[Paper Review] Network Traffic Prediction Using Long Short-Term Memory

2023-2 URP | SKKU AAI 2021311828 Park Su Yeon

INTRODUCTION

- The network traffic prediction helps in controlling and using the computer network optimally.
- To get fine results and decrease the error compared to earlier algorithms, various methodologies are used.

* Network traffic: The data packet travelling across the network via devices in the form bits

* Time series: The data collected at time intervals

LSTM

- LSTM improved results as compared to ARIMA and RNN
- A type of recurrent neural network
- Solve the problem of long-term dependencies, gradient exploding/vanishing

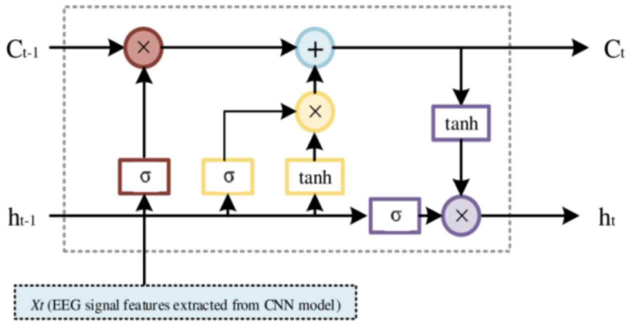


Figure 1. LSTM Framework [14][15].

$$
\begin{aligned}
\mathbf{f}^{(t)} &= \sigma\left(\mathbf{W}_f\mathbf{x}^{(t)} + \mathbf{U}_f\mathbf{h}^{(t-1)} + \mathbf{b}_f\right) && (1)\\
\mathbf{i}^{(t)} &= \sigma\left(\mathbf{W}_i\mathbf{x}^{(t)} + \mathbf{U}_i\mathbf{h}^{(t-1)} + \mathbf{b}_i\right) && (2)\\
\tilde{\mathbf{c}}^{(t)} &= \tanh\left(\mathbf{W}_c\mathbf{x}^{(t)} + \mathbf{U}_c\mathbf{h}^{(t-1)} + \mathbf{b}_c\right) && (3)\\
\mathbf{c}^{(t)} &= \mathbf{i}^{(t)} \odot \tilde{\mathbf{c}}_{(t)} + \mathbf{f}^{(t)} \odot \mathbf{c}^{(t-1)} && (4)\\
\mathbf{o}^{(t)} &= \sigma\left(\mathbf{W}_o\mathbf{x}^{(t)} + \mathbf{U}_o\mathbf{h}^{(t-1)} + \mathbf{b}_o\right) && (5)\\
\mathbf{h}^{(t)} &= \mathbf{o}^{(t)} \odot \tanh(\mathbf{c}^{(t)}) && (6)
\end{aligned}
$$

TIME SERIES CLUSTERING

- The heterogeneity of the network traffic
- Raw data-based methods, Feature-based methods, Model-based methods
- K-means (k=20) & DBSCAN clustering algorithms

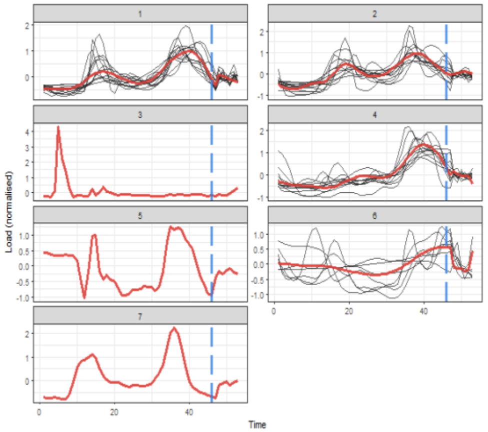
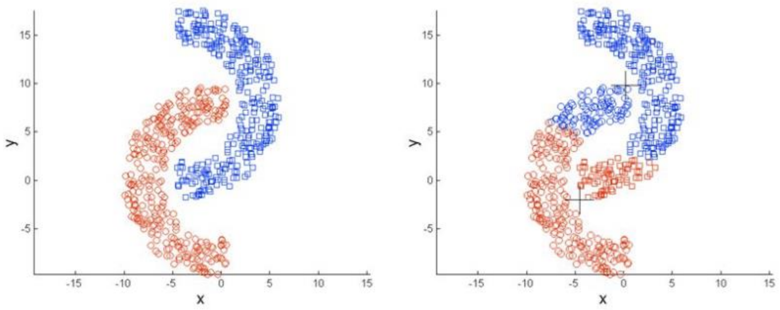


Figure 2. Time Series Modelling

Table1. Network Time Series Clustering Features

Feature Name	Description
Mean	Mean traffic size
Variance	The variance of traffic size
Min	Smallest traffic size
Max	Largest traffic size
Median	Median of traffic size
Ptp	Range of values
Skew	The skewness of time series
Epoch	Epoch size used for aggregation
Mask	Mask size for aggregation
Entropy	Sample entropy of data
Acf10	Lag-10 autocorrelation coefficient
Acf1	Lag-1 autocorrelation coefficient



DATA TRANSFORMATION

- Normalization of the time series
- Modelling the deltas instead of the actual values
- Modelling the logarithm value of time series

EXISTING WORK(1)

- Discrete wavelet transform, ARIMA, and RNN
- LST-TP(long Term Span Traffic Prediction Model): LSTM + attention
- Traffic matrix prediction based on LSTM RNN
- LSTM-TPDTNS(new LSTM based traffic predict dynamic transport network slicing framework)
- ConvLSTM(convolutional LSTM model), STL(seasonal and trend decomposition)

EXISTING WORK(2)

- Short term traffic prediction based on the LSTM neural network
- LSTM model + DNN, introduce autocorrelation coefficient with the model
- Real-time network traffic prediction based on the LSTM

EXPERIMENT ANALYSIS

- 4 variations of LSTM
  - Vanilla LSTM(vlstm): 50 unit LSTM layer + 50 unit dense layer, dropout 0.2, look back window 3, epoch 20, batch size 8, standard scaler
  - Delta LSTM(dlstm): vlstm + data preprocessing to deltas 계산
  - Cluster LSTM(clstm): 20개 LSTM 모델, clustered data로 학습, dropout 0.3, batch size 128
  - Clustered Delta LSTM(cdlstm): clstm + data preprocessing to deltas 계산

EXPERIMENT ANALYSIS

- Time series dataset
  - 6 different time series given by RJ Hyndman
  - Data recorded per 5 min / hour / day

Table2. Details of Time Series Dataset

Series Name	Total Size in MB	Time Interval
Daily-1	51	1day
Daily-2	69	1Day
Hourly-1	1231	1hr
Hourly-2	1657	1hr
5min-1	14772	5 min
5min-2	19888	5 min

- NRMSE(Normalized Root Mean Squared Error)

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{t=n}(y' - y)^2}$$

RESULT

- CDLSTM has lowest RMSE compared to other models.
- The obtained value of NRMSE for Daily 1 and Daily 2 time series are .109 and .179 respectively.

Table3. NRMSE for Various Algorithms

Time series name	ARIMA+ RNN	VLS TM	DLS TM	CLS TM	CDLS TM
Daily 1	.115	.120	.114	.112	.109
Daily 2	.191	.198	.188	.184	.179
Hour ly1	.022	.030	.021	.0214	.018
Hour ly2	.024	.032	.023	.022	.019
5Min -1	.012	.019	.011	.010	.008
5Min -2	.008	.011	.0074	.007	.006

CONCLUSION

- The proposed work performs well in comparison with the previous work like Arima and RNN.
- There is a remarkable decrement in the NRMSE of the system, which results in efficient network traffic prediction.
- The feature-based clustering used for the time series data provided a valuable role in improving the results.

THANK YOU

Q&A