Bandwidth prediction for volatile networks with Informer

Birkan Denizer, Olaf Landsiedel

Angaben zur Veröffentlichung / Publication details:

Denizer, Birkan, and Olaf Landsiedel. 2023. "Bandwidth prediction for volatile networks with Informer." In 2nd Workshop on Machine Learning & Networking (MaLeNe), co-located with the 5th International Conference on Networked Systems (NetSys 2023), Potsdam, Berlin, September 4, 2023: proceedings, edited by Michael Seufert, Andreas Blenk, and Olaf Landsiedel. Augsburg: Universität Augsburg.

2ND WORKSHOP
"MACHINE LEARNING & NETWORKING" (MaLeNe) PROCEEDINGS

SEPTEMBER 4, 2023

CO-LOCATED WITH
THE 5TH INTERNATIONAL CONFERENCE ON NETWORKED SYSTEMS (NETSYS 2023)
POTSDAM, GERMANY

Bandwidth Prediction for Volatile Networks with Informer

Birkan Denizer
bde@informatik.uni-kiel.de
Kiel University
Kiel, Germany

Olaf Landsiedel
ol@informatik.uni-kiel.de
Kiel University
Kiel, Germany

Abstract

5G networks provide high throughput and low latency connections, crucial for remote monitoring and control of mission-critical operations. Managing buffer levels and accurate bandwidth estimations are essential for low-latency applications. However, wireless networks are susceptible to fluctuations in quality metrics due to mobility and interference, impacting link utilization. Sudden quality deterioration can lead to lower Quality-of-Experience (QoE). To address this, we propose a neural network-based bandwidth prediction system. Our system utilizes historical data for time-series forecasting using the Informer model. It achieves 10% lower errors on a publicly available LTE dataset and 51% lower errors on a publicly available 5G dataset. Future work includes multivariate predictions and the creation of a new 5G dataset.

Keywords: Bandwidth prediction, LTE, 5G, Informer

1 Introduction

5G networks enable high throughput and low latency connections. Remote monitoring and control of the vital infrastructure of mission-critical operations such as surgeries or production lines depend on 5G to provide a good Quality-of-Experience (QoE). To achieve low latency communication, such operations rely on accurately filling buffers to send data to minimize the latency of the end-to-end connection. Applications calculate estimations for the average bandwidth of a connection and generate packets. Low-latency applications keep buffer sizes to a minimum. Minimal buffer sizes result in higher dependence on accurate knowledge about future buffer states. However, applications experience low QoE when sudden quality deterioration happens. In this paper, we introduce a neural network-based approach for bandwidth prediction. It allows adjustment of application-specific parameters to increase QoE. Such a prediction system is not only vital for better link utilization but also for low latency.

Inherently, wireless networks are highly susceptible to fluctuations in quality metrics such as Signal-to-noise Ratio (SNR) or Received Signal Strength Indicator (RSSI). User mobility or sources of interference and blockage may cause variable bandwidth on the connection. This variance results in sub-optimal usage of the existing infrastructure as applications require time to adjust to newly attainable bandwidth.

Our performance metrics include data received from hardware and software sources such as SNR, RSSI, bandwidth, or round-trip time. Established approaches use historical data to generate the bandwidth predictions for a horizon. Time-series forecasting (TSF) is the process of analyzing and predicting future values based on historical data. Our approach builds on TSF as the main task to predict the bandwidth of a connection. In this domain, Long Short-term Memory (LSTM) and Transformer models gained popularity in extracting meaningful information [3] [5].

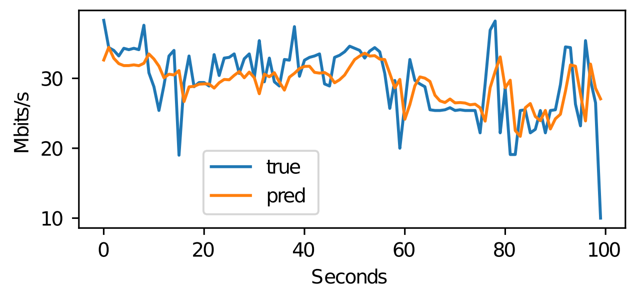

Figure 1. Bandwidth prediction on the LTE dataset

This paper makes the following contributions:

1. Design and development of a neural network-based approach that utilizes the Informer architecture to make bandwidth predictions based on historical metrics
2. Analysis of datasets and implementation of the MinMaxScaler for data normalization to a given range
3. Evaluation of our bandwidth prediction approach which shows up to 10% lower errors on the LTE dataset and up to 51% lower errors on the 5G dataset

2 Methodology

Transformer gained popularity since its introduction to Natural Language Processing (NLP) tasks [5]. One of the key challenges with Transformer-based approaches is their complexity in terms of time and memory during training.

Informer [6] is an efficient Transformer-based model for Long Sequence TSF tasks. It introduces the ProbSparse self-attention mechanism to reduce time and memory complexity during training, self-attention distilling to improve performance in the presence of extremely long input sequences, and a generative decoder to improve inference efficiency.

In this paper, we utilize the Informer architecture for bandwidth predictions. Namely, we use multivariate channel context information to predict several seconds of future bandwidth. Additionally, we implement the MinMaxScaler as our

Horizon	Metric	TPA-LSTM	Informer	Informer-M
1	RMSE	4.0038	4.0901	4.1281
1	MAE	2.9043	2.9786	3.0126
2	RMSE	4.6102	4.3954	4.4021
2	MAE	3.2362	3.1906	3.16
3	RMSE	5.0779	4.716	4.6033
3	MAE	3.5488	3.4356	3.2702

Table 1. Comparison on the LTE dataset: TPA-LSTM is better on horizon 1, Informer-M is better on Horizon 2&3

analysis of the datasets shows features with different ranges. The MinMaxScaler provides normalization of the data to a given range, often between zero and one. It scales each feature separately to the given range. Separate scaling prevents features with larger ranges from affecting the bias of scaling.

3 Datasets

We evaluate our approach on public LTE [3] and 5G [4] datasets. The LTE dataset contains several transportation scenarios like bus and subway lines. The LTE dataset contains bandwidth, LTE-neighbors, RSSI, Reference Signal Received Quality (RSRQ), change in ENodeB compared to the previous second, time advance to the next ENodeB, speed, and band. We choose "downloading" and "video streaming" use cases while driving from the 5G dataset. The 5G dataset includes metrics such as Reference Signal Received Power (RSRP, RSRQ), Channel Quality Indicator (CQI), SNR, RSSI, and download and upload bandwidths.

Traces include regularly sampled data. The 5G dataset has missing data points. Therefore, we use the forward-fill imputation method to remedy missing points.

4 Evaluation

In this work, we use the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) to evaluate prediction quality. We implement the MinMax Scaler and test its performance with the Informer (noted as Informer-M) on prediction horizons ranging from 1 to 24. We use bus line 15 and train line 7 of the LTE dataset, and "downloading while driving" as the mobility scenario in the 5G dataset.

Azarid et al. [1] compare Autoregressive Integrated Moving Average (ARIMA) and LSTM models and find that LSTM outperforms ARIMA for predictions. Mei et al. [2] show that the Temporal Pattern Attention LSTM (TPA-LSTM) model outperforms Recursive Least Squares (RLS), Random Forest (RF), and LSTM. We select TPA-LSTM as one of our baselines.

As Table 1 shows, TPA-LSTM performs better for horizon length 1 on the LTE dataset. Also in Table 1, Informer performs between 1-8 % lower on MAE, and 5 to 10 % lower on RMSE for horizon lengths 2 and 3.

On the 5G dataset, our solution with imputation (noted as Informer-I) achieves a 45 to 48% lower RMSE and 49 to 51% lower MAE compared to prior Informer-based solutions on different horizons, see Table 2.

Horizon	Metric	Informer + Lasso	Informer-I
1	RMSE	0.72	0.3801
1	MAE	0.35	0.1764
6	RMSE	1.19	0.6151
6	MAE	0.63	0.2994
24	RMSE	1.33	0.7318
24	MAE	0.73	0.3613

Table 2. Comparison on the 5G dataset: Informer-I performs overall better compared to prior Informer-based solution

5 Conclusion

This work introduces a neural network-based design that applies a state-of-the-art Transformer-based Informer model to predict bandwidth. Our bandwidth prediction system aids in helping users to have better QoE. In this work, we apply our solution to various datasets. Our results show up to 10% lower errors on the LTE dataset, and up to 51% lower errors on the 5G dataset compared to state-of-the-art approaches.

For future work, we plan to make multivariate predictions based on multivariate data. We also plan to generate a new regularly-sampled 5G dataset for new mobility scenarios.

6 Acknowledgements

This research has been partially funded by the German Ministry of Transport and Digital Infrastructure within the projects CAPTN Förde Areal (45DTWV007B) and CAPTN Förde Areal II (45DTWV08D).

References

[1] Amin Azari, Panagiotis Papapetrou, Stojan Denic, and Gunnar Peters. 2019. Cellular traffic prediction and classification: A comparative evaluation of LSTM and ARIMA. In *Discovery Science: 22nd International Conference, DS 2019, Split, Croatia, October 28–30, 2019, Proceedings 22*. Springer, 129–144.

[2] Lifan Mei, Jinrui Gou, Yujin Cai, Houwei Cao, and Yong Liu. 2022. Realtime mobile bandwidth and handoff predictions in 4G/5G networks. *Computer Networks* 204 (2022), 108736.

[3] Lifan Mei, Runchen Hu, Houwei Cao, Yong Liu, Zifa Han, Feng Li, and Jin Li. 2019. Realtime mobile bandwidth prediction using lstm neural network. In *Passive and Active Measurement: 20th International Conference, PAM 2019, Puerto Varas, Chile, March 27–29, 2019, Proceedings 20*. Springer, 34–47.

[4] Darijo Raca, Dylan Leahy, Cormac J Sreenan, and Jason J Quinlan. 2020. Beyond throughput, the next generation: a 5G dataset with channel and context metrics. In *Proceedings of the 11th ACM multimedia systems conference*. 303–308.

[5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[6] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35. 11106–11115.