

## Smoothing-aided long-short term memory neural network-based LTE network traffic forecasting

Mohamed Khalafalla Hassan<sup>1,2</sup>, Sharifah Hafizah Sayed Ariffin<sup>1</sup>, Sharifah Kamilah Syed-Yusof<sup>1</sup>, Nurzal Effiyana Ghazali<sup>1</sup>, Mohammed Eltayeb Ahmed Kanona<sup>2</sup>, Mohammad Rava<sup>3</sup>

<sup>1</sup>School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

<sup>2</sup>Faculty of Telecommunication Engineering, Future University, Khartoum, Sudan

<sup>3</sup>Department of Software Engineering, School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

### Article Info

#### Article history:

Received Oct 6, 2020

Revised Oct 14, 2021

Accepted Nov 5, 2021

#### Keywords:

Forecasting

Long-short term memory

Network traffic

Smooth

### ABSTRACT

There is substantial demand for high network traffic due to the emergence of new highly demanding services and applications such as the internet of things (IoT), big data, blockchains, and next-generation networks like 5G and beyond. Therefore, network resource planning and forecasting play a vital role in better resource optimization. Accordingly, forecasting accuracy has become essential for network operation and planning to maintain the minimum quality of service (QoS) for real-time applications. In this paper, a hybrid network- bandwidth slice forecasting model that combines long-short term memory (LSTM) neural network and various local smoothing techniques to enhance the network forecasting model's accuracy was proposed and analyzed. The results show that the proposed hybrid forecasting model can effectively improve the forecasting accuracy with minimal data loss.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



### Corresponding Author:

Mohamed Khalafalla Hassan

School of Electrical Engineering, University Technology Malaysia

81310, Johor Bahru, Johor, Malaysia

Email: memo1023@gmail.com

## 1. INTRODUCTION

Internet traffic has grown tremendously in the past decade due to vast network deployments in various domains and new emerging technologies with application-centric services. This growth was primarily driven by the surge in mobile traffic resulting from the substantial increase in mobile users and the emergence of new data-greedy technologies such as 5G, cloud services, the internet of things (IoT), and artificial intelligence (AI) based applications. Consequently, this has created a scalability issue for the service providers because network elements and resources must scale up spontaneously and dynamically to cope with the rising network resource demands.

Moreover, network management has become a complex task due to strong dependencies between the different service layers at which congestion may develop and spread horizontally and vertically. Network congestion caused by poor management will eventually affect the quality of service (QoS) and service levels. Therefore, proactive approaches to bandwidth and network resource management have become crucial. Such approaches would necessarily entail forecasting future network demands and planning accordingly to ensure a dynamic and timely reactive response. Hence, the accuracy of the predictive approaches has become a significant factor to apply predictive frameworks in production environments.

Preprocessing has also become crucial in data science, signal processing, and machine learning (ML) because of self-similarity, strong long-range dependence, and burstiness of network traffic. Data

preprocessing is also important to address the issue of incomplete, inconsistent (containing errors or outlier values), and varying noise patterns existing and embedded in collected data, which eventually lead to service providers being incapable of satisfying the minimum QoS requirements of dynamic bandwidth allocation. Therefore, preprocessing is required before applying network forecasting techniques to enhance data quality [1]–[5].

The significance of noise processing or removal has been addressed previously in [6]–[8] where the authors performed various time series breakdown using techniques such as loess, Hilbert Huang transformation (HHT), and wavelet processing to eliminate data fluctuation, seasonality and noise components. However, noise removal must be performed carefully after assessing the statistical properties of the data under investigation because the process itself can eliminate a significant portion of the data itself. Generally, two broad approaches based on different models have been used for developing bandwidth forecasting algorithms: i) statistical analysis models and ii) supervised ML models. Statistical analysis models are based on the generalized autoregressive integrated moving average (ARIMA) model, while most traffic forecasting models are based on supervised ML, in particular, artificial neural networks (ANNs) [1], [2]. ARIMA-based models fall short when dealing with nonlinear and non-stationary data as ARIMA requires a stationary property to be imposed [6]–[16] unlike ANN. This paper presents the results of the study on the effect of hybrid long-term short-term memory (LSTM) neural network and local smoothing techniques.

## 2. RELATED WORK

The literature review reveals that several models have been applied for time series analysis and bandwidth forecasting. An analysis of the performance of various machine learning techniques for forecast performance assessment of video over the internet, techniques such as neural network, decision trees, and support vector machine (SVM) was undertaken in [17]. The study concluded that modeling through time series method is more suitable and produces more stable results. Also, the study revealed that ANN outperformed other discussed machine learning techniques used as benchmarks.

Alawe *et al.* [18] proposed a novel mechanism to scale out the access management functions (AMF) in a 5G virtualized environment. The mechanism, which is based on forecasting mobile traffic using LSTM neural network to estimate the user attach request rate, makes it possible to predict the exact number of AMF instances required to process the upcoming user traffic. By being proactive, the proposed solution allows the deployment latency, which may degrade network performance, to be avoided when scaling up resources. Simulation results confirmed the efficiency of the LSTM-based solution compared with a threshold-based solution. The proposed approach applied the LSTM directly on the request rate data without any preprocessing, which may lead to forecast accuracy degradation as discussed and demonstrated in [7], [8].

Dyllon *et al.* [19] developed a nonlinear autoregressive exogenous neural (NARX) network model for time series network traffic analysis. The study implemented a neural network model to predict the future trends of the London South Bank University (LSBU) bandwidth data traffic. Dataset was collected using the paessler router traffic grapher (PRTG) tool. The results showed that NARX neural network is a good method for predicting time series data.

In another study, Dalgkitis *et al.* [20] compared LSTM performance for 4G traffic forecast against seasonal ARIMA (SARIMA) and support vector regression (SVR). The dataset was collected for 122 days and divided into two subsets: training and testing. The study found that LSTM performance was superior to SARIMA and SVR.

A deep traffic predictor (DeepTP) model to forecast long-period cellular network traffic was proposed in [21]. The results showed that the DeepTP model outperformed other traffic forecast models by more than 12.3%. The proposed approach applied a neural network directly to the dataset without any preprocessing. This may lead to forecast performance degradation due to data inconsistency, burstiness of the network traffic, and noise fluctuations, affecting QoS, network management, and security. In this study, we extend the reported research by studying the effects of the hybrid LSTM neural network and local smoothing techniques.

To resolve noise issues through preprocessing before applying traffic and other time series analysis, Yoo and Sim [6] proposed a hybrid Loess-ARIMA-based forecast model. Authors claimed that such a model has the potential of enhancing the efficiency of resource utilization, especially in high-speed networks, to accommodate the rapid increase in rising demands for scientific data applications. A seasonal decomposition of time series by Loess (STL) and ARIMA was applied on simple network management protocol (SNMP). The results revealed that the proposed forecast model was resilient against abrupt changes in network usage provided that the multistep forecast was used as the primary scenario.

In [7], the significance of the interference-less machine learning approach in a time series forecast as a crucial component of prediction performance was discussed, especially when forecasting multiple steps ahead. Afolabi *et al.* [7] used HHT as the noise elimination technique. The results were then compared with conventional and state-of-the-art approaches. In their study, Joo and Kim [8] discussed a wavelet-based prediction method to analyze the time series in the time and frequency domains. The study presented several scenarios. The results concluded that the proposed method outperformed the other approaches. However, based on our analysis, because the proposed hybrid models are static and do not react to the dynamic nature of traffic loads since the underlying defined functions work on local scales, it would be difficult to capture the rapidly varying noise fluctuations. Moreover, some techniques such as the wavelet transform method can aggressively eliminate parts of the original data if not implemented carefully.

### 3. METHOD

The proposed research method is depicted in Figure 1. Data used was collected from a premier internet service provider (ISP) representing LTE (4G) aggregated bandwidth slice. The dataset is divided into two subsets. The first part represents the training dataset that was used to train the forecasting models, while the second subset was used for testing purposes. LSTM neural network was used as the forecasting technique. LSTM neural network was chosen in this study because of its superiority in finding correlations between current and previous states due to its unique cell-based structure, unlike other shallow/deep multilayer perceptron networks. Therefore, it is considered one of the most suitable time series candidates despite noise smoothing not being part of its feature extraction capability. The bandwidth slice as a time series data contains an inherent temporal correlation (non-zero temporal autocorrelation). LSTM is superior for learning temporal dependencies in sequential data [22]. The effectiveness of the LSTM neural network for resource forecasting in communication networks has already been analyzed in [18], [20]–[23]. In this work two forecast time scales were used, one day and one week. A hybrid forecasting model that combines LSTM neural network and various local smoothing techniques is used to enhance forecast accuracy. Local smoothing techniques allow the removal of noise and fluctuation in short scales. Compared with other wavelet-based techniques, local smoothing techniques react more dynamically to noise level and short-term variations due to flexible window sizes that can be applied throughout the dataset. A similar approach has been utilized in [24] whereby the Li *et al.* investigated the superior nonlinear approximation capacity of using SVM compared with “classical” local smoothing techniques such as moving average, Gaussian smoothing, and Savitzky-Golay filter. The results showed that the proposed model outperformed the state-of-the-art model, i.e., logistic regression. The effectiveness of the proposed method was verified through available real network traffic datasets.

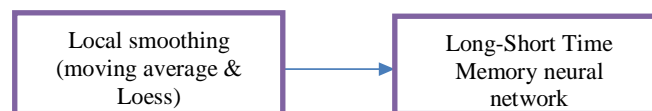


Figure 1. Proposed method

#### 3.1. Moving average

Generally, moving averages are usually calculated to identify trends. As discussed in the related work section above, ARIMA has been used extensively in time series and network traffic forecast. Moving average is a type of real-time filter that removes high frequencies from data. In signal processing, moving averages are also called “low pass filters” [25]. These filters have calculated coefficients that are equal to the reciprocal of the span or bandwidth.

Moving average is also known as “exponential smoothing”. Let  $C_i$  be defined as throughput at the time  $i$ . Let  $c = \{C_i\}$ ,  $i=1 \dots p$  be the time series where  $p$  is the time series length. Therefore, the moving average of the period  $q$  at the time  $l$  can be calculated as depicted in (1) [25]. The computational complexity as shown in (1).

$$m_l^q = \frac{1}{q} \sum_{i=1}^q c_{l-i+1} \quad (1)$$

#### 3.2. Local regression technique (Loess)

The Loess method [26] is based on fitting simple models to localized data subsets to form a curve that approximates the original data. Loess can be used as the sole forecasting method for traffic modeling in

very limited scenarios [27] or as a preprocessing technique for seasonal decomposition in time series forecasting with ARIMA [13]. However, it is not suitable for long-range forecast due to the window size optimal selection dilemma as explained below. The observations  $(x_i, y_i)$  are assigned neighborhood weights using the tricube weight function shown in (6). Let  $\Delta_i(x) = |x_i - x|$  be the distance from  $x$  to  $x_i$  and let  $\Delta_i(x)$  be these distances ordered from smallest to largest. Then, the neighborhood weight for the observation  $x_i, y_i$  is defined by the function  $w_i(x)$ .

$$w_i(x) = \left( \frac{\Delta_i(x)}{\Delta_q(x)} \right) \tag{2}$$

For  $x_i$  such that  $\Delta q(x) < \Delta_i(x)$  where  $q$  is called bandwidth and represents the number of observations in the subset of data localized around  $x$ . In the proposed algorithm, this approach is applied to fit a trend polynomial to the last  $k$  observations of the resource utilization. Accordingly, for each new observation, a new trend line  $\hat{g}(x) = \hat{a} + \hat{b}(x)$  is found. This trend line is used to estimate the next observation  $\hat{g}(xk + 1)$ . The new observation can be considered as bandwidth slice utilization [27].

**3.3. Smoothing windows selection**

The smoothing window in the smoothing process is represented by  $q$ . If  $q$  is selected to be small, insufficient data will fall within the smoothing window and as a result, a noisy fit will be produced. On the other hand, if  $q$  is selected to be large, a substantial amount of data will be eliminated. Therefore,  $q$  is selected to produce the least mean squared error (MSE). For moving average,  $q$  is found to be 0.003, and MSE is 2.4155e+07. As for Loess,  $q$  is found to be 0.002, and MSE is 6.4096e+04. For comparison purposes, wavelet decomposition, also performed at level one with MSE, was found to be 2.155e+7. Table 1 illustrates the steps for the local smoothing process. The algorithm computation complexity is  $O(n2)$ , where the computational complexity of Loess and moving average is  $O(q)$ .

Table 1. Algorithm 1

Input	Bandwidth time series	Series length	smoothing window size	Output
$\hat{y}$	$Z$	$q$	MSE	
$\hat{y}$ as locally fitted (predicted) value using local smoothing technique				
1- For $n=q$ to $Z-q$ do				
2- Initialize $K []$				
3- for $j= n-q$ to $n+q$ Do				
4- Smooth $(\hat{y}(j))$ with minimum MSE				
5- Assign $(\hat{y} \wedge (j))$ into $K []$				
6- return $\hat{y}$				

**3.4. LONG-SHORT TERM MEMORY (LSTM)**

LSTM is a neural network with modified structural components and is composed of chained units called cells, which are the most basic units of LSTM [28], [29]. Each cell is made up of three gates, namely an input gate, an output gate, and a forget gate. The function of the input gate is to save or memorize the current state; the output gate is used for the output; the forget gate is used to dismiss some information from the past. The relationships between the gates are sigmoid, dot product, and tanh functions. Figure 2 depicts the architecture of an LSTM cell and the functional relationships between the different gates. In Figure 2,  $f_t$  represents the forget gate,  $i_t$  represents the input gate and  $o_t$  is the output gate.

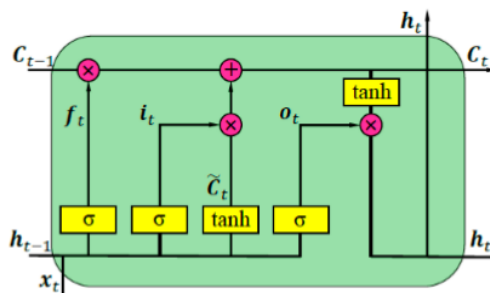


Figure 2. LSTM structure [21]

Figure 2 describe the logical relationships between input and output map it to the resource network variables along with the smoothing variables.  $x_t$  and  $h_t$  represent the input and the output respectively while  $W$  and  $b$  are weights and bias, respectively. Similar to the approach used in [23], hyperparameter selection was made by grid search, as depicted in Table 2. The learning computational complexity of LSTM  $O(W)$  per time step where  $W$  is equal to  $4ncnc + 4ninc + ncno + 3nc$  where  $nc$  is the number of memory cells,  $ni$  is number of inputs units and  $no$  is number of output gates.

Table 2. LSTM hyperparameters

Parameter	Name
Library	TensorFlow [30], Keras, NumPy, Sklearn
Batch size	1
Epochs	20
Optimizer/Learning rate	Adam
Loss function	RMSE
Neurons	2
Hidden layer	1
Activation function	ReLU

To enhance forecast accuracy, the hybrid LSTM and local smoothing forecast model is proposed. The mathematical representation of the hybrid LSTM and moving average model can be represented as (3) and (4).

$$C_t = \sigma(W_f \cdot [h_{t-1}, m_l^q] + b_f) * C_{t-1} + \sigma(W_i \cdot [h_{t-1}, m_l^q] + b_i) \tanh(W_c \cdot [h_{t-1}, m_l^q] + b_c) \quad (3)$$

$$h_t = \sigma(W_o \cdot [h_{t-1}, m_l^q] + b_o) * \tanh(\sigma(W_f \cdot [h_{t-1}, m_l^q] + b_f) * C_{t-1} + \sigma(W_i \cdot [h_{t-1}, m_l^q] + b_i) * \tanh(W_c \cdot [h_{t-1}, m_l^q] + b_c)) \quad (4)$$

The mathematical expressions of the hybrid LSTM and Loess model may be represented as (5) and (6).

$$C_t = \sigma(W_f \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_f) * C_{t-1} + \sigma(W_i \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_i) * \tanh(W_c \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_c) \quad (5)$$

$$h_t = \sigma(W_o \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_o) \tanh(\sigma(W_f \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_f) * C_{t-1} + \sigma(W_i \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_i) * \tanh(W_c \cdot [h_{t-1}, \hat{g}(x_k + 1)] + b_c)) \quad (6)$$

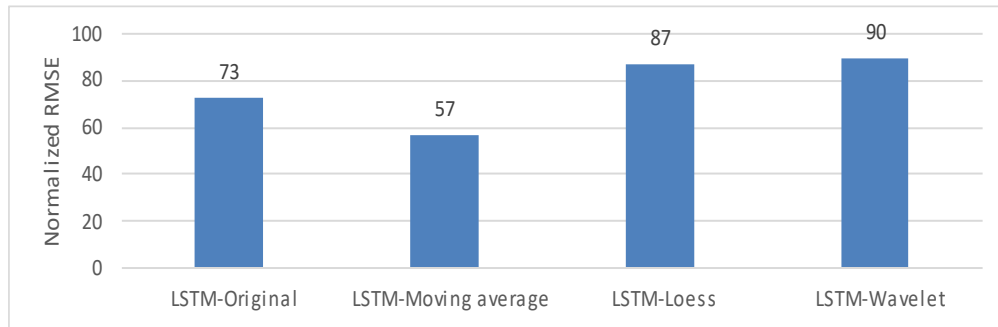
#### 4. RESULTS AND DISCUSSION

To enhance the time series forecast models, the augmented dicky-fuller (ADF) [31] test was used to confirm the stationarity of the time series while LSTM was used to model a non-stationary time series, as mentioned in the method section. Previous studies have recommended examining the stationarity of regression models as stationarity could lead to misleading results [32]. Real-time live trace was used as a dataset and modeled as a time series problem. Loess and moving average were applied successfully as a preprocessing technique with minimum smoothing window  $q$ . The higher the  $q$  values, the better the smoothing and the larger the amount of data loss. Certainly, in today's data-centric world, losing even small amounts of data could lead to QoS degradation. Also, the dataset has been divided equally into two sub-sets: 50% for the training and 50% for the testing with cross-validation, with Adam as the learning optimizer.

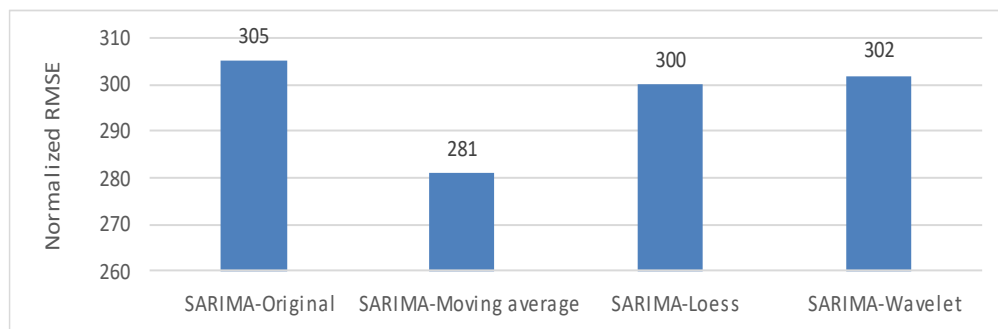
Figures 3(a) to 3(d) shows the performance comparison of the hybrid LSTM-moving average, LSTM-Loess, and LSTM-wavelet for the 50-time step ahead forecast and the 300-time step ahead forecast, with the LSTM-wavelet, used as a benchmark. In addition to that, to validate the effectiveness of our hybrid model, well-known hybrid techniques such as SARIMA with Loess, moving average, and wavelet were used as benchmarks as well. Results show that that the hybrid LSTM techniques outperform other non-hybrid techniques. For the 50-time step forecast, LSTM-moving average had the lowest root mean square error (RMSE) values compared with the other techniques. The LSTM-moving average RMSE value was the best with 57, while LSTM-Loess and LSTM-wavelet came in second and third, respectively. LSTM-moving average is evidently the best choice.

Similar results were found for the 300-time step forecast, as depicted in Figures 3(c) and 3(d). LSTM outperformed SARIMA by almost 76% even without smoothing techniques. The performance-improvement percentages were enhanced even further when comparing LSTM-moving average with

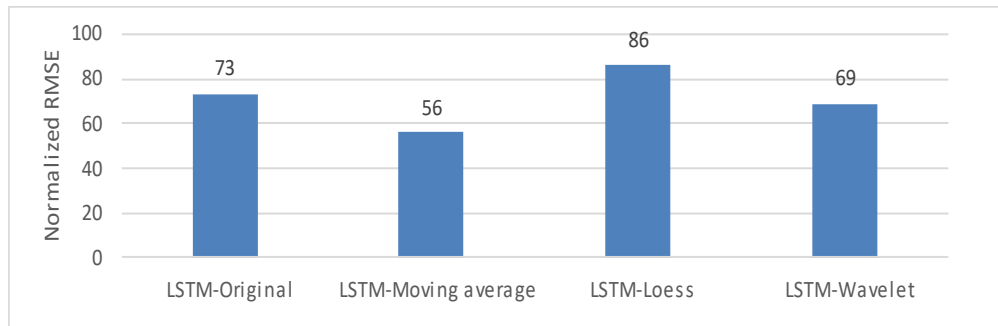
SARIMA-moving average, LSTM-Loess with SARIMA-Loess, and LSTM-Wavelet with SARIMA-wavelet. The performance-improvement percentages were 79%, 71%, and 70% for the 50-time steps ahead forecast; and 76%, 80%, and 77% for the 300-time steps ahead forecast.



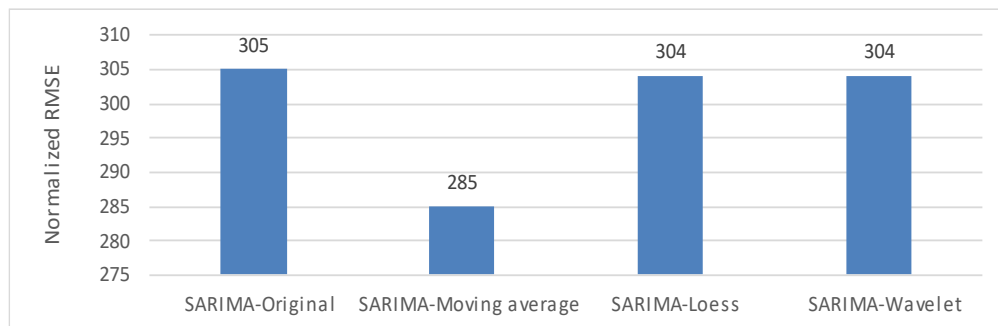
(a)



(b)



(c)



(d)

Figure 3. Performance comparison (a) LSTM-hybrid for 50-time step, (b) SARIMA-hybrid for 50-time step, (c) LSTM-hybrid for 300-time step, and (d) SARIMA-hybrid for 300-time step

Overall, it is evident that the moving average-based hybrid forecast technique outperforms other techniques due to the high smoothing capability that moving average can provide. Taking into consideration Algorithm 1 is used to minimize unnecessary data loss, the moving average-based hybrid forecast technique produced the highest smoothing MSE of  $2.4155e+07$ , compared with  $6.4096e+04$  for Loess and  $2.155e+7$  for wavelet smoothing. An explanation for this is that the smoothing window  $q$  in moving average has the greatest impact on the smoothing process, as depicted in (1) where  $q$  is constant for all  $x_i$ . This can be contrasted with Loess where  $q$  is embedded in weighing tricube functions that reduce the direct significant impact on the smoothing process, as a result of which less smoothing will be applied, as depicted in (2). In contrast, the wavelet-based hybrid smoothing technique is influenced more by the selected mother wavelet coefficients than the smoothing windows.

The Diebold-Mariano test [33] was then applied to check the statistical significance of the obtained results. The findings prove the superiority of the LSTM hybrid techniques. Based on the RMSE results, LSTM hybrid techniques are better and statistically different from other SARIMA hybrid techniques. These findings coincide with the previous findings of [23] except that in that study, SVM was used. Moreover, this results is considered as enhancement of the previous work in [6], [8], [18]. Figures 4(a) to 4(b) shows LTE bandwidth forecast using LSTM-Loess for 50-time steps ahead and 300-time steps ahead. As can be seen from Figure 4, our model is consistent with the actual series. The results confirm the hybrid LSTM's capability to capture seasonality and data fluctuations with minimal data loss (minimum MSE) arising from the smoothing process.

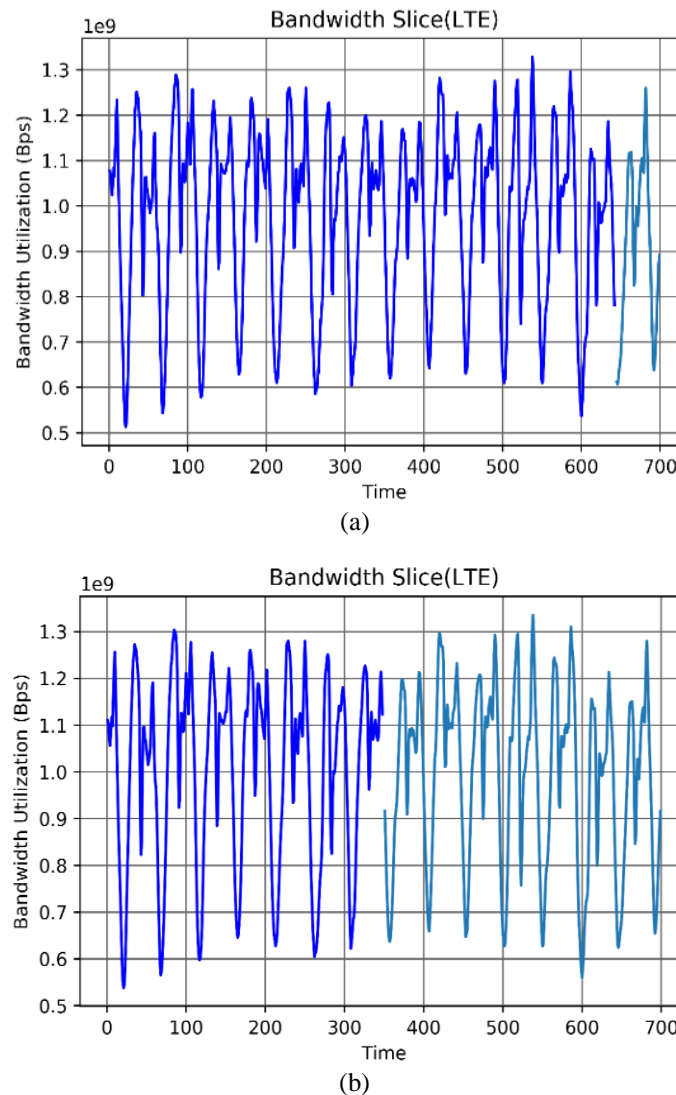


Figure 4. LTE original and forecasted bandwidth (a) 50-time step and (b) 300-time step

## 5. CONCLUSION

In this paper, hybrid local smoothing and LSTM modeling approaches were used to forecast LTE bandwidth slice utilization. Three local smoothing techniques, namely moving average, Loess and wavelet, were investigated in this study. The results reveal that the hybrid LSTM and moving average model showed better forecast performance in terms of RMSE. The significant performance advantages are attributed mainly to the prepositive smoothing mechanism controlled by the time window, which can eventually alleviate the high variability of local noise patterns. The results were verified by statistical significance tests and by undertaking a comparison with other similar state-of-the-art approaches. We believe that the method proposed can be used for slice traffic forecast in 4G/5G slice resource management. Further, this work can be extended to be applied in the automatic resource allocation algorithm as a part of slice allocator or orchestrator in 5G and beyond.

## REFERENCES




- [1] L. Nie, D. Jiang, L. Guo, and S. Yu, "Traffic matrix prediction and estimation based on deep learning in large-scale IP backbone networks," *Journal of Network and Computer Applications*, vol. 76, pp. 16–22, Dec. 2016, doi: 10.1016/j.jnca.2016.10.006.
- [2] D. D. Ningombam and S. Shin, "Traffic offloading in multicast device-to-device cellular networks: a combinatorial auction-based matching algorithm," *Sensors*, vol. 20, no. 4, Feb. 2020, doi: 10.3390/s20041128.
- [3] A. Svigelj, R. Serbec, and K. Alic, "Network traffic modeling for load prediction: a user-centric approach," *IEEE Network*, vol. 29, no. 4, pp. 88–96, Jul. 2015, doi: 10.1109/MNET.2015.7166196.
- [4] R. Li, Z. Zhao, J. Zheng, C. Mei, Y. Cai, and H. Zhang, "The learning and prediction of application-level traffic data in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3899–3912, Jun. 2017, doi: 10.1109/TWC.2017.2689772.
- [5] T. P. Oliveira, J. S. Barbar, and A. S. Soares, "Computer network traffic prediction: a comparison between traditional and deep learning neural networks," *International Journal of Big Data Intelligence*, vol. 3, no. 1, 2016, doi: 10.1504/IJBDI.2016.073903.
- [6] W. Yoo and A. Sim, "Time-series forecast modeling on high-bandwidth network measurements," *Journal of Grid Computing*, vol. 14, no. 3, pp. 463–476, Jun. 2016, doi: 10.1007/s10723-016-9368-9.
- [7] D. Afolabi, S.-U. Guan, K. Man, P. Wong, and X. Zhao, "Hierarchical meta-learning in time series forecasting for improved interference-less machine learning," *Symmetry*, vol. 9, no. 11, Nov. 2017, doi: 10.3390/sym9110283.
- [8] T. W. Joo and S. B. Kim, "Time series forecasting based on wavelet filtering," *Expert Systems with Applications*, vol. 42, no. 8, pp. 3868–3874, May 2015, doi: 10.1016/j.eswa.2015.01.026.
- [9] G. Abdoli, "Comparing the prediction accuracy of LSTM and ARIMA models for time-series with permanent fluctuation," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3612487.
- [10] M. F. Iqbal, M. Zahid, D. Habib, and L. K. John, "Efficient prediction of network traffic for real-time applications," *Journal of Computer Networks and Communications*, vol. 2019, pp. 1–11, Feb. 2019, doi: 10.1155/2019/4067135.
- [11] R. Gu, Z. Yang, and Y. Ji, "Machine learning for intelligent optical networks: a comprehensive survey," *Journal of Network and Computer Applications*, vol. 157, May 2020, doi: 10.1016/j.jnca.2020.102576.
- [12] R. Boutaba *et al.*, "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *Journal of Internet Services and Applications*, vol. 9, no. 1, Jun. 2018, doi: 10.1186/s13174-018-0087-2.
- [13] M. K. Hassan *et al.*, "Dynamic learning framework for smooth-aided machine-learning-based backbone traffic forecasts," *Sensors*, vol. 22, no. 9, March. 2022, doi.org/10.3390/s22093592
- [14] C. Li, Y. Han, Z. Sun, and Z. Wang, "A novel self-similar traffic prediction method based on wavelet transform for satellite internet," *ICST Transactions on Ambient Systems*, vol. 4, no. 14, Aug. 2017, doi: 10.4108/eai.28-8-2017.153306.
- [15] R. Madan and P. S. Mangipudi, "Predicting computer network traffic: a time series forecasting approach using DWT, ARIMA and RNN," in *2018 Eleventh International Conference on Contemporary Computing (IC3)*, Aug. 2018, pp. 1–5, doi: 10.1109/IC3.2018.8530608.
- [16] M. K. Hassan *et al.*, "Analysis of hybrid non-linear autoregressive neural network and local smoothing technique for bandwidth slice forecast," *TELKOMNIKA*, vol. 19, no. 4, April. 2021, doi.org/10.12928/telkomnika.v19i4.17024.
- [17] J. You, H. Xue, L. Gao, G. Zhang, Y. Zhuo, and J. Wang, "Predicting the online performance of video service providers on the internet," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19017–19038, Sep. 2017, doi: 10.1007/s11042-017-4460-0.
- [18] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, and P. Bertin, "Improving traffic forecasting for 5G core network scalability: a machine learning approach," *IEEE Network*, vol. 32, no. 6, pp. 42–49, Nov. 2018, doi: 10.1109/MNET.2018.1800104.
- [19] S. Dyllon, T. Hong, O. A. Oumar, and P. Xiao, "Educational bandwidth traffic prediction using non-linear autoregressive neural networks," in *Robotics Transforming the Future*, Sep. 2018, pp. 431–441, doi: 10.13180/clawar.2018.10-12.09.37.
- [20] A. Dalgkitis, M. Louta, and G. T. Karetos, "Traffic forecasting in cellular networks using the LSTM RNN," in *Proceedings of the 22nd Pan-Hellenic Conference on Informatics*, Nov. 2018, pp. 28–33, doi: 10.1145/3291533.3291540.
- [21] J. Feng, X. Chen, R. Gao, M. Zeng, and Y. Li, "DeepTP: an end-to-end neural network for mobile cellular traffic prediction," *IEEE Network*, vol. 32, no. 6, pp. 108–115, Nov. 2018, doi: 10.1109/MNET.2018.1800127.
- [22] B. Soni, D. K. Patel, and M. Lopez-Benitez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97437–97451, 2020, doi: 10.1109/ACCESS.2020.2995633.
- [23] U. Singh, J.-F. Determe, F. Horlin, and P. De Doncker, "Crowd forecasting based on WiFi sensors and LSTM neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6121–6131, Sep. 2020, doi: 10.1109/TIM.2020.2969588.
- [24] Y. Li, J. Wang, X. Sun, Z. Li, M. Liu, and G. Gui, "Smoothing-aided support vector machine based nonstationary video traffic prediction towards B5G networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7493–7502, Jul. 2020, doi: 10.1109/TVT.2020.2993262.
- [25] A. Raudys and Ž. Pabarškaitė, "Optimising the smoothness and accuracy of moving average for stock price data," *Technological and Economic Development of Economy*, vol. 24, no. 3, pp. 984–1003, May 2018, doi: 10.3846/20294913.2016.1216906.
- [26] M. K. Hassan *et al.*, "SLA management for virtual machine live migration using machine learning with modified kernel and statistical approach," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, Feb. 2021






- [27] A. Beloglazov and R. Buyya, "Energy efficient resource management in virtualized cloud data centers," in *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, 2010, pp. 826–831, doi: 10.1109/CCGRID.2010.46.
- [28] S. Haider *et al.*, "LSTM neural network based forecasting model for wheat production in Pakistan," *Agronomy*, vol. 9, no. 2, Feb. 2019, doi: 10.3390/agronomy9020072.
- [29] X.-H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, Jul. 2019, doi: 10.3390/w11071387.
- [30] M. Abadi *et al.*, "TensorFlow: large-scale machine learning on heterogeneous distributed systems," *arXiv:1603.04467*, Mar. 2016, *arXiv:1603.04467*.
- [31] T. Kulaksizoglu, "Lag order and critical values of the augmented dickey-fuller test: a replication," *Journal of Applied Econometrics*, vol. 30, no. 6, pp. 1010–1010, Sep. 2015, doi: 10.1002/jae.2458.
- [32] S. Bunrit, N. Kerdprasop, and K. Kerdprasop, "Multiresolution analysis based on wavelet transform for commodity prices time series forecasting," *International Journal of Machine Learning and Computing*, vol. 8, no. 2, pp. 175–180, Apr. 2018, doi: 10.18178/ijmlc.2018.8.2.683.
- [33] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, vol. 13, no. 3, Jul. 1995, doi: 10.2307/1392185.

## BIOGRAPHIES OF AUTHORS






**Mohamed Khalafalla Hassan**    is a researcher and ICT specialist with 16 years of wide range of research and ICT experience. Received his BSc in computer engineering in 2004 from future university Sudan and MSc from University Putra Malaysia (UPM) in 2009 in communication network engineering, currently he is a PhD candidate in communication engineering at university technology Malaysia (UTM), his interests are but not limited to software defined networks, application of machine learning in communication networks and network function virtualization. He can be contacted at email: memo1023@gmail.com.






**Sharifah Hafizah Sayed Ariffin**    received her B.Eng. (Hons) from London, in 1997, and obtained her MEE and Ph.D. in 2001 and 2006 from University Technology Malaysia, and Queen Mary, University of London, London respectively. She is currently an Associate Professor with the Faculty of Electrical Engineering, University Technology Malaysia. Her current research interest is on Internet of Things, Ubiquitous computing and smart devices, wireless sensor networks, IPV6, Handoff management, network and mobile computing system. She had published 116 papers, 17 copyrights, 1 integrated circuit, and 1 trademark. She can be contacted at email: sharifah@fke.utm.my.



**Sharifah Kamilah Syed-Yusof**    received B.Sc. degree in Electrical Engineering from George Washington University, USA in 1988 and obtained her M.E.E. and Ph.D. degrees from UTM in 1994 and 2006, respectively. She is currently working as associate professor with Faculty of Electrical Engineering, UTM. Her research interest includes wireless communication, software defined radio and cognitive radio. She is a member of Eta Kappa Nu (HKN). She can be contacted at email: kamilah@utm.my.






**Nurzal Effiyana Ghazali**    received M.S. degree in Electrical Engineering from Shibaura Institute of Technology and Ph.D. degrees from UTM in 2016, she is doing research in mobile computing, mobility management, network communication protocol, and sport monitoring system. She can be contacted at email: nurzal@utm.my.



**Mohammed Eltayeb Ahmed Kanona**    received the B.Sc., M.Sc., and Ph.D. degrees in Telecommunication Engineering from Future University-Sudan. He is currently the Deputy Dean, Faculty of Telecommunication and Space Technology and He is also head of IoT research center. He actively involved in research about forward scattering radar. His research interests include information theory, SDN, IoT, cloud computing, machine learning and neural networks, and mobile communication. He can be contacted at email: mohamed\_kanona@hotmail.com.



**Mohammad Rava**    is a researcher and software engineer with a specialization in machine learning and metaheuristic algorithms. He completed his PhD in Software Engineering in Universiti Teknologi Malaysia. His current research interests are machine learning solution, metaheuristic algorithms and pattern recognition and analysis. He can be contacted at email: morava.ir@gmail.com.