IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 3, 2021, accepted April 13, 2021, date of publication April 16, 2021, date of current version April 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3073670

Deep Learning-Based Prediction of Resource Block Usage Rate for Spectrum Saturation Diagnosis

HAN SEUNG JANG[1], (Member, IEEE), HOON LEE[2,3], (Member, IEEE), HYEYEON KWON[4], AND SEUNGKEUN PARK[4], (Member, IEEE)

[1]School of Electrical, Electronic Communication, and Computer Engineering, Chonnam National University, Yeosu 59626, South Korea
[2]Department of Information and Communications Engineering, Pukyong National University, Busan 48513, South Korea
[3]Department of Smart Robot Convergence and Application Engineering, Pukyong National University, Busan 48513, South Korea
[4]Radio Resource Research Group, Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, South Korea

Corresponding authors: Hoon Lee (hlee@pknu.ac.kr) and Hyeyeon Kwon (hykwon@etri.re.kr)

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant through the Korean Government [Ministry of Science and ICT (MSIT)] (Development of technology for securing and supplying radio resources) under Grant 2021-0-00092.

ABSTRACT Strict restrictions on spectrum utilization and the rapid increases in mobile users have brought fundamental challenges for mobile network operators in securing sufficient spectrum resources. In designing reliable cellular networks, it is essential to predict spectrum saturation events in the future by analyzing the past behavior of base stations, especially their frequency resource block (RB) utilization states. This paper investigates a deep learning-based forecasting strategy of the future RB usage rate (RBUR) status of hundreds of LTE base stations deployed in Seoul, South Korea. The dataset consists of real measurement RBUR samples with a randomly varying number of base stations at each measurement time. This poses a difficulty in handling variable-length RBUR data vectors, which is not trivial for state-of-the-art deep learning estimation models, e.g., recurrent neural networks (RNNs), developed for handling fixed-length inputs. To this end, we propose a two-step RBUR estimation approach. In the first step, we extract a useful feature of the RBUR dataset that accurately approximates the behavior of the top quantile base stations. The feature parameters are carefully designed to be fixed-length vectors regardless of the dimensions of the raw RBUR samples. The fixed-length feature parameter vectors are readily exploited as the training dataset of RNN-based prediction models. Thus, in the second step, we propose a feature estimation strategy where the RNN is trained to predict the future RBUR from the input feature parameter sequences. With the estimated RBUR at hand, we can easily predict the spectrum saturation of the future LTE systems by examining the resource utilization states of the top quantile base stations. Numerical results demonstrate the performance of the proposed RBUR estimation methods with the real measurement dataset.

INDEX TERMS LTE, recurrent neural network, resource block usage rate, spectrum saturation.

I. INTRODUCTION

Exploding mobile traffic demands, which mostly are stemmed from the rapid increase in mobile communication subscribers and the requirements for high-resolution multimedia supports, have become severe problems in designing wireless communication systems. Strict restrictions on the utilization of wireless resources prohibit mobile network operators (MNOs) from simply exploiting a wide range of frequency spectrum. With limited frequency resources, base stations deployed in hotspot areas can hardly serve mobile users rushed into busy hours. In cellular network systems, this poses a spectrum saturation issue [2], [3] where the communication of mobile devices is no longer supported due to the shortage of the capacity. A fundamental solution is to perform an efficient spectrum allocation across distributed base stations such that base stations with bursty traffic requirements can get access to more frequency resources. To proceed with this dynamic network optimization task, MNOs need to accurately predict the behavior of cellular networks such as changes in mobile traffic and resource usages of each base station.

The associate editor coordinating the review of this manuscript and approving it for publication was Ruofei Ma.

Predicting short-term changes in wireless networks has been intensively studied in various scenarios. A recursive partitional clustering algorithm is presented in [7] to estimate patterns of mobile network traffic. It characterizes geographic models of spatial traffic. The work in [8] designed a mobility model for wireless multimedia networks to obtain analytical expressions for traffic distribution. Although these seminal works have opened new research opportunities to network traffic predictions, such model-based approaches lack generality for adapting real-world datasets that are typically unstructured and highly randomized with unavailable distributions and unknown correlation measures. To handle such a chaotic nature of mobile communication systems, it is essential to develop a more sophisticated model-free estimation method that does not require any prior knowledge of networking models.

A. RELATED WORKS

Such a challenge can be addressed by regression techniques in the machine learning field, where estimation rules of arbitrary systems are identified in a data-driven manner. The time-varying properties of mobile networks make a traffic estimation problem as a task of handling time series data. Shallow structures, e.g., the support vector regression (SVR), have been exploited to address such estimation formulations [9], and they have been shown to be effective in certain applications including user tracking scenarios [10]. However, practical cellular networks consist of numerous base stations distributed over a wide range of coverage area, and thus the prediction of the mobile traffic flows need a more powerful and intelligent architecture.

Recent successes of deep learning (DL) techniques [11], [12] have brought intensive researches on various prediction applications in wireless systems. Among diverse neural network (NN) structures, recurrent NNs, in particular, the long-short term memory (LSTM) [13] and the gated recurrent unit (GRU) [14] have been proven to be suitable for time series prediction problems. The feasibility of the LSTM for the security of the internet-of-things environment has been studied in [15]. A binary classification technique is provided to detect attack events from network trace datasets including packet payload. The work in [16] employs the LSTM to forecast traffic and user location. It has been revealed that the LSTM structure indeed outperforms shallow networks such as the SVR and the autoregressive integrated moving average model. However, it is confined to the short-term prediction of five mobile users, and such small-scale results cannot provide any insights into the design of the large-scale network deployment. To handle the spectrum saturation issue effectively, it is necessary to focus on the overall cellular network, which typically consists of hundreds of base stations.

Traffic estimation techniques of multiple base stations have been investigated in [17]. To capture the correlations among spatially distributed base stations, the concept of multi-task learning is employed which constructs an individual LSTM for the traffic prediction of each base station. These individual LSTMs are connected by a common shared NN that extracts the spatial correlation features of distributed base stations. The dataset of [17] comprises the traffic volumes of sixteen base stations, and the multi-task LSTM architecture has been proven to be efficient in their scenarios. For the network-wide prediction tasks, however, a massive number of cells should be considered, and the dataset of [17] is far from the practical network deployment problems. Therefore, the approach in [17] would entail the scalability issue since it requires to train and execute multiple LSTM modules dedicated to particular base stations. Traffic loads of thousands of base stations in China have been studied in [18]. Spatio-temporal correlation features of the network-wide traffic dataset are extracted by using autoencoder units. A similar multi-task learning architecture is adopted so that multiple autoencoders, each developed for the prediction of each base station, should be executed for the traffic estimation of the interested network region, still lacking the scalability. To this end, [19] presents a novel DL-based traffic prediction framework that converts the traffic data of Milan into graphical images called traffic heat maps. The traffic of base stations is represented as intensities of a pixel in an image capturing the networking area. The authors in [19] combined convolutional NNs and LSTM to handle spatially and temporally correlated images.

Despite these recent progresses, the spectrum saturation issue has not yet been addressed adequately since existing researches are dedicated to predicting the mobile traffic data. Base stations in LTE-A systems are assigned by heterogeneous capacity, i.e., frequency channels and bandwidth, according to their target coverage. Although the mobile traffic is suitable for quantifying the absolute capacity of base stations, it cannot provide intuition on heterogeneous network deployments and diagnosis of spectrum saturation. To this end, our focus should be on the network-wide trends of the capacity of base stations, rather than simply forecasting the absolute capacity changes in the traffic curves. Hence, the dataset needs to involve the utilization status of the frequency resources of busy-hour base stations. Such a measure is referred to as *resource block usage rate* (RBUR) [2]–[6], which defines the ratio of the spectrum utilization to the total available frequency bandwidth. The RBUR has been recently analyzed in [2]–[6] as a crucial measure of the network behaviors including the spectrum saturation, spectral efficiency, and upper limits on the mobile traffic. The RBUR measures the margin of frequency resources to the maximum allowable bandwidth, i.e., the relative capacity of heterogeneous base stations. Thus, it can provide important guidelines to the MNOs on how much spectrum should be reserved for future communication services. Therefore, the RBUR is suitable for the diagnosis of spectrum saturation issues [3].

For the target of analyzing the spectrum saturation, the RBUR dataset only includes a subset of base stations with bursty traffic demands that exhibit the top quantiles of the RBUR. The number of base stations involved in the RBUR dataset varies according to the measurement time, the MNOs, and the operating frequency band. Thus, the dimension of the

resulting training samples is not the same. Existing NN-based approaches cannot be directly applied to this scenario since they require the identical size of input data. This drawback is mainly stemmed from the fact that NNs, especially the RNN structure, are designed to accept inputs with fixed dimensions. Therefore, to predict the busy-hour RBUR correctly, it is essential to develop a novel DL technique that is flexible to the input data size.

B. CONTRIBUTIONS

This paper proposes a DL-based RBUR prediction model focusing on the long-term evolution (LTE) measurements and LTE-advanced (LTE-A) systems in Seoul, South Korea. The RBUR samples are measured at hundreds of base stations deployed in hotspot areas in Seoul. The considered RBUR dataset has recently been investigated in [3] collected for 24 months. Compared to this existing study, this work handles a more high-dimensional RBUR dataset measured for 38 months. Due to the randomness in the number of base stations operating at each month, the RBUR dataset consists of sample vectors with lengths varying for each measurement time. This blocks us to straightforwardly employ conventional RNN-based prediction models as in [3], [15]–[19]. To address this difficulty, the authors in [3] a probabilistic approach that identifies an empirical distribution of the RBUR dataset. With properly chosen probability density functions, the RBUR samples of unavailable base stations are randomly generated so that all the data vectors have an identical length. The resulting artificial realizations are readily utilized for the training of RNNs. However, the method in [3] degrades the prediction performance due to the randomness incurred in the artificial training dataset. Furthermore, training RNN models with high-dimensional input features is difficult, which typically leads to significant performance degradation. Since we consider a more number of samples, the conventional approach cannot be directly applied to our scenario.

To handle this issue, we propose an efficient RBUR estimation approach, which consists of two consecutive machine learning procedures: *feature extraction* step and *feature estimation* step. The feature extraction step identifies a low-dimensional characterization of the high-dimensional RBUR samples, referred to as a feature parameter. The feature parameter is carefully designed such that it can accurately approximate the real measurement RBUR dataset. We optimize the feature parameter based on the nonlinear regression concept. The resulting feature parameter is shown to be a fixed dimensional vector for all measurement time, and thus it is readily processed by conventional RNN structures. This is achieved by the feature estimation step, whose target is to predict the extracted feature parameters, rather than forecasting the high-dimensional RBUR samples. Notice that the existing work [3] has focused on estimating RBUR realizations artificially sampled from the empirical distribution found from the measurement data. Such an empirical sampling approach incurs the random imperfection in a training dataset, thereby leading to performance degradation. Moreover, this conventional method directly tackles the estimation of the high-dimensional dataset, requiring prohibitive training complexity. The proposed feature estimation step addresses this difficulty by predicting the extracted low-dimensional feature parameters. For the feature estimation step, the RNN-based prediction model is presented for predicting the time series of the extracted features. To diagnosis the spectrum saturation issue efficiently, a new performance metric is designed, which measures the network-wide occupation states of frequency resources by carefully aggregating the RBURs of each frequency band. Numerical results verify the effectiveness of the proposed RBUR prediction method.

The rest of this paper is organized as follows. Section II explains the RBUR dataset and its pre-processing strategy for the target of the spectrum saturation analysis. The feature extraction step is presented in Section III, and it is followed by the RNN-based feature estimation method in Section IV. Numerical results for demonstrating the proposed approach are given in Section V. The paper is terminated with concluding remarks in Section VI.

II. DATA DESCRIPTION

To analyze the network stability and the spectrum saturation issues, MNOs collect their RBUR data [2], [3] of all base stations. The RBUR is defined as the ratio of the number of occupied RBs to the total number of available RBs within an hour. In particular, the *busy-hour RBUR* indicates the highest RBUR in a single day. In what follows, the RBUR implies the busy-hour RBUR.

Base stations deployed in densely populated cells exhibit a significantly higher RBUR than those in rural areas, which, in general, have sufficient frequency resources for supporting only a small number of subscribers. We thus focus on sampled cells showing the top 10% RBURs in hotspot areas of Seoul, South Korea. More precisely, among total L cells in Seoul, the RBURs of the top 10% quantiles, i.e., $0.1L$ cells, are available in the considered dataset. Two anonymous MNOs measure the RBUR for 38 months from March 2017 to April 2020. For convenience, we denote these MNOs as operators A and B. We focus on investigating the long-term behavior of base stations to effectively address the spectrum saturation diagnosis in a long period, e.g., a half year. Thus, each RBUR sample stands for the average resource utilization rate over weekdays in the third week of each month. Such a long-term network prediction helps the MNOs to handle the spectrum saturation problem in a network-wide perspective [20], [21]. The estimated long-term RBUR can be exploited to investigate the demands of mobile users at a certain spectrum band, helping the MNOs set their strategies for the frequency auction in the future.

Two MNOs operate in distinct frequency bands, each having different bandwidths. The bandwidths of the LTE and LTE-A systems are given by 10 MHz and 20 MHz, respectively, each of which contains 50 and 100 RBs [1]. For anonymity requested from the MNOs, we do not specify

the number of frequency bands and their bandwidths as well as carrier frequencies. The b-th operating frequency band of operator $z \in \{A, B\}$ is referred to as band z_b. Without loss of the generality, operator z is assumed to have F_z distinct frequency bands, i.e., $b = 1, \ldots, F_z$. The number of cells in the dataset is not fixed but exhibits highly randomized properties. It changes for measurement month, operating frequency band, and the MNOs. Such a variation is incurred due to some practical issues, e.g., installing new base stations or removing existing ones. These replacements occur monthly. Therefore, predicting the RBUR of a specific base station is not meaningful. Instead, we need to focus on analyzing the behavior of the overall cellular network. In our dataset, the numbers of top 10% cells vary 347 to 596 and 165 to 796 for operators A and B, respectively. This heterogeneous data size blocks us to employ NN-based prediction models, such as the LSTM, as in existing studies [3], [15]–[19].

To handle this issue, we first pre-process the raw RBUR dataset. Let $L_{z_b,m}$ be the number of cells of band z_b, i.e., the b-th frequency band of operator z ($z \in \{A, B\}$), measured at month m ($m = 1, \ldots, M$) where $M = 38$ is the total number of the measurement months. The RBUR data of particular z_b and m can be represented by a $L_{z_b,m}$-dimensional vector. The elements of the RBUR vector are first sorted in the decreasing order, and the resulting ordered vector is denoted by $\mathbf{y}_{z_b,m}$. Such an ordering is essential to facilitate the diagnosis of the spectrum saturation issues. For instance, if the i-th element of $\mathbf{y}_{z_b,m}$ can be predicted accurately, we can get the resource utilization status of the base station occupying the i-th most spectrum resources in the top 10% cells.

Since the number of the cells $L_{z_b,m}$ highly fluctuates for each month m, we first normalize its domain into the bounded range in the unit of percentage. Recalling that $\mathbf{y}_{z_b,m}$ contains the RBUR of the top 10% cells, its i-th element becomes the RBUR of the top $\frac{10i}{L_{z_b,m}}$% base station. These collectively form the quantile vector $\mathbf{x}_{z_b,m}$ of length $L_{z_b,m}$ defined as

$$\mathbf{x}_{z_b,m} \triangleq \left[\frac{10}{L_{z_b,m}}, \frac{20}{L_{z_b,m}}, \ldots, \frac{10(L_{z_b,m}-1)}{L_{z_b,m}}, 10\right]^{\mathrm{T}}. \quad (1)$$

Finally, the pre-processed RBUR data on frequency band b of operator z measured at month m consists of a pair of vectors $(\mathbf{x}_{z_b,m}, \mathbf{y}_{z_b,m})$. The pre-processed ordered statistics can be efficiently handled by machine learning techniques that capture regular patterns in input data samples.

In what follows, we propose an efficient prediction method of the RBUR dataset. A two-step approach is presented. First, in the feature extraction step, we extract an important characterization of the RBUR dataset that can accurately recover the groundtruth samples with a low-dimensional feature parameter set. Next, in the feature estimation step, a RNN-based prediction method is provided to predict the feature parameters of the future months.

III. FEATURE EXTRACTION

The pre-processed data tuples $(\mathbf{x}_{z_b,m}, \mathbf{y}_{z_b,m})$ have different dimension for each m, b, and z, and thus it is not trivial to handle them with existing NN-based prediction models [3], [17]–[19] that only accept fixed-length input vectors. For the successful prediction performance, the temporal correlations of the network-wide data samples typically require to consider a group multiple units of the LSTM where each LSTM module is dedicated to each operator and frequency band [3], [17]–[19]. Since we have a total of eight distinct spectrum bands for two MNOs, such a direct approach would not be practical in our dataset. Although fully convolutional neural networks could handle variable-length input data, it typically requires a deep architecture with a large number of trainable parameters. Furthermore, high-dimensional data vectors entail fundamental challenges in the training of very deep NN structures. In particular, we may encounter the curse of dimensionality issue [12] which incurs sparsely distributed datapoints in high-dimensional vector space. Therefore, a very large number of training data samples are needed to optimize NNs with high-dimensional quantile vectors.

A simple but effective way of addressing the curse of dimensionality issue is to compress input data into low-dimensional vector spaces. To this end, this section presents feature extraction methods that characterize the high-dimensional RBUR data by using few parameters with a fixed size. The proposed feature extraction method employs three trainable parameters which are optimized to generate the RBUR samples accurately. As will be shown later, the accuracy of the feature extraction approaches the upperbound R-square performance. Consequently, the proposed feature extraction method is suitable for handling the variable-length input data as well as extracting low-dimensional feature of the RBUR measurement samples.

The pre-processed RBUR vectors become monotonically decreasing functions with respect to the quantile percentage index. Based on this fact, we employ a generative model $f(x; \boldsymbol{\theta})$ with a learnable *feature parameter* $\boldsymbol{\theta}$ which characterizes the RBUR of the top $x \in [0, 10]$ percentage cell. In this work, the following power series model is considered.

$$f(x; \boldsymbol{\theta}_{z_b,m}) = \alpha_{z_b,m} x^{\beta_{z_b,m}} + \gamma_{z_b,m}, \quad (2)$$

where $\alpha_{z_b,m}$, $\beta_{z_b,m}$, and $\gamma_{z_b,m}$ are parameters to be trained. They collectively form a three-tuple feature parameter $\boldsymbol{\theta}_{z_b,m} \triangleq (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$. We refer (2) as a *feature model*. A three-tuple of these trainable parameters $\boldsymbol{\theta}_{z_b,m} = (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$ acts as a *feature* of the RBUR samples of a certain band z_b measured on month m. With the properly optimized $\boldsymbol{\theta}_{z_b,m}$ at hand, we can create the RBUR data of the top x percentage base station by using the proposed feature model (2). The feature parameter $\boldsymbol{\theta}_{z_b,m}$ can also be regarded as a compression of the raw RBUR data that can generate numerous realization of the RBUR samples from the generative model (2). As a result, the original RBUR data of size $L_{z_b,m}$, which is larger than 347 for operator A, can be reduced into three-dimensional feature parameter. As will be explained in Section IV, this feature parameter becomes the target of our RNN-based prediction model. Processing with the three-dimensional feature

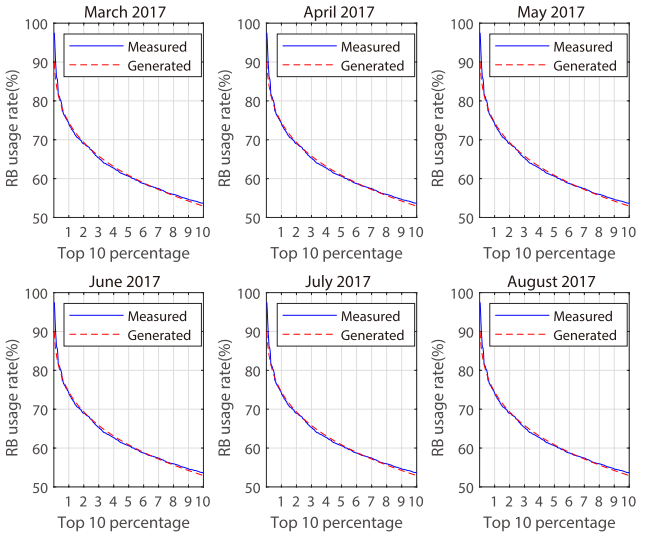


FIGURE 1. Comparison of generated and measured RBUR samples of band A_1.

parameter $\boldsymbol{\theta}_{z_b,m} = (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$ is more efficient than handling the high-dimensional real measurement data directly.

The remaining work is to determine a good feature parameter so that it can accurately characterize the stochastic properties of the RBUR data. To this end, the regression technique is adopted to train the feature parameter in a supervised manner. Let (x_i, y_i), $\forall i = 1, \ldots, L_{z_b,m}$, be the i-th element of the groundtruth RBUR tuple $(\mathbf{x}_{z_b,m}, \mathbf{y}_{z_b,m})$, i.e., x_i and y_i respectively stand for the i-th elements of $\mathbf{x}_{z_b,m}$ and $\mathbf{y}_{z_b,m}$. The regression loss function is constructed as the mean-square-error (MSE) written by

$$\Phi(\boldsymbol{\theta}_{z_b,m}) = \frac{1}{L_{z_b,m}} \sum_{i=1}^{L_{z_b,m}} (f(x_i; \boldsymbol{\theta}_{z_b,m}) - y_i)^2 \tag{3}$$

$$= \frac{1}{L_{z_b,m}} \sum_{i=1}^{L_{z_b,m}} \left(f\left(\frac{10i}{L_{z_b,m}}; \boldsymbol{\theta}_{z_b,m} \right) - y_i \right)^2, \tag{4}$$

where (4) comes from the definition (1). An efficient solution to the MSE minimization problem can be efficiently attained from non-linear least square methods [24].

The result for a particular dataset of band A_1 measured at six different months is plotted in Fig. 1. First, it is observed that the RBUR curves generated by the proposed feature extraction model effectively characterize the non-increasing property of the pre-processed RBUR samples. Also, the generated RBUR data well matches with the real measurement samples for all simulated months and top 10 quantile. This implies that high-dimensional RBUR samples can be characterized by three-dimensional feature parameter $\boldsymbol{\theta}_{z_b,m} = (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$ with high accuracy, which are bridged by the feature extraction model (2) and the data pre-processing strategy presented in Section II. To see the estimation accuracy of the feature extraction model more clearly, Tables 1 and 2 present the R-squared (coefficient of determination) performance for the proposed feature model. The R-squared values of all cases are close to the unity, verifying the accuracy of the proposed feature extraction approach. Therefore, we can conclude that the proposed feature $\boldsymbol{\theta}_{z_b,m}$ can accurately characterizes the real measurement data regardless of the months and the operators.

TABLE 1. R-squared performance for proposed feature model for operator *A*.

Month	Band A_1	Band A_2	Band A_3
April 2017	0.98726	0.99007	0.92084
June 2017	0.99392	0.98608	0.68983
August 2017	0.99645	0.98276	0.90982
October 2017	0.99658	0.99133	0.89648
December 2017	0.98705	0.96864	0.99398
February 2018	0.98871	0.98071	0.94126
April 2018	0.99753	0.99239	0.96324
June 2018	0.99884	0.98989	0.98094
August 2018	0.99877	0.99793	0.98226
October 2018	0.99478	0.99231	0.98763
December 2018	0.99512	0.98061	0.98953
February 2019	0.99561	0.97752	0.95592
April 2019	0.99625	0.98798	0.98342
June 2019	0.99869	0.97545	0.95218
August 2019	0.99856	0.98511	0.96388
October 2019	0.99955	0.98676	0.96943
December 2019	0.99508	0.98419	0.9429
February 2020	0.99894	0.98771	0.96515
April 2020	0.99749	0.99219	0.9628

TABLE 2. R-squared performance for proposed feature model for operator *B*.

Month	Band B_1	Band B_2	Band B_3
April 2017	0.99328	0.98938	0.99382
June 2017	0.99052	0.96603	0.98149
August 2017	0.99749	0.95053	0.96785
October 2017	0.98227	0.9186	0.9115
December 2017	0.98382	0.99657	0.98886
February 2018	0.98311	0.99774	0.99384
April 2018	0.99014	0.99754	0.99404
June 2018	0.99561	0.99019	0.98725
August 2018	0.9977	0.98528	0.97889
October 2018	0.99315	0.99201	0.99452
December 2018	0.9903	0.9876	0.97881
February 2019	0.99806	0.99165	0.9859
April 2019	0.99565	0.99563	0.9932
June 2019	0.9978	0.99485	0.99097
August 2019	0.99723	0.99691	0.98513
October 2019	0.99843	0.99737	0.9886
December 2019	0.99685	0.98955	0.99696
February 2020	0.99742	0.98931	0.99612
April 2020	0.99636	0.99301	0.98085

IV. FEATURE ESTIMATION

The three-dimensional feature parameter $\boldsymbol{\theta}_{z_b,m} = (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$ successfully models the RBUR dataset which contains high-dimensional vectors $\mathbf{y}_{z_b,m}$ with variable length $L_{z_b,m}$ for each m, b, and z. Hence, in the feature estimation step, we alternatively tackle the estimation procedure of the feature parameter $\boldsymbol{\theta}_{z_b,m}$ instead of handling the raw RBUR data $\mathbf{y}_{z_b,m}$. As discussed, the proposed feature model accurately characterizes the real-measurement RBUR samples, which indeed form the time series dataset. This implies that the well-trained feature parameter $\boldsymbol{\theta}_{z_b,m}$ inherits time-varying properties of the RBUR data for measurement month $m = 1, \ldots, M$. Therefore, the trained feature parameters $\boldsymbol{\theta}_{z_b,m}$ for $m = 1, \ldots, M$ entails temporal correlations of the real-measurement RBUR samples. To address this time series data, we propose a RNN-based prediction approach which is powerful for capturing the temporal correlations of input data. For the implementation of the RNN, we examine both the LSTM [13] and GRU architecture [14] in the simulation which have been proven to be effective than a vanilla RNN structure.

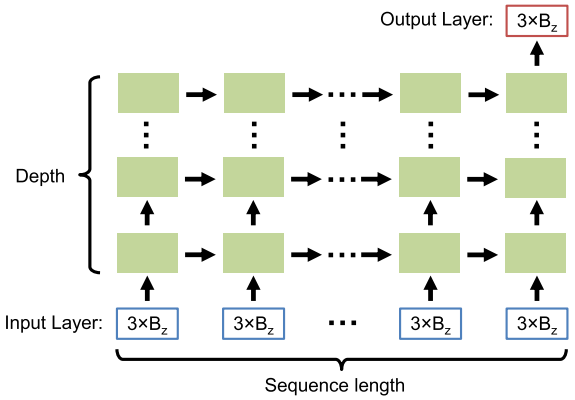


FIGURE 2. Proposed RNN structure.

Fig. 2 depicts the proposed multi-layer RNN for predicting future values of the feature parameter. Two distinct RNNs are employed for each operator. By observing the past S months' features $\boldsymbol{\theta}_{z_b,m}, \ldots, \boldsymbol{\theta}_{z_b,m+S-1}$, $\forall b = 1, \ldots, F_z$, we jointly estimate a group of future parameters $\boldsymbol{\theta}_{z_b,m+S+Q}$, $\forall b = 1, \ldots, F_z$, in Q months. Here, Q denotes the difference between the target future month and the latest month in the input sequence. Hence, the input is a sequence of the feature parameters of length S months, and the corresponding output becomes the Q-month forward future parameters. To this end, a many-to-one RNN structure is employed as illustrated in Fig. 2. As will be discussed in Section V, the choices of S and Q significantly affect the prediction performance for the considered RBUR dataset. Thus, the sequence length S and the target difference Q are regarded as hyperparameters should be carefully optimized through trial-and-error-based grid search processes. In this work, we consider the maximum bound of Q, denoted by V, as $Q \leq V$ with $V < M$ since we only have $M = 38$ month-long data.

A. TRAINING AND TEST SET

This subsection describes the construction strategies of the training and testing datasets. Let us define the collection of the feature parameters $\boldsymbol{\theta}_{z_b,m}$ for each operator $z \in \{A, B\}$. For total $M = 38$ months $m = 1, \ldots, M$, the optimized feature parameters $\boldsymbol{\theta}_{z_b,m} = (\alpha_{z_b,m}, \beta_{z_b,m}, \gamma_{z_b,m})$, $\forall m, b, z$, collectively form an augmented data matrix P_z of size M-by-$3F_z$ denoted by

$$P_z = \begin{bmatrix} \boldsymbol{\theta}_{z_1,1} & \cdots & \boldsymbol{\theta}_{z_{F_z},1} \\ \boldsymbol{\theta}_{z_1,2} & \cdots & \boldsymbol{\theta}_{z_{F_z},2} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{z_1,M} & \cdots & \boldsymbol{\theta}_{z_{F_z},M} \end{bmatrix} \tag{5}$$

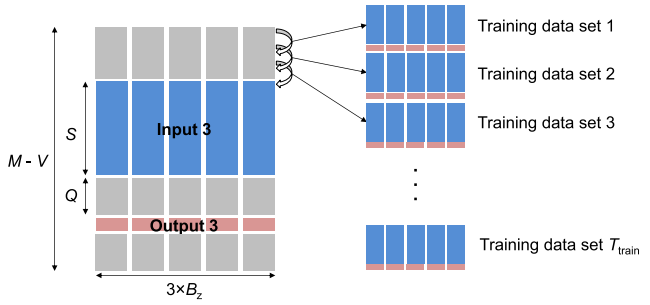
The n-th row vector of P_z denoted by $P_z[n,:]$ becomes an input to the proposed RNN-based prediction model. The augmented matrix P_z is split into two sub-matrices P_z^{train} of size $(M-V)$-by-$3F_z$ and P_z^{test} of size V-by-$3F_z$ which are utilized for generating train and test datasets, respectively. More precisely, P_z^{train} and P_z^{test} are composed of the first $M-V$ rows and the last V rows of P_z, respectively. These are written by

$$P_z^{\text{train}} = \begin{bmatrix} P_z[1,:] \\ \vdots \\ P_z[M-V,:] \end{bmatrix} \quad \text{and} \quad P_z^{\text{test}} = \begin{bmatrix} P_z[M-V+1,:] \\ \vdots \\ P_z[M,:] \end{bmatrix}. \tag{6}$$

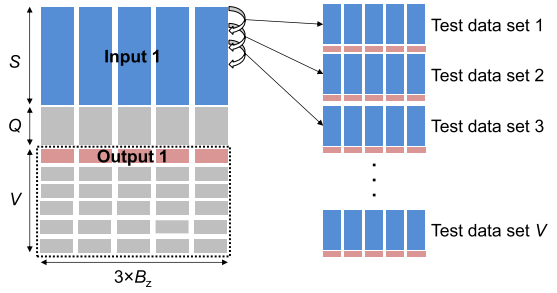
We now discuss the generation steps of train and test datasets illustrated in Fig. 3. Each training sample is needed to be constructed with an input sequence, i.e., the set of the feature parameters for the past S months, and the corresponding groundtruth label, i.e., the Q-month forward feature parameters. Let $P_z^{\text{train}}[n,:]$ be the n-th row of the augmented data matrix P_z^{train}. The k-th training sample of operator z is represented by a tuple $(\mathbf{W}_z^{(k)}, \mathbf{t}_z^{(k)})$ of an input sequence matrix $\mathbf{W}_z^{(k)}$ of size S-by-$3F_z$ and the corresponding Q-month forward groundtruth vector $\mathbf{t}_z^{(k)}$ of length $3F_z$. These are defined as

$$\mathbf{W}_z^{(k)} = \begin{bmatrix} P_z^{\text{train}}[k,:] \\ P_z^{\text{train}}[k+1,:] \\ \vdots \\ P_z^{\text{train}}[k+S-1,:] \end{bmatrix}, \quad \mathbf{t}_z^{(k)} = P_z^{\text{train}}[k+S+Q,:]. \tag{7}$$

Since P_z^{train} contains $M-V$ sequences of the feature parameters as row vectors, we can obtain total $T_{\text{train}} = (M-V) - (S+Q)+1$ training samples by carefully combining the inputs and their corresponding groundtruth values (see Fig. 3 (a)). For example, when $M = 38$, $V = 6$, $S = 4$, and $Q = 2$, we obtain $T_{\text{train}} = (38-6) - (4+2) + 1 = 27$ training data sets. The first training sample in this example is composed of first $S = 4$ months input row vectors (March, April, May, and June 2017) and $Q = 2$ months later output row vector (August 2017). In the same way, the second training sample is obtained by shifting one month compared to the first training



(a) Train dataset



(b) Test dataset

FIGURE 3. **Generation of train and test datasets.**

data set, and then it is composed of $S = 4$ months input row vectors (April, May, June, and July 2017) and $Q = 2$ months later output vector (September 2017). Similarly, we can generate V test samples from P_z^{test} as depicted in Fig. 3 (b).

B. TRAINING STEP

The training strategy of the proposed RNN with given T_{train} samples is presented. Let $\mathbf{g}(\cdot; \boldsymbol{\Theta}_z)$ be a vector-valued mapping which characterizes the input-output relationship of the RNN specified by a trainable parameter set $\boldsymbol{\Theta}_z$ for operator z. The result of the RNN forward computation $\hat{\mathbf{t}}_z^{(k)} = \mathbf{g}(\mathbf{W}_z^{(k)}; \boldsymbol{\Theta}_z)$ for the k-th training input $\mathbf{W}_z^{(k)}$ is exploited as the estimate $\hat{\mathbf{t}}_z^{(k)}$ of the corresponding groundtruth of the feature parameter $\mathbf{t}_z^{(k)}$ in (7). The estimated feature parameter vector $\hat{\mathbf{t}}_z^{(k)}$ is represented as

$$\hat{\mathbf{t}}_z^{(k)} = \mathbf{g}(\mathbf{W}_z^{(k)}; \boldsymbol{\Theta}_z) \tag{8}$$

$$\triangleq [\hat{\boldsymbol{\theta}}_{z_1,k+S+Q} \quad \cdots \quad \hat{\boldsymbol{\theta}}_{z_{F_z},k+S+Q}], \tag{9}$$

where $\hat{\boldsymbol{\theta}}_{z_b,m} \triangleq (\hat{\alpha}_{z_b,m}, \hat{\beta}_{z_b,m}, \hat{\gamma}_{z_b,m})$ and $\hat{\alpha}_{z_b,m}$, $\hat{\beta}_{z_b,m}$, and $\hat{\gamma}_{z_b,m}$ respectively stand for the estimates of the groundtruth feature parameters $\alpha_{z_b,m}$, $\beta_{z_b,m}$, and $\gamma_{z_b,m}$ that have been optimized through the regression task of (4). The cost function $\text{Cost}_z^{(k)}(\boldsymbol{\Theta}_z)$ of the k-th training data $(\mathbf{W}_z^{(k)}, \mathbf{t}_z^{(k)})$ is designed as the MSE between the proposed feature model in (2) evaluated with the groundtruth $\mathbf{t}_z^{(k)}$ in (7) and its estimate obtained by the RNN $\hat{\mathbf{t}}_z^{(k)} = \mathbf{g}(\mathbf{W}_z^{(k)}; \boldsymbol{\Theta}_z)$ in (8), i.e., the MSE between $f(\cdot; \boldsymbol{\theta}_{z_b,m})$ and $f(\cdot; \hat{\boldsymbol{\theta}}_{z_b,m})$ for $b = 1, \ldots, F_z$. It is written by (10) as shown at the bottom of this page.

As a result, the total MSE cost $C_z(\boldsymbol{\Theta}_z)$ for operator z is given by

$$C_z(\boldsymbol{\Theta}_z) = \frac{1}{T_{\text{train}}} \sum_{k=1}^{T_{\text{train}}} \text{Cost}_z^{(k)}(\boldsymbol{\Theta}_z).s \tag{11}$$

The RNN parameter $\boldsymbol{\Theta}_z$ is optimized such that the MSE cost $C_z(\boldsymbol{\Theta}_z)$ is minimized. This can be carried out by the standard gradient descent (GD) methods, e.g., the Adam algorithm [23]. Denoting $\delta > 0$ as the learning rate, the GD update rule for minimizing $C_z(\boldsymbol{\Theta}_z)$ can be expressed as

$$\boldsymbol{\Theta}_z \leftarrow \boldsymbol{\Theta}_z - \delta \nabla C_z(\boldsymbol{\Theta}_z), \tag{12}$$

where ∇ indicates the gradient operator. Such an update is repeated until the predefined training epoches.

C. ESTIMATION STEP AND SPECTRUM SATURATION PREDICTION

Once the RNN $\mathbf{g}(\cdot; \boldsymbol{\Theta}_z)$ is trained for specific input sequence length S and target difference Q, we can readily obtain the estimate of the feature parameter as in (8). With the RNN outputs $\hat{\boldsymbol{\theta}}_{z_b,m}, \forall b$, at hands, the estimated RBUR for month m can be created from the proposed feature model in (2). In particular, the RBUR of the top $x \in [0, 10]$ percentage base station operating on frequency band b of operator z at month m can be predicted as $f(x; \hat{\boldsymbol{\theta}}_{z_b,m}) = \hat{\alpha}_{z_b,m} x^{\hat{\beta}_{z_b,m}} + \hat{\gamma}_{z_b,m}$. By observing the estimated RBUR of a certain quantile x, the MNOs can get the information regarding the level of the resource occupation at the future month m. The results would be exploited for the diagnosis of the spectrum saturation events.

Still, however, it is not easy to infer the status of the overall network from the estimated RBUR data since it focuses only on the a certain frequency band b. Furthermore, the performance of the LTE and LTE-A systems typically vary for the frequency bands due to several implementation issues. For instance, the bandwidths assigned for each operating spectrum are not the same, thereby resulting in the different number of the RBs for each band z_b. Also, the radius of the cell coverage region is highly dependent on the frequency bands [25]. This is due to the fact that the propagation environments of radio-frequency signals rely on their carrier frequencies. For this reason, the quality-of-service experienced

$$\text{Cost}_z^{(k)}(\boldsymbol{\Theta}_z) = \frac{1}{L_{z_b,k+S+Q} F_z} \sum_{i=1}^{L_{z_b,k+S+Q}} \sum_{b=1}^{F_z} \left\{ \left(\alpha_{z_b,k+S+Q} \left(\frac{10i}{L_{z_b,k+S+Q}} \right)^{\beta_{z_b,k+S+Q}} + \gamma_{z_b,k+S+Q} \right) - \left(\hat{\alpha}_{z_b,k+S+Q} \left(\frac{10i}{L_{z_b,k+S+Q}} \right)^{\hat{\beta}_{z_b,k+S+Q}} + \hat{\gamma}_{z_b,k+S+Q} \right) \right\}^2 \tag{10}$$

TABLE 3. Model parameters and values.

Parameters	Values
Depth (layers) of model	5
Number of neurons in each layer	[50 50 50 50 50]
Learning rate	0.01
Iterations	10,000
Target month Q	$1 \sim 6$
Sequence length S	$1 \sim 6$
Activation function	tanh, ReLU, LReLU

by mobile users significantly depends on their locations since the available bandwidth varies according to the distance from a base station.

To analyze the spectrum saturation effectively, we introduce a new performance metric referred to as an weighted average RBUR (WA-RBUR) which is designed to compensate for the asymmetry in the bandwidths and cell coverages. The WA-RBUR is defined as the weighted sum of the RBUR vectors $\mathbf{y}_{z_b,m}$ with a predefined weight of φ_{z_b} different frequency bands $b = 1, \ldots, F_z$. The weight φ_{z_b} with unit sum $\sum_{b=1}^{F_z} \varphi_{z_b} = 1$ represents the contribution of frequency band b to the overall network capacity, i.e., the number of the RBs assigned to band b. It is designed such that it can normalize the effect of heterogeneous bandwidths and coverages. Let BW_{z_b} in MHz and ξ_{z_b} be the bandwidth and the coverage of frequency band b of operator z, respectively. Since there are five RBs per 1 MHz, the number of the RBs assigned for frequency band z_b is given by $5\mathsf{BW}_{z_b}$. Then, the effective number of the RBs assigned to frequency band z_b can be calculated as

$$N_{z_b} = \xi_{z_b} \times 5\mathsf{BW}_{z_b}. \tag{13}$$

The weight φ_{z_b} can then be defined as

$$\varphi_{z_b} = \frac{N_{z_b}}{\sum_{i=1}^{F_z} N_{z_i}}. \tag{14}$$

Finally, the WA-RBUR of operator z at month m is written by $\sum_{b=1}^{F_z} \varphi_{z_b} \mathbf{y}_{z_b,m}$. The WA-RBUR infers the network-wide resource utilization states by aggregating the RBUR of all operating spectrum bands with careful considerations of individual bandwidths and coverages. Therefore, it is more suitable to predict the spectrum saturation of an operator. Notice that with the WA-RBUR can be easily evaluated with the estimated feature parameter $\hat{\boldsymbol{\theta}}_{z_b,m}$ obtained by the proposed RNN.

V. NUMERICAL RESULTS

This section assesses the prediction performance of the proposed RBUR estimation process. Table 3 lists the model parameters and their values. For the RNN constructions, we employ both the LSTM and GRU models. Various hidden activations such as the hyperbolic tangent (tanh), rectified linear unit (ReLU), and leaky ReLU (LReLU) are investigated. The prediction error of the trained RNNs is measured by the mean absolute error (MAE) evaluated over the test dataset in the unit of percentage.

TABLE 4. Prediction error (%) of operator *A*.

tanh	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	**4.39**	**2.12**	**4.13**	1.58	11.59	**8.20**
$S=2$	10.34	4.20	8.34	11.74	**5.43**	12.65
$S=3$	10.49	6.79	9.44	9.35	8.96	11.74
$S=4$	12.62	3.97	8.18	9.96	11.51	10.12
$S=5$	9.67	3.80	7.12	4.41	6.51	11.67
$S=6$	4.90	3.40	9.28	8.20	10.45	10.47
ReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	8.58	3.75	14.86	1.44	11.23	11.96
$S=2$	12.45	8.29	9.05	10.97	8.29	9.50
$S=3$	13.14	7.58	11.81	10.35	10.04	11.09
$S=4$	11.60	7.41	8.58	10.67	10.79	11.00
$S=5$	9.34	5.95	9.24	7.81	9.65	12.99
$S=6$	8.24	4.46	7.63	8.18	8.79	10.23
LReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	6.10	2.96	12.76	**0.96**	10.83	14.61
$S=2$	11.25	7.01	12.76	10.09	7.81	10.22
$S=3$	14.97	7.19	11.80	7.57	13.18	14.82
$S=4$	10.48	7.94	8.02	10.87	10.11	15.55
$S=5$	10.42	3.93	7.68	7.29	9.55	13.33
$S=6$	7.80	5.18	7.79	7.74	10.65	12.59

(a) LSTM

tanh	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	6.95	**0.78**	**0.93**	2.67	10.62	6.62
$S=2$	13.01	3.69	7.84	9.02	**0.69**	13.36
$S=3$	11.36	6.16	9.06	8.97	7.64	12.85
$S=4$	12.11	6.29	6.23	10.67	8.63	10.64
$S=5$	12.46	2.32	6.91	5.22	9.05	14.09
$S=6$	9.66	3.37	7.45	7.86	10.89	12.99
ReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	**2.91**	4.27	14.44	1.72	11.17	50.26
$S=2$	10.37	8.20	12.24	11.54	7.46	8.95
$S=3$	12.18	8.67	9.10	9.13	10.64	10.69
$S=4$	13.94	8.30	8.29	13.09	10.66	13.77
$S=5$	12.84	4.67	8.72	8.66	7.79	9.74
$S=6$	9.17	2.93	8.27	9.99	10.16	11.90
LReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	3.50	2.47	6.60	**1.00**	8.81	19.25
$S=2$	12.69	5.67	12.13	9.65	7.59	12.56
$S=3$	10.94	6.28	10.01	8.42	8.72	12.56
$S=4$	11.32	4.84	7.01	11.06	11.58	**2.12**
$S=5$	13.92	2.96	5.38	6.67	6.98	10.36
$S=6$	10.34	2.84	8.91	8.08	10.03	11.46

(b) GRU

Table 4 presents the prediction error of the WA-RBUR of operator A with the LSTM and GRU models for various combinations of the input length S and the target difference Q. The best performance is highlighted in boldface blue letters for each given Q. Regardless of the RNN structures, the tanh activation exhibits good performance for all simulated Q. We observe that the GRU performs better than the LSTM.

The prediction error is less than only 1% when Q is small, and it slightly increases as Q grows. This is because predicting the Q-month forward RBUR becomes more difficult as the target difference Q gets larger. It is interesting to see that the best input sequence length S is generally given by $S = 1$ or $S = 2$ for all simulated Q. This implies that the temporal correlation of the RBUR of operator A is not that strong, and one or two-month samples are sufficient for successful prediction. Operator A has sufficient RBs since it occupies a wide range of bandwidth compared to operator B. Therefore, the spectrum saturation of operator A is dependent only on few-month past data and becomes independent with its long-term history.

Operator B exhibits the opposite results. The WA-RBUR prediction error of operator B is summarized in Table 5. We can see that at all simulated Q, using more past samples, i.e., a larger S, is beneficial for achieving a good prediction performance. The spectrum saturation issue becomes more severe for operator B as it has insufficient frequency resources. This leads to the time-dependent nature of the RBUR whose trends highly rely on the changes in time-varying features such as the number of subscribers. Thanks to the temporal correlations, the prediction error of operator B is generally smaller than that of operator A. In particular, the prediction error of the $Q = 6$-month forward future is only given as 0.81% for the GRU model, whereas that of operator A is large as 2.12%.

Fig. 4 shows the WA-RBUR prediction results of operator A (Fig. 4 (a)) and operator B (Fig. 4 (b)) with the best hyperparameters chosen from Tables 4 and 5. Six consecutive months from November 2019 to April 2020 are considered. The blue solid and red dotted lines represent the ground-truth and estimated WA-RBUR, respectively. From the figure, we can see that the WA-RBUR of both operators is accurately estimated in all simulated months. There is a gap in the prediction results of operator A in January 2020 and April 2020, but the gap is small as discussed in Table 4. It is worth noting that for the correct diagnosis of the spectrum saturation issue, the behaviors of the top few base stations are more important than other cells. In this regard, the proposed method still provides a good prediction result for the top 5% quantiles in all simulated cases.

One possible approach to diagnose spectrum saturation is to examine whether top 5% WA-RBUR is greater than 70%. Fig. 5 shows the prediction results of the top 5% WA-RBUR for operators A (Fig. 5 (a)) and B (Fig. 5 (b)), respectively. Among total 38 months, the WA-RBUR of the last 6 months, i.e., from the 33-rd month to the 38-th month, are predicted by the proposed frameworks. For comparison, the following baseline schemes are considered.

- *FNN:* The feature estimation is carried out by a fully-connected neural network (FNN) consisting of 9 dense layers each having 15/9, 30/18, 60/36, 120/72, 240/144, 120/72, 60/36, 30/18, and 15/9 neurons for operators A/B. Similar to the proposed RNN, the hyperparameters of the FNN baseline, i.e., the activations, the input length S, and the target difference Q, are carefully optimized as in Tables 4 and 5.

TABLE 5. Prediction error (%) of operator *B*.

tanh	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	**0.76**	**0.67**	27.67	12.48	27.77	19.44
$S=2$	4.36	7.34	6.74	3.58	5.64	5.54
$S=3$	13.39	3.49	2.21	**0.67**	4.86	2.45
$S=4$	10.85	2.36	5.33	6.62	5.34	**0.84**
$S=5$	2.11	2.00	6.41	5.27	7.54	4.23
$S=6$	3.96	12.15	5.31	8.42	3.49	3.66
ReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	6.28	10.56	58.23	16.78	33.67	10.83
$S=2$	3.81	14.24	91.12	5.17	5.26	4.18
$S=3$	7.04	11.30	1.94	4.46	**2.11**	4.14
$S=4$	24.44	12.27	5.27	7.75	4.39	4.58
$S=5$	1.82	12.12	7.14	6.45	3.61	4.23
$S=6$	6.99	3.48	8.14	8.69	2.46	3.28
LReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	4.15	134.44	29.43	20.46	9.77	1.87
$S=2$	5.35	22.92	26.41	6.82	2.18	4.91
$S=3$	4.46	6.36	**1.63**	3.78	2.25	2.36
$S=4$	3.49	5.33	5.06	5.27	3.66	4.17
$S=5$	1.11	2.53	123.51	7.48	4.39	3.64
$S=6$	5.34	4.59	4.73	9.45	3.72	1.65

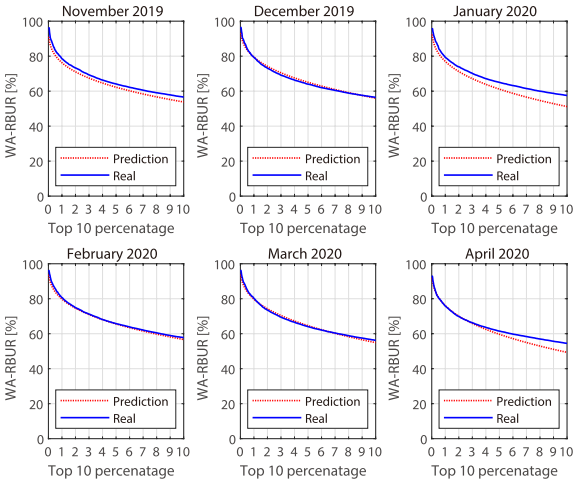
(a) LSTM

tanh	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	5.95	1.73	86.75	6.13	68.87	7.37
$S=2$	3.70	22.55	4.20	3.38	1.53	7.11
$S=3$	3.77	1.86	**0.56**	**0.67**	4.05	3.72
$S=4$	7.49	1.07	5.50	4.75	4.46	4.90
$S=5$	0.80	8.49	3.29	15.13	7.36	7.89
$S=6$	3.86	1.92	4.49	7.78	**0.80**	5.08
ReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	**0.29**	9.64	39.64	21.52	11.84	6.49
$S=2$	4.50	17.89	15.13	4.63	2.47	5.28
$S=3$	4.44	**0.52**	0.60	5.25	1.92	2.11
$S=4$	3.96	11.19	8.46	7.28	6.70	4.76
$S=5$	4.46	2.57	3.28	6.06	3.63	7.48
$S=6$	6.61	3.31	1.69	4.29	3.19	4.13
LReLU	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	$Q=6$
$S=1$	2.55	22.86	120.37	7.56	47.42	16.77
$S=2$	3.41	3.37	4.64	2.24	1.49	7.02
$S=3$	5.08	73.13	1.45	2.88	1.81	1.59
$S=4$	6.56	3.57	8.22	5.36	4.30	3.54
$S=5$	3.78	10.83	6.12	5.94	3.62	4.69
$S=6$	4.13	4.97	2.82	2.76	4.29	**0.81**

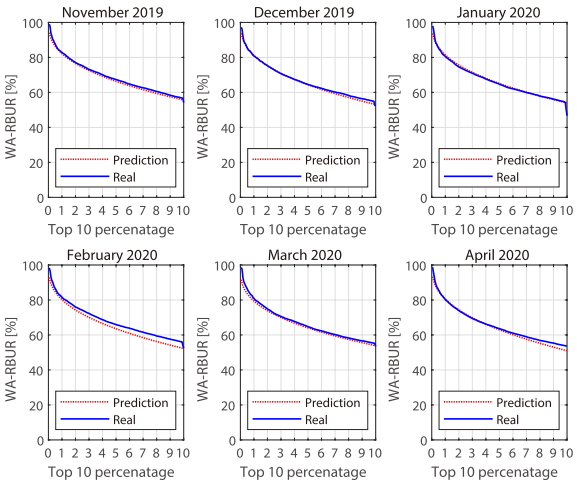
(b) GRU

- *VAR:* We employ a vector autoregression (VAR) method [26], which is one of the popular machine learning techniques addressing time series data. The VAR predicts the feature parameter at a certain month by taking weighted moving average of the last samples. Trainable variables, i.e., weighting matrices and a bias vector, are optimized through the linear least square method.

It is observed that regardless of the operators, the prediction of the proposed method is more accurate compared to the FNN

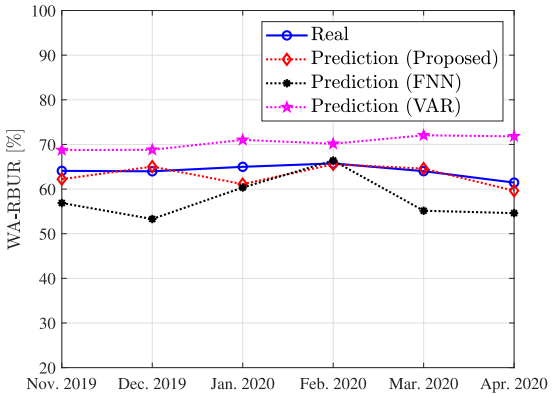


(a) Operator A

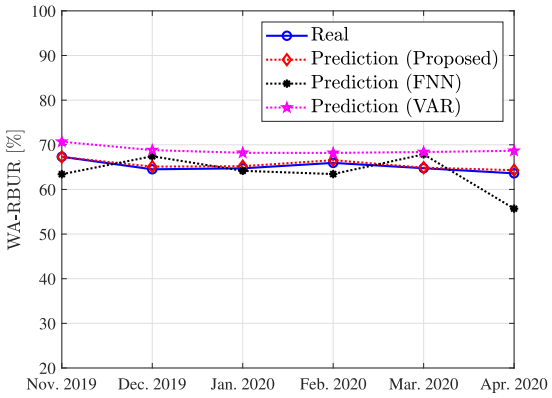


(b) Operator B

FIGURE 4. Top 10% WA-RBUR prediction results.



(a) Operator A



(b) Operator B

FIGURE 5. WA-RBUR prediction results for 6 months.

TABLE 6. WA-RBUR prediction error (%) for 6 months.

Months	33	34	35	36	37	38
FNN	7.20	10.68	4.63	0.65	8.86	6.84
VAR	4.22	4.34	5.60	3.95	7.44	9.83
Proposed	**1.86**	**1.04**	**3.87**	**0.19**	**0.57**	**1.82**

(a) Operator A

Months	33	34	35	36	37	38
FNN	3.91	2.88	0.52	2.54	3.14	7.90
VAR	3.33	3.87	3.53	1.98	3.33	4.83
Proposed	**0.04**	**0.59**	**0.47**	**0.60**	**0.12**	**0.70**

(b) Operator B

and VAR baselines. This implies that the RNN structure is suitable for estimating a sequence of the feature parameters. According to the proposed method, we would expect that the spectrum saturation does not occur for target months since the estimated WA-RBUR does not exceed 70%. However, the prediction values of both operators are quite close to the threshold of 70% WA-RBUR. Therefore, it is recommended for the operators to secure more spectrum resources not to encounter the spectrum saturation problem. On the contrary, the FNN and VAR baselines incorrectly diagnose the spectrum saturation events, especially for operator A, since their prediction values are much lower than 70%. Such a failure might cause difficulty in providing reliable communication services in practice.

Finally, Table 6 compares the prediction error of the WA-RBUR of the last six months investigated in Fig. 5. The lowest prediction error is highlighted in boldface blue letters. As expected from Fig. 5, the proposed RNN model is superior to other baselines in all months and operators. In particular, the WA-RBUR prediction error of operator B is smaller than 1%. The result validates the effectiveness of the proposed RNN-based approach. The FNN is not intended for analyzing sequential data and may fail to estimate the future RBUR correctly. The VAR generally performs better than the FNN baseline, but compared to the RNN model, its prediction error is not sufficiently small. This becomes pronounced for operator A. The estimation processes of the VAR rely on linear filtering operations to the past input sequences, posing difficulty for extracting the inherent nonlinearity of the RBUR dataset. On the contrary, the FNN can effectively

capture complicated properties of the RBUR measurements by adding nonlinear activations and hidden layers. For this reason, the prediction error of the FNN is lower than that of the VAR baseline in some cases. Both the nonlinearity and temporal correlation can be addressed by the RNN architecture, leading to the superiority of the proposed RNN approach to other baselines. Therefore, we can conclude that the RNN model is effective in handling the considered RBUR dataset.

VI. CONCLUSION

This paper investigated a deep learning-based forecasting strategy of the future RBUR status of hundreds of LTE base stations deployed in Seoul, South Korea. We solved the fixed-length input problem in traditional RNN-based forecasting methods by proposing a two-step approach: a feature extraction step and a feature estimation step. Based on the proposed feature extraction method, we can easily handle a large number (hundreds) of RBUR data only with three-dimensional feature parameters. The effectiveness of the proposed feature extraction strategy is verified by the R-squared values. The proposed RBUR estimation methods were then trained with LSTM and GRU models with the three-dimensional feature parameters of each operating band. Also, we introduced a new performance metric referred to as a weighted average RBUR (WA-RBUR), which is designed to compensate for the asymmetry in the bandwidths and cell coverages. We evaluated the performance of the proposed RBUR estimation methods in terms of WA-RBUR. The prediction accuracy of the proposed methods is less than only 3% and 1% for operators *A* and *B*, respectively. Hopefully, the proposed RBUR estimation methods can be utilized to diagnose spectrum saturation for mobile network operators with limited frequency resources.

REFERENCES

[1] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, 2011.

[2] S. Park, M. Agiwal, H. Kwon, and H. Jin, "An evaluation methodology for spectrum usage in LTE—A networks: Traffic volume and resource utilization perspective," *IEEE Access*, vol. 7, pp. 67863–67873, 2019.

[3] Y. Kim, H. Kwon, and S. Park, "Prediction of LTE spectrum saturation using quantiles of busy hour RB usage rates," *IEEE Commun. Lett.*, vol. 23, no. 8, pp. 1427–1431, Aug. 2019.

[4] Y. Lee, Y. Kim, Y. Park, and S. Park, "Probabilistic analysis of spectral efficiency for LTE based on PDCCH measurement data," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1626–1630, Sep. 2019.

[5] T. Kim, S. Park, H. Kwon, and Y.-K. Kim, "Estimation of mobile frequency saturation: Traffic volume and frequency resource utilization perspective," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2019, pp. 1454–1456.

[6] S. Hwang and S. Park, "On the effects of resource usage ratio on data rate in LTE systems," in *Proc. 19th Int. Conf. Adv. Commun. Technol. (ICACT)*, 2017, pp. 78–80.

[7] K. Tutschku and P. Tran-Gia, "Spatial traffic estimation and characterization for mobile communication network design," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 5, pp. 804–811, Jun. 1998.

[8] F. Ashtiani, J. A. Salehi, and M. R. Aref, "Mobility modeling and analytical solution for spatial traffic distribution in wireless multimedia networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1699–1709, Dec. 2003.

[9] N. Sapankevych and R. Sankar, "Time series prediction using support vector machines: A survey," *IEEE Comput. Intell. Mag.*, vol. 4, no. 2, pp. 24–38, May 2009.

[10] S. Gezici, H. Kobayashi, and H. V. Poor, "A new approach to mobile position tracking," in *Proc. IEEE Sarnoff Symp. Adv. Wired Wireless Commun.*, Mar. 2003, pp. 204–207.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[14] K. Cho, B. Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. Conf. Empiricial Methods Natural Lang. Process. (EMNLP)*, Oct. 2014, pp. 1724–1734.

[15] A. Diro and N. Chilamkurti, "Leveraging LSTM networks for attack detection in fog-to-things communications," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 124–130, Sep. 2018.

[16] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Commun. Mag.*, vol. 57, no. 6, pp. 114–119, Jun. 2019.

[17] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, "Spatio-temporal wireless traffic prediction with recurrent neural network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 554–557, Aug. 2018.

[18] J. Wang, J. Tang, Z. Xu, Y. Wang, G. Xue, X. Zhang, and D. Yang, "Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, May 2017, pp. 1–9.

[19] C.-W. Huang, C.-T. Chiang, and Q. Li, "A study of deep learning networks on mobile traffic forecasting," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.

[20] R. Wald, T. M. Khoshgoftaar, R. Zuech, and A. Napolitano, "Network traffic prediction models for near- and long-term predictions," in *Proc. IEEE Int. Conf. Bioinf. Bioeng.*, Nov. 2014, pp. 362–368.

[21] Y. Yu, J. Wang, M. Song, and J. Song, "Network traffic prediction and result analysis based on seasonal ARIMA and correlation coefficient," in *Proc. Int. Conf. Intell. Syst. Design Eng. Appl. (ISDEA)*, Oct. 2010, pp. 980–983.

[22] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[23] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.

[24] P. E. Gill and W. Murray, "Algorithms for the solution of the nonlinear least-squares problem," *SIAM J. Numer. Anal.*, vol. 15, no. 5, pp. 977–992, Oct. 1978.

[25] *Study on Channel Model for Frequencies From 0.5 to 100 GHz, Release 14*, document 3GPP TR 38.901, 2020.

[26] H. Lutkepohl, *New Introduction to Multiple Time Series Analysis*. Hoboken, NJ, USA: Wiley, 2005.

HAN SEUNG JANG (Member, IEEE) received the B.S. degree in electronics and computer engineering from Chonnam National University, Gwangju, South Korea, in 2012, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute for Science and Technology (KAIST), Daejoen, South Korea, in 2014 and 2017, respectively.

He was a Postdoctoral Fellow with Chungnam National University, Daejoen, from September 2017 to April 2018, and the Information Systems Technology and Design (ISTD) Pillar, Singapore University of Technology and Design (SUTD), Singapore, from May 2018 to February 2019. He is currently an Assistant Professor with the School of Electrical, Electronic Communication, and Computer Engineering, Chonnam National University, Yeosu, South Korea. His research interests include cellular Internet-of-Things (IoT)/machine-to-machine (M2M) communications, machine learning, smart grid, and energy ICT.

HOON LEE (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2012 and 2017, respectively. In 2018, he was a Postdoctoral Fellow with the Singapore University of Technology and Design, Singapore. Since 2019, he has been with Pukyong National University, Busan, South Korea, where he is currently an Assistant Professor with the Department of Information and Communication Engineering. His research interests include optimization, machine learning, and signal processing for wireless networks.

HYEYEON KWON received the B.S. degree in computer science and statistics and the M.S. and Ph.D. degrees in computer science from Chung Nam National University, Daejeon, South Korea. She is currently a Principal Member of Engineering Staff with the Electronics and Telecommunications Research Institute, Daejeon. Her research interests include mobile communications and data analysis for spectrum management.

SEUNGKEUN PARK (Member, IEEE) received the B.S. and M.S. degrees in applied statistics from Korea University, Seoul, South Korea, in 1991 and 1993, respectively, and the Ph.D. degree in information communication engineering from the Chungbuk National University, Cheongju, South Korea, in 2004. He is currently a Principal Member with the Electronics and Telecommunications Research Institute, Daejeon. His current research interests include communication theory and spectrum management.