Spatial Movement Pattern Analysis in Public Transportation Networks in Seoul

Kyuhyup Oh
Kyung Hee Univ.
Youngin, S.Korea
k8383@khu.ac.kr

Kwanho Kim
Incheon Nat. Univ.
Incheon, S.Korea
khokim@inu.ac.kr

Aekyung Kim
Kyung Hee Univ.
Youngin, S.Korea
akim1007@khu.ac.kr

Youngkyu Lee
Seoul Metro.
Rapid Transit Corp.
Seoul, S.Korea
magpie@smrt.co.kr

Jae-Yoon Jung
Kyung Hee Univ.
Young-in, S.Korea
jyjung@khu.ac.kr

Abstract

Recently, many studies have tried to find travel patterns to understand citizen's movement behaviours and improve transportation services based on the big data in public transportation systems. This research introduces a method of discovering and evaluating spatial movement patterns from smart card transaction data of multi-modal transportation networks such as subway and bus. The transaction data contains the origin and destination subway stations (or bus stops), boarding and alighting times, traveller's type, and so on. To abstract the location data of origin and destination, the GeoHash coding system is adopted, in this paper, which is a hierarchical geo-coding scheme based on rectangle grid. The proposed method discovers the spatial movement patterns, which express frequent movement behaviours between two zones on the map. The zone is a set of adjacent grids on the map which have similar movement characteristics. The extracted movement patterns between zones are evaluated by three measures from different viewpoints. Smart card transaction data of subway and bus networks in Seoul was used to illustrate the spatial movement pattern analysis and the result was visualized on the Google Map. It is expected that this kind of movement pattern discovery can be exploited for wide applications such as transportation network planning and location-based target marketing.

Keywords: Public transportation, Smart card data, Movement pattern analysis

1 Introduction

It is a significant work to discovery geographically adjacent zones having a large number of people's movements for urban development and public transportation planning [5]. A zone is regarded as a set of the regions which have similar function [2]. In this paper, the zone is extracted to discover the movement patterns in transportation networks.

Many previous studies suggest analytic methods that discover zones and analyse the relationship between the zones using smart cards transaction data [4]. In this paper, a method for analyzing spatial movement patterns in multiple transportation networks is proposed by extending the previous methods [3].

2 Spatial movement pattern analysis

This paper proposes the approach to analyse the spatial movement patterns in the regions with multi-modal transportation means such as subway and bus by extending the previous method that discovers the movement patterns in a subway network.

The proposed method identifies the movement patterns by combining regions, and determining adjacent zones. In particular, the result of the proposed method is movement patterns on the two-dimensional map, which is different from that of the previous study [3]. The extracted movement patterns are evaluated by three metrics such as support, lift, and cosine metrics, which aim to assess the extracted movement patterns in terms of the number of covered instances, the correlation between sets of features, and the similarity between sets of features, respectively.

2.1 Spatial movement pattern algorithm

The proposed method of extracting and combining of the spatial movement patterns applies the regional adjacency proximity to the previous method in step of merging a pair of movement patterns. Initially, each movement is considered as an individual movement pattern, and then the patterns are merged based on their adjacent proximity for a given origin and destination. Specifically, a pair of movement patterns which yield the most frequent movements on the average is selected to be merged into a single movement pattern iteratively until no remaining movement pattern to merge. Finally, each discovered pattern expresses a movement group from a merged origin zone to a merged destination zone.

The discovered movement patterns is then assessed in terms of coverage, accuracy and support [3] to evaluate their effectiveness from different viewpoints.

2.2 Geo-coding method

In order to split and combine the partial regions, we adopt the GeoHash system based on the lattice structure [1]. Depending on station's unique number, the lines of latitude-longitude of stations are converted GeoHash of 12 digits and the size of lattice could be converted and used according to the analysis objectives. For instance, 8 buckets which are located at top, bottom, right, left and diagonal sides of GeoHash lattice are the candidates for merging of adjacent regions.

3 Expermiment

We collected a smart card payment dataset (2,587,618 trips) of Seoul public transportation network (subways and buses) for a weekend, from AM 04:00 17 March, 2012 to AM 01:00 19 March, 2012 in Seoul, Korea. The dataset is used to discover and analyze the spatial movement patterns. In this paper, 6-digit GeoHash code was used to map subway stations and bus stops to the grids in the city. As a result, each grid has 1.2km width and 609.4m height. Total 740 grids were considered as initial zones for movement patterns of the algorithm. Table 1 shows a sample of aggregated OD data based on 6-digit GeoHash code.

The result of the movement pattern discovery is visualized web program on a google map of Seoul, as shown in Fig.1. In this figure, the red rectangles and blue rectangles depict the origins GeoHash grid and the destinations GeoHash grid respectively and the ranks of movement patterns by support value are designated by the numbers.
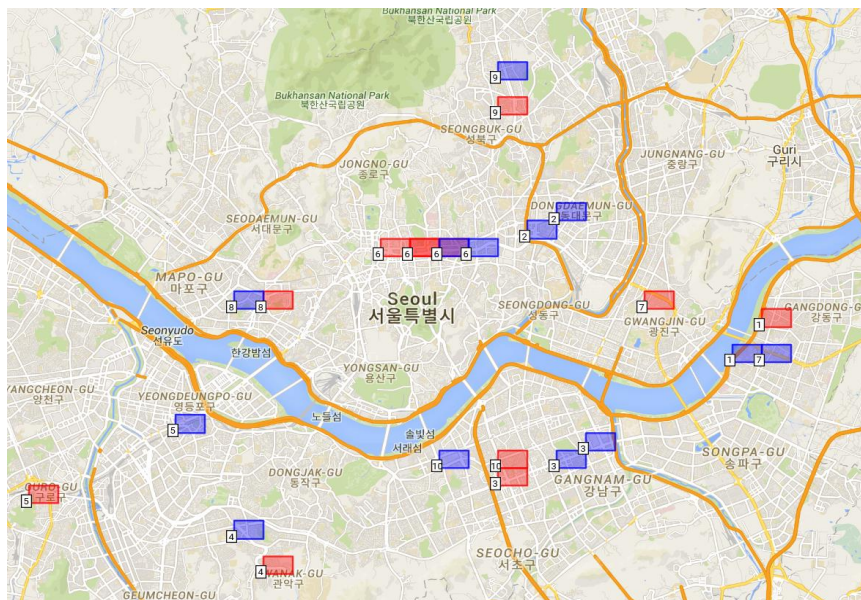
Table 1: Aggregation of OD data based on GeoHash code

Day	Hour	Origin	Destination	Freq.
17	7	Wydqh	wydmg	48
17	7	Wydqh	wydmu	2
17	8	Wydhz	wydjn	6
17	8	Wydhz	wydjp	364

4 Conclusion

In this research, we suggest an extended approach for analyzing partial movement patterns based on the existing approach for analyzing the movement pattern. We adopted the proposed approach in a real-world dataset obtained from smart card transaction data in a subway and a bus network in Seoul, Korea. By applying the proposed approach, first, the movement patterns of passengers are extracted, and the stations are mapping with GeoHash, finally the partial movement patterns are discovered.

Since our approach is based on partial position, a variety of transportation can be applied for partial movement pattern analysis. Moreover, it is believed that the proposed method can also be applied to analysis for correlations between passengers using the regional transportations and characterize the regional transportations.

References

[1] F. Anthony, E. Chris, H. James and L. Skylar, Spatio-temporal Indexing in Non-relational Distributed Databases, In *Proc. IEEE Int. Conf. on Big Data*, 2013

[2] G. Fusco, and M. Caglioni, Hierarchical Clustering through spatial interaction data. The case of commuting flows in south-eastern France, *Lecture Notes in Computer Science*, 6782:135-151, 2011.

[3] K.H. Kim, K.H. Oh, Y.K. Lee, S.H. Kim, and J.Y. Jung, An Analysis on Movement Patterns between Zones Using Smart Card Data in Subway Networks, *International Journal of Geographical Information Science*, 28(9):1781-1801, 2014

[4] J.S. Park, K.S. Lee, Mining Trip Patterns in the Large Trip-Transaction Database and Analysis of Travel Behavior, *Journal of the Economic Geographical Society of Korea*, 10(2):44-63, 2006.

[5] J. Yuan, Y. Zheng, and X. Xie, Discovering regions of different functions in a city using human mobility and POIs, In *Proceedings of the18th ACM SIGKDD International Conference on Discovery and Data Mining*, 12:186-194, 2013.

Figure 1: Spatial movement patterns between zones discovered from subway and bus networks in Seoul.