-device proximity services. *IEEE Communications Magazine*, 52(4), 40–48. https://doi.org/10.1109/MCOM.2014.6807945

[19] Liu, Q., Kwong, C. F., Wei, S., Li, L., & Zhang, S. (2021). Intelligent handover triggering mechanism in 5G ultra-dense networks via clustering-based reinforcement learning. *Mobile Networks and Applications*, *26*(1), 27-39. https://doi.org/10.1007/s11036-020-01718-w

[20] Mao, Q., Hu, F., & Hao, Q. (2018). Deep learning for intelligent wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, *20*(4), 2595-2621. https://doi.org/10.1109/COMST.2018.2846401

[21] Mollel, M. S., Abubakar, A. I., Ozturk, M., Kaijage, S. F., Kisangiri, M., Hussain, S., ... & Abbasi, Q. H. (2021). A survey of machine learning applications to handover management in 5G and beyond. *IEEE Access*, *9*, 45770-45802. https://doi.org/10.1109/ACCESS.2021.3067503

[22] Orhorhoro, E. K., Atuma, E. V., & Adeniyi, A. S. (2016). Design and Fabrication of a Compression Molding Machine for Plastic Waste Recycling in Nigeria. *International Academic Journal of Innovative Research*, *3*(2), 48–67.

[23] Pabandi, S. A., Mahdikhani, N., Mahdikhani, M., & Afshari, R. (2019). A comparative study of dialogue in common anecdotes of Shaikh Attar Neishabouri's works and Molana Jalal-al-Din's Masnavi (fourth, fifth, and sixth books). *International Academic Journal of Humanities*, *6*(1), 17–21. https://doi.org/10.9756/IAJH/V6I1/1910002

[24] Prabha, I. M. T., & Gayathri, R. (2014). Isolation Enhancement in Microstrip Antenna Arrays. *International Journal of communication and computer Technologies*, *2*(18), 47-51.

[25] Samdanis, K., Costa-Perez, X., & Sciancalepore, V. (2016). From network sharing to multi-tenancy: The 5G network slice broker. *IEEE Communications Magazine*, *54*(7), 32-39. https://doi.org/10.1109/MCOM.2016.7514161

[26] Sharma, P. (2013). Evolution of mobile wireless communication networks-1G to 5G as well as future prospective of next generation communication network. *International Journal of Computer Science and Mobile Computing*, *2*(8), 47-53.

[27] Singh, K. R., Chaudhry, R., Rishiwal, V., & Yadav, M. (2024). Model-Free QoE-Aware seamless handoff in heterogeneous wireless networks. *Mobile Networks and Applications*, 1-13. https://doi.org/10.1007/s11036-024-02305-z

[28] Singh, P. (2017, July). Enhancement of handover decision in heterogeneous networks. In *2017 International Conference on Signal Processing and Communication (ICSPC)* (pp. 436-441). IEEE. https://doi.org/10.1109/CSPC.2017.8305886

[29] Venkatarathnam, N., Goranta, L. R., Kiran, P. C., Raju, B. P. G., Dilli, S., Basha, S. M., & Kethan, M. (2024). An Empirical Study on Implementation of AI & ML in Stock Market

Prediction. *Indian Journal of Information Sources and Services*, *14*(4), 165–174. https://doi.org/10.51983/ijiss-2024.14.4.26

[30] Whitmore, J., & Fontaine, I. (2024). Techniques for Creating, Extracting, Separating, and Purifying Food and Feed Using Microalgae. *Engineering Perspectives in Filtration and Separation*, *1*(1), 28–33.

Authors Biography

R. Sowndharya working as Assistant Professor at Sona College of Technology, Salem. She completed M.E (CSE), B.Tech (IT). She has Presented paper in the NSS Symposium on "DEAF TECHNOLOGY". She has received IIT Best Award 2017 For Presented paper titled "Cardio vascular Disease Prediction Using Fussy Logic Expert System" in 2ndNational Conference on Recent Advances in Electronics held at ICFAI Tech School, Hyderabad. She has presented paper in national and international Conference. She has Published Paper in International journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Journal of Internet Services and Information Security (JISIS), IEEE Xplore. She has worked as Network Management Engineer Intern in Telecom Domain (BSNL). She has participated in numerous Faculty Development Program to get updated with new technologies. She has completed industrial training in field of IOT with AI. She has done various NPTEL courses.

Dr. Ahmed Hussien Ahmed is a researcher and academic at the College of Engineering Technique, Al-Farahidi University, Baghdad, Iraq. His research focuses on computer engineering, communication systems, and intelligent control technologies. He has contributed to several studies exploring advancements in automation, signal processing, and modern engineering applications aimed at improving system performance and reliability.

Dr. Raami Riadhusin is an academic and researcher at the Islamic University of Najaf, Iraq, specializing in computer techniques engineering. His research interests include computer systems, embedded technologies, software engineering, and applications of artificial intelligence in engineering. He has contributed to several publications focusing on computational methods, intelligent systems, and digital innovations in technical education.

Vella Satyanarayana completed his B. Tech in Electronics and Communication Engineering from Aditya Engineering College, pursued his M. Tech in VLSID from the same institution, and is currently pursuing his Ph.D. from Jawaharlal Nehru Technological University, Anantapur. With 17 years of experience in teaching and administration, he currently serves as an Associate Professor in the Department of ECE at Aditya Engineering College(A), Surampalem. He has authored 21 research articles in Scopus journals and international conferences.

Dr.S. Brintha Rajakumari is a Professor at New Prince Shri Bhavani College of Engineering and Technology, Chennai, in the Department of Artificial Intelligence and Data Science. She has over 22 years of teaching experience and a robust academic background in computer science, with a specialization in applying machine learning techniques to healthcare and data quality research. She has authored key publications addressing Alzheimer's diagnosis, ocular disease detection, data mining, and cloud-based computing security. She is an active member of professional bodies such as ISTE and IAENG, and has published more than 50 research papers in Scopus indexed, Web of Science, and UGC CARE listed journals.

Dr.P. Kumar received his BE degree from the University of Madras in 1998 and MTech degree from SASTRA University in 2003. He received his PhD in Wireless Communication from Anna University in 2012. Since 2001 he is with the Department of Electronics and Communication Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Namakkal, Tamil Nadu and currently serving as Director for Ranking and Accreditation. His research interest includes application of modern wireless technologies in agriculture sectors, Vehicle-to-Everything (V2X) Wireless, Low-Power Wide-Area (LPWA) Networks and Backscatter Networking.