무선 네트워크에서 시퀀스-투-시퀀스 기반 모바일 궤적 예측 모델

Sammy Yap Xiang Bang *, 양희규*, Syed M. Raza**, 추현승**
*성균관대학교 수퍼인텔리전스학과
**성균관대학교 전자전기컴퓨터공학과
sammyyap98@skku.edu, huigyu@skku.edu, s.moh.raza@skku.edu, choo@skku.edu

Sequence-to-Sequence based Mobile Trajectory Prediction Model in Wireless Network

Sammy Yap Xiang Bang*, Huigyu Yang*, Syed M. Raza**, Hyunseung Choo**
* Dept. of Superintelligence, Sungkyunkwan University
**Department of Electrical and Computer Engineering, Sungkyunkwan University

요 약

In 5G network environment, proactive mobility management is essential as 5G mobile networks provide new services with ultra-low latency through dense deployment of small cells. The importance of a system that actively controls device handover is emerging and it is essential to predict mobile trajectory during handover. Sequence-to-sequence model is a kind of deep learning model where it converts sequences from one domain to sequences in another domain, and mainly used in natural language processing. In this paper, we developed a system for predicting mobile trajectory in a wireless network environment using sequence-to-sequence model. Handover speed can be increased by utilize our sequence-to-sequence model in actual mobile network environment.

1. Introduction

Mobile network traffic is continuously increasing with the increase of smart phone user number worldwide. 5G wireless technology has been developed to provide higher speed, lower latency and greater capacity services compare to 4G LTE networks. The 5G network is being commercialized in mobile devices and the demand for ultra-low latency of 5G is increasing especially in high-performance services such as autonomous driving and 8K video streaming service[1]. Proactive mobility management is essential for 5G mobile networks as 5G mobile networks provide new services with ultra-low latency through dense deployment of small cells.

One of the crucial aspects in mobility management is handover management. Handover management covers aspects of users' access point or base station switching. Mobile trajectory handover prediction is one of the most widely used approaches in handover management for cellular networks since it allows proactive radio resource allocation. Accurate handover prediction can significantly reduce handover latency, signaling overhead, and call drop rate[2].

Deep Neural Networks (DNNs) are powerful models that have achieved excellent performance on difficult learning tasks in various fields. However, DNNs work well whenever large labeled training sets are available, they cannot be used to map sequences to sequences. A general end-to-end approach to sequence learning, sequence-to-sequence model was introduced by Google in 2014[3]. This model has the power to map sequences of different lengths to each other even if the inputs and outputs are not correlated and their lengths can differ.

In this paper, we propose a sequence-to-sequence model designed for mobile trajectory prediction. We are able to achieve 87.6% accuracy for 7 steps mobile trajectory prediction. The structure of this paper is organized as follows: In the section 2, we present the related work. Next, we present the training data and our proposed model in section 3 and section 4. Then, we present the evaluation of our model in section 5. Finally, we summarize our work and outline the future research direction.

2. Related Work

A. Long Short-Term Memory

Recurrent Neural Networks is the neural network structure for supervised learning that captures the features of continuous data for prediction. Additional hidden layers between the input and output layers learn properties of sequential data recurrently to predict the next step of sequence. However, it has a long-term dependency problem where the past information exponentially decays as the sequence length becomes long. Long Short-Term Memory (LSTM) solves this problem by introducing cell state which saves more information of past sequences [4]. LSTM is made up of a memory cell, an input gate, an output gate and a forget gate. The memory cell is responsible for remembering the previous state while the gates are responsible for controlling the amount of memory to be exposed. LSTM based prediction models have shown good performance for long sequential data and that is why it has been widely used in different applications such as speech recognition.

B. GRU model

Complex structure of LSTM cell makes it computationally expensive. GRU simplifies the LSTM cell by combining cell state and hidden state, and uses two gates (reset, update) instead of three [5]. Similar to LSTM, GRU uses additional gates to control the memory of past sequences. However, GRU has fewer parameters to calculate comparing to LSTM and this makes GRU simpler and computationally less expensive. Moreover, GRU shows even better performance than LSTM for less complex datasets. This enables GRU to be used for applications with less complex data like music composition analysis.

3. Data

The data used in this paper was collected in a wireless network environment with a total of 12 Access Points (APs) at Sungkyunkwan University Pangyo Campus. The network controller saves a handover log based on the user mobility in the campus. The acquired log files consist of mobile device MAC (Media Access Control) addresses and the IDs of the handover source and destination APs. Sequences of mobile devices movement are created using the handover log. For example, if handover occurs for a device with MAC address of 0111.0222.0333 from AP number 1 to 5, 5 to 6, 6 to 3, 3 to 8, mobile trajectory sequence of {1, 5, 6, 3, 8} is created. To ensure each of the 12 APs are treated equally without any bias, we further preprocess our mobile trajectory sequence with one-hot encoding. The index in the vector equivalent to the ID of the AP has the value one, and all other values are zero. For example, AP number 5 will be represented as {0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0} where fifth value is one and remaining 11 values are zero. In this paper, we work on 7 steps mobile trajectory prediction with 22 previous APs inputs. Total of 2575 mobile trajectory sequences from 289 users are divided into 80% for training data and 20% for testing data.

4. Mobile Trajectory Prediction Model

In this section, we present our sequence-to-sequence based mobile trajectory prediction model. Our model predicts mobile trajectory sequence of length q (7 APs) from historical trajectory sequence of length p (22 APs) through encoder network and decoder network as shown in Figure 1. The Encoder network processes the input sequence through a fully connected neural network and single sublayers of LSTM cells to create Encoder state that contains compressed representation of all the elements. The Encoder state is passed to the Decoder network that simultaneously predicts q elements of target sequence by using two FC neural networks and single sublayers of LSTM cells.

The target sequence prediction of length q from the Decoder network improves with the teacher forcing method in training. Teacher forcing is a strategy for training sequence-to-sequence model that uses ground truth as input, instead of model output from a prior time step as an input. Without teacher forcing, trajectory prediction may run off target easily while trajectory prediction is corrected at each step with teacher forcing.

In our model, single sublayers of LSTM cells are replaced with single sublayers of GRU cells in Encoder and Decoder networks. Performance comparison for usage of LSTM and GRU cells will be present in Section 5.

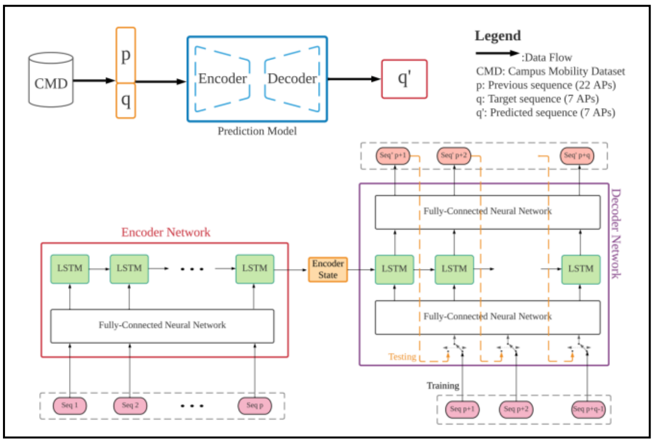

Figure 1. Mobile Trajectory Prediction Model

5. Evaluation

Each result presented in this paper is an average of ten experiments which are conducted on a system with Intel Core i7-11700KF CPU, NVIDIA GeForce RTX 3070 Ti graphic cards, and 32GB RAM. We evaluated our model using LSTM or GRU cells in term of prediction accuracy where prediction is considered correct only if all 7 individual predicted steps are correct. A wrong prediction at any individual step of the predicted sequence is unacceptable for proactive mobility management as it is handled on per hop basis.

Prediction accuracy results of seven steps mobile trajectory predictions for sequence-to-sequence model using LSTM cells and GRU cells are showed in Figure 2 as a function of training epochs. Prediction with model using LSTM cells shows highest accuracy of 87.96% while model using GRU cells shows highest accuracy of 86.99%.

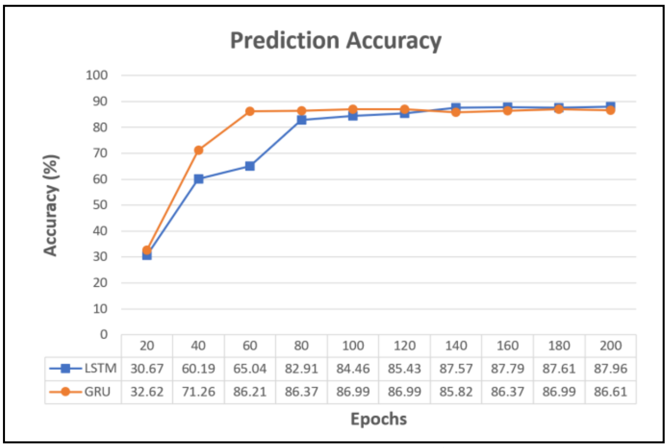

Figure 2. Accuracy for Prediction Model

Sequence-to-sequence model using LSTM cells outperform sequence-to-sequence model using GRU cells by around one percent. The results in Figure 2 show a direct correlation between accuracy increment and increasing epoch number. It is worth noting that model using LSTM cells take around 140 epochs to achieve 87% accuracy while model using GRU cells take only 60 epochs to achieve 86% accuracy. This makes sense as GRU has simpler structure and less parameters compare to LSTM. Furthermore, model using LSTM cells takes around 135 milliseconds to complete one epoch while model using GRU cells takes only 105 milliseconds which is 23% less than model using LSTM cells.

The results in Figure 2 establishes that model using LSTM cells yields slight better overall prediction accuracy, whereas model using LSTM cells trains faster and can achieve high accuracy. In case of real live wireless networks for mobility or resource management purposes, sequence-to-sequence model using GRU cells is preferable as it achieve high accuracy in significant less time.

6. Conclusion

In this paper, we present a sequence-to-sequence based mobile trajectory prediction model in wireless network. We also compared the performance of the model using LSTM cells and GRU cells. We are able to achieve accuracy of 87% for 7 steps mobile trajectory prediction. Mobile trajectory prediction enables mobile operators to not only proactively manage the user mobility but also assign, reassign, or scale the network, and compute resources in MECs proactively to enhance user experience. For future research, we will investigate other advanced predictive models to achieve higher prediction accuracy for our work.

ACKNOWLEDGEMENT

This work was partly supported by the MSIT (Ministry of Science and ICT), Korea, under the ICT Creative Consilience program (IITP-2022-2020-0-01821) and Grand Information Technology Research Center support program (IITP-2022-2015-0-00742) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation). This work was partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2020R1A2C2008447). This work was partly supported by the High-Potential Individuals Global Training Program (IITP-2021-0-02132).

References

[1] Dangi, R., Lalwani, P., Choudhary, G., You, I., & Pau, G. (2022). Study and Investigation on 5G Technology: A Systematic Review. Sensors (Basel, Switzerland), 22.

[2] Öztürk, M., Gogate, M., Onireti, O., Adeel, A., Hussain, A., & Imran, M.A. (2019). A novel deep learning driven, low-cost mobility prediction approach for 5G cellular networks: The case of the Control/Data Separation Architecture (CDSA). Neurocomputing, 358, 479-489.

[3] Sutskever, I., Vinyals, O., & Le, Q.V. (2014). Sequence to Sequence Learning with Neural Networks. NIPS.

[4] Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. Neural computation. 9. 1735-80. 10.1162/neco.1997.9.8.1735.

[5] Cho, K., Merrienboer, B.V., Gülçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. EMNLP.