Gold digging from smart card data in understanding users' travel behaviors

Chih-Chieh Hung

Department of Management Information Systems, National Chung Hsing University

Abstract

Understanding users' travel behaviors from smart card transactions Smart card is an emerging research topic that have attracted the attention of transport researchers and planners during the last two decades. Although known as a rich and continuous source of public transit data, smart card transactions reveal limited information about users' movement behavior so that it is challenging to mine knowledge from those transactions, just like gold digging. In this talk, we are going to share our "gold digging" experience on 1) inferring commuters' route preference in a metro network and 2) predicting supply and demand in a docked bike system.

In the first topic, to recover the exact routes taken by commuters inside a metro system, we strategically propose two inference tasks to handle the recovering, one to infer the travel time of each travel link that contributes to the total duration of any trip inside a metro network and the other to infer the route preferences based on historical trip records and the travel time of each travel link inferred in the previous inference task. As these two inference tasks have interrelationship, most of existing works perform these two tasks simultaneously. However, our solution TripDecoder adopts a totally different approach. TripDecoder fully utilizes the fact that there are some trips inside a metro system with only one practical route available. It strategically decouples these two inference tasks by only taking those trip records with only one practical route as the input for the first inference task of travel time and feeding the inferred travel time to the second inference task as an additional input, which not only improves the accuracy but also effectively reduces the complexity of both inference tasks.

In the second topic, to predict the demand and supply of bikes at any station, we propose STGNN, a novel data-driven Spatial-Temporal Graph-Neural Network to solve the bike demand and supply prediction problem by unifiedly embedding the dynamic ST information in two novel ST graphs. Given station locations and historical rental data on bike flow over the past time slots 0 to $t-1$, we seek to predict online the bike demand and supply at any station at time t. To extract joint spatial-temporal dependency,

STGNN employs a generator to construct, at the beginning of time t, two graphs which embed the flow relationships between stations at various time slots(flow-convoluted graph) and dynamic demand-supply pattern correlation between stations (pattern correlation graph), respectively. Given the two spatial-temporal graphs, STGNN subsequently employs a graph neural network with novel flow-based and attention-based aggregators to generate embedding of each node for docked bike prediction.