Forecasting the network traffic with PROPHET*

Paweł Cembaluk, Jakub Aniszewski, Aleksandra Knapińska[0000−0003−2654−4893], and Krzysztof Walkowiak[0000−0003−1686−3110]

Department of Systems and Computer Networks, Wrocław University of Science and Technology, Wrocław, Poland
aleksandra.knapinska@pwr.edu.pl

Abstract. With the constant development of networking technologies and the increase of internet userbase, traffic prediction is becoming a vital part of today's network optimization. In this paper, we propose a method for network traffic prediction based on the PROPHET model. We examine its different parameters find their best configuration for diverse traffic types. Our research has shown that PROPHET is an accurate solution for backbone optical network traffic forecasting for a 14-day horizon.

Keywords: Traffic prediction · Application-aware network · Timeseries.

1 Introduction

The overall traffic in today's backbone optical networks is seeing tremendous growth in the last few years, especially during the COVID-19 pandemic. Various network-based services are widely used in many areas, including education, business, and entertainment. In turn, the summary traffic consists of multiple low-bitrate flows and thus is characterized by strong daily and weekly seasonality with periodically recurring patterns [3]. In light of the inevitable *capacity crunch*, various solutions are being developed for more efficient use of the existing network resources. Multilayer application-aware network optimization [6] is seen as a promising approach, in which different types of traffic related to various services and applications are treated according to their specific quality of service (QoS) requirements. The knowledge regarding future volumes of traffic used in network optimization algorithms can improve their efficiency and decrease bandwidth blocking [7].

In this article, we present a traffic prediction method based on PROPHET – a solution proposed by Facebook [5]. The PROPHET is a forecasting procedure for time data series based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. This method proved its effectiveness in the prediction of cellular network traffic [1,2,8] but, to the best of our knowledge, has not been studied in the context of backbone optical

* This work was supported by National Science Centre, Poland under Grant 2019/35/B/ST7/04272.

networks. To fill this research gap, we propose a throughout study of various parameters of PROPHET and evaluate it on datasets corresponding to diverse types of backbone optical network traffic. The rest of the paper is organized as follows. In Sec. 2, we define the problem and describe the chosen method. Sec. 3 describes conducted numerical experiments. Sec. 4 concludes this work.

2 Problem definition and method description

The problem considered in this work concerns the prediction of traffic in backbone optical networks. The created regression model attempts to predict future volumes of a specific traffic type with a given sampling rate based on historical data. The chosen forecast horizon is 14 days.

The PROPHET algorithm [5] provides automated timeseries forecasts that can be tuned. In this paper, we explore the impact of three of the PROPHET parameters on the prediction quality of various network traffic types. The first parameter is the changepoint prior scale, and it determines how much the trend changes at changepoints. Its value is a tradeoff between trend under- and overfitting. The second parameter is the seasonality prior scale. Its large value allows the seasonality to fit large variations, and a lower one decreases the magnitude of the seasonality. Because of the density of input data, as the last parameter, we chose the number of automatically placed changepoints.

As a reference approach, we propose a Linear Regression model, which proved to be an accurate method for network traffic forecasting [4]. Our implementation, adjusted for long-term traffic forecasting, takes four input features: the hour of the day, minute of the day, second of the day, and weekday.

3 Numerical experiments

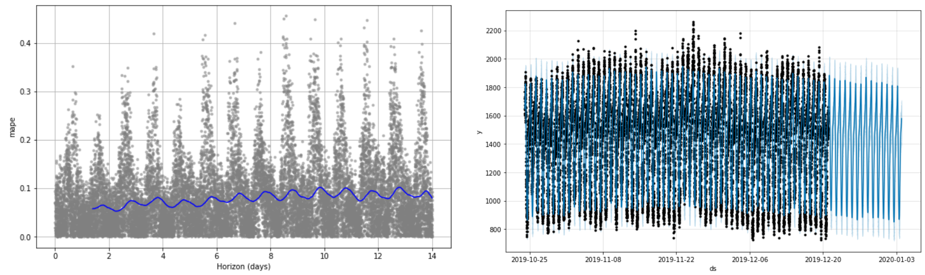
The datasets in our experiments contain data based on the information from the Seattle Internet Exchange Point (SIX), collected between 22 X 2019 and 23 XII 2019, with a sampling rate of 5 minutes. To simulate diverse traffic types in a network, we added some fluctuations into the original data. To measure how the traffic in created datasets differs from the collected aggregated Seattle measurements, we use the mean absolute percentage error (MAPE). In this paper, we consider three datasets of diverse profiles: *traffic a* (MAPE = 3.39%), *traffic b* (MAPE = 8.21%) and *traffic c* (MAPE = 13.35%). Low MAPE values imply fewer fluctuations since the traffic is the most similar to the original aggregated SIX traffic. Intuitively, high MAPE values mean more fluctuations. The provided MAPE values are averaged across all the samples in each dataset. For more details regarding the datasets creation, we refer to [4].

We explore the impact of PROPHET parameters on each traffic type, to obtain the most accurate prediction. For the changepoint prior scale, we investigate the values of 0.0005, 0.001, 0.01, 0.1, and 0.5. For the seasonality prior scale, the default value in the PROPHET is 10.0, which means almost no tuning. In our experiments, the tested values are 0.1, 1.0, and 10.0. For the number of automatically placed changepoints, the default value in PROPHET is only 25, and in our experiments, the tested values are 25, 50, and 100. Overall, for each traffic type, we examine 45 parameter configurations.

Table 1: MAPE values for three best PROPHET configurations and the reference approach for each traffic type

	changepoint prior scale	seasonality prior scale	number of changepoints	MAPE
traffic a	0.001	1.0	25	0.07777
	0.001	0.1	100	0.07780
	0.001	10.0	25	0.07797
		linear regression		0.22504
traffic b	0.001	1.0	100	0.11837
	0.001	0.1	100	0.11845
	0.001	10.0	100	0.11874
		linear regression		0.23897
traffic c	0.001	1.0	25	0.15992
	0.001	10.0	25	0.16004
	0.0005	10.0	25	0.16007
		linear regression		0.25212

In Table 1, we present the results of the three best PROPHET configurations for each traffic type together with their parameters, alongside the reference Linear Regression model. Intuitively, the most accurate forecasts were obtained for the least fluctuating *traffic a*, and the highest prediction errors were noted for the most difficult *traffic c*. Performing the experiments, we noticed that the most significant PROPHET parameter is the changepoint prior scale. Its smallest value resulted in the lowest MAPE across traffic types. In particular, for *traffic b*, the difference between the worst result for the changepoint prior scale of 0.001 and the best for 0.5 was as high as four percentage points. After setting its value too low, i.e., 0.0001, the optimization algorithm failed – the model had to fall back to the Newton algorithm, yielding worse results than its default one. The seasonality prior scale did not seem to affect the results as much. In general, lower values, i.e., 0.1 and 1.0, yielded lower errors than 10.0. The differences between them were, however, only fractions of a percent. Furthermore, the number of changepoints did not appear to have any significant impact on the results either. Its influence was the most noticeable for cases with a changepoint prior scale of 0.01, where a higher number of changepoints resulted in more accurate forecasts. Once again, the results differ by only a fraction of percent, whereas more changepoints require more computational power for calculating the traffic prediction. Thus, it may not be worth it to increase this parameter. Overall, the predictions made by the PROPHET were of significantly higher quality than the reference Linear Regression. Though this approach appears in the literature as a prominent solution for short-term traffic forecasting (e.g., [4]), the PROPHET is versatile for long-term network traffic prediction.



(a) (b)

Fig. 1: PROPHET performance for *traffic a*. (a) Rolling window MAPE for best prediction model; (b) Prediction of the best model made for a 14 days horizon

An illustration of the results can be found in Fig. 1a and 1b where we present plots for the best model that we received during the experiments for *traffic a*. Fig. 1a shows MAPE with a 10% rolling window over the results. The grey points are the errors for each predicted point in the horizon. Figure 1b shows the input traffic data (in black) along with its predictions (in blue), and a 14-day forecast.

4 Conclusion

In this paper, we investigated the topic of optical backbone network traffic prediction. The developed model based on Facebook's PROPHET proved its high performance making 14-day forecasts for diverse types of network traffic. For each traffic type, various parameter configurations were tested to find the most accurate model. We found that the most important parameter in the PROPHET model for network traffic prediction is changepoint prior scale. In future work, we plan to use traffic forecasts for the optimization of application-aware networks.

References

1. Chakraborty, P., et al.: A comparative study for time series forecasting within software 5g networks. In: 14th Int. Conf. on Signal Proc. and Comm. Systems (2020)
2. Li, Y., et al.: Prophet model and gaussian process regression based user traffic prediction in wireless networks. Sc. China Information Sciences **63**(4) (2020)
3. NOKIA: The Nokia Deepfield Network Intelligence Report 2020. Tech. rep. (2020)
4. Szostak, D., et al.: Machine learning classification and regression approaches for optical network traffic prediction. Electronics **10**(13), 1578 (2021)
5. Taylor, S.J., et al.: Forecasting at scale. The American Statistician **72**(1), 37–45 (2018)
6. Tomkos, I., et al.: Application aware multilayer control and optimization of elastic wdm switched optical networks. In: 2018 OFC Conference (2018)
7. Troia, S., et al.: Deep learning-based traffic prediction for network optimization. In: 2018 ICTON Conference. IEEE (2018)
8. Xu, M., et al.: Hybrid holiday traffic predictions in cellular networks. In: NOMS 2018 IEEE/IFIP Network Operations and Management Symposium. IEEE (2018)