In partial fulfillment of the terms for obtaining the PhD degree, Sean Bin Yang will give a lecture on the following subject:

**Path Representation Learning in Road Networks**

*on Wednesday 21th of December 2022, 9:00, 100% online*

**Abstract:**
With the continued digitization of transportation systems, there is an increasing demand for path-based smart city applications which require appropriate path representations. To support path-based applications, different techniques have been proposed to learn path representations. However, existing studies suffer from the following limitations: (1) Supervised methods learn a task-specific path representation and require a large amount of labeled training data, which works well on the labeled task, but generalizes poorly on other tasks; (2) Although graph representation learning based methods learn a task-unspecific path representation, they cannot capture sequential dependencies and fail to introduce the temporal information into the learned path representations; (3) Existing works mainly focus on accuracy improvement, ignoring the mode scalability and size, which plays a critical role in resource-constrained environments.

In this thesis, we investigate the task-unspecific path representation learning approaches that are able to generalize well for different downstream tasks. More specifically, we focus on the following specific works. (1) Context-aware path ranking in road networks. (2) Unsupervised path representation learning. (3) Weakly-supervised contrastive curriculum path representation learning. (4) Lightweight and scalable path representation learning.

First, we study the path ranking framework PathRank. This framework learns task-specific path representations, which are used to rank candidate paths. In particular, we propose a training data enrichment strategy to enhance the learning process. Subsequently, we propose an end-to-end context-aware multi-task framework to enable the PathRank. We conduct extensive experiments on one real-world dataset to verify the effectiveness of the PathRank.

Second, we study the unsupervised path representation learning framework Path InfoMax (PIM) by maximizing the mutual information. PIM takes as input a path and output task-unspecific path representations. In particular, we first propose a curriculum negative sampling strategy to enhance the PIM training. Then, we propose a path-path discriminator and path-node discriminator to jointly learn task-unspecific path representations by capturing the global and local information of the path. Finally, we conduct extensive experiments on two downstream tasks under two real-world datasets. The results illustrate our PIM is more effective than other baseline methods. In addition, the pre-trained PIM can enhance supervised learning methods.

Third, we study a weakly-supervised contrastive curriculum temporal path representation learning framework by leveraging the information from weak labels and considering both spatial and temporal correlations. This framework takes as input a temporal path (a path with a departure time) and returns task-unspecific temporal path representations. We first introduce the weak label to capture the temporal variation of traffic. Then, we study the weakly-supervised contrastive learning method to enable the temporal path encoder training. Subsequently, we combine weakly-supervised contrastive learning with a two-stage curriculum learning strategy to improve the performance of weakly-supervised contrastive learning. Finally, we conduct extensive experiments on three downstream tasks under three real-world datasets. The results show the effectiveness of our proposals.

Finally, we study the lightweight and scalable path representation framework LightPath. This framework aims at learning task-unspecific path representations by reducing resource consumption and enabling model scalability with respect to path length. In particular, we first propose a sparse auto-encoder that guarantees LightPath with good scalability of path length. Then, we propose cross-network and cross-view relational reasoning to train sparse path encoders jointly. Subsequently, we propose global-local knowledge distillation to reduce the model size and improve the performance of the learned path representations.
representations. Finally, extensive experiments are conducted, and the results demonstrate the efficiency and scalability of the LightPath.

Members of the assessment committee are Associate Professor Álvaro Torralba, Aalborg University, Professor Gao Cong, Nanyang Technological University, and Professor Baihua Zheng, Singapore Management University. Professor Bin Yang is Sean Bin Yang’s supervisor. The moderator Associate Professor Álvaro Torralba.

All interested parties are welcome.