A Data Driven Framework for QoE-Aware Intelligent EN-DC Activation

Syed Muhammad Asad Zaidi*, Marvin Manalastas*, Muhammad Umar Bin Farooq*, Haneya Qureshi*,
Adnan Abu-Dayya†, and Ali Imran*

*AI4Networks Research Center, University of Oklahoma-Tulsa, USA
†Department of Electrical Engineering, Qatar University, Doha Qatar
Email: {asad, marvin, umar.farooq, haneya, ali.imran}@ou.edu, adnan@qu.edu.qa

Abstract—In emerging 5G networks, User Equipment camps traditionally on 4G network. Later, if the user requests a 5G service, it can simultaneously camp on 4G and 5G using E-UTRAN New-Radio Dual-Connectivity (EN-DC) approach. In EN-DC, poor radio-conditions in either 4G or 5G network can be detrimental to user Quality-of-Experience (QoE). Although operators want to maximize EN-DC activation to fully utilize the 5G network, sub-optimal parameter configuration to turn on EN-DC can compromise key-performance-indicators due to excessive radio-link-failures (RLFs) or voice-muting. While the need to maximize the EN-DC activation is obvious for maximizing the 5G network's utility, RLF and mute avoidance are vital to maintain the QoE requirements. To achieve aforementioned tradeoff, this paper presents the first solution to optimally configure the EN-DC activation parameters. We collect two datasets from real network to develop machine-learning-models to predict RLF and muting, respectively. We also investigate and compare the potential of various under-sampling, oversampling, and synthetic data generation techniques including Tomek-Links and Generative Adversarial Networks for their potential to address the data-imbalance problem inherent in the real network training data. Leveraging these models, we formulate and solve two QoE-aware optimization problems that can maximize EN-DC activation while minimizing RLF or voice-muting. System-level simulation-based results show that compared to state-of-the-art solution that does not take into account RLF or voice-muting risk in EN-DC activation, the proposed solution can intelligently determine EN-DC activation criteria that minimize the risk of RLF and voice-muting while giving the operator's desired level of priority to maximize 5G network utilization.

Index Terms—5G, New Radio, EN-DC, Radio Link Failure, Voice Call Muting, Artificial Intelligence

I. INTRODUCTION

5G New Radio (NR), with innovative use cases of enhanced Mobile Broadband (eMBB) for large volume transmissions, massive Machine Type Communications (mMTC) for sensors and Internet of Things (IoT) devices, and Ultra Reliable Low Latency Communications (URLLC) for mission critical applications come with unprecedented Quality of Experience (QoE) goals. Given the magnitude of these use cases and applications, studies project that 5G subscriptions will top 2.6 billion by the end of 2025 [1]. While the capacity crunch will be addressed primarily by ultra-dense Base Station (BS) deployment and mmWave band utilization [2], ensuring QoE with a conglomeration of 5G and legacy technologies, i.e., 4G-LTE, remains an open challenge of utmost importance.

As per 3GPP Release 15 specification 37.863 [3], E-UTRAN New Radio Dual Connectivity (EN-DC) allows 5G capable User Equipment (UE) to simultaneously connect to a 4G-LTE eNodeB (eNB) and 5G gNodeB (gNB). Although EN-DC concept is first conceived in 3GPP Release 15, the process stated in the more recent releases such as in Rel. 16 and 17 is similar without any significant changes. Therefore, we focus our discussion on the EN-DC procedure standardized in 3GPP Release 15. In EN-DC, the LTE eNB acts as a master node, playing a crucial role in the EN-DC session setup, while the gNB acts as a secondary node providing the 5G data path as illustrated in Fig. 1. This non-standalone 5G network deployment enabled by EN-DC is aimed to address the extreme capacity demand challenge, reduces the capital expenditures (CAPEX) of network operators, and accelerate the penetration of 5G networks. However, this additional complexity leads to new challenges including the increased signaling overhead and the need to determine the optimal parameters for EN-DC activation/deactivation.

EN-DC activation comes with an intrinsic trade-off between 5G network utilization and potential QoE degradation due to Radio Link Failure (RLF) and voice call muting. To the best of the authors' knowledge, there does not exist a study in open literature to quantitatively analyze this trade-off and offer a solution to optimally configure EN-DC activation parameters. A high number of EN-DC activations are desirable for fully utilizing the 5G network. However, an unintelligent maximization of EN-DC activations can result in several Quality of Experience (QoE) related issues, including excessive amount of ping-pong EN-DC activation/deactivation, recurrent RLFs, and exasperating voice call muting. RLF is the radio interface disruption between BS and UE, and is typically caused by coverage hole or poor signal quality as a result of high interference. UE observes high interference either during handover (HO) process due to sub-optimal HO parameter configuration [4]–[7], or due to the interference from the neighboring cells usually at the cell-edge. On the other hand, voice call muting refers to the instance when a UE is unable to receive audio packets during an active voice call. Similar to RLF, voice call muting is observed mostly due to poor radio conditions.

EN-DC operations require the control plane to be setup between UE and LTE eNB, while the user plane can be transmitted by 5G gNB or both LTE eNB and 5G gNB simultaneously as per the configuration. As a result, UE must maintain a strong connection both with LTE and 5G nodes simultaneously. Poor radio conditions either on 4G or 5G can subsequently result in RLF or voice muting, which will be detrimental to the user experience.

Table I: Literature work on dual connectivity approaches.

Reference	Energy Efficiency	Resource Scheduling	Throughput Enhancement	Reliability	Load Balancing	Video Transmission	User Association	ML/AI	Parameters for Dual-Connectivity
[8]		✓	✓						
[9]		✓			✓		✓		
[10], [11]							✓		
[12]				✓			✓		
[13]	✓				✓		✓		
[14]			✓				✓		
[15]		✓	✓						
[16]							✓		TimeToTrigger only
[17]	✓			✓					
[18]			✓					SVM	
[19]	✓								
[20]	✓		✓				✓		
[21]				✓					
[22]	✓	✓	✓						
[23]		✓				✓			
Proposed Framework	✓		✓	✓			✓	Deep Learning	5G b1-thresholds: RSRP & SINR, 4G RSRP & SINR thresholds

By accelerating the EN-DC activation in an attempt to increase network efficiency, EN-DC may be triggered at poor radio frequency (RF) conditions at either 4G or 5G network. This can result in call disconnect, and service disruption. Following the service disruption, repeated re-accessibility attempts not only increase signaling but drain UE battery as well. Thus, optimal configuration to activate/deactivate EN-DC is essential to maintain the expected QoE and network efficiency of the 5G network. This optimal set of EN-DC activation parameters will also lower the signaling overhead by preventing the unnecessary EN-DC activation attempts e.g., in poor 4G or 5G radio conditions. Thus, the proposed solution eliminates the additional signaling generated due to RLFs caused by sub-optimal EN-DC activation parameters.

As quantifying the relationship between the signal condition (i.e., RSRP and SINR) vs RLF and voice muting is essential for optimization of EN-DC activation, the proposed framework leverages a data-driven approach to quantify this relationship in the absence of analytical models due to system-level complexity. Existing analytical models to address the dual connectivity issues are based on numerous assumptions to make the equations tractable, however, the results become impractical in real world scenario. On the other hand, existing simulators do not capture the intricacies of a live mobile network, and the simplification renders the results practically unusable in a live network. With the widespread deployment of 5G mobile networks, there is an urgent need to develop a real network-based data-driven approach to address the aforementioned challenges. To the best of authors' knowledge, there does not exist any data driven work based on real network data that focuses on maximizing 5G network utilization through EN-DC and minimizing potential QoE degradation due to RLF and voice call muting.

A. Related Work

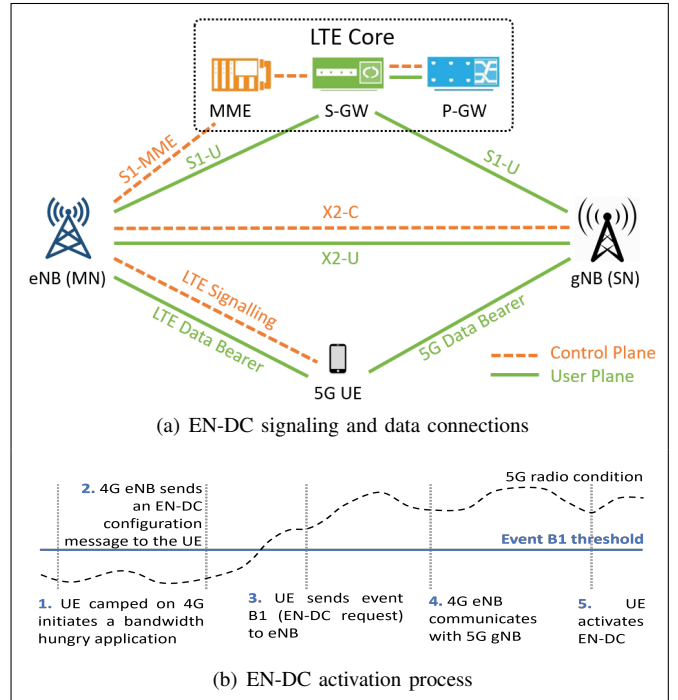The concept of dual-connectivity has been studied well over the past years [24]–[26]. A detailed review of these studies can be found in a recent mobility management survey in emerging cellular networks [27]. Meanwhile, more specific studies of dual-connectivity gain in terms of delay and throughput [8], [14], [15], [18], [20], [22], mobility [9]–[12], energy efficiency [13], [17], [19], [20], [22], reliability [12], [17], [21], and latency [28] exist in literature as well. However, to the best of the authors' knowledge, no study in the existing literature addresses the QoE-aware criteria to activate dual-connectivity between two different mobile technologies viz a viz 4G and 5G. Particularly, there does not exist a study on RLF and muting instances in the context of EN-DC.

(a) EN-DC signaling and data connections

(b) EN-DC activation process

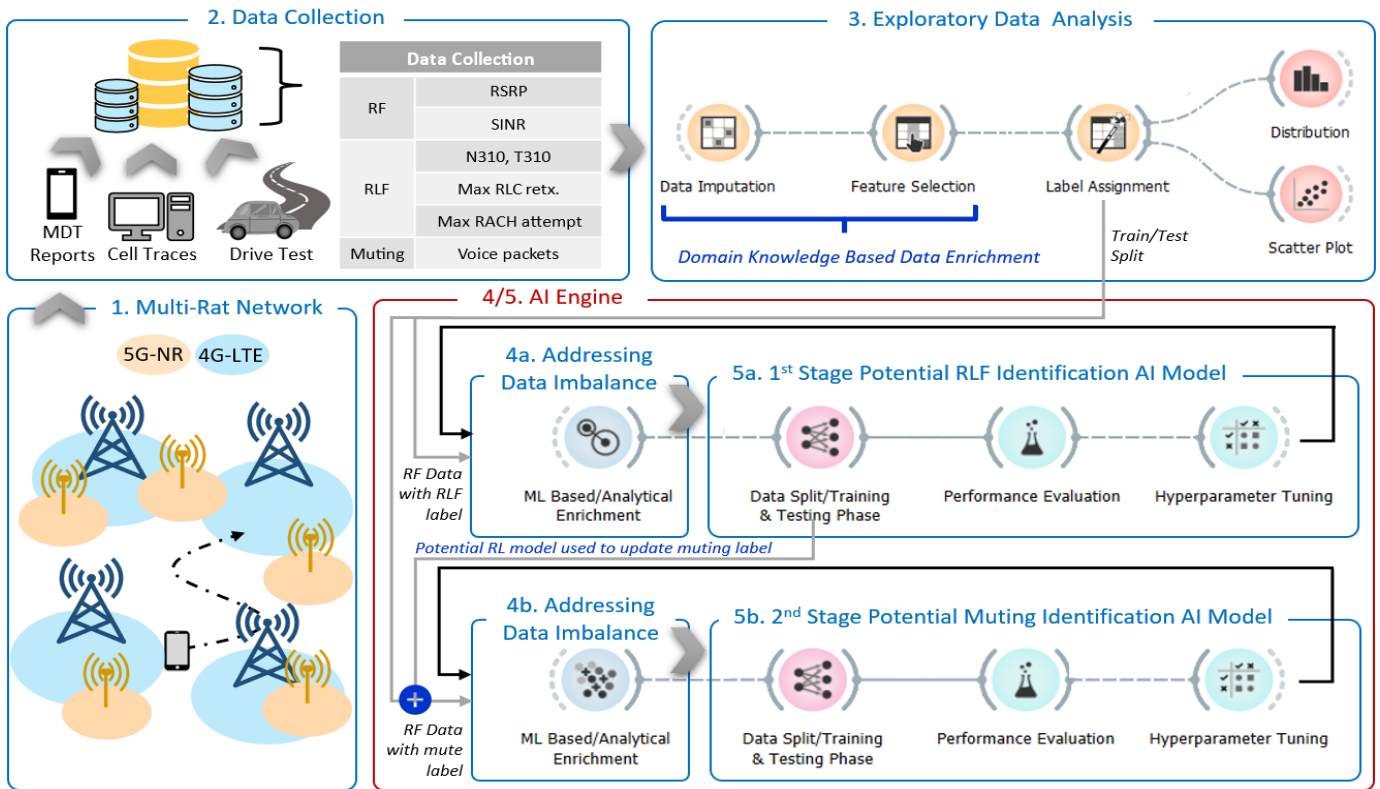Figure 1: EN-DC activation and signaling process for 5G NR

Figure 2: Overview of the proposed AI-based prediction models for potential RLF and potential muting.

Most of the RLF-related literature [29]–[33] addresses intra-frequency HO issues by controlling the system's common parameters. For instance, in [29], time-to-trigger (TTT) and handover margin (HOM) are tuned based on the type of RLF observed during the HO. Similarly, in [30], authors propose to tune A3-offset to prevent RLF between intra-frequency neighbors. Authors in [31] categorized HO failure into too early, too late, and wrong cell HO to adjust TTT and A3-offset accordingly. Apart from optimizing intra-frequency HO parameters, authors in [32] proposed transmission power adjustments to eliminate coverage holes in an attempt to avoid RLF. RLF detection approach in [33] used a radio frequency (RF) threshold to detect possible instances of RLF and accelerated the HO to a better cell if available. However, no scheme to determine optimal RF thresholds is not presented in the study.

On the other hand, the academic literature on voice call muting, specifically IP-based Voice over LTE (VoLTE) muting, is rather scarce. The primary reason for this is bi-fold; a) the low penetration rate of VoLTE calls due to the incapability of mobile handset, the inability of eNBs to cater VoLTE calls, or the reluctance of the mobile network operators to enforce VoLTE calls; b) the voice muting prevention is often treated as a separate problem on its own. Instead, the traditional hit-and-trial-based tuning of various parameters for coverage hole avoidance, SINR improvement, seamless handover, and resource availability is often assued to indirectly minimize voice muting. The RLF avoidance approaches discussed above [29]–[33] may minimize voice call muting as well, However,

the optimization techniques aimed specifically at voice muting prevention need to meet more stringent criteria than the aforementioned approaches. This is because, unlike traditional HTTP/FTP traffic, voice call requires real-time low-latency packet transfer for high definition and jitter-free voice communication. Hence, in a bid to camp the UE on the best available frequency band, network operators use different set of mobility parameters for VoLTE and ordinary data traffic.

In the context of voice call muting, the study of HO between WiFi access points [34] and radio resource scheduling [35] exist in literature. However, none of the existing studies aim to investigate a scheme for a QoE-aware dual-connectivity (EN-DC). Furthermore, as concluded earlier, most of the RLF prevention approaches proposed in literature target intra-frequency HO optimization and do not identify actual measurement thresholds to detect possible RLF. Therefore, there is a dire need for a framework that can detect potential RLF threshold and potential muting threshold considering the signal strength and quality, and utilize that information to configure the optimal inter-Random Access Technology (inter-RAT) parameters for resource-efficient and QoE-aware EN-DC activation.

This paper is the extension of our previously published conference paper [36]. In contrast to [36] the new contributions in this work include a AI model for voice muting minimization, and deep hyper-parameter optimization both in the development of AI models and when minimizing the impact of class imbalance. Moreover, in [36] the results shown were based on random values of B1 RSRP thresholds. On

the contrary, this work presents the optimization function formulation for RLF and muting aware EN-DC activation to get the optimal parameters required for EN-DC activation.

The proposed model has been shown in Fig. 2. Though there is clear similarity between the RLF and muting prediction models, they are distinct in two key aspects. 1) Different input data: RLF model relies only on RF data with past RLF labels as input to it, whereas muting prediction model relies on RF data with muting labels as well as the output of the RLF model as input to it. 2) Different class imbalance techniques: RLF prediction model works best with a different imbalance technique (Tomek Link) compared to the muting prediction model which performs best with GAN. Reason for this difference in performance of class imbalance handling techniques for the two models is explained in detail in Section III-D and III-E.

B. Contributions

The main contributions of this paper can be summarized as follows:

1) This is the first study to quantify and optimize the trade-off between 5G network utilization and user QoE degradation due to RLF or muting during EN-DC activation, leveraging real network data measurements. To optimize this tradeoff, we propose a two-stage AI model trained on real data from a live commercial network. The developed model is capable of accurately predicting QoE degradation in terms of potential RLF and muting instances, as shown in Fig. 2.

2) We present a domain knowledge-based scheme to enrich the minority RLF and muting samples by identifying potential RLF occasions using low-level counters specified by 3GPP. To address the data imbalance problem, we investigate the potential of a large number of data balancing approaches, including state-of-the-art oversampling and undersampling techniques, as well as non-traditional techniques such as Generative Adversarial Network (GAN).

3) In order to enable network operators to configure two different sets of mobility parameters for voice bearer activated UEs and non-voice active UEs, we formulate and solve two separate optimization problems that maximize EN-DC activation and minimize RLF or voice muting risk. Given the non-convexity of the optimization problems, we solve them using Genetic Algorithm (GA). To evaluate the efficacy of GA to yield a near optimal solution, we benchmark GA's performance against a brute force-based solution.

4) We perform extensive system-level simulations to evaluate the proposed AI-based QoE aware EN-DC activation framework while comparing it with the state-of-the-art industry solution for EN-DC activation (i.e., EN-DC activation without taking into account RLF or voice muting risk). Results show that the proposed solution can intelligently configure EN-DC activation criteria such that RLF or voice muting can be reduced to practically zero.

The rest of the paper is organized as follows. In Section II, we briefly describe the 3GPP based EN-DC activation procedure, RLF trigger conditions, and the support of voice calls over cellular networks. Real LTE network measurement data collection, exploration, and data imbalance issue are described in Section III. The two stage AI model to predict potential RLF and potential muting is also presented in Section III. Optimization problem formulation for an efficient RLF and mute-aware EN-DC activation criteria is discussed in Section IV. In Section V we show that the Minimization of Drive Test (MDT) data can be used to determine the suitable EN-DC activation configuration parameters while minimizing the chances of RLF and voice muting. Finally, we conclude the paper in Section VI.

II. BACKGROUND

In this section, we briefly describe the 3GPP standard based procedures for EN-DC activation, RLF trigger criteria, and the support of voice calls over cellular networks.

A. EN-DC in 3GPP Release 15

A major focus of 3GPP Release 15 [3] is to get the first incarnation of 5G NR into the field that complements the existing 4G LTE network. To make this happen, a solution in the form of Dual Connectivity option 3X or EN-DC is crafted. EN-DC enables UEs to connect simultaneously to 4G and 5G NR base stations. Under this solution, a UE first camps on 4G eNB and then initiate activation of EN-DC. Fig. 1(a) illustrates EN-DC signaling and data connections.

Master node (MN), LTE eNB in this case, starts the EN-DC activation process by sending the EN-DC configuration message to the UE. This message contains the event B1 measurement criteria that define the 5G RF threshold. 5G capable UE sends event B1 to the MN if the Reference Signal Received Power (RSRP) or Signal to Interference and Noise Ratio (SINR) of the 5G cell becomes better than the B1-threshold as shown in Fig. 1(b). The entering condition of event B1 can be expressed as:

$$M_n + O_{fn} + O_{cn} - hyst > B1_{thres} \qquad (1)$$

where M_n is the measurement result, either RSRP or SINR, of the 5G gNB, $hyst$ is the hysteresis parameter, O_{fn} and O_{cn} are the optional frequency and cell offset parameters, respectively.

Once the MN receives the event B1 from the UE, it communicates with the 5G gNB for admission control check, and capability enquiry. 5G gNB is referred to as secondary node (SN) after the EN-DC activation.

B. Radio Link Failure in 3GPP

Radio link failure is an instance when a UE abnormally detaches its connection with the serving cell. RLF procedure is the same for 4G and 5G networks and is observed when either of the following three conditions is met continuously for a certain period.

- When timer $T310$ expires after consecutive out-of-sync indication parameter $N310$ has expired.
- When the configured number of consecutive unsuccessful Random Access Channel (RACH) attempts have been reached.
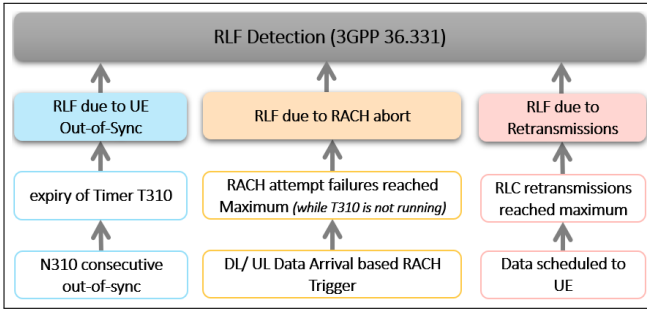
Figure 3: Summary of 3GPP-defined RLF trigger criteria.

- When the number of consecutive Radio Link Control (RLC) retransmissions equals the value of the parameter $maxRLCretransmissions$.

These conditions are also summarized in Fig. 3.

C. Voice Over Cellular Networks

The legacy 4G networks and the latest 5G NR networks support voice services through VoLTE and Voice over NR (VoNR) respectively. The packet switch-based VoLTE and VoNR deliver high definition voice with much lesser jitter and delay than the traditional circuit switch networks. 3GPP [37] has standardized QoS Class Identifier (QCI) value of 5 for voice call signaling, and QCI of 1 for actual voice call packets. Resource scheduling for a voice activated user is achieved through Semi-Persistent Scheduling (SPS) where a fixed number of resources are allocated with high priority in a periodical manner and at predefined location within the bandwidth. This is done to minimize the service interruption for an active voice call due to resource congestion.

D. Relationship Between RLF and Muting

Although VoLTE and VoNR are given higher resource allocation priority, they are still susceptible to muting under poor RF conditions due to the drop or loss of voice packets. Under worst circumstances, the voice muting can extend for several seconds, which can be detrimental to user experience. Unlike RLF which has underlying counters that trigger RLF, voice muting is not dependent on any underlying parameters. This is due to the real-time flow of packets between the two participants, and call muting can be observed almost instantly with deteriorating signal strength (RSRP) or quality (SINR). Given the time sensitive nature of voice packets, call muting can happen even if actual RLF does not happen and only the conditions that can lead to potential RLF occur.

III. AI MODEL FOR RLF AND MUTING PREDICTION TO ENABLE SMART EN-DC ACTIVATION

This section describes the process of collecting actual measurement data from a real 4G network to the stage where we develop AI models. The AI based prediction models are designed to predict if a given set of RSRP and SINR conditions are indicator of potential RLF and muting instance. These models are trained by using classification approaches on the historic data as explained in Fig. 2 and Section II. Note here that we cannot employ a rule-based mechanism to identify the

RLF or voice muting, as the UE is limited to measure the RLF counters only when it camps on the respective cell. The AI models developed in this article can therefore proactively avoid poor QoE due to potential RLF or voice muting before the UE camps on the 5G cell, i.e., before the UE starts monitoring 3GPP based RLF counters. In other words, 3GPP rule-based mechanism allows only reactive observation of the RLF after EN-DC and thus cannot help in minimizing the RLF or muting during EN-DC, whereas proposed AI based scheme allows proactive prediction of the potential RLF and muting takes that into account for optimal EN-DC activation. In this section, we also highlight the challenge of severe class imbalance in the collected data and analyze several methods to address the issue.

A. Data Collection, Cleansing and Pre-Processing

1) Potential RLF: A drive test in a commercially deployed 4G network is conducted for a total of 13 hours. Measurements including RSRP and SINR are recorded together with the low level RLF related parameters mentioned in Section II-B at a time interval of 100ms. Out of the 460,000 data samples recorded, the observed instances of actual RLF are only 543 ($\sim$7 RLF every 10 minutes). Using this highly imbalanced raw data with very few RLF occurrences for training the AI-models will give results biased towards the dominant class i.e., no RLF. To address this problem, we enhance the data set by incorporating all the chances of possible RLF for making the model more robust in detecting RLF. Using domain knowledge, we identify sets of RSRP and SINR combinations where the corresponding underlying RLF related parameters ($T310$, $N310$, $N311$, $maxRACHattempts$, $maxRLCretransmissions$) show abnormality. We designate these points as potential RLF. The processed RF data with potential RLF instances label is shown in the Fig. 4.

The data pre-processing involving the integration of the low-level RLF related parameters leading to the actual RLF brings the total RLF samples to 27,794. However, this number of RLF samples still constitutes a very small fraction of the 460,000 total samples. This data when used for training the Machine Learning (ML) models gives varying performance depending on the ML tool used. However, due to severe class imbalance between non-RLF and RLF classes, the performance remains poor for all ML models. The performance of the state-of-the-art ML models trained using an imbalanced dataset is shown under the Raw Data column of Table II. A key observation to be made from results in column 1 of the Table II that due to extreme class imbalance, simpler ML models may overfit more thus giving high accuracy by classifying almost all samples as normal. More advanced models with a larger number of training parameters such as XGBoost and deep learning do not overfit to that extent but still, have high miss-classification rate as indicated by their lower accuracy. This observation highlights the significance of the class imbalance problem in real data. This problem is addressed in a later section. Accuracy metric in Table II is defined as below:

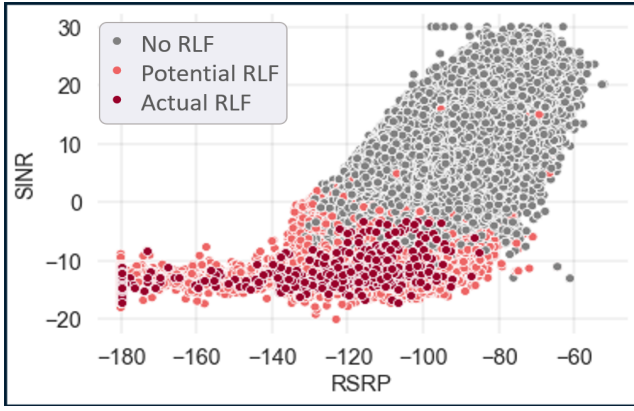$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (2)$$

Figure 4: Potential RLF occurrences versus the UE RSRP and SINR measurements.

where TP and TN are the the true positive and true negative rate, and FP and FN refer to false positive and false negative predictions.

2) Potential Muting: VoLTE call based drive test is conducted for eight hours and call muting related data such as RSRP and SINR measurement are recorded every 100ms. To accurately identify the RF condition that can lead to call muting, we place one static call participant under good RF conditions. The other participating user is placed in a moving vehicle with continuously changing RF conditions. Distinguishing voice muting is not as straightforward as identifying RLF. Unlike RLF wherein users send flag to the base station when it occurs, there does not exist any flag to identify voice muting. Instead, using domain knowledge, we look into the Real-time Transfer Protocol (RTP) packets transmission to determine clues for the voice muting. RTP packets are continuously exchanged between the UE and the BS during the call period, and the absence of RTP packets can suggest muting occasions. However, the absence of RTP packets exchange can also be observed during the call setup phase and moments after call termination which, if not identified properly, can lead to erroneous muting identification. For this reason, we designate the data points as voice muting only if the RTP packets are absent while the voice call is in the established phase. With this setup, out of the 0.3 million data samples recorded, we observe 2092 actual voice muting instances ($\sim$4.36 muting instances per minute). Similar to RLF, the fraction of actual muting is far lesser than the normal case.

For both the potential RLF and potential mute cases, we first scaled the input data, and then performed a train-test split of 80%-20%. On the training data, we addressed data imbalance by applying several techniques as discussed below.

B. Addressing Data Imbalance

As observed from results in column 1 of Table II, a key challenge in creating an RLF and muting prediction model is the training data class imbalance. If used without a class balancing technique, ML models trained on the data will either be highly biased towards the majority class mostly missing the minority class instances, or will have low overall accuracy. Misclassification of the minority class will be detrimental to the fidelity of the model as in our context, the minority

class (potential RLF/muting class) is the class of interest. For that reason, data imbalance problem must be addressed to have meaningful results. Common techniques used for data balancing fall under two categories namely oversampling and undersampling. The former augment the minority class to match the size of the majority class while the latter performs the opposite. In the following, we briefly discuss the approaches we leverage to address the data imbalance problem. In the following discussion, we denote the minority class and majority class as C_{min} and C_{maj}, respectively.

1. *Oversampling Techniques:*
   - Random over sampling randomly duplicates observations from the C_{min} to reinforce its signal.
   - Synthetic Minority Oversampling Technique (SMOTE) synthesizes new minority instances.
   - In Generative Adversarial Network (GAN), two neural networks contest with each other in the training phase. The goal of the first neural network is to befool the second neural neural network by generating synthetic data that resembles the input training data. The role of the second neural network is to correctly identify the synthetically produced data. In this context, we use GAN to create new synthetic new samples of the minority class i.e., oversample the C_{min}.

2. *Undersampling Techniques:*
   - Random under-sampling randomly removes observations from the C_{maj}.
   - Near miss algorithm eliminates the majority class data point after identifying the two nearest samples in the distribution belonging to different classes, thereby trying to balance the distribution.
   - Condensed Nearest Neighbor Rule (CNN) works by classifying each sample of C_{maj} using kNN (k-nearest neighbor) with k=1, and misclassified samples are reassigned to C_{min}.
   - A pair of data instances (x_i, x_j) where $x_i \in C_{min}$, $x_j \in C_{maj}$ and $d(x_i, x_j)$ is the distance between x_i and x_j, is called a Tomek link if there is no data instance x_k ($x_k \in C_{min}$ or $x_k \in C_{maj}$) such that $d(x_i, x_k) < d(x_i, x_j)$ or $d(x_j, x_k) < d(x_i, x_j)$. The tomek link algorithm removes the unwanted overlap between C_{min} and C_{maj} by removing majority class
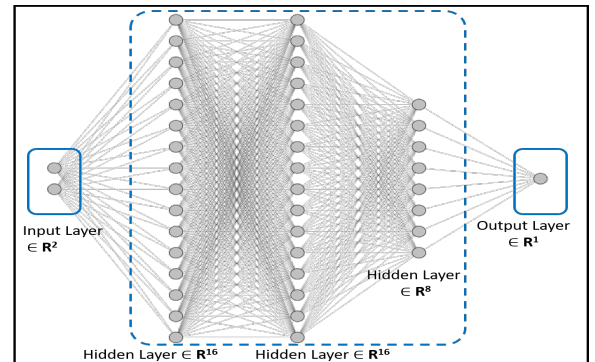

Figure 5: Structure of the deep learning based model for predicting potential RLF. The model is trained, tested and validated after addressing data imbalance using Tomek link.

Table II: Accuracy and F-1 score of the potential RLF models against various data-imbalance resolution techniques.

| Classification Algorithm	Metric	Raw Data	Random over sampling	Smote	Random under sampling	Near Miss	CNN	Tomek Links	ENN	NCL	Cluster Centroids	GAN
Logistic Regression	Accuracy	97%	88%	89%	88%	95%	95%	97%	98%	97%	90%	97%
KNN	Accuracy	98%	98%	95%	98%	88%	97%	97%	96%	97%	98%	97%
SVM	Accuracy	97%	89%	89%	89%	89%	97%	89%	89%	97%	94%	97%
Naive Bayes	Accuracy	97%	88%	90%	97%	95%	95%	96%	95%	96%	90%	96%
Decision Trees	Accuracy	97%	93%	90%	90%	48%	93%	97%	96%	96%	88%	96%
Random Forest	Accuracy	79%	94%	93%	93%	57%	97%	98%	97%	97%	94%	97%
XGBoost	Accuracy	78%	93%	91%	91%	78%	97%	98%	97%	97%	94%	97%
Deep Learning	Accuracy	74%	89%	89%	89%	72%	97%	99%	72%	98%	94%	97%
Logistic Regression	F1	0.75	0.88	0.49	0.88	0.68	0.67	0.74	0.74	0.74	0.53	0.75
KNN	F1	0.78	0.78	0.68	0.78	0.44	0.75	0.78	0.72	0.77	0.69	0.79
SVM	F1	0.73	0.88	0.88	0.88	0.88	0.75	0.88	0.88	0.74	0.63	0.76
Naive Bayes	F1	0.70	0.88	0.50	0.70	0.62	0.69	0.70	0.66	0.69	0.51	0.72
Decision Trees	F1	0.75	0.89	0.90	0.90	0.16	0.57	0.75	0.73	0.74	0.47	0.75
Random Forest	F1	0.86	0.90	0.92	0.92	0.20	0.77	0.79	0.78	0.79	0.66	0.80
XGBoost	F1	0.88	0.91	0.91	0.91	0.31	0.76	0.88	0.78	0.74	0.64	0.79
Deep Learning	F1	0.87	0.88	0.88	0.88	0.10	0.76	0.93	0.10	0.80	0.62	0.76
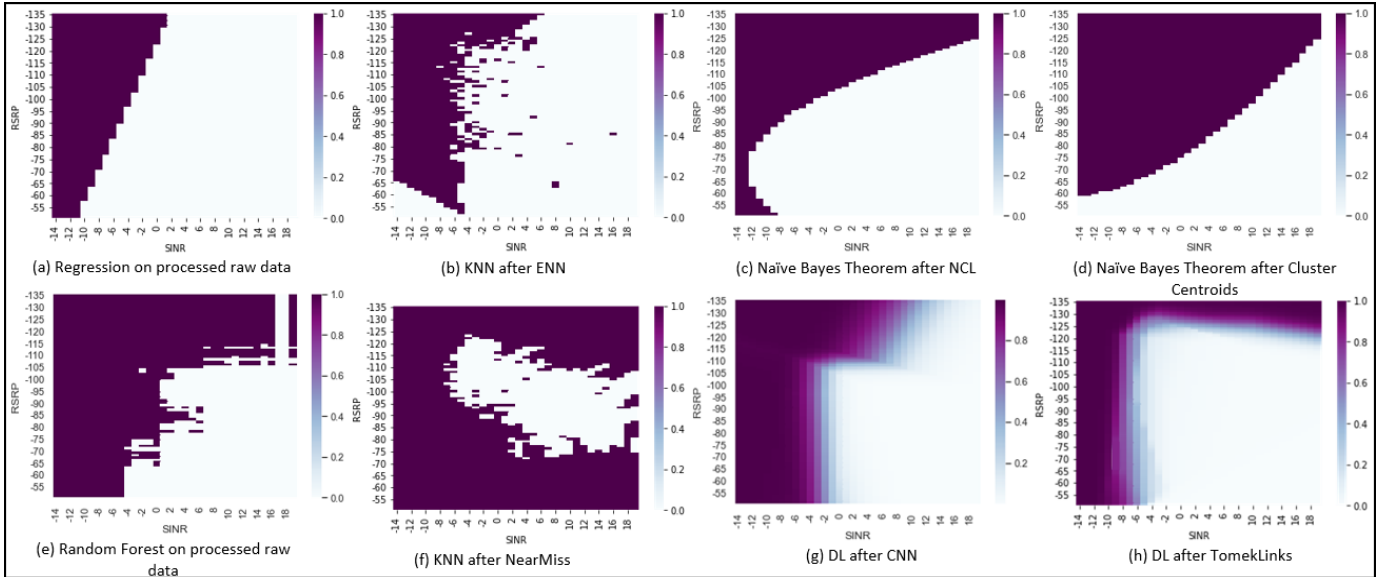
Figure 6: Decision boundary of the potential RLF models shown in Table II.

Table III: Deep learning hyper-parameters for potential RLF model.

Hyperparameter Name	Search Range/Value
DNN depth d	{1,2,3,5}
DNN width w	{5,8,10,16}
Activation Function (Hidden Layers)	Relu
Activation Function (Output Layers)	Sigmoid
Optimizer	Adam (Gradient Descent)
Loss Metric	Binary Cross Entropy

sample from Tomek link data pair. This is done based on the assumption that for the data points that form a Tomek link, either one of them is a noise or both are in the borderline.

- Edited Nearest Neighbor Rule (ENN) removes any instance whose class label is different from the class of at least two of its three nearest neighbors.
- Neighborhood Cleaning Rule (NCL) modifies the ENN

where three neighbors of each data point are found. If the classification of the data point $x_j \in C_{maj}$ given by its three neighbors contradicts the original class of x_i, then x_i is removed. Conversely, if the data point $x_i \in C_{min}$ and the three neighbors miss-classify x_i as a majority class sample, then the nearest neighbors that belong to the majority class are removed.

- In cluster centroids under-sampling, we find the clusters of the majority class with K-mean algorithms. Then it replaces the cluster points with cluster centroids as the new majority samples.

C. AI Model for Potential RLF

After addressing the class imbalance problem, we design to train and evaluate models using a range of state-of-the-art ML techniques including logistic regression, KNN, SVM,

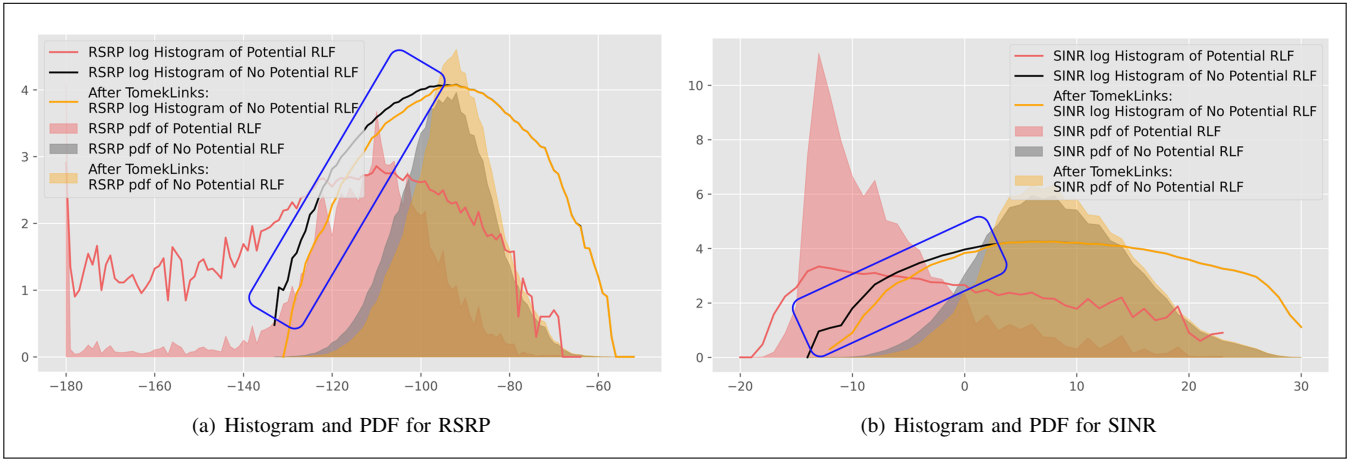(a) Histogram and PDF for RSRP   (b) Histogram and PDF for SINR

Figure 7: Effect of Tomek Links in addressing data imbalance and improving class isolation (highlighted by blue rectangle).

Naive Bayes, decision trees, random forest, XGBoost, and deep learning based models.

To achieve optimal performance, we perform hyperparameter optimization for each ML algorithm. To avoid under or over-fitting, we investigate a variety of deep learning neural network architectures with a range of hyper-parameters as shown in Table III. Our experiments show that a deep learning model with fully connected three hidden layers with 16, 16 and 8 neurons, respectively as shown in the Fig. 5, yields the best results. The model is trained using epoch size of 100 and batch size of 10.

Table II shows the accuracy and F1-score for various ML models. Results show that deep learning with data imbalance problem addressed by Tomek links outperforms others with accuracy and F1 score of 99% and 0.93, respectively. The superior performance of Tomek links stems from its ability to delineate the class boundaries and remove noise and thus make data less ambiguous. It helps to improve the isolation between the overlapped classes by removing the majority samples at the border area as illustrated in Fig. 7.

The decision boundaries of the ML models trained using data from different data balancing techniques are shown in Fig. 6. As illustrated in Fig. 6(h), the decision boundary created by the deep learning model using balanced data from Tomek links indicates the expected increase in potential RLF with deterioration of either RSRP or SINR.

D. AI Model for Potential Muting

We follow a similar approach to build an AI model for potential muting prediction as in III-C. Table IV shows the performance of various ML algorithms in predicting the potential muting instances.

Unlike the potential RLF model where Tomek link based preprocessing combined with deep learning worked best, here deep learning model trained with data using GAN shows the best results in terms of F1 score. For this model, we also investigate a range of deep learning architectures with a variety of hyper-parameters to prevent under- or over-fitting as shown in table III. Our experiments show that a deep learning model with fully connected three hidden layers of 8, 8 and

4 neurons, respectively yields the best performance for voice muting prediction. The model is trained using epoch size of 100 and batch size of 10.

Since data generation using GAN is more complex compared to other SOTA undersampling and oversampling approaches, we conduct an additional layer of sanity check by comparing the real data and synthetic data generated by GAN. Kullback–Leibler (KL) and Jensen-Shannon (JS) divergence of the GAN generated samples from the real data, along with the probability density function (PDF) of the original minority class and the synthetically generated minority class data is shown in Fig. 9. The results show that GAN has produced synthetic data that closely resembles real data.

E. Why GAN Based Data Augmentation has Positive Gain for Voice Muting Data and Negative for RLF Data

Fig. 8 shows that the class separation in the voice muting data is much more pronounced than the RLF prediction shown in Fig. 7. This difference in the class distribution stems from the fact that: a) both classes belong to different metrics i.e., potential RLF and potential muting, b) network configures voice bearer activated UE with a different and more aggressive set of mobility parameters to keep the UE in a better RF condition at all times. In the potential RLF case, the class distribution in Fig. 7 is more overlapped. As Tomek links is an approach designed to improve the class border isolation, it leads to much better performance compared to other approaches, including GAN. In the case of RLF data, the use of GAN to augment minority class worsens the performance. This is because the GAN-generated synthetic data further blurs the boundary between the classes.

On the contrary, in the case of call muting data, where boundaries are relatively well defined, GANs generated data addresses the class imbalance successfully without increasing the boundary overlap to a level that would undermine the performance of the model.

This is an insightful finding that suggest that despite their popularity for augmenting training data, GANs should be used cautiously as their gain can be negative instead of positive
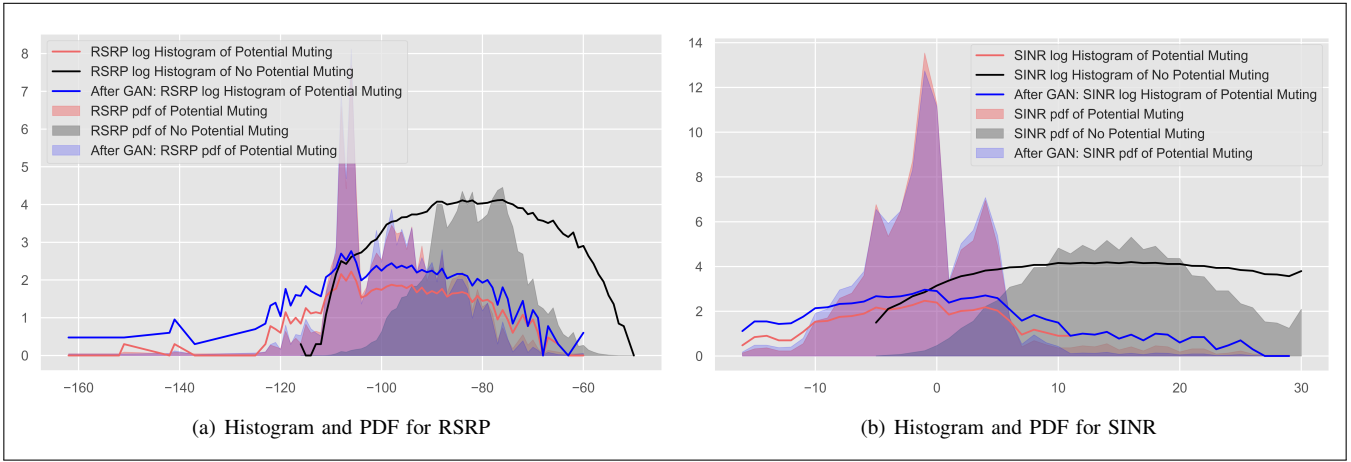
(a) Histogram and PDF for RSRP

(b) Histogram and PDF for SINR

Figure 8: Effect of GAN in mitigating the class imbalance issue.

depending on the class distributions and overlap in the training data.

IV. QoE Aware EN-DC Activation

In this section, we describe the optimization function formulation for RLF and muting aware EN-DC activation.

A. EN-DC Activation, RLF, and Muting Formulation

Current industry practice is to maximize the EN-DC activation instances for maximum utilization of the 5G network. However, simply maximizing the EN-DC activation without taking into consideration of the underlying trade-offs can be detrimental to the user QoE due to higher RLF and voice muting instances. Hence, mobile network operators should take into account the following objectives when enabling EN-DC in their network:

- Maximize EN-DC request by the EN-DC capable UE to fully leverage the 5G NR features.
- Facilitate EN-DC activation for each EN-DC request, which is essentially minimizing the difference between the number of EN-DC requests and EN-DC activations.
- Avoid degradation in retainability due to RLF at either 4G or 5G network after EN-DC activation.
- Prevent voice muting after activating EN-DC for UEs with voice service demands.

Using the notations defined in Table V, we can define the number of EN-DC activations α as:

$$\alpha(\mathbf{\Delta}^u, \mathbf{\Theta}, U_c) = \sum_{u \in U_c} 1\left[\mathbf{\Delta}_i^u > \mathbf{\Theta}_i \ \forall \ i\right] \quad (3)$$

where $1\{\cdot\}$ is the indicator function, and the subset $U_c \subseteq U$ is the set of EN-DC capable UEs configured with B1 measurement report. $\mathbf{\Delta}_i^u$ is the i-th element of the set of RF condition of user $u \in U_c$, i.e., for any user u, $\mathbf{\Delta}_1^u = \delta_{5R}^u$, $\mathbf{\Delta}_2^u = \delta_{4R}^u$, $\mathbf{\Delta}_3^u = \delta_{5S}^u$, $\mathbf{\Delta}_4^u = \delta_{4S}^u$. Similarly, the i-th element of the set of thresholds is $\mathbf{\Theta}_i$, where $\mathbf{\Theta}_1 = \theta_{B1}$, $\mathbf{\Theta}_2 = \theta_{4R}$, $\mathbf{\Theta}_3 = \theta_{5S}$, $\mathbf{\Theta}_4 = \theta_{4S}$.

(a) PDF, KL and JS Divergence for RSRP

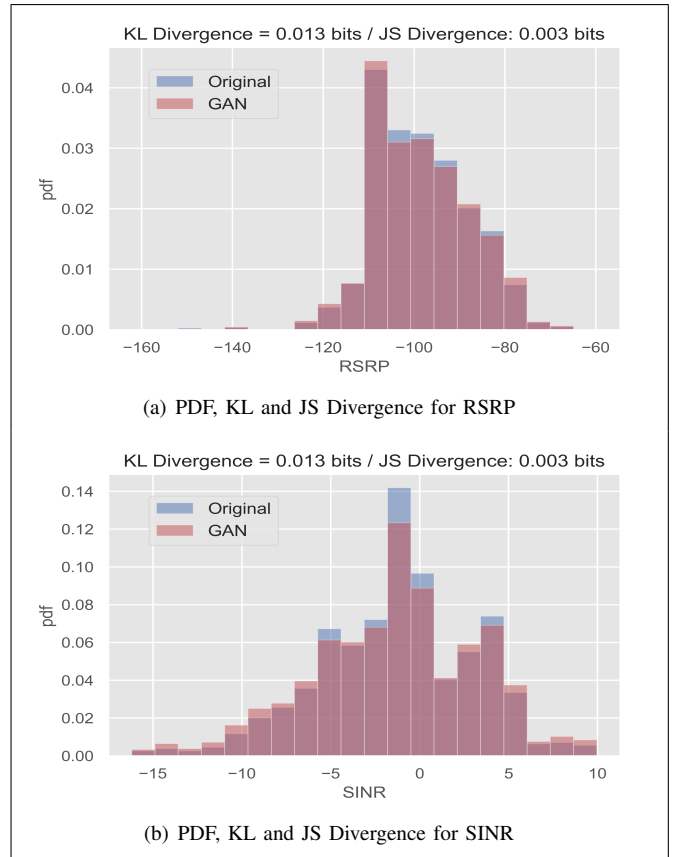(b) PDF, KL and JS Divergence for SINR

Figure 9: Comparison of the original minority class (muting instances) and the synthetic data generated from GAN.

UEs may experience RLF after EN-DC activation due to poor RF conditions. The number of RLF instances denoted here by β can be defined as:

$$\beta(\mathbf{\Delta}^u, \zeta, U_a) = \sum_{u \in U_a} max\left(\zeta(\mathbf{\Delta}^{u,4}), \zeta(\mathbf{\Delta}^{u,5})\right) \quad (4)$$

where $U_a \subseteq U_c$ is the set of EN-DC activated UEs, and ζ is the potential RLF AI-Model, which takes in $\mathbf{\Delta}^u$ as input and outputs a prediction of 1 or 0 representing the occurrence

Table IV: Accuracy and F-1 score of the potential voice muting models against various data-imbalance resolution techniques.

Classification Algorithm	Metric	Raw Data	Random over sampling	Smote	Random under sampling	Near Miss	CNN	Tomek Links	ENN	NCL	Cluster Centroids	GAN
Logistic Regression	Accuracy	99%	91%	89%	91%	98%	99%	99%	99%	99%	91%	99%
KNN	Accuracy	99%	95%	97%	93%	95%	99%	99%	99%	99%	98%	99%
SVM	Accuracy	99%	94%	96%	93%	94%	99%	99%	99%	99%	87%	99%
Naive Bayes	Accuracy	99%	89%	86%	90%	91%	99%	99%	99%	99%	87%	99%
Decision Trees	Accuracy	99%	97%	97%	89%	88%	97%	99%	99%	99%	90%	99%
Random Forest	Accuracy	99%	97%	98%	92%	92%	99%	99%	99%	99%	97%	99%
XGBoost	Accuracy	99%	96%	97%	93%	91%	99%	99%	99%	99%	29%	99%
Deep Learning	Accuracy	99%	93%	96%	93%	94%	99%	99%	99%	99%	96%	99%
Logistic Regression	F1	0.41	0.61	0.11	0.71	0.30	0.45	0.41	0.44	0.43	0.12	0.51
KNN	F1	0.43	0.59	0.21	0.73	0.28	0.45	0.48	0.45	0.48	0.39	0.57
SVM	F1	0.41	0.63	0.27	0.73	0.28	0.43	0.44	0.47	0.49	0.24	0.76
Naive Bayes	F1	0.40	0.60	0.09	0.70	0.27	0.37	0.39	0.42	0.40	0.09	0.79
Decision Trees	F1	0.41	0.60	0.20	0.69	0.29	0.17	0.45	0.46	0.44	0.11	0.46
Random Forest	F1	0.43	0.59	0.33	0.62	0.31	0.42	0.44	0.49	0.49	0.41	0.84
XGBoost	F1	0.41	0.60	0.33	0.63	0.35	0.45	0.45	0.50	0.48	0.47	0.86
Deep Learning	F1	0.45	0.63	0.22	0.62	0.32	0.47	0.51	0.47	0.48	0.26	0.89

Table V: List of symbols used in optimization problem formulation.

Symbol	Description	Symbol	Description
U	Set of all UEs	u	Any user $u \in U$
U_c	Set of UEs with EN-DC configuration	U_a	Set of EN-DC activated UEs
δ_{5R}^u	5G RSRP of u	θ_{B1}	5G RSRP threshold
δ_{4R}^u	4G RSRP of u	θ_{4R}	4G RSRP threshold
δ_{5S}^u	5G SINR of u	θ_{5S}	5G SINR threshold
δ_{4S}^u	4G SINR of u	θ_{4S}	4G SINR threshold
$\mathbf{\Delta}^u$	$[\delta_{5R}^u, \delta_{4R}^u, \delta_{5S}^u, \delta_{4S}^u]$	Θ	$[\theta_{B1}, \theta_{4R}, \theta_{5S}, \theta_{4S}]$
$\mathbf{\Delta}^{u,4}$	$[\delta_{4R}^u, \delta_{4S}^u]$	$\mathbf{\Delta}^{u,5}$	$[\delta_{5R}^u, \delta_{5S}^u]$
α	EN-DC Activation function	ζ	Potential RLF AI-Model
β	RLF function	η	Potential Muting AI-Model
γ	Muting function	-	-

of potential RLF and no RLF, respectively. The output of the potential RLF AI-model is represented as $\zeta(\mathbf{\Delta}^u)$. We consider the output of potential RLF AI-model for both 4G and 5G settings as the UE will experience RLF if either of the 4G or 5G RF conditions are bad.

Similarly, for the set of UEs requiring voice services, the number of muting instances γ can be written as:

$$\gamma(\mathbf{\Delta}^u, \eta, U_a) = \sum_{u \in U_a} max\left(\eta(\mathbf{\Delta}^{u,4}), \eta(\mathbf{\Delta}^{u,5})\right) \quad (5)$$

where the potential muting AI-Model η takes in $\mathbf{\Delta}^u$ as input and outputs a prediction of 1 or 0 representing the occurrence of potential muting and no muting, respectively. The output of the potential muting AI-model is represented as $\eta(\mathbf{\Delta}^u)$.

Operators can increase EN-DC activations by configuring lower values of EN-DC thresholds Θ. However, this can lead to RLF or voice muting after EN-DC activation, rendering the dual connectivity procedure useless. Keeping in view this tradeoff, the optimization problem in subsection IV-B and IV-C is formulated to achieve maximum utility and resource efficiency.

B. RLF Aware EN-DC Optimization

We formulate a multi-objective optimization problem such that it maximizes EN-DC activations while minimizing the chances of RLF occurrences. This RLF aware EN-DC optimization function is given as:

$$\underset{\Theta^r = [\theta_{B1}^r, \theta_{4R}^r, \theta_{5S}^r, \theta_{4S}^r]}{\operatorname{argmax}} \quad \frac{\alpha^w}{\beta^{(1-w)}}$$

$$\text{subject to} \quad \theta_{B1,low}^r \leq \theta_{B1}^r \leq \theta_{B1,high}^r,$$
$$\theta_{4R,low}^r \leq \theta_{4R}^r \leq \theta_{4R,high}^r, \quad (6)$$
$$\theta_{5S,low}^r \leq \theta_{5S}^r \leq \theta_{5S,high}^r,$$
$$\theta_{4S,low}^r \leq \theta_{4S}^r \leq \theta_{4S,high}^r,$$
$$w < 1.$$

where w is the operator defined weight that can be used to adjust the relative importance of EN-DC activations (α), and RLF (β). θ_{B1}^r, θ_{4R}^r, θ_{5S}^r and θ_{4S}^r are the optimization variables for RLF aware EN-DC activation. The first four constraints limit the parameters in the 3GPP defined range. The range of optimization variables and constraints indicate that (6) is a large-scale non-convex NP-hard problem due to the inherent coupling of optimization parameters and the EN-DC requests. Non-convexity stems from the fact that we are dealing with four integer metrics (RSRP and SINR of 4G and 5G) in a heterogeneous multi-RAT network deployment, where the randomness in UE location and resource requirement result in variable cell loads that affect 4G and 5G SINR metrics differently. In addition, the 4G and 5G RSRP also change with the distance from the BS, however, non-uniform BS deployment along with user mobility makes RSRP non-deterministic.

C. Voice Muting Aware EN-DC Optimization

As voice users require low-latency and jitter-free communication unlike FTP/HTTP users, network operators can configure a different sets of mobility parameters for UEs with active voice bearers. This allows operators to keep the
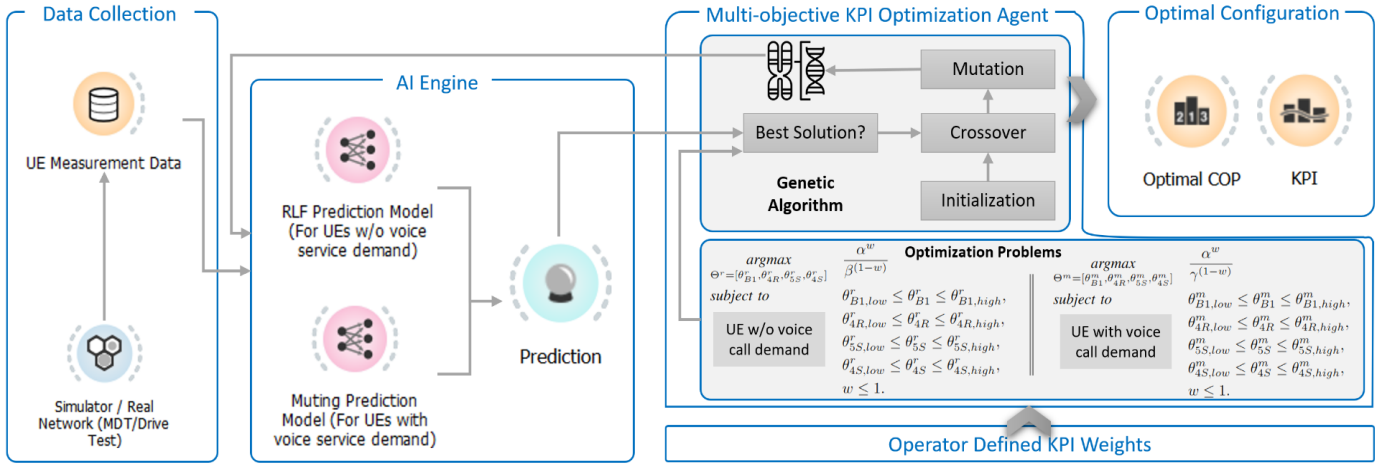
Figure 10: Proposed smart EN-DC activation framework.

UE undergoing a voice call in good radio conditions, and other factors like load balancing and handover rate are given less priority. To enable intelligent exploitation, we design a separate optimization problem for the voice UEs. The EN-DC activation parameters returned by the voice muting aware EN-DC optimization problem will be used to configure only voice UEs. The multi-objective optimization for voice muting aware EN-DC activation can be formulated as follows:

$$
\begin{aligned}
\underset{\Theta^m=[\theta_{B1}^m,\theta_{4R}^m,\theta_{5S}^m,\theta_{4S}^m]}{\mathrm{argmax}} \quad & \frac{\alpha^w}{\gamma^{(1-w)}} \\
\text{subject to} \quad & \theta_{B1,low}^m \leq \theta_{B1}^m \leq \theta_{B1,high}^m, \\
& \theta_{4R,low}^m \leq \theta_{4R}^m \leq \theta_{4R,high}^m, \qquad (7) \\
& \theta_{5S,low}^m \leq \theta_{5S}^m \leq \theta_{5S,high}^m, \\
& \theta_{4S,low}^m \leq \theta_{4S}^m \leq \theta_{4S,high}^m, \\
& w < 1.
\end{aligned}
$$

where w is the operator defined weight that can be used to adjust the relative importance of EN-DC activations (α), and muting (γ). θ_{B1}^m, θ_{4R}^m, θ_{5S}^m and θ_{4S}^m are the optimization variables for voice muting aware EN-DC activation and the first four constraints keep their values in the 3GPP defined range.

V. PROPOSED SMART EN-DC ACTIVATION FRAMEWORK AND SIMULATION RESULTS

Fig. 10 illustrates the high level overview of the proposed AI powered EN-DC activation framework. We develop a cascaded 2-stage AI model where stage-1 works to predict RLF while muting prediction is done by the second stage. The predictions obtained from stage-1 are used alongside actual voice muting samples obtained from real world measurements. Finally, the second stage AI model is trained to predict potential voice muting. Next, the optimization agent evaluates the objective function in the multi-objective Key Performance Indicator (KPI) optimization problem formulated in the previous section. This is done keeping in view the operator defined weightage to the number of EN-DC activations and the number of RLF/mute. As the problems in (6) and (7) are non-convex, they can be solved using brute force (BF) algorithm. However, the large convergence time of brute force is not suitable for time

sensitive problems of RLF, and muting explored in this paper. Heuristic optimization algorithms can be explored to deal with the large convergence time of BF. The conventional algorithms such as greedy algorithm can produce a local optimal solution but sometimes may fail to produce global optimal solution. In this backdrop, we analyze genetic algorithm for solving eq. (6) and (7)), which though known to be less efficient compared to simpler greedy algorithm has higher chances of converging to global optima in iterations considerably less than the BF. In future extension of this study where more advanced objective function with many more KPIs and constraint will be considered, an extensive investigation of optimization algorithm might be needed to find the most suitable solution approach.

Our analysis reveals that GA can converge faster than the BF approach, which makes the solution agile [38]. The faster convergence is particularly useful for rapidly changing network conditions. As illustrated in Fig. 10, the optimization agent iteratively gets the RLF/muting prediction from the AI engine for given EN-DC Configuration and Optimization Parameters (COPs), and the optimal COPs that yield the maximum utility function are obtained.

A. Simulation Setup

As discussed in earlier sections, the AI models for RLF and mute detection are developed from the insights drawn from the data preprocessing, model building, and testing process, and these models are based on real network data. However, as operators do not allow to experiment on a live commercial network, we evaluate the performance of our proposed EN-DC activation framework using the simulated data obtained from a state-of-the-art 3GPP compliant simulator called Synthetic-NET [39]. SyntheticNET is chosen as it has the key features that are needed for this study but are missing in most other simulators including 3GPP-based detailed HO and mobility management.

A multi-RAT network with nine macro 4G eNBs each having three sectors, and sixteen higher frequency omni directional 5G gNBs are deployed in a square of 25km^2 area. LTE eNBs are laid out uniformly in a grid form, while 5G small cells are deployed randomly representing hotspot locations. A total of 300 mobile UEs traverse the area following random way point mobility model. The speed of the users is set to

Table VI: Simulation details for Smart EN-DC activation.

Technology	4G LTE	5G NR
Frequency	2.1GHz	3.5GHz
Cell Type	Macro Cell	Small Cell
Antenna Type	Directional	Omni
Number of Transmitters	27	16
Transmit Power	40dBm	30dBm
Base Station Height	30m	20m
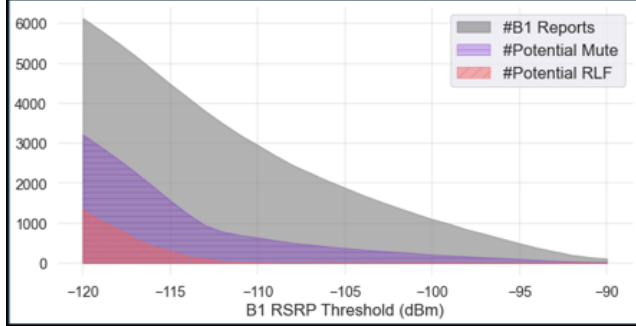
Figure 11: Number of UE generated B1 reports (EN-DC activation requests) against RSRP threshold.

Table VII: Optimal parameters obtained from GA for a UE with a data call requirement.

w	Algo	Iterations	Utility	Optimal Parameters $\Theta^r = [\theta_{B1}^r, \theta_{4R}^r, \theta_{5S}^r, \theta_{4S}^r]$
1	BF	741,321	1246.5	-120dBm, -120dBm, -6dB, -7dB
	GA	335	1225.5	-120dBm, -119dBm, -8dB, -6dB
0.5	BF	741,321	47.2	-112dBm, -118dBm, -3dB, -2dB
	GA	2890	46.1	-112dBm, -118dBm, -7dB, -2dB
0	BF	741,321	2.2	-108dBm, -118dBm, -1dB, -2dB
	GA	5543	2.1	-108dBm, -118dBm, -2dB, -2dB

Figure 12: Number of EN-DC activations and RLF observed when using optimal parameters in Table VII.

120km/h and the simulations are run for 12,000ms (equivalent to 12,000 Transmission Time Intervals 'TTIs'). After running the simulation, using the RSRP range [-120dBm, -90dBm], and the SINR range [-10dB, 10dB], we generate KPIs for a total of 741,321 distinct combinations of the four optimization parameters. More detail about the network configuration can be found in Table VI.

UEs are configured to measure RF condition of 5G gNB every 0.5s, and an event B1 measurement report is sent to the MN if the B1 criteria are met. Fig. 11 shows the effect of changing the B1 threshold on the number of B1 reports (EN-DC requests), potential RLF, and potential mute occurrences. Fig. 11 signifies the need for a smart EN-DC activation scheme i.e., the importance of optimally assigning B1 threshold. An incorrect B1 threshold may deteriorate retainability KPI or integrity KPI through a large number of RLF instances, and voice muting. Note that the legacy mobile networks use -120dBm B1 RSRP threshold to maximize the 5G network utilization without taking into consideration the ensuing QoE degradation.

B. Performance Evaluation

We implement the proposed framework in SyntheticNET. Table VII shows a comparison between GA and BF in solving the optimization problem presented in (6). We also evaluate the proposed framework and the robustness of GA to solve the optimization problem for different weights of EN-DC activation and RLF. 1, 0.5, and 0 values of w correspond to the very high importance of EN-DC activation, equal importance for both EN-DC activation and RLF, and high importance for RLF, respectively. It can be observed that GA converges 2212, 256, and 133 times faster compared to the BF for the three cases, respectively. The longer time needed for GA to find an optimal solution for 2nd and 3rd case is intuitive as the solutions space becomes more complex when EN-DC activation has to take into account minimization of RLF and muting in addition to

maximizing the 5G network utilization only. The values of the objective function and optimal parameters returned by GA and BF for different weights indicate that GA can converge very close to the optimal value.

We implement the optimal parameters reported in Table VII in the SyntheticNET and observe the number of EN-DC requests, EN-DC activations, and the potential RLF occurrences. Fig. 12 shows that the number of EN-DC requests, EN-DC activations as well as the number of potential RLF is highest for w=1 and all three of them decrease as the value of w decreases. With w=1 in (6), the optimization function maximizes EN-DC activations and disregard RLF. This is shown in Fig. 12 where 1328 of the 6025 EN-DC activations results in RLF. Existing mobile networks use the EN-DC configuration that gives maximum weightage to the number of EN-DC activations (w=1). Fig. 12 also shows that we can reduce RLFs by decreasing w, and can totally eliminate the chances of RLF with w=0. This however comes at ∼50% loss of EN-DC activations. Thus, the proposed framework offers a solution for 5G operators to optimize the tradeoff between network utilization and QoE.

For w=0, the BF solution results in 3501 EN-DC requests and 2413 EN-DC activations indicating that 1088 EN-DC requests were not entertained due to chances of RLF from poor RF condition. On the contrary, only 914 EN-DC requests were rejected when the optimal parameters obtained from GA were used which resulted in 2295 EN-DC activations.

Table VIII compares the performance of GA with BF for solving the voice muting aware EN-DC optimization given in (7). A similar trend is observed for voice muting aware EN-DC optimization, where GA converges significantly faster than the BF at the cost of a slightly less optimal objective function. Fig. 13 shows the number of EN-DC requests, EN-DC activations, and potential muting instances when the optimal EN-DC

Table VIII: Optimal parameters obtained from GA for UEs requiring voice call services.

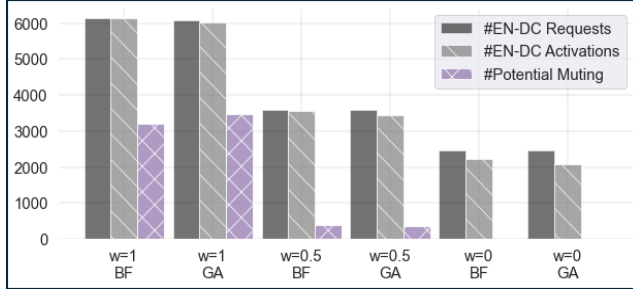| w	Algo	Iterations	Utility	Optimal Parameters $\Theta^r = [\theta_{B1}^r, \theta_{4R}^r, \theta_{5S}^r, \theta_{4S}^r]$		
1	BF	741,321	1142	-120dBm, -120dBm, -6dB, -7dB		
	GA	969	1130.4	-120dBm, -120dBm, -7dB, -10dB		
0.5	BF	741,321	49	-115dBm, -110dBm, -5dB, -1dB		
	GA	5607	46.6	-115dBm, -110dBm, -10dB, 0dB		
0	BF	741,321	2.2	-112dBm, -110dBm, 0dB, -1dB		
	GA	10987	2.1	-111dBm, -110dBm, -6dB, -1dB		

Figure 13: Number of EN-DC activations and muting observed when using optimal parameters in Table VIII.

parameters from Table VIII are deployed in SyntheticNET with different weights. It is shown that zero chances of muting instances can be achieved by assigning more weightage to mute using w=0. However, since UE is more susceptible to muting rather than RLF, zero mute occasions can be obtained at the cost of even lower EN-DC activations (2085) compared to the similar case with w=0 in Fig. 12.

VI. CONCLUSION

EN-DC mode addresses strict capacity requirements of the UE by enabling dual-connectivity to 4G and 5G cells. However, dual-connectivity can be beneficial only if it can be retained for the required time duration.

Currently, no EN-DC mode selection scheme exists in the literature that takes into account the risk of RLFs and voice muting. In this paper, we propose a data driven framework for intelligent EN-DC activation that can minimize RLF and muting instances. The core idea of the proposed framework is to use prior RSRP, SINR, RLF, and voice muting data, gathered either through drive test or MDT reports to train RLF and voice muting prediction models. These models are then used as part of the objective function in an optimization problem to determine optimal B1 threshold to determine EN-DC activation criteria that can offer the desired trade-off between utilization of 5G network and QoE deterioration caused by RLF or voice muting.

A key challenge in building RLF and voice mute prediction models from real data is extremely imbalanced training data, as a number of RLF and voice muting events would be far less compared to total observations. Our investigation of a large number of data balancing techniques shows that different techniques work best for different types of data. For example, for RLF data where there is significant overlap between the two (RLF and no RLF) classes, training data augmentation

techniques such as GAN that aim to increase the number of samples in minority class by generating synthetic minority samples, may worsen the situation. Instead, for such data, class balancing techniques such as Tomek Link work best as they tend to remove the samples from the majority class that are at the boundary. On the other hand, for voice muting data, where class imbalance might be even more extreme, but boundaries are relatively less overlapping, GANs outperform all other techniques. The performance evaluations of the proposed solution, using a 3GPP compliant simulator shows that the proposed scheme can be used to either totally eliminate RLF and voice muting if a calculated reduction in EN-DC activation can be tolerated, or it can be used to achieve any operator policy based trade-off between the 5G network utilization and risk of QoE deterioration caused by RLF or call muting.

REFERENCES

[1] Ericsson, "Ericsson Mobility Report: 5G subscriptions to top 2.6 billion by end of 2025," Tech. Rep., 2019. [Online]. Available: https://www.ericsson.com/4acd7e/assets/local/reports-papers/mobility-report/documents/2019/emr-november-2019.pdf
[2] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: How to Empower SON with Big Data for Enabling 5G," IEEE Network, 2014.
[3] 3GPP, "37.863 - E-UTRA (Evolved Universal Terrestrial Radio Access) - NR Dual Connectivity (EN-DC) of LTE 1 Down Link (DL) / 1 Up Link (UL) and 1 NR band," Tech. Rep., 2019.
[4] M. U. Bin Farooq, M. Manalastas, W. Raza, A. Ijaz, S. M. A. Zaidi, A. Abu-Dayya, and A. Imran, "Data driven optimization of inter-frequency mobility parameters for emerging multi-band networks," in GLOBECOM 2020 - 2020 IEEE Global Communications Conference, 2020, pp. 1–6.
[5] M. Manalastas, M. U. B. Farooq, S. M. A. Zaidi, A. Abu-Dayya, and A. Imran, "A data driven framework for inter-frequency handover failure prediction and mitigation," IEEE Transactions on Vehicular Technology, 2022.
[6] M. U. B. Farooq, M. Manalastas, W. Raza, S. M. A. Zaidi, A. Rizwan, A. Abu-Dayya, and A. Imran, "A data-driven self-optimization solution for inter-frequency mobility parameters in emerging networks," IEEE Transactions on Cognitive Communications and Networking, 2022.
[7] M. U. B. Farooq, M. Manalastas, S. M. A. Zaidi, A. Abu-Dayya, and A. Imran, "Machine learning aided holistic handover optimization for emerging networks," in 2022 IEEE International Conference on Communications (ICC). IEEE, 2022.
[8] C. Pupiales, D. Laselva, and I. Demirkol, "Capacity and congestion aware flow control mechanism for efficient traffic aggregation in multi-radio dual connectivity," IEEE Access, vol. 9, pp. 114 929–114 944, 2021.
[9] K. Qi, T. Liu, C. Yang, S. Suo, and Y. Huang, "Dual connectivity-aided proactive handover and resource reservation for mobile users," IEEE Access, vol. 9, pp. 36 100–36 113, 2021.
[10] T. Mumtaz, S. Muhammad, M. I. Aslam, and N. Mohammad, "Dual Connectivity-Based Mobility Management and Data Split Mechanism in 4G/5G Cellular Networks," IEEE Access, vol. 8, pp. 86 495–86 509, 2020.
[11] S. Mondal, S. Al-Rubaye, and A. Tsourdos, "Handover prediction for aircraft dual connectivity using model predictive control," IEEE Access, vol. 9, pp. 44 463–44 475, 2021.
[12] P.-J. Hsieh, W.-S. Lin, K.-H. Lin, and H.-Y. Wei, "Dual-connectivity prevenient handover scheme in control/user-plane split networks," IEEE Transactions on Vehicular Technology, vol. 67, no. 4, pp. 3545–3560, 2018.

[13] X. Huang, S. Tang, Q. Zheng, D. Zhang, and Q. Chen, "Dynamic femtocell gNB on/off strategies and seamless dual connectivity in 5g heterogeneous cellular networks," *IEEE Access*, vol. 6, pp. 21 359–21 368, 2018.

[14] M. A. Lema, E. Pardo, O. Galinina, S. Andreev, and M. Dohler, "Flexible dual-connectivity spectrum aggregation for decoupled uplink and downlink access in 5g heterogeneous systems," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2851–2865, 2016.

[15] J. Sun, S. Zhang, S. Xu, and S. Cao, "High throughput and low complexity traffic splitting mechanism for 5G non-stand alone dual connectivity transmission," *IEEE Access*, vol. 9, pp. 65 162–65 172, 2021.

[16] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmwave mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2069–2084, 2017.

[17] M. Moltafet, R. Joda, N. Mokari, M. R. Sabagh, and M. Zorzi, "Joint Access and Fronthaul Radio Resource Allocation in PD-NOMA-Based 5G Networks Enabling Dual Connectivity and CoMP," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6463–6477, 2018.

[18] Y. Yang, X. Deng, D. He, Y. You, and R. Song, "Machine Learning Inspired Codeword Selection For Dual Connectivity in 5G User-Centric Ultra-Dense Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8284–8288, 2019.

[19] L. Sharma, B. B. Kumar, and S.-L. Wu, "Performance analysis and adaptive drx scheme for dual connectivity," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 289–10 304, 2019.

[20] Q. Han, B. Yang, and X. Wang, "Queue-aware cell activation and user association for traffic offloading via dual-connectivity," *IEEE Access*, vol. 7, pp. 84 938–84 951, 2019.

[21] Z. Gu, H. Lu, P. Hong, and Y. Zhang, "Reliability enhancement for VR delivery in mobile-edge empowered dual-connectivity sub-6 ghz and mmwave hetnets," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.

[22] H. Cui and F. You, "User-centric resource scheduling for dual-connectivity communications," *IEEE Communications Letters*, pp. 1–1, 2021.

[23] G. S. Park and H. Song, "Video quality-aware traffic offloading system for video streaming services over 5G networks with dual connectivity," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5928–5943, 2019.

[24] F. B. Tesema, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Mobility modeling and performance evaluation of multi-connectivity in 5g intra-frequency networks," in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–6.

[25] D. Öhmann, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Achieving high availability in wireless networks by inter-frequency multi-connectivity," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–7.

[26] D. S. Wickramasuriya, C. A. Perumalla, K. Davaslioglu, and R. D. Gitlin, "Base station prediction and proactive mobility management in virtual cells using recurrent neural networks," in *2017 IEEE 18th Wireless and Microwave Technology Conference (WAMICON)*, 2017, pp. 1–6.

[27] Syed Muhammad Asad Zaidi, Marvin Manalastas, Hasan Farooq, and A. Imran, "Mobility Management in 5G and Beyond: A Survey and Outlook," *IEEE ACCESS*, 2020.

[28] M. Centenaro, D. Laselva, J. Steiner, K. Pedersen, and P. Mogensen, "Resource-efficient dual connectivity for ultra-reliable low-latency communication," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5.

[29] A. Alhammadi, M. Roslee, M. Y. Alias, I. Shayea, and S. Alraih, "Dynamic handover control parameters for lte-a/5g mobile communications," in *2018 Advances in Wireless and Optical Communications (RTUWO)*, 2018, pp. 39–44.

[30] Y. Mal, J. Chen, and H. Lin, "Mobility robustness optimization based on radio link failure prediction," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2018, pp. 454–457.

[31] M. T. Nguyen, S. Kwon, and H. Kim, "Mobility robustness optimization for handover failure reduction in lte small-cell networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4672–4676, 2018.

[32] M.-h. Song, S.-H. Moon, and S.-J. Han, "Self-optimization of handover parameters for dynamic small-cell networks," *Wireless Communications and Mobile Computing*, vol. 15, no. 11, pp. 1497–1517, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcm.2439

[33] J. Puttonen, J. Kurjenniemi, and O. Alanen, "Radio problem detection assisted rescue handover for lte," in *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2010, pp. 1752–1757.

[34] S. K. Srivastava, M. R. Kanagarathinam, S. Diggi, and H. Natarajan, "Cleh — cross layer enhanced handover for ims sessions," in *2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC)*, 2018, pp. 1–4.

[35] A. Łukowa and V. Venkatasubramanian, "Performance of strong interference cancellation in flexible ul/dl tdd systems using coordinated muting, scheduling and rate allocation," in *2016 IEEE Wireless Communications and Networking Conference*, 2016, pp. 1–7.

[36] S. M. A. Zaidi, M. Manalastas, A. Abu-Dayya, and A. Imran, "Ai-assisted rlf avoidance for smart en-dc activation," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[37] 3GPP, "3GPP TS 23.501 - System Architecture for the 5G System," Tech. Rep., 2020.

[38] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 336–361, 2013.

[39] S. M. A. Zaidi, M. Manalastas, H. Farooq, and A. Imran, "SyntheticNET: A 3GPP Compliant Simulator for AI Enabled 5G and Beyond," *IEEE Access*, pp. 1–1, 2020.