Reinforcement-supported Artificial Neural Network-based Trajectory Prediction

Topic: DS algorithms with a view towards Machine Learning and Artifical Intelligence

Negar Emami, Torsten Braun
Communications and Distributed System Group,
Institute of Computer Science,University of Bern,
Neubrückstrasse 10, Bern, Switzerland
Email: {*negar.emami, torsten.braun*}*@inf.unibe.ch*

Abstract. With the advent of 5G, Beyond 5G, and 6G cellular networks, mobility prediction has become a crucial task to enable a wide range of services such as handover and resource management. Mobility predictors analyze users' historical mobility data, extract meaningful information, and learn moving patterns to forecast future locations and trajectories of users' end systems. A trajectory is defined as the sequence of GPS coordinates or cellular antennas' IDs over time.

We present a trajectory predictor based on Long Short Term Memory (LSTM), a variant of Recurrent Neural Networks (RNNs), to improve the performance of network applications, namely handover management and service migration for the Context Awareness Engine project with Orange SA [1]. When a mobile user moves within a mobile cellular network, transferring the cellular connection from one base station antenna to another is called handover, which requires many signaling operations. The knowledge of the user's future location can offer a proactive handover mechanism so that the signaling procedure can be done before the mobile user's arrival at a target antenna. In this way, the total system delay can remarkably decrease. Modern network services aim for reliable communication, low latency, and high throughput, especially in the case of user mobility, where services might easier experience interruption and quality degradation. Distributed technologies such as Multi-access Edge Computing (MEC) and Service Migration are introduced to offload requested data from the cloud to edge servers closer to users. In this context, trajectory prediction enables designing efficient service migration that guarantees service continuity and Quality of Service (QoS) in a proactive manner. Hence, attaining a reliable, accurate, and optimal trajectory predictor is a pivotal task.

Within various machine learning (ML) and artificial neural network (ANN) approaches, LSTM models have achieved a remarkable performance in the trajectory prediction task. However, most state-of-the-art works choose the neural networks heuristically and apply the same model to every user type, which does not guarantee optimal performance. We use Reinforcement Learning (RL) as a self-learning approach that is able to automate the LSTM architecture search process and hyperparameter optimization to explore the best neural architecture for each mobile user individually. To accelerate RL's optimization process, we apply a Transfer Learning (TL) method. TL is a technique for reusing one task's developed model as the starting point for another task in order not to initialize the second task from scratch. This way, we transfer the knowledge from a pre-trained RL-suggested neural network to a newly RL-suggested neural network. Although LSTMs have achieved excellent prediction accuracy, high training time and computational power still stay as their main bottlenecks. LSTMs are extremely slow by nature due to the fact that they learn spatio-temporal dependencies in sequential order. As an extension for our work [1], we propose the combination of RL and Convolutional Neural Networks (CNNs) as the optimal solution to save more computational resources. CNNs support parallelism and are computationally faster. One-dimensional CNNs concentrate on sequential data from a global perspective and extract features by applying efficient convolutional operations.

To evaluate the proposed predictors' performance, we have done our computations on the High-Performance Computing Cluster of the University of Bern known as HPC Cluster - UBELIX, which supports parallel execution of multiple user predictions. Our RL-LSTM achieves on average 69.7% accuracy, which is almost 10% better than other ML approaches. The suggested RL-LSTM predictor consumes 176 minutes to explore-exploit the best neural architecture and to train data, which is only 28% of the time that a Grid Search-based LSTM requires (625 minutes) for performing the same task. On the other hand, our RL-CNN predictor achieves 67.7% accuracy, which is slightly less than RL-LSTM's accuracy, while its optimization and training time is only 55 minutes, 31% of the time RL-LSTM already had spent. Moreover, in a real distributed scenario, applying thousands of individual RL-CNNs is impossible due to limited computational resources. Therefore, we detect similar trajectory users, build an exclusive RL-CNN per cluster based on a few users' data, and transfer the pre-trained neural network knowledge between group members. With this approach, we can save up to 90% of computational resources while losing a few couples of percentages of the average accuracy.

References

[1] Z. Zhao, N. Emami, H. Santos, L. Pacheco, M. Karimzadeh, T. Braun, A. Braud, B. Radier, and P. Tamagnan, "Reinforced-lstm trajectory prediction-driven dynamic service migration: A case study," *https://boris.unibe.ch/id/eprint/155232*, p. 15, 15-Dec-2020.