

## Modeling Activity Pattern Using XGBoost for Mining Smart Card Data

**Authors :** Eui-Jin Kim, Hasik Lee, Su-Jin Park, Dong-Kyu Kim

**Abstract :** Smart-card data are expected to provide information on activity pattern as an alternative to conventional person trip surveys. The focus of this study is to propose a method for training the person trip surveys to supplement the smart-card data that does not contain the purpose of each trip. We selected only available features from smart card data such as spatiotemporal information on the trip and geographic information system (GIS) data near the stations to train the survey data. XGboost, which is state-of-the-art tree-based ensemble classifier, was used to train data from multiple sources. This classifier uses a more regularized model formalization to control the over-fitting and show very fast execution time with well-performance. The validation results showed that proposed method efficiently estimated the trip purpose. GIS data of station and duration of stay at the destination were significant features in modeling trip purpose.

**Keywords :** activity pattern, data fusion, smart-card, XGboost

**Conference Title :** ICTEP 2018 : International Conference on Transportation Engineering and Planning

**Conference Location :** Bangkok, Thailand

**Conference Dates :** January 18-19, 2018