ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/370487655

Dual-Step Transfer Learning-Based Prediction Model for Next-Generation Intelligent Cellular Networks

Conference Paper · May 2023

CITATIONS
0

READS
5

4 authors:

Waqar Ali Aziz
national university of science and technology islamabad
3 PUBLICATIONS **1** CITATION
SEE PROFILE

Iacovos Ioannou
University of Cyprus
21 PUBLICATIONS **81** CITATIONS
SEE PROFILE

Marios Lestas
Frederick University
87 PUBLICATIONS **817** CITATIONS
SEE PROFILE

Vasos Vassiliou
University of Cyprus
141 PUBLICATIONS **1,668** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project
Distributed Artificial Intelligence in D2D Communication. View project

Project
IDS4IoT - Computational and Artificial Intelligence Solutions for Intrusion Detection in Internet of Things View project

All content following this page was uploaded by Iacovos Ioannou on 03 May 2023.

The user has requested enhancement of the downloaded file.

Dual-Step Transfer Learning-Based Prediction Model for Next-Generation Intelligent Cellular Networks

Waqar Ali Aziz*†, Iacovos Ioannou *†, Marios Lestas‡, Vasos Vassiliou*†
*Department of Computer Science, University of Cyprus, Nicosia, Cyprus
†CYENS Center of Excellence, Nicosia, Cyprus
‡Department of Electrical Engineering, Frederick University, Nicosia, Cyprus
aziz.waqar-ali@ucy.ac.cy, ioannou.iakovos@ucy.ac.cy, eng.lm@frederick.ac.cy, vasosv@ucy.ac.cy

***Abstract*—Deep learning-enabled cellular traffic modelling and prediction is an indispensable part of next-generation intelligent mobile networks. It can help proactive resource management operations in big data-driven cellular networks. Existing approaches usually adopt a well-known prediction model for univariate time series forecasting and require a huge amount of training data. In this work, we propose a dual-step transfer learning-based prediction model for multivariate spatiotemporal cellular traffic prediction. This framework groups the base stations into different clusters based on the traffic correlation. A base model is trained on the aggregated dataset of one cluster by using a combination of Recurrent Neural Network and Bidirectional Long-Short Term Memory (RNN-BLSTM). By using parameter-based transfer learning, the base model is provided to the other clusters as a starting point where it undergoes the process of fine-tuning on an intra-cluster aggregated dataset. After that, the obtained intra-cluster prediction model is transferred to all the base stations inside the cluster, where simple training is required with few parametric adjustments. We conduct extensive experiments on a real cellular traffic dataset and results show that the proposed dual-step transfer learning-based prediction model can achieve good prediction accuracy while reducing the data, time and computation requirements.**

***Index Terms*—Deep learning, dual-step transfer learning, multivariate spatio-temporal cellular traffic prediction, RNN-BLSTM**

I. Introduction

The internet and its applications have become a primary communication tool for performing daily online activities. As predicted in a recent Cisco Report [1], the global mobile devices are expected to increase up to 13.1 billion by 2023, of which 1.4 billion will be 5G capable, thus increasing the complexity of the modern wireless cellular networks. Furthermore, the research community has reached a consensus that the integration of Artificial Intelligence (AI) with cellular networks can reduce the network overhead and elevate the Quality of Service (QoS) of both the access and the core networks [2]. More specifically, cellular traffic prediction is indispensable for the fusion of AI and future cellular networks (e.g. 6G). Although the traditional Machine Learning (ML) based prediction models have achieved great success, the training of these prediction models requires abundant labelled training data and is usually an expensive, time-consuming, and sometimes unrealistic task [3].

Transfer learning (TL) [4] is a promising machine learning technique for solving the traffic prediction problem by transferring knowledge from the source to the target domain. However, most of the work is done on univariate time series prediction. Therefore, this paper investigates the potential of dual-step TL on multivariate spatio-temporal cellular traffic. Unlike [5], in which the authors use a cloud service at a cloud server to store the univariate spatio-temporal dataset from the BSs, we first divide all the BSs into clusters based on the traffic distribution of each BS. Within a cluster, instead of combining the traffic of BSs, we take the aggregate traffic for each time sample by computing the average traffic of all the BSs. This aggregated dataset contains the traffic pattern of all the BSs within a cluster and significantly reduces the computation time for the base model training. After that, the model will be transferred in two steps: (1) the base model is transferred to the target cluster, and it will go through the process of fine-tuning to fit the model on the aggregated dataset of the target cluster, and (2) once the model is built on the aggregated dataset within each cluster, then it can be transferred to each BS. Hence, this technique prevents the BSs from learning from scratch. In other words, an intra-cluster model is formed for each cluster using the pre-trained model of another cluster, which reduces a significant amount of data, time, and computational resources when transferred to the BSs as a starting point.

Following the proposed framework and description given above, the contributions of this paper are summarized as follows:

1) We use a traffic similarity strategy to cluster all the BSs.
2) We train the base model using the aggregated traffic pattern to reduce the base model training time.
3) We use a dual-step TL technique to reduce the computation time at the BSs.
4) We propose the use of an RNN-BLSTM neural network

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 739578, the ADROIT6G project of the SNS-JU under Grant Agreement No. 101095363, and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

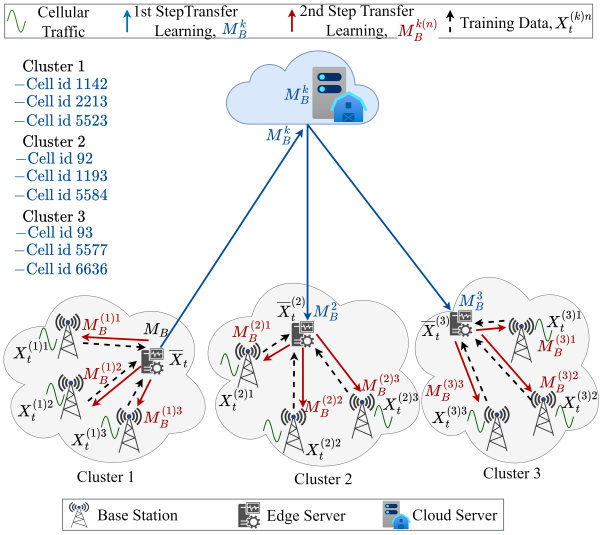


Fig. 1: Illustration of dual-step transfer learning-based prediction model in a cellular network having clusters based on the similarity in traffic distribution of BSs.

for the prediction of multivariate spatio-temporal cellular traffic.

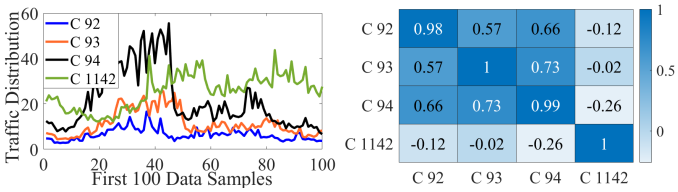
According to our analysis, the RNN-BLSTM has never been employed for heterogeneous multivariate spatio-temporal cellular traffic prediction except in speech gender classification, daily activity recognition, etc. [6], [7].

In this paper, we use a dataset generated by Telecom Italia cellular network over the city of Milan [8]. The Milan city cellular coverage is partitioned into 100×100 grids, called cells. The BS serves and logs the user's telecommunication activities in each cell. Thus, we use BS or cell interchangeably to denote a cell [16]. The rest of the paper is organized as follows. Section II discusses the related work. A problem formulation and a detailed proposed framework are presented in Section III and IV. The results are presented in Section VI based on the experimental setup in Section V. Finally, we conclude the paper in Section VII.

II. Related Work

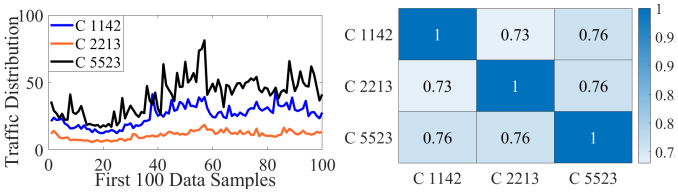
In TL, a pre-trained source model is transferred into a target domain. More specifically, the parameter-based TL approach is commonly applied for time series prediction problems where the knowledge is transferred at the model/parameter level [4]. TL can be jointly applied with many machine learning algorithms, e.g. deep neural networks. The reader may refer to the latest survey [9] to read the detailed information about the time series prediction algorithms.

In [10], the authors propose using CNN along with TL to improve prediction accuracy. However, we understand from the relevant literature that time series prediction problems can be better addressed with RNN rather than CNN. In [11], [12], the authors use CNN to model the spatial dependencies and LSTM neural network to model the time series information of the cellular traffic. In the same work, the inter-clustering TL strategy is then proposed to analyze and exploit the



(a) Time series pattern. (b) Traffic pattern correlation.

Fig. 2: (a) shows the visual difference in internet traffic patterns of different BSs, (b) shows the Pearson correlation of internet traffic among the selected BSs.



(a) Time series pattern. (b) Traffic pattern correlation.

Fig. 3: (a) shows the similarity in internet traffic patterns of different BSs within cluster 1, (b) shows the Pearson correlation of internet traffic among the BSs of cluster 1.

similarities in the spatial and temporal domains. Similarly, in [12], the authors use CNN to classify the cells, and LSTM neural network is used to train the prediction model. In [13], the authors use a convolutional LSTM (convLSTM) neural network to train the prediction model and use the Pearson correlation coefficient to find the correlation between the cross-domain datasets and different cellular traffic. In [14], the authors use an inter-cluster TL approach to train the model. In [5], the authors use a regularization-based TL method to share the knowledge with the target domain, but they use cloud servers to store the data from the BSs situated within the area of interest. This approach increases the burden on the network side. Furthermore, abundant cloud resources are required to store the data and train the prediction model. In [15], the authors use the LSTM-based transfer learning approach for univariate data.

However, in all the above approaches, the work is restricted to one input feature, i.e. internet traffic, as the multivariate spatio-temporal traffic is quite heterogeneous and requires a complex prediction model. In our paper, we propose the use of parameter-based dual-step TL with the fusion of the RNN-BLSTM neural network to predict the multivariate spatio-temporal cellular traffic. In our framework, first, the BSs are divided into clusters based on the traffic pattern similarity. After that, the prediction model is trained on the aggregated dataset within a cluster. The trained model is shared with the target cluster, and the model is further tuned on the aggregated dataset of the target cluster. Once the intra-cluster model is generated, then it can be provided to BSs for cellular traffic prediction with minor adjustments.

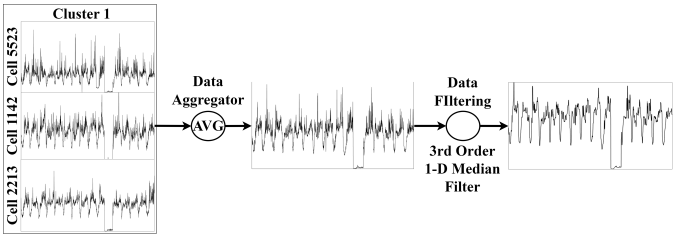


Fig. 4: Data augmentation process within each cluster.

TABLE I: Mean-variance analysis of multivariate spatio-temporal dataset.

		Cell ID 1142	Cell ID 2213	Cell ID 5523	Aggregated Data	Filtered Data
SMS	Mean	0.89	0.6	1.71	1.1	1.1
	Variance	0.28	0.126	1.23	0.34	0.3
Calls	Mean	0.81	0.67	2.23	1.24	1.23
	Variance	0.26	0.19	2.27	0.59	0.57
Internet	Mean	24.8	10.9	36.2	23.94	23.6
	Variance	88.1	13.6	294.6	81.94	72.28

III. Problem Formulation

Assume a source cluster has an aggregated dataset $\overline{X}_t$, and $\overline{X}_t^{(k)}$ are the aggregated datasets of the target clusters. The datasets of the target BSs inside each cluster is $X_t^{(k)_n}$. We aim to develop one base model for a source cluster $\{M_B\}$ and N transferred models for the target clusters $\{M_B^k\}$. After receiving the transferred model $\{M_B^k\}$, the target clusters fit the model on their aggregated datasets and generate n transferred models for the BSs $\{M_B^{(k)n}\}$ inside the cluster. When a new data sample $X_{(t+1)}^{(k)n}$ comes at n^{th} BS inside cluster k, the model needs to predict $Y_{(t+1)}^{(k)n}$ as accurate as possible. Subsequently, the loss function can be expressed as

$$L^{(k)n} = \frac{1}{|M_B^{(k)n}|} \sum_{Y_{(t+1)}^{(k)n} \in X_t^{(k)_n}} d(\hat{Y}_{(t+1)}^{(k)n}, Y_{(t+1)}^{(k)n})$$

Where d is the error metric, $|M_B^{(k)n}|$ is the size of dataset, and $\hat{Y}_{(t+1)}^{(k)n}$ is the prediction of actual value $Y_{(t+1)}^{(k)n}$.

IV. Proposed Framework

In this section, we give a detailed description of our proposed multivariate spatio-temporal cellular traffic prediction framework. The system model with nine BSs divided into three clusters is shown in Fig. 1. Our proposed framework can accommodate any number of clusters. Increasing the number of clusters based on traffic similarity among the BSs will enhance the prediction accuracy but at the cost of intra-cluster model training. The proposed model is mainly comprised of four important steps:

A. Cellular Network Clustering

In order to understand the diverse traffic patterns in different BSs, let us examine cells 92, 93, 94, and 1142. Fig. 2a presents the internet traffic distribution of the chosen BSs. A set of 100 samples are taken from the Telecom Italia dataset file having the date 01-11-2013. It can be seen that these BSs have different time series patterns, which can be verified by calculating the Pearson correlation coefficient, presented in Fig. 2b. Although cell 92, cell 93, and cell 94 are geographically close, they do not share similar traffic patterns. Similarly, the aforementioned cells have significant traffic pattern diversity with the cell 1142, which has no spatial relationship with them. This pattern diversity can influence the accuracy of the TL-based prediction model, and it is not possible to get full advantage of TL without addressing these factors.

In the examination done in our paper, we divide all the BSs into three clusters based on their internet traffic pattern similarity. For example, the internet traffic pattern similarity of cluster 1 can be observed in Fig 3a, whereas Fig 3b presents the Pearson correlation coefficient matrix. For simulation purposes, we are taking a total of 9 BSs, which are divided into three clusters, as shown in Fig. 1.

B. Cellular Traffic Data Aggregation

As we have seen, the spatially distributed BSs have diverse traffic patterns. These traffic patterns are highly heterogeneous and essentially considered to be non-IID [16], which can be analyzed from Fig. 2b and from Table I. Training a prediction model on a non-IID dataset is an extensive time, data, and computation task. In our paper, we propose inter-cluster learning where the prediction model is first trained within one cluster and then transferred to the other clusters. Data samples are taken from all the BSs within a cluster and averaged over the time step. This aggregated data includes the traffic pattern of each BS and reduces the model training time. After data aggregation, a 3^{rd} order 1-Dimensional median filter is applied to remove the spikes in the dataset. It can be observed from Table I that after removing the spikes from the augmented cellular traffic pattern, there is an insignificant change in the mean values between the aggregated and the filtered time series, while there is a slight reduction in the variance of the considered time series dataset. The data augmentation process is shown in Fig. 4.

C. Prediction Model

RNN-BLSTM combines two learning algorithms, (1) Real-Time Recurrent Learning (RTRL), which is based on the gradient-based update rule, and (2) Bidirectional LSTM that consists of two separate hidden layers, the forward (resp. backwards) layer, which learns on past (resp. future) trend. These two hidden layers are connected with the output layer to fuse past and future trends. This combination achieves better prediction accuracy on long-term dependencies. The proposed neural network is shown in Fig. 5.

1) Input layer: The dataset is comprised of $n = 1600$ data samples, having $p = 3$ traffic types (sms, calls, and internet) which are equivalent to 3 input features. A batch size of $n_b = 22$ is used for the training of the proposed model by keeping a balance between the estimation of accurate gradient and the computation cost. Moreover, the timestamp $K = 7$ is used.

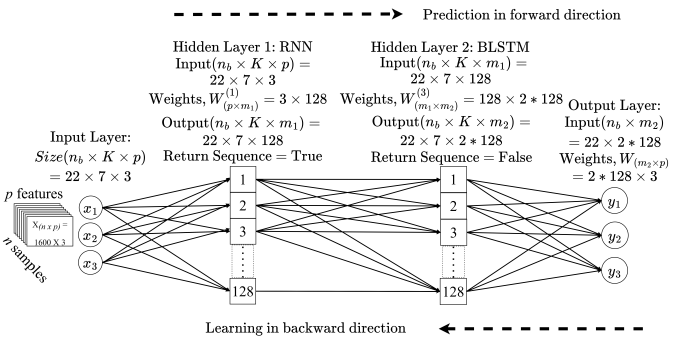


Fig. 5: RNN-BLSTM neural network model with one RNN layer and one BLSTM layer, each having 128 units.

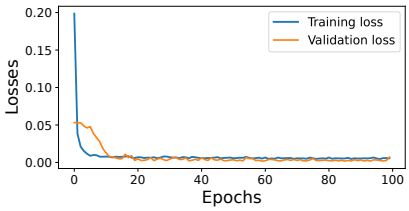


Fig. 6: Training and validation loss vs epochs.

2) Recurrent Neural Network: In our proposed mechanism, a single-layered RNN is used which is comprised of $m_1 = 128$ hidden units. The input size of the RNN layer is $n_b \times K \times p$ = $22 \times 7 \times 3$. This RNN layer produces the output equal to $n_b \times K \times m_1$ = $22 \times 7 \times 128$, which becomes the input to the single-layered BLSTM layer with the return sequence = True.

3) Bidirectional Long Short-Term Memory Neural Network: BLSTM is designed in such a way that it consists of 3 gates to control the flow of information, which are:

1) Forget gate
2) Input gate
3) Output gate

In the BLSTM technique, the input sequence is first fed to LSTM forward cell and then to the backward cell (which can be seen in the weight vector of the BLSTM layer, $2*128$). The output is the combination of the output of two cells. This forward and backward flow of the sequential data results in better and faster learning. BLSTM layer takes input of size $n_b \times K \times m_1$ = $22 \times 7 \times 128$ and produces output of size $n_b \times K \times (2 * m_2)$ = $22 \times 7 \times (2 * 128)$.

4) Output Dense Layer: The return sequence = False in the BLSTM layer shows that it will only return the last hidden state of the BLSTM layer. The BLSTM layer is followed by the dense output layer of shape $(2 * 128) \times 3$, which gives the desired predicted value $\tilde{Y_{t+1}}$.

The RNN-BLSTM model has trained over 100 epochs, taking 603 msec on average per epoch (11 msec per step). Fig. 6 presents the training and the validation loss for the proposed RNN-BLSTM model. It can be observed that the model performance is improving with experience. Furthermore, the summary of the neural network is presented in Table II.

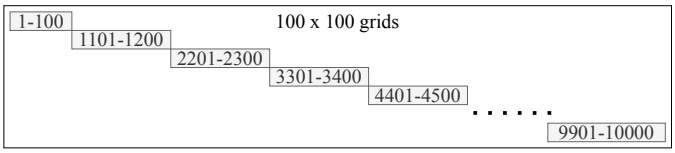


Fig. 7: Cells used to collect the dataset for clustering and model training.

D. Dual-Step Transfer Learning

It can be seen in Fig.1 that the base model M_B is trained on the aggregated data $\overline{X}_t$ of cluster 1, which reduces the need for big dataset to train the base model as well as reduces the base model training time. When this model $M_B^{(1)3}$ is transferred to the Cell id 5523, it needs less tuning because $\overline{X}_t$ and $X_t^{(1)3}$ ($X_t^{(1)3}$ = Cell id 5523) have a correlation of more than 95% for sms, call and internet traffic distribution. Whereas, if $M_B^{(1)3}$ is transferred to Cell id 5577 ($X_t^{(3)2}$= Cell id 5577), which is inside cluster 3, then the model needs more refinement on $X_t^{(3)2}$ because the correlation of $\overline{X}_t$ and $X_t^{(3)2}$ is 0.66, 0.68 and 0.51 for sms, call and internet traffic. Thus, in order to avoid extra training computation and to reduce the model training time at each BS, we propose a dual-step TL framework. The base model M_B is first trained on the intra-cluster aggregated data $\overline{X}_t$, as described in Section IV-B. After training, the copy of the base model M_B^k is provided to the cloud server. In the second step, cluster 2 and 3 fetch a copy of the base model M_B^2 and M_B^3 from the central cloud. These clusters first tune the fetched model on their aggregated data $\overline{X}_t^{(2)}$ and $\overline{X}_t^{(3)}$. After training, the model copies $M_B^{(k)n}$ can be provided to each BS i.e., $M_B^{(3)2}$ is the transferred model for the Cell id 5577 having data $X_t^{(3)2}$. The aforementioned operation prevents the second cluster from learning from scratch and prevents each BS from learning from a transferred model which is trained on a completely different dataset. This dual-step TL framework for cellular traffic prediction is of great benefit in terms of time, data, and computation resources. This technique will also reduce the burden on the core links and on the cloud servers as the model is trained on the edge servers.

V. Experimental Setup

This section provides details about the dataset collection for clustering and model training and the performance evaluation metrics.

TABLE II: Summary of the RNN-BLSTM model.

Layer (type)	Output Shape	Param #
simple_rnn (SimpleRNN)	(None, 7, 128)	16896
dropout (Dropout)	(None, 7, 128)	0
bidirectional (Bidirectional)	(None, 256)	263168
batch_normalization	(None, 256)	1024
dropout (Dropout)	(None, 7, 128)	0
dense_1 (Dense)	(None, 3)	771
Total params: 281,859 Trainable params: 281,347 Non-trainable params: 512		

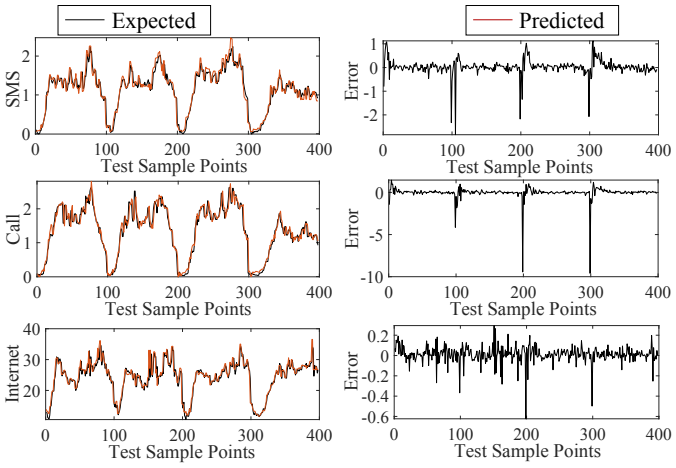
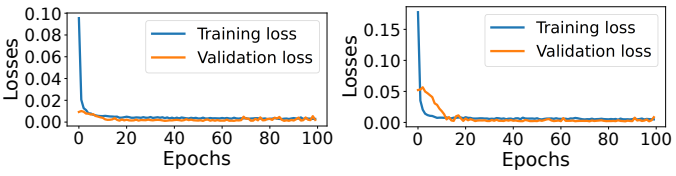


Fig. 8: Prediction performance of the base model on sms, call and internet traffic using the aggregated data of cluster 1.



(a) Cluster 2. (b) Cluster 3.

Fig. 9: Training and the validation loss of the transferred model vs epochs.

A. Dataset Collection

The Telecom Italia Milan city cellular coverage is partitioned into $10,000$ cells [8]. For simulation purposes, we take a total of 1000 cells, and a dataset of the first 16 days from $01-11-2013$ to $16-11-2013$ is considered. The chosen cells are situated on the diagonal of the 100×100 grids. These cells have sufficient required heterogeneity in their traffic patterns to make distinct clusters, as shown in Fig. 7. For the TL-based prediction model, 75% data is used for training, while 25% data is used for testing.

B. Performance Evaluation Metrics

Following the literature, four metrics are considered to evaluate the performance of the proposed framework, (1) Mean Absolute Error, $MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n}$, (2) Mean Absolute Percentage Error, $MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$, (3) Mean Square Error, $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ and (4) Root Mean Square Error, $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$. These metrics compare the actual y_i and the predicted values $\hat{y}_i$ of the time series. RMSE is often compared with the MAE, where a large gap shows that the error size is erratic.

VI. Performance Evaluation

In this section, we provide the results of the extensive experiments which were conducted on the real-world dataset described in Section IV-B, and by using the proposed prediction model discussed in Section IV-C. Mostly, we use internet traffic as a reference to explain the graphs and figures, but the same explanation is applied to sms and call traffic.

A. Base Model Performance

For simulation purposes, we choose cluster 1 to train the base model using RNN-BLSTM neural network. The results of the multivariate spatio-temporal prediction model trained on the aggregated intra-cluster datasets are presented in Fig. 8. Graphs on the left side show the expected and predicted traffic patterns for sms, call, and internet traffic, whereas graphs on the right side show the per-sample absolute error. The higher spikes in the error graph are due to the sudden insignificant change of traffic pattern, which seems harder for the prediction model to predict accurately, e.g. there is a 12 times sudden decrease in the cellular call traffic at test sample 299 as compared to 298 (decrease from 1.12261 to 0.10).

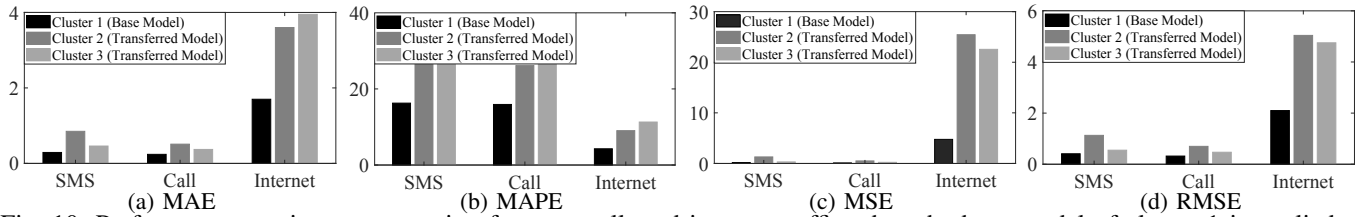
The performance metrics of the base model can be analyzed in Fig. 10. MAPE for sms, call, and internet traffic is 16.27, 15.92, and 4.27, respectively. MAPE values are higher for sms and call traffic because of the presence of sudden spikes in the dataset, which affect the average prediction accuracy of the model. For example, at test sample 299, the expected value is 0.1017, but the predicted value is 1.217, which produces a high MAPE value of 1190. It can also be observed from Fig. 10 that the RMSE and the MAE values are not diverging from each other but are close to each other, which shows the effectiveness of the prediction model. MSE is the average of the square of the difference between the actual and the predicted value, which shows the accuracy of the regression model. In our scenario, the MSE for sms, call, and internet traffic is 0.17, 0.1, and 4.76, respectively, for the base model.

B. Transfer Learning Performance

A point of significance in this work is the use of a dual-step TL strategy. Specifically, the parameter-based TL technique is investigated for the prediction of multivariate spatio-temporal cellular traffic.

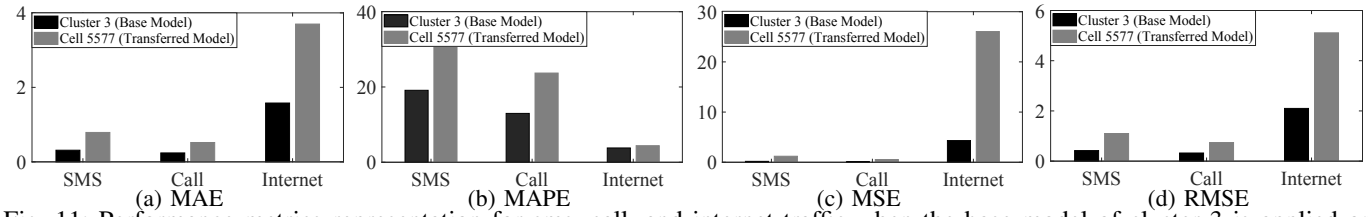
1) First Step Transfer Learning: In the first step, the pre-trained model of cluster 1 is transferred to the other clusters via the cloud server. When the pre-trained base model is implemented on the aggregated data of cluster 2 and 3 without any fine-tuning, although the losses are higher, it can be observed from Fig. 10 that the TL-based prediction model is capable of predicting the target cluster traffic with minor adjustments. For example, the RMSE of internet traffic for clusters 2 and 3 is 5.04 and 4.75, while it is 2.1 for cluster 1.

Fig. 9 shows the training and the validation loss when the transferred model is trained on the aggregated data of cluster 2 and 3, respectively. After training, clusters 2 and 3 will update the model parameters at the edge server. Fig. 11 shows the prediction losses for cluster 3 after minor adjustments. For example, the MAE, MAPE, MSE, and RMSE for the internet traffic in cluster 3 are improved by 219%, 244%, 562%, and 225%, respectively. These updated models become the intra-cluster base model for clusters 2 and 3.



(a) MAE (b) MAPE (c) MSE (d) RMSE

Fig. 10: Performance metrics representation for sms, call, and internet traffic when the base model of cluster 1 is applied on the aggregated dataset of cluster 2 and 3.



(a) MAE (b) MAPE (c) MSE (d) RMSE

Fig. 11: Performance metrics representation for sms, call, and internet traffic when the base model of cluster 3 is applied on the aggregated dataset of Cell 5577.

TABLE III: Error metric for the dual-step transfer learning-based prediction model for cell 5577.

Cluster 3	Traffic Type	MAE	MAPE	MSE	RMSE
Cell id 5577 (intra-cluster model training on BS dataset)	SMS	0.313	19.12	0.18	0.42
	Call	0.24	12.98	0.1	0.324
	Internet	1.58	3.8	4.32	2.1

2) Second Step Transfer Learning: In the second step, the intra-cluster base model is transferred to the BSs inside each cluster via the edge server. For simulation purposes, we are taking one BS from cluster 3 with cell id 5577. When the intra-cluster base model has applied to the BS 5577 dataset, due to the similarity in the traffic patterns, the MAPE value for internet traffic is just 4.4%.

Finally, the Intra-cluster base model is trained on the traffic pattern of the BS. After minor adjustments, it can be observed in Table III that the prediction accuracy improves significantly. MAE, MAPE, MSE, and RSME for the internet traffic are 1.58, 3.8, 4.32, and 2.1, respectively. It can be seen that the MAPE value reduces 12%, from 4.4 to 3.8 for internet traffic. Thus, by applying a dual-step TL-based prediction model, the data, time, and computational resources can be reduced for the model training and prediction without degrading the prediction performance.

VII. CONCLUSION

In this work, we investigated the use of a dual-step transfer learning-based prediction model for the prediction of heterogeneous multivariate spatio-temporal cellular traffic. We divided a cellular coverage area into different clusters based on the similarity in traffic distribution. A base model is developed using RNN-BLSTM neural network and trained on the aggregated dataset of one cluster. The parameter-based transfer learning approach is used to transfer the model to the other clusters, where it goes through the process of fine-tuning. Finally, the intra-cluster model is obtained, which is transferred to each BS within the cluster where it again goes through lightweight training with few parametric adjustments. The proposed technique achieved a good accuracy for a multivariate spatio-temporal dataset while utilizing less time, data, and computational resources. In the future, we will investigate the number of clusters required to achieve minimum upper-bound prediction error.

REFERENCES

[1] Cisco Annual Internet Report (2018–2023) White Paper, Cisco, San Jose, CA, USA, 2020.

[2] Letaief, Khaled B., et al. "The roadmap to 6G: AI empowered wireless networks." IEEE Communications Magazine 57.8 (2019): 84-90.

[3] Paleyes, Andrei, Raoul-Gabriel Urma, and Neil D. Lawrence. "Challenges in deploying machine learning: a survey of case studies." ACM Computing Surveys 55.6 (2022): 1-29.

[4] Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.

[5] Li, Hang, et al. "Data-Efficient Communication Traffic Prediction With Deep Transfer Learning." ICC 2022-IEEE International Conference on Communications. IEEE, 2022.

[6] Xue, Shaofei, et al. "Rapid speaker adaptation based on D-code extracted from BLSTM-RNN in LVCSR." 2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP). IEEE, 2016.

[7] Zhang, Jianshu, Jian Tang, and Li-Rong Dai. "RNN-BLSTM Based Multi-Pitch Estimation." Interspeech. Vol. 2016. 2016.

[8] Barlacchi, Gianni, et al. "A multi-source dataset of urban life in the city of Milan and the Province of Trentino." Scientific Data 2.1 (2015): 1-15.

[9] Jiang, Weiwei. "Cellular traffic prediction with machine learning: A survey." Expert Systems with Applications (2022): 117163.

[10] Fawaz, Hassan Ismail, et al. "Transfer learning for time series classification." 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018.

[11] Zhang, Chuanting, et al. "Deep transfer learning for intelligent cellular traffic prediction based on cross-domain big data." IEEE Journal on Selected Areas in Communications 37.6 (2019): 1389-1401.

[12] Dridi, Aicha, et al. "Transfer learning for classification and prediction of time series for next-generation networks." ICC 2021-IEEE International Conference on Communications. IEEE, 2021.

[13] Zeng, Qingtian, et al. "Traffic prediction of wireless cellular networks based on deep transfer learning and cross-domain data." IEEE Access 8 (2020): 172387-172397.

[14] Wu, Qiong, et al. "Deep transfer learning across cities for mobile traffic prediction." IEEE/ACM Transactions on Networking (2021).

[15] Wan, Xianbin, et al. "Network Traffic Prediction Based on LSTM and Transfer Learning." IEEE Access 10 (2022): 86181-86190.

[16] Zhang, Chuanting, et al. "Dual attention-based federated learning for wireless traffic prediction." IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, 2021.

View publication stats