Resource Allocation for Virtualized Mobile Core Networks: Databased Network Traffic Prediction for Bandwidth Allocation

Zhu Xuanyu

2016/6/25

1 Background

To research on resource allocation for virtualized mobile Core Networks, the NFV group recorded 20TB data from ChinaUnicom. With the data, a user behaviour model has been built to allocate CPU and memory resource, using first-order markov model to predict users' state. The model performs well, but it only considers number of users in different states, ignoring location information.

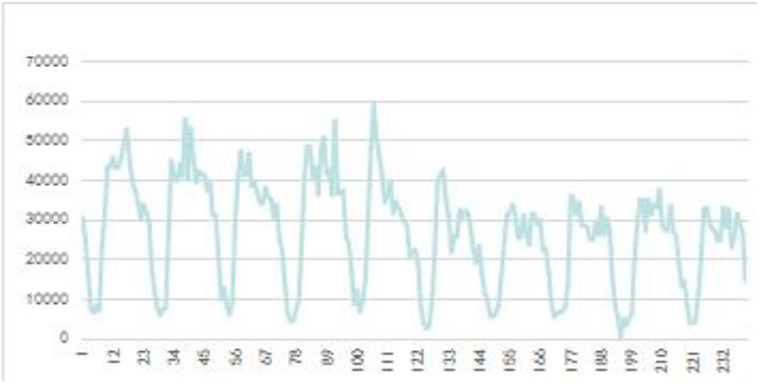
To allocate bandwidth resource, and to make full use of the data, we have to build a new model predicting the traffic of each cell. The new model will also be helpful to improve the user experience.

2 Purpose

Find a method to predict the traffic of each cell.

3 Data Structure

- Record_structure: MINUTE CELL:NUM,CELL:NUM,
- Record_example: 2016-01-11 04:02 40611:1,21862:1,20951:1,22541:1,61421:1,0:3
- About 2GB data in total, I choose some of the cells to test.

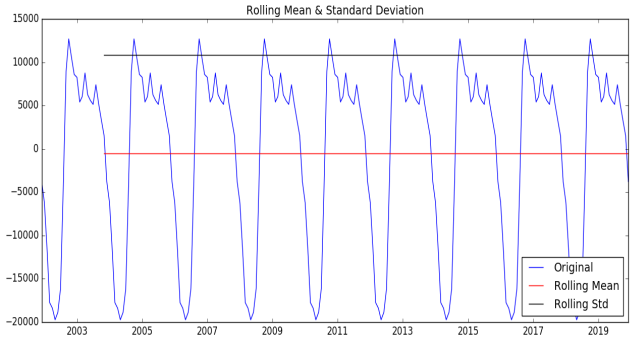


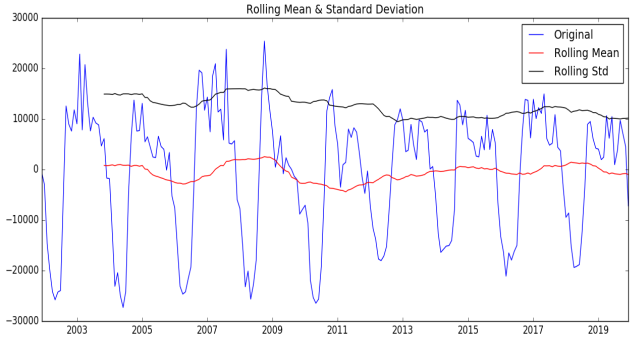
4 Contents

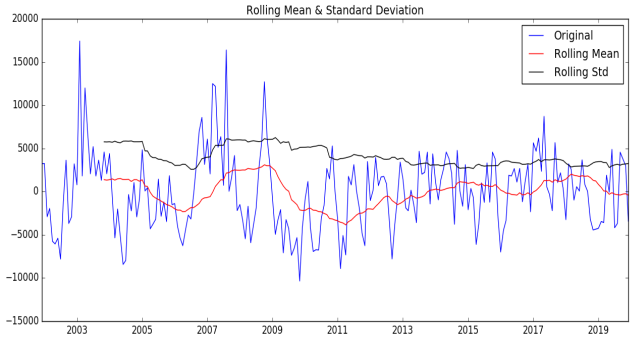
4.1 ARMA

- In the statistical analysis of time series, autoregressivemoving-average (ARMA) models provide a parsimonious description of a (weakly) stationary stochastic process in terms of two polynomials, one for the autoregression and the second for the moving average.

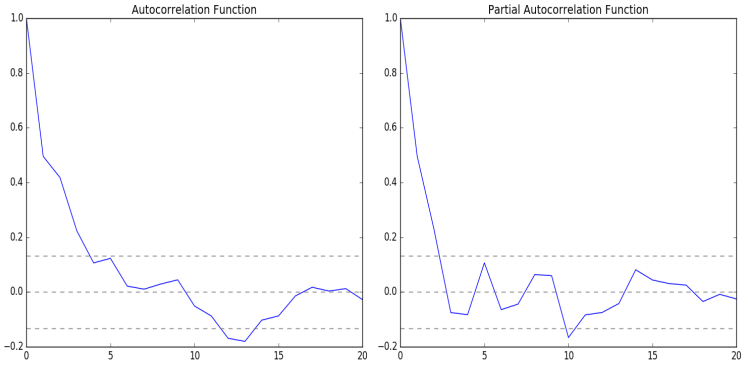
- Need to make the time process a (weakly) stationary stochastic process before using ARMA model. Difference, decompose and some other ways can make this. I choose decompose:



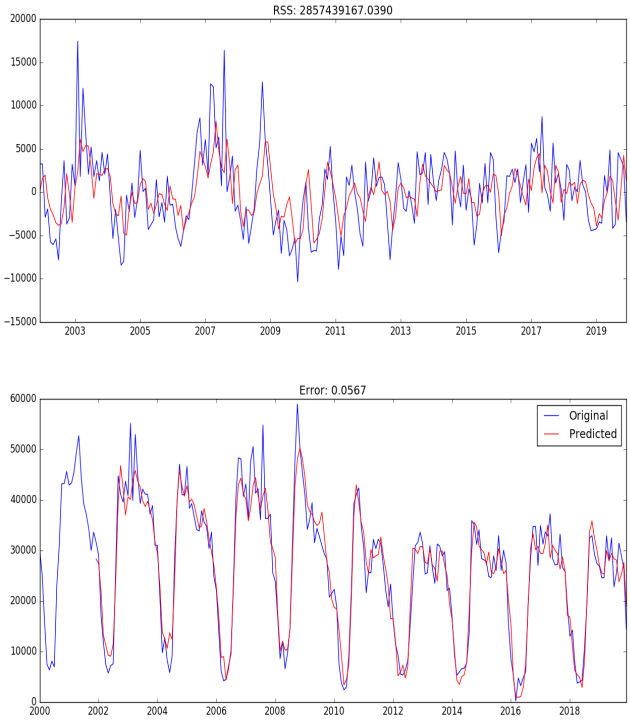




- After checking, the residual can be seen as a stationary stochastic process, so we can now figure the ACF and PACF to choose parameters of ARMA model.



- We can learn from the figure that $p = 3$ and $q = 4$ may be a good choice. With help of Pandas and Statsmodels, I can now give my prediction:



- The RMSE is acceptable, and this method is valid.

4.2 Neural networks

- Simple neural networks or RNN/LSTM methods can also used in time series prediction. A recurrent neural network (RNN) is a class of artificial

neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior.

- Training a good neural network needs a lot of time and it's hard to choose good parameters, especially when there's much data. I tried LSTM in small datasets with PyBrain, but result is bad.

5 Results

- ARMA model is appropriate for the problem, partly because the time series have obvious periodicity.
- RNN/LSTM have the potential to perform well, but hard to adjust parameters and training takes much time.
- I have not tried SVR and other ways, in the future I can try them.

6 Reflections

In this term's project, I have learnt a lot. By adding a group and discuss about ways to improve the model, I learnt to work with others, and by solving the problem, I learnt to survey for solution, and learnt about different kind of prediction methods. The most important thing I learnt is that, in solving the problems, I am now familiar about research, and confident to perform well in future work and study.