https://doi.org/10.1051/itmconf/20257003021

Advanced Network Traffic Prediction Using Deep Learning Techniques: A Comparative Study of SVR, LSTM, GRU, and Bidirectional LSTM Models

Yuxin Wang

College of Arts and Sciences, New York University, New York, United States

Abstract. Accurate prediction of network traffic patterns is essential for optimizing network resource allocation, managing congestion, and strengthening cybersecurity. This study examines the effectiveness of four machine learning models—Support Vector Regression (SVR), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Bidirectional Long Short-Term Memory (Bi-LSTM)—in forecasting traffic patterns using both web-based and real-world datasets. The models are evaluated based on their generalization accuracy, as measured by Mean Absolute Percentage Error (MAPE), computational efficiency, and their ability to capture underlying traffic dynamics. Results indicate that GRU surpasses SVR and LSTM in terms of prediction accuracy and computational speed, while Bidirectional LSTM demonstrates superiority in capturing long-term dependencies across extended periods. These findings underscore the significant potential of deep learning models, particularly GRU and Bidirectional LSTM, in improving the precision and reliability of network traffic predictions. The study offers insights into the strengths and limitations of each model, contributing to the ongoing development of more robust and efficient network traffic forecasting methods.

1 Introduction

Accurate and efficient network traffic prediction ensures optimal network performance and prevents congestion. Network administrators rely on predictive models to allocate resources, identify potential bottlenecks, and enhance the quality of service. Traditionally, statistical models like AutoRegressive Integrated Moving Average (ARIMA) and Holt-Winters have been employed for prediction. However, they often fail to capture the non-linearities and long-term dependencies present in network traffic data.

With the advent of machine learning, especially deep learning, models such as Recurrent Neural Networks (RNNs) have become prominent due to their capability to process time-series data. Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) offer enhanced performance in managing sequential data by retaining relevant information over

Corresponding author: yw6054@nyu.edu

longer time frames. In contrast, simpler models like Support Vector Regression (SVR) are better suited for smaller datasets but struggle with larger, more complex patterns.

This paper aims to evaluate the difference in performance of SVR, LSTM, GRU, and Bidirectional LSTM (Bi-LSTM) in forecasting network traffic. The models are tested based on their prediction accuracy and computational efficiency, focusing on capturing the underlying patterns in the network traffic.

2 Related Work

Network traffic prediction has been extensively studied due to its role in managing resources, mitigating congestion, and enhancing security. Traditional methods, such as statistical models, usually face challenges when attempting to capture the non-linear and intricate patterns found in modern network traffic [1, 2]. The introduction of machine learning, particularly RNN-based architectures like LSTM and GRU, has significantly improved long-term traffic forecasting, especially in scenarios with high variability [3].

Innovations like Diffusion Convolutional Recurrent Neural Networks (DCRNN) have further advanced the field by combining graph learning with RNNs, improving the prediction of both temporal and topological features [4]. Other deep learning approaches, such as multilayer perceptrons, have been applied to large-scale traffic matrix estimation, enhancing accuracy and resource allocation by capturing long-term dependencies across multiple network links [5].

Hybrid models that integrate artificial neural networks (ANNs) with statistical methods have shown success in volatile, high-traffic environments, with improvements in real-time anomaly detection [6, 7]. Additionally, the combination of dimensionality reduction techniques like Principal Component Analysis (PCA) with classification algorithms such as SVM and Random Forest has improved classification accuracy by focusing on key features and minimizing noise [8].

Recurrent models, particularly LSTM and GRU, have demonstrated notable success in capturing long-term dependencies in network traffic [9, 10]. Bidirectional LSTM (Bi-LSTM) models, optimized with metaheuristic algorithms, have further improved short-term traffic flow prediction by capturing dependencies in both forward and backward sequences [11].

In summary, while traditional statistical models serve as a baseline, modern deep learning techniques—especially RNNs and hybrid models—have proven more effective in capturing non-linear and long-term patterns in network traffic.

3 Proposed Methodology

This part outlines the methodology for network traffic prediction, focusing on SVR, LSTM, GRU, and Bi-LSTM models. The dataset employed is derived from the WIDE Project and captures real-world traffic patterns.

3.1 Data Preprocessing

The WIDE Project dataset includes traffic data collected at 15-minute intervals over 24 hours, with features such as packet length, timestamps, and protocol types. The dataset was divided into smaller segments using Editcap to improve computational efficiency. Additionally, standardization was applied, normalizing the data to a mean of 0 and a standard deviation of 1 to ensure model convergence.

 https://doi.org/10.1051/itmconf/20257003021

3.2 SVR

SVR serves as a baseline model, minimizing prediction error within a predefined tolerance margin. Hyperparameter tuning was conducted using GridSearch, optimizing variables like the regularization term (C), kernel type, and epsilon. Predictions were rescaled using the 'reverse_scale' function, and Mean Absolute Percentage Error (MAPE) was selected to evaluate the performance.

3.3 LSTM

LSTM models, known for handling time-series data, use memory cells controlled by forget, input, and output gates to capture both long-term and short-term patterns, which makes them suitable for recognizing traffic patterns like congestion (see Fig. 1).

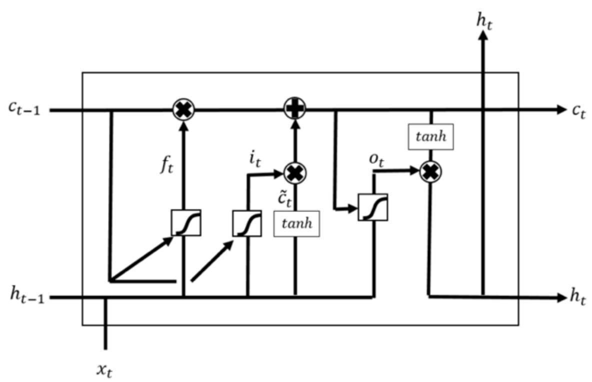

Fig. 1. LSTM Cell Structure.

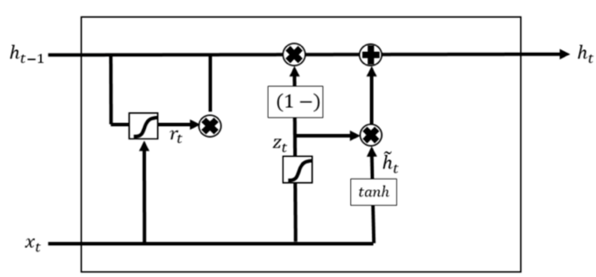

Fig. 2. GRU Structure.

3.4 GRU

GRU models simplify the LSTM architecture by combining gates, reducing computational complexity while maintaining similar accuracy. This makes GRU efficient for real-time network traffic prediction, especially in resource-constrained environments (see Fig. 2).

3.5 Bidirectional LSTM

Bi-LSTMs analyze input sequences in forward and reverse directions, enabling the model to capture complex traffic patterns that depend on both past and future states. This is particularly useful in detecting traffic bursts or drops.

3.6 Model Evaluation

The models were evaluated using MAPE, which increased with extended look-ahead periods, reflecting higher uncertainty in future predictions. GRU demonstrated better efficiency, while Bidirectional LSTM excelled in scenarios requiring more complex time-dependent processing.

4 Results and Discussion

This section evaluates the performance of SVR, LSTM, GRU, and Bidirectional LSTM models for network traffic prediction, with MAPE serving as the primary evaluation metric.

4.1 Analysis of MAPE Trends for All Models

Across all models (see Fig. 3, 4, 5), MAPE increases with extended time-step predictions, a typical pattern in time-series forecasting due to two factors:

1. Increased Uncertainty: Longer time horizons result in more uncertain predictions, especially in non-linear systems like network traffic.
2. Cumulative Error: Errors made in earlier predictions compound as the forecast extends further into the future.

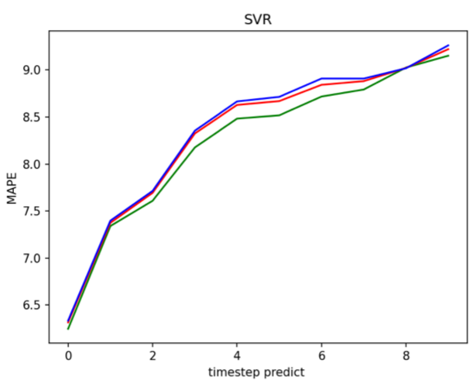

Fig. 3. MAPE trend for SVR model.

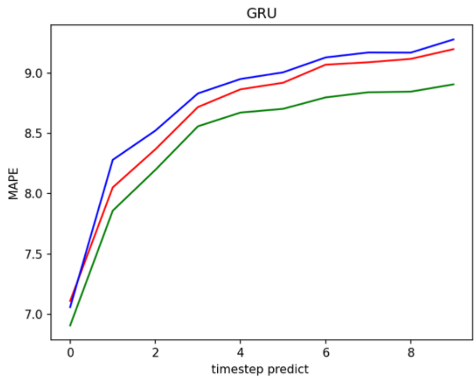

Fig. 4. MAPE trend for GRU model.

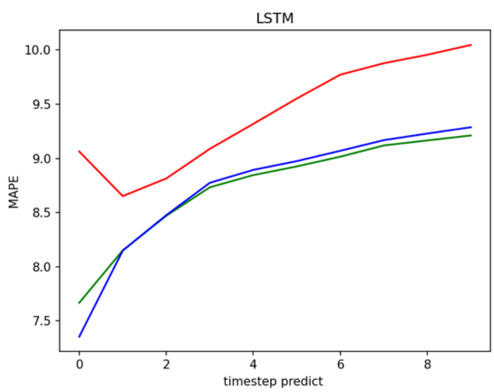

Fig. 5. MAPE trend for LSTM model.

4.2 SVR, GRU, and LSTM Model Performances

SVR performed best in capturing simple traffic patterns, with a first-step MAPE of 6.23%. GRU struck a balance between accuracy and computational efficiency, achieving a MAPE of 6.57%. LSTM, while designed for long-term dependencies, struggled with overfitting, resulting in a higher MAPE of 7.13% (see Table 1).

Table 1. First-step MAPE results for SVR, GRU, and LSTM models.

Look Back	Look Ahead	Neural Layer 1	Neural Layer 2	Neural Layer 3	MAPE: Single step (%)	Model
40	15				6.23318062	SVR
60	10	80	60	100	6.57251721	GRU
30	10	40	40	100	7.13031672	LSTM

4.3 Bi-LSTM Model Performance

Bi-LSTM, with its bidirectional architecture, improved slightly over standard LSTM, achieving a MAPE of 7.04%. Its ability to process input sequences in both directions helps capture complex dependencies, but it still lagged behind GRU and SVR in overall accuracy (see Table 2).

Table 2. First-step MAPE results for Bi-LSTM model.

Look Back	Look Ahead	Neural Layer 1	Neural Layer 2	Neural Layer 3	MAPE: Single Step (%)	Model
40	10	60	40	80	7.0466621	Bi-LSTM

4.4 General Observations and Recommendations

1. Simplicity vs. Complexity: SVR outperformed more complex models like LSTM and Bi-LSTM for simpler traffic patterns, suggesting that simpler models can sometimes be more effective when deep architectures are unnecessary.
2. GRU's Middle Ground: GRU offers a balance between complexity and performance, making it suitable for tasks involving moderately intricate data patterns.
3. Contextual Use of Models: While LSTM and Bi-LSTM did not perform as well in this study, they may excel in more complex, noisy datasets where long-term dependencies play a greater role.

5 Conclusion

This study provides a comparative analysis of three prominent models for network traffic prediction: SVR, LSTM, GRU, and Bidirectional LSTM. GRU demonstrated a balanced performance with superior accuracy and computational efficiency, particularly in scenarios with moderately complex traffic patterns. In contrast, SVR proved more effective in more straightforward datasets with prevalent periodic traffic patterns. Bidirectional LSTM outperformed the other models in cases requiring the capture of long-term dependencies, although its computational complexity limits its applicability in real-time environments.

Future work could explore the use of more complex and volatile datasets and additional network features such as latency, protocol types, and packet flow direction. Further optimization of hyperparameters, leveraging automated techniques such as AutoML, could also enhance model performance. The results highlight the significance of selecting the most suitable model based on specific characteristics of the dataset and the intended application scenario. As network environments evolve, integrating these machine learning techniques with real-time monitoring systems presents promising opportunities for improved network resource management and cybersecurity measures.

References

1. M. Joshi, T. Hassn Hadi, A Review of Network Traffic Analysis and Prediction Techniques, *arXiv e-prints*, Art. no. arXiv:1507.05722 (2015).
2. G. O. Ferreira, C. Ravazzi, F. Dabbene, G. C. Calafiore, M. Fiore, Forecasting Network Traffic: A Survey and Tutorial With Open-Source Comparative Evaluation, IEEE Access. **11**, 6018–6044 (2023).

3. N. Ramakrishnan, T. Soni, Network Traffic Prediction Using Recurrent Neural Networks, in *17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL, USA (2018), pp. 187–193.
4. D. Andreoletti, S. Troia, F. Musumeci, S. Giordano, G. Maier, M. Tornatore, Network Traffic Prediction based on Diffusion Convolutional Recurrent Neural Networks, in *IEEE INFOCOM Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Paris, France (2019), pp. 246–251.
5. D. Jiang, G. Hu, Large-Scale IP Traffic Matrix Estimation Based on the Recurrent Multilayer Perceptron Network, in *IEEE International Conference on Communications*, Beijing, China (2008), pp. 366–370.
6. S. Chabaa, A. Zeroual, J. Antari, Identification and Prediction of Internet Traffic Using Artificial Neural Networks, *J. Intell. Learn. Syst. Appl.* **2**, 147–155 (2010).
7. G. Zheng, W. K. Chai, J.-L. Duanmu, V. Katos, Hybrid deep learning models for traffic prediction in large-scale road networks, *Inf. Fusion* **92**, 93–114 (2023).
8. W. Xing, Time Series Traffic Analysis based on PCA and GMM Algorithm, in *3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, Xi'an, China (2022), pp. 706–710.
9. S. Wang, Q. Zhuo, H. Yan, Q. Li, Y. Qi, A Network Traffic Prediction Method Based on LSTM, *ZTE Commun.* **17**, 19–25 (2019).
10. H.-C. Jang, C.-A. Chen, Urban Traffic Flow Prediction Using LSTM and GRU, *Eng. Proc.* **55**, 86 (2023).
11. B. Naheliya, P. Redhu, K. Kumar, MFOA-Bi-LSTM: An optimized bidirectional long short-term memory model for short-term traffic flow prediction, *Physica A* **634**, 129448 (2024).