Predicting Internet Path Dynamics and Performance with Machine Leanring

Zhenghui Wang
Shanghai Jiao Tong University
Shanghai, China
felixwzh@outlook.com

Hao Wang
Shanghai Jiao Tong University
Shanghai, China
?@?.com

Yuheng Zhi
Shanghai Jiao Tong University
Shanghai, China
?@?.com

Shukai Liu
Shanghai Jiao Tong University
Shanghai, China
?@?.com

ABSTRACT

We study the problem of predicting internet path dynamics and performance. We use `traceroute` measurement and machine learning models.

KEYWORDS

TODO

ACM Reference Format:
Zhenghui Wang, Hao Wang, Yuheng Zhi, and Shukai Liu. 2017. Predicting Internet Path Dynamics and Performance with Machine Leanring. In *Proceedings of SJTU Computer Network Workshop (CNW).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[[Zhenghui says "I think we should focus on the process not the final result, which is also important."]]

1 INTRODUCTION

1. introduce the original paper: their task, method, and dataset

2. we find some drawbacks in their data. For instance, the route change times problem discussed in github. And the way they process the data.

3. we try some other models to get better performance like xgboost and LSTM.

2 RELATED WORKS

1. we introduce the original paper in details

2. we can introduce some other machine learning methods applied in computer network scenarios. **[[Zhenghui says "I'll take this part"]]**

3. very briefly introduce xgboost and lstm

CNW, December 2017, Shanghai, China

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

3 DATA ANALYSIS

[[Zhenghui says "I think this is a important part, who will take this part?"]]

1. We can plot some figures of the statistics of data, like the distribution of the route duration and avgRTT.

2. We can further discuss the relation between routes in one path or in different paths.

3. Then we could discuss why the authors of the paper process the data in a wrong way

4. We show our solution for data process. three ways for random forest models.

4 EXPERIMENT

4.1 Classic Models

We show the experiment results of the 3 different data we obtained, namely **K&fix**,**K&update**,**timeslot&update**. We need to find some difference between our 3 data processing methods and the authors', i.e.,**origin**.

The experiments we need to conduct are as follows:

- 1. **K&fix**+RF
- 2. **K&update**+RF
- 3. **timeslot&update**+RF
- 4. **origin**+RF
- 5. **K&fix**+xgBoost
- 6. **K&update**+xgBoost
- 7. **timeslot&update**+xgBoost
- 8. **origin**+xgBoost

[[Zhenghui says "Note"]]We first predict the route duration instead of resLife. Further study on the difference between these two predicting objective could be conduct if we have time. Currently, we predict the route duration for task 1, because we can only predict this when we use LSTM.

4.2 Deep Models

There are two kinds of input for the LSTM at each timestep, (i) one simple scalar (ii) a vector. The experiments we need to conduct are as follows:

- scalar input + LSTM
- vector input + LSTM

5 CONCLUSIONS

TODO

ACKNOWLEDGMENTS

TODO