Journal of the Korean Geographical Society, Vol. 44, No. 2, 2009(176-186)

Daily Travel Pattern using Public Transport Mode in Seoul: An Analysis of a Multi-Dimensional Motif Search

Chang-Hyeon Joh*

핵심정보배열 추출에 의한 서울시 대중교통 통행패턴 분석

조창현*

Abstract : Transportation policy to facilitate the public mode use is of the foremost importance to the local governments of Metropolitan Seoul, regarding the economic and environmental consequences of the increasing use of car. Understanding the travel behaviour is essential to the establishment of proper policy to guide more people to the use of public modes instead of private. The paper reports a result of sequential analysis of individual travel behaviour in Metropolitan Seoul, using a multi-dimensional motif search technique applied to Smart Card data that integrates individuals' different public mode uses. Groups of travel patterns with similar sequential information identified distinctive travel behaviour between Seoul north and south and between metro and bus uses. Travel patterns are more bounded within north Seoul and south Seoul respectively than crossing Han River between north and south. Within north and south, travel patterns visiting northern CBD and southern CBD, respectively, as well as their local neighbour in north and south, often use metro and metro-local bus combination, while travel patterns visiting only the north and south locals without CBDs more use only the local bus line and even only the areal bus line.
Key Words : travel behaviour, public mode, CBD, sequential information, multi-dimensional motif search

요약 : 경제적, 환경적 이유에서, 승용차 이용을 억제하고 대중교통수단 분담율을 제고하는 교통정책은 수도권 지자체에게 있어 매우 중요한 과제이다. 통행행태를 근본적으로 이해하는 것은 이러한 정책의 수립과 집행에 매우 중요한 필요조건이다. 소규모 샘플자료를 이용, 특정한 통행 행태와 몇몇의 사회경제적 혹은 지리공간적 변수들 간의 상관관계에 대한 횡단면적 분석을 행하는 것이 이 분야 전통적인 연구주제였다. 연구결과는 스냅샷과 유사한, 시공간적으로 제한된 정보만을 제공한다. 그러나, 통행은 공간적으로 상이한 장소에서 벌어지는 일상활동 참여로부터 파생되며, 일상 활동계획 상에서의 활동–통행 간 순서 관계는 통행 행태의 중요한 틀이다. 본 연구는 다차원 정보배열비교법과 핵심정보배열추출기법을 이용, 서울시민의 일일 대중교통 이용 자료인 스마트 카드 자료를 분석하여 서울시내 대중교통 이용 행태의 일반 특성을 탐구한다. 분석 결과 서울 강남/북 간, 버스-전철 연계통행과 간선버스 통행 간 통행 행태의 중요한 차이를 확인할 수 있었다. 즉 통행패턴은 보통 강북과 강남 권역 안에서 형성이 되어 있어, 자연 장애물인 한강을 건너 강남/북을 가로지르는 통행패턴은 상대적으로 적었다. 또한 강북과 강남 각각의 권역 안에서 CBD를 들르는 통행패턴은 보통 전철 혹은 지선버스-전철의 연계교통수단을 이용하는 데 반해 지선버스 혹은 간선버스만을 이용하는 통행패턴은 CBD가 아닌 지역으로 목적지가 국한되는 경향이 있다.
주요어 : 통행행태, 대중교통, CBD, 정보배열, 다차원 핵심정보추출

* Assistant Professor, Department of Geography, Kyung-Hee University, bwchjoh@khu.ac.kr

1. Introduction

Increasing use of the fuel energy threatens our environmental quality significantly. Global crisis in the world economy and consequently fluctuating world oil and food markets again demand further consideration on the less oil-consuming transportation system. Apart from the logistics of goods, individual travel also contributes significantly to the fossil fuel consumption and pollution, which gives rise to economic and environmental problems. About 30 % of the total area of the US cities is occupied by the transportation-related facilities, about 18% of GNP is spent on transport, and about 22% of personal expenditure goes for transportation in US (Hanson, 2004). Roberts (2003) even insists that war in Iraq was inevitable in these regards, and the urban planner was responsible for the breakout of the war.

With every consideration (monetary and environmental), the importance of facilitating the public mode use and discouraging the car use is paramount. Changing the modal split for more public mode use in the society requires proper policy measures. They may cover a range of policy concerns, such as 'hard' measure of the transportation infrastructure and 'soft' measure of the demand management. Given the transportation facilities and service infrastructure, short-term and cost-effective measures focus on the demand management, including toll price, congestion charge, public mode subsidy, car pull, tele-work and flexible commuting hour, to name a few (Roorda *et al.*, 2005). These all aim to increase the share of public mode use while decreasing the car use, and to reduce the transport cost in the society by reducing the peak-hour traffic congestion.

Another striking example of public mode use policy is found in Metropolitan Seoul, Korea. The IT-strong country finds that the local governments in the central province co-work to address an integrated payment system called Smart Card. Since July 1st of 2004, Seoul introduced the system that offers travelers almost free transfer between public modes including bus and metro (Shin *et al.*, 2008). The integrated public transportation system provides passengers with better transfer connections between public modes, and the travelers pay only the additional amount of money that is mostly small. For example, the traveler pays only KRW 300 when she transfers local bus to metro because local bus costs KRW 500 and metro costs KRW 800. She even does not need to pay additionally when transferring metro to local bus for another example. In this way, the traveler pays the basic (terminal) cost only once for any number of transfers in a travel. The integrated payment system has recently been extended to the integration of areas that covers the overall Seoul Metropolitan Area, including Gyunggi province and Inchon city. Some cities of Choongnam province are also included when connected by the Seoul metro line.

The system is well recognized by the citizens and becomes popular among the public mode users. It is yet to further be improved in various aspects to better encourage the public mode use, given the fact that the private car use is still of significantly high proportion for commuting, and also the peak-hour congestion remains terrible in the main freeways. Although a better urban structure is the ultimate solution, the demand management measures are still necessary to be investigated. A pre-requisite for any demand management measure is the understanding of the travel behaviour. In this regards, the current paper focuses on the analysis of travel behaviour observed in the Smart Card.

Understanding travel behaviour has been an important research area in the field of travel demand forecasting. Conventional approach to travel behaviour analysis is typically compositional and cross-sectional in the sense that, from travel behaviour of a small sample of population, particular travel behaviour is correlated with some a-priori selected socioeconomic/land-use variables. Moreover, most research implicitly assumes the trip conducted for its own purpose. Examples are abundant, mostly from transportation sciences and some from urban and transportation geographies. Readers are referred to Timmermans *et al.* (2002) for the details. The result of such research typically provides the information of travel behaviour limited in particular time and space with inappropriate assumption on trips as purpose. Travel is however derived from activity participation, the implementation of which has sequential implication in daily schedule. A sequential analysis of travel behaviour hence is necessary to better understanding of the structural (not only compositional but also sequential) information embedded in travel behaviour.

The current research aims at obtaining the structural information of travel behaviour by sequence-analysing individual travel behaviour in Metropolitan Seoul, using a multi-dimensional motif search technique applied to Smart Card data that integrates individuals' different public mode uses. To this end, the paper organized as follows. The following section briefly introduces the data analysis methods. Section 3 describes the data in detail. Section 4 provides the analysis results and discussions on implications. The paper ends with some concluding remarks.

2. Methods

1) Multi-dimensional sequence alignment

Sequence alignment (Wilson, 1998) has recently been applied to a wide range of social sciences research including analysis of daily life pattern in time geographic study (Shoval and Issacson, 2007), travel behavior analysis in transportation sciences (Rindsfuer *et al.*, 2003; Recker *et al.*, 2008), web mining behavior using web-log data (Hay *et al.*, 2004) and consumers' supermarket choice over time using scanner panel data (Joh *et al.*, 2003) in marketing science, time allocation decisions in time use research (Wilson, 2006), and tourists' overseas destination choice behavior (Bargeman *et al.*, 2002) and behavioral pattern of theme park use (Kemperman *et al.*, 2004) in tourism study, to name a few.

The relevance of the methods comes from the context where the sequential information is important, in addition to the cross-sectional information, embedded in the data. Sequential information involves not only the first-order Markovian relationships between elements in a sequence but may also include more structural, distant relationships between elements. In the context of transport mode choice, for example, using public mode back to home after work may be an outcome of using public mode going to work in the morning, the mode choice of which took place far earlier than going-back mode. The resultant measure will identify the degree of similarity and dissimilarity among travel sequences regarding the sequential as well as cross-sectional information.

A multi-dimensional extension of the original sequence alignment is suggested by Joh *et al.* (2002) to further identify the degree of similarity and dissimilarity among travel sequences regarding the information of interdependency between simultaneous choices such as destination and mode choices for the implementation of a trip. Structural measurement could then include all three kinds of information such as cross-sectional, sequential and interdependency information embedded in multi-dimensional sequential patterns. An example is that destination choice and travel mode choice are in general accompanying each other. The current study applies multi-dimensional sequence alignment method (MD-SAM) to measure the similarity and dissimilarity among travel patterns observed in Seoul Smart Card data, on which they are clustered. The readers are referred to Joh *et al.* (2002) for the details of MD-SAM.

2) Motif search

Generally speaking, motif is the information common to the individuals of a group. A population could in general be divided into a small number of sub-population groups based on a certain set of homogeneous characteristics that the individuals of a group share with each other. Such a set of homogeneous characteristics is called consensus, conserved, common part, skeleton or motif (McClure *et al.*, 1994). In our context here, motif is the sub-pattern common to the individual travel patterns of a group.

Once identified the groups of travel patterns by means of MD-SAM and a clustering algorithm, one could easily detect the motifs of each travel pattern group by examining which parts of sequential information are common to travel patterns of the group. Joh *et al.* (2008) applied motif search technique to the heavy-vehicle travel patterns to identify the motif trip chains common to particular groups of freight transport. The current paper employs the technique to profile each travel-pattern groups that were identified by MD-SAM discussed above. We will search the motifs of common information in terms of travel destination and mode. As the data size of the current study is fairly large, the motif search here is conducted with the help of a heuristic sampling strategy, to be detailed in Section 4.1.

3. Data

Smart Card data archives the details of trips and tours that individuals make. Every single day produces about ten million trip records. The information included in the records is shown in Table 1, where a number of transportation indicators are available. As for the terminology, trip and tour are to be distinguished in the sense that tour has the original departure location and final destination and consists of one or more trips. If a tour has a transfer before reaching the final destination, the tour has two trips. Therefore, a trip's origin and destination may be the original departure location and final destination or transfer location. Travel is a more general term, and in terms of its complexity, it is called a simple travel pattern when a tour has no transfer, and a complex travel pattern when one or more transfers.

For the current study, we arbitrarily chose a normal weekday that had no particular nationwide event. Our data collected on a Tuesday, November 13th of 2007, recorded 4,011,703 individuals' 10,081,677 trips. The trip data area covers the Seoul Metropolitan Area including Seoul, Gyeonggi and Inchon for bus and metro trips and Choongnam for metro trips. The data at hand however provides partial information in the sense that many of bus trips originated from or destined to the regions outside Seoul were not able to specify the exact location of departure or arrival. The analysis of the current study therefore includes the trips by metro for all regions of the study area and by bus for regions within Seoul.

The study aims at identifying representative types of travel patterns in terms of similar sequences of travel destination and mode by means of motif search. To this end, the study distinguishes five destination regions (Figure 1),

Table 1. Content of the Smart Card data

	Variable	Information
1	Individual ID	Individual information (Not for use)
2	Card ID	Smart Card ID (Not for use)
3	Card company ID	Card service provider ID
4	Departure time	-
5	Transaction ID	Identification for transfer
6	Travel mode	Metro, local bus, areal bus, intercity bus
7	Transfer frequency	For each tour
8	Bus line ID	(Not for metro)
9	Bus company ID	Bus company ID and Metro line ID
10	Bus ID	Bus vehicle ID (Not for Metro)
11	User Classification	Primary school, Middle-level school, Adults, Free for Old
12	Initial departure time	Start time of the day for Bus (Not for Metro)
13	Departure station ID	Bus and Metro
14	Arrival time	-
15	Arrival station ID	Bus and Metro
16	Size of travel party	Number of persons that a card paid
17	Departure fare	Amount of money paid when departing
18	Arrival fare	Amount of money additionally paid when arriving

including old (or northern) CBD of Seoul, new (or southern) CBD of Seoul, northern Seoul, southern Seoul and outer Seoul (Inchon, Gyeonggi and Choongnam) in the metropolitan area. North and south parts of Seoul are distinguished by the Han River running through the city from the east to the west. Northern CBD, denoted as CBD1 in the figure, is the traditional central area of the capital city Seoul and has the offices of central government and the old commercial districts lasting from Korea's old dynasty. Southern CBD, denoted as CBD2, has settled its development only in the late 80s, and the land value, particularly in the residential areas, has been up much more than other parts of Seoul in both terms of the price level and the increasing speed. Northern and southern CBDs are located in the middle of North and south parts of Seoul, respectively. They are characterized by the land use for commercial and public service, in addition to residential purpose, whereas other areas are recognized as the land use for residential district and industry complex.

Public transport mode is three kinds for the analysis, including local bus, areal bus and metro. Local bus line is short and often used for connecting metro stations to the villages that do not have a metro station in a walking distance. Areal bus is main bus line in the city. Besides, intercity connects between cities and provinces, but the system was not completed at the time of the data and the records are rarely detected. The current study does not include the intercity bus trips, and only metro trips are included for the trips outer Seoul.

Some of important averaged characteristics of

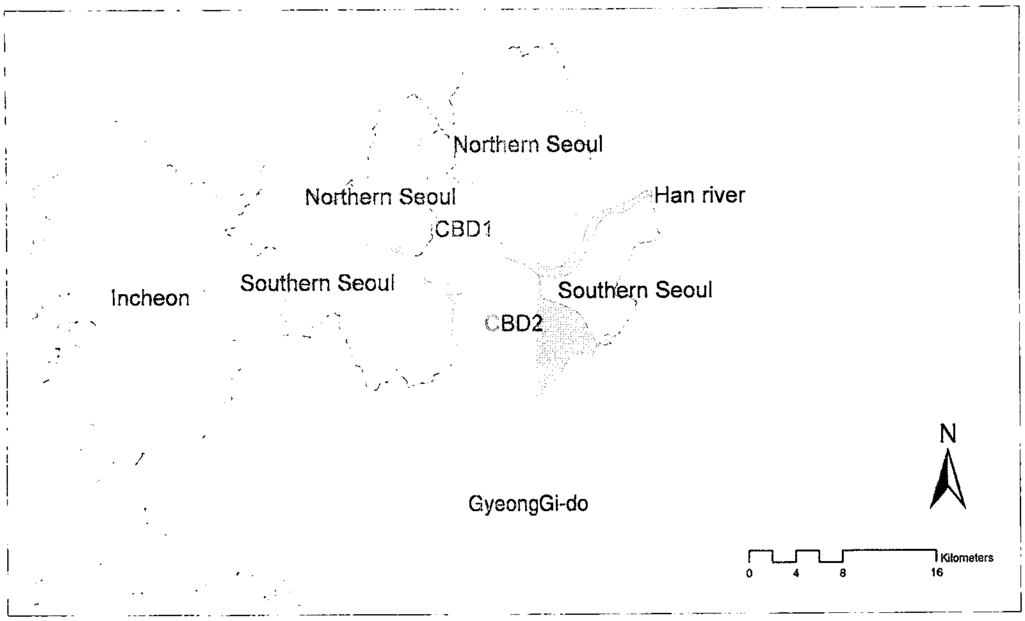

Figure 1. Trip destinations in the study area

the Seoul Metropolitan citizen's daily travel patterns by public modes are summarized in Table 2. In the table, travel time is measured in minute, and travel distance in km. Transfer is counted between different buses and between bus and metro. Transfer between metro is not counted. A tour may include several trips for transfer, as discussed earlier.

The average number of total trips 2.51 however indicates that almost one extra trip is involved for a transfer or an activity destination additional to the coming and going to the work or school. In terms of the number of mode uses, metro takes largest share of the public modes. Modal split is indicated as 30.16, 19.05 and 50.40 % for local bus, main bus and metro, respectively. The importance of metro becomes even more when considering that many of the local bus uses is for accessing to the metro station. The importance of metro is again shown by travel time and travel distance for each mode, implying that the main public mode is by far the metro. The number of tours slightly smaller than 2 means that a daily travel by public modes often involves car and taxi as well, which are not detected by the Card system.

Trip distribution between five regions as the trip destination in the study area shows that high frequency of travel destination in CBDs. Two CBDs and two residential parts each distinguished by the Hand River show approximately even proportions of trip distribution. Much small proportion of trips to the outer Seoul regions is mostly contributed by the detection limitation of the Card data at the time of the data date.

4. Analysis

1) Analysis scheme

The analysis of the current research aims to

Table 2. Descriptive characteristics of personal travel by public modes in Seoul

	mean	std.	trip destination	%
total number of trips	2.51	1.434	CBD1	12.33
total travel time	58.22	44.992	CBD2	13.03
total travel distance	28.57	19.655	northern Seoul	32.08
number of transfers	0.68	0.992	southern Seoul	29.46
number of tours	1.90	0.815	outer Seoul	13.10
number of local bus uses	0.76	1.065		
number of areal bus uses	0.48	0.896		
number of metro uses	1.27	1.027		
local bus time spent	7.83	14.415		
areal bus time spent	7.72	18.632		
metro time spent	42.67	45.676		
local bus distance travelled	8.66	13.108		
areal bus distance travelled	5.43	10.906		
metro distance travelled	14.48	13.235		

show the classification of large data providing the expected types of travel patterns given the regional configuration of the Seoul Metropolitan Area. Given the limited time of the current research, we decided to proceed with about 10 % of the total individuals, which amounts to the travel patterns of 400,000 people. A travel pattern in the analysis consists of two sequences, trip destination and transportation mode, the categories of which are described in Section 3.

Even with the 10 % of the total individuals of the card data, however, the computation time for classification and motif search is very high. The actual procedure of the current analysis therefore employed an efficient heuristic sampling strategy for large data classification addressed by Hay (2003) and adjusted it for our purpose to let the sampled set better represent the population. The quintessence of the sampling is that we assume a certain number of individual travel patterns sampled evenly across total sample space, select one most unlike individual (we call this 'the anchor pattern, to be explained later), and select the remaining individuals to be sampled one by one on the basis of the dissimilarity from the anchor pattern to cover the whole range of sample space. The detailed procedure is following.

(1) A total of 400,000 individuals' travel patterns consist of the analysis set, randomly selected from the data of 4,000,000 patterns.

(2) A set of 100 travel patterns are randomly selected from the 400,000 patterns.

(3) A travel pattern is selected from the 100 patterns, such that the pattern has the largest sum of pairwise dissimilarities against other 99 individuals' travel patterns. The dissimilarity is computed in terms of the dissimilarity of two dimensional sequential information including destination and transport mode. This selected pattern is named 'the anchor travel pattern' because the remaining random selection proceeds on the basis of this pattern.

(4) A total of 400 travel patterns from 400,000 patterns are sampled by finding every 1000^{th} travel pattern distanced away from the anchor travel pattern until finding the most dissimilar

400,000th travel pattern, the order of which is computed in terms of the dissimilarity from the anchor travel pattern.

(5) The resultant 400 travel patterns sampled are defined as the representative travel patterns and are clustered into a small number (*C*) of groups of similar travel patterns. Ward clustering algorithm based on the matrix of pairwise dissimilarities between the 400 two-dimensional travel patterns finds C clusters.

(6) A center travel pattern is identified for each cluster, which has the smallest sum of pairwise distances with all other travel patterns of the cluster.

(7) All remaining 399,600 travel patterns are added to one of the clusters on the basis of the distance to the center travel patterns.

(8) Each cluster is profiled by means of the search of the motifs or common sub-patterns.

The pairwise comparison employs MD-SAM (Joh *et al.*, 2002) and motif identification uses a variant of MD-SAM (Joh *et al.*, 2008), which all use the sequential and interdependency information embedded in the travel pattern.

Our expectation is that the resultant classification of travel patterns likely reflects the regional configuration of Seoul Metropolitan Area. That is for example, each region has stronger relationship with itself for the destination, and northern Seoul destination may well be connected more often with CBD1 destination, whereas southern Seoul destination with CBD2 destination. The current research provides the profile information obtained from the multi-dimensional motif search instead of conventional cross-sectional summary of travel behavior because the multidimensional sequential analysis provides sequential relationship between travel stops and interdependency relationship between trip destination and travel mode, in addition to the conventional cross-sectional information.

2) Analysis results

The sampled 400,000 individuals' travel patterns are fairly representative with regard to the major travel characteristics. The total number of trips is 1,004,957 which gives the average number of trips per person 2.51 (std. = 1.44). The trip destination distribution among regions of the study area is 12.34, 13.04, 32.12, 29.41 and 13.09 % for CBD1, CBD2, northern Seoul, southern Seoul and outer Seoul, respectively. Modal split is 30.26, 18.92 and 50.55 % for local bus, areal bus and metro, respectively.

The anchor travel pattern identified has six trips as in the following, where Nseoul and Sseoul denote northern Seoul and southern Seoul, respectively.

$$\begin{bmatrix} \text{Nseoul} & \text{Nseoul} & \text{Nseoul} & \text{CBD1} & \text{Nseoul} & \text{Nseoul} \\ \text{local} & \text{metro} & \text{local} & \text{metro} & \text{metro} & \text{local} \end{bmatrix}$$

The pattern has many destinations and is often switching the transport modes. The 400 travel patterns sampled from the 400,000 travel patterns were cluster-analysed by using the MD-SAM pairwise comparisons and Ward clustering algorithm, which distinguishes four pattern clusters ($C = 4$). The size of each cluster is 63 patterns for Cluster 1, 152 patterns for Cluster 2, 99 patterns for Cluster 3 and 86 patterns for Cluster 4. The four center patterns of the clusters are identified respectively as in the following.

$$\begin{bmatrix} \text{Nseoul} \\ \text{local} \end{bmatrix}, \begin{bmatrix} \text{Sseoul} \\ \text{metro} \end{bmatrix}, \begin{bmatrix} \text{Sseoul} \\ \text{local} \end{bmatrix} \text{ and } \begin{bmatrix} \text{Nseoul} \\ \text{local} \end{bmatrix}$$

These center clusters well represent non-CBDs and most frequent public modes in different combinations. Given the clusters of representative patterns, the remaining travel patterns of Seoul metropolitan citizens are conducted by adding each pattern to one of four clusters on the basis

Table 3. Cluster profile by motifs

Cluster 1 (1,953)	freq.	Cluster 2 (11,476)	freq.
[Nseoul / local]	810	[Sseoul / metro]	3,323
[CBD1 / metro]	373	[Sseoul / *]	1,123
[Nseoul / *]	298	[CBD1 / *]	938
[CBD1 Nseoul / metro metro]	5	[CBD2 / metro]	255
[Nseoul CBD1 Nseoul / local metro metro]	1	[outer / metro]	224
		[Nseoul / metro]	66
		[CBD2 Nseoul / * *]	1
Cluster 3 (4,851)	**freq.**	**Cluster 4 (3,655)**	**freq.**
[Sseoul / *]	1,817	[Nseoul / *]	2,168
[Sseoul / local]	1,596	[Nseoul / local]	1,027
[Sseoul / areal]	14	[Nseoul / areal]	75
[Sseoul / metro]	11	[Nseoul / metro]	1
		[CBD1 / *]	1
		[CBD1 Nseoul / * areal]	1

of similarity measure against the cluster centers.

Finally, the profile of each cluster using motif search is summarized as in Table 3. In the table, parenthesis next to the cluster heading denotes the number of pairwise comparisons of the cluster. Also, the asterisk '*' denotes 'don't care' implying that no common element found for the dimension at the position. Again, Nseoul and Sseoul denote northern Seoul and southern Seoul regions, respectively. The profiles of Custer 1 and 4 are mostly concerned with northern Seoul and northern CBD regions, whereas Cluster 2 and 3 are related with southern Seoul and southern CBD regions. Northern CBD is connected with northern Seoul region, whereas southern CBD is connected with southern Seoul region. This result was expected earlier, which shows that the cluster centers successfully distinguish among multi-dimensional travel patterns.

Cluster 1 is characterized by more metro uses and destination choices of both northern Seoul and northern CBD, whereas Cluster 4 is characterized by frequent bus uses and mostly limited its travel destinations to northern Seoul. Likewise, Cluster 2's profiles are concerned with metro travel to both southern Seoul and southern

CBD, whereas Cluster 3 mainly shows southern Seoul profile using bus services. The motif identification more clearly reveals the structural characteristics of the cluster profiles that are not easy for ordinary cross-sectional analysis to provide.

In words, the results can be summarized in two representative aspects. First, the most clearly identifiable characteristics of Seoul travel patterns using public modes is that the trip destinations are divided by the natural barrier, Han River, and north and south parts of Seoul cover the geographical range of travel-pattern motifs, rather exclusively to each other. This suggests that the distance decay rule still strongly influence the range of travel despite of the well-developed network of intra-Seoul public transportation. Secondly, in each sphere, that is north and south parts of Seoul, complex travel using metro and local bus has tendency to cover CBD as well as local area, whereas travel using bus only limits its range of travel within local area. This implies that the long-distance travel prefers the combination of local bus and metro to the long-range areal bus, which shows some modal split between short and long ranges of the destination.

5. Concluding remarks

The current research aims at obtaining the structural information of travel behaviour by sequence-analysing individual travel behaviour in Metropolitan Seoul, using a multi-dimensional motif search technique applied to Smart Card data that integrates individuals' different public mode uses. Major works conducted are following. First, a huge set of travel patterns of about half million individuals were successfully classified on the basis of multi-dimensional sequence alignment methods by means of effective heuristics using a randomized selection scheme. Secondly, travel-pattern groups clustered well identified distinctive travel behaviour between north and south parts of Seoul and between metro and bus uses.

Further research efforts are required in the following aspects. First, spatial analysis correlating travel pattern profiles with land use using GIS is necessary for further implications to the better understanding of Seoul's spatial structure. Secondly, the spatial resolution of the analysis should be improved. The current analysis divided Seoul into four big regions, but the original data provides the travel events information at the level of the exact geographical location. Furthermore, CBD definition was arbitrary in the current analysis. All together, a better approach defining the spatial unit for the analysis is required for a more meaningful application results. Thirdly, the socioeconomic characteristics of travelers also need to be analyzed to have the better traveler segmentation (Lee and Oh, 2008). Finally, the whole population collected in the data needs to be aligned and motif-profiled to have the entire picture of Seoul move in everyday life.

Acknowledgement: The author thanks Dr. Shin, S.I. at Seoul Development Institute for making Seoul Smart Card data accessible and Mr. Jung, W.H. at Kyung-Hee University for his work on data preparation.

References

Bargeman, B., Joh, C. H., and Timmermans, H. J. P., 2002, A typology of tourist vacation behavior using a sequence alignment method, *Annals of Tourism Research*, 29, 320-337.

Hanson, S., 2004, Getting there, in Hanson, S. (ed.), *Geography of Urban Transportation*, Guilford

Press, 3-18.

Hay, B., 2003, *SAM heuristic, Sequence Alignment Methods in Web Usage Mining*, Ph.D. Thesis, LUC, 299-330.

Hay, B. H., Wets, G., and Van Hof, K., 2004, Mining navigation patterns using a sequence alignment method, *Knowledge and Information Systems*, 6, 150-163.

Joh, C. H., Kim, C. S., and Sung, H. M., 2008, An activity-based analysis of heavy-vehicle trip chains, *Journal of the Economic Geographical Society of Korea*, 11, 192-202. (in Korean)

Joh, C. H., Arentze, T. A., and Timmermans, H. J. P., 2008, Improved motif identification of activity sequences: application to interactive computer experiment data, *Transportation Research Record*, 2054, 93-101.

Joh, C. H., Arentze, T. A., Hofman, F., and Timmermans, H. J. P., 2002, Activity pattern similarity: a multidimensional alignment method, *Transportation Research B*, 36, 385-403.

Joh, C. H., Popkowski, P., and Timmermans, H. J. P., 2003, Identifying purchase-history sensitive shopper segments using scanner panel data and sequence alignment methods, *Journal of Retailing and Consumer Services*, 10, 135-144.

Kemperman, A., Joh, C. H., and Timmermans, H. J. P., 2004, Comparing first-time and repeat visitors' activity patterns in a tourism environment, *Tourism Analysis*, 8, 159-164.

McClure, M. A., Vasi, T. K., and Fitch, W. M., 1994, Comparative analysis of multiple protein sequence alignment methods, *Molecular Biology and Evolution*, 11, 571-592.

Lee, B. J. and Oh, S. H., 2008, User-oriented information services of public transportation, *Transportation Technology and Policy*, 5(3), 132-135. (in Korean)

Recker, W., Duan, J., and Wang, H., 2008, Development of an estimation procedure for an activity-based travel demand model, *Computer? Aided Civil and Infrastructure Engineering*, 23, 483-501.

Rindsfüser, G., Mühlhans, H., Doherty, S. T., and Beckmann, K. J., 2003, Tracing the planning and execution of activities and their attributes: design and application of a hand-held scheduling process survey, in *Proceedings of the 10^th^ International Conference on Travel Behaviour Research*, Lucerne, Switzerland.

Roberts, I., 2003, Car wars, http://www.guardian.co.uk/world/2003/jan/18/iraq.usa.

Roorda, M. J., Doherty, S. T., and Miller, E. J., 2005, Operationalising household activity scheduling models, in Lee-Gosselin, M., and Doherty, S. T. (eds.), *Integrated Land-Use and Transportation Models: Behavioural Foundations*, Elsevier, Oxford.

Shin, S. I., Lee, C. J., and Cho, Y. C., 2008, A methodology to evaluate regional public transportation services using public transport use data, *Transportation Technology and Policy*, 5(1), 111-127. (in Korean)

Shoval, N. and Isaacson, M., 2007, Sequence alignment as a method for human activity analysis in space and time, *Annals of the Association of American Geographers*, 97, 282-297.

Timmermans, H. J. P., Arentze, T. A., and Joh, C. H., 2002, Analyzing space-time behavior: new approaches to old problems, *Progress in Human Geography*, 26, 175-190.

Wilson, C., 1998, Activity-travel pattern analysis by means of sequence alignment methods, *Environment and Planning A*, 30, 1017-1038.

Wilson, C., 2006, Reliability of sequence-alignment analysis of social processes: Monte Carlo tests of ClustalG software, *Environment and Planning A*, 38, 187-204.

Correspondence: Chang-Hyeon Joh, Department of Geography, Kyung Hee University, 1 Hoegi-dong, Dongdaemun-gu, Seoul, 130-701, Republic of Korea (e-mail: bwchjoh@khu.ac.kr, phone: +82-2-961-9264)

교신: 조창현, 서울시 동대문구 회기동 1 경희대학교 지리학과 (이메일: bwchjoh@khu.ac.kr, 전화: 02-961-9264)

Recieved April 28, 2009
Revised May 26, 2009
Accepted May 28, 2009