Application of the L-index to the Delineation of Market Areas of Retail Businesses

Lee, Sang−Kyeong[1] · Lee, Byoungkil[2]

Abstract

As delineating market areas of retail businesses has become an interesting topic in marketing field, Lee and Lee recently suggested a noteworthy method, which applied the hydrological analysis of geographical information system (GIS), based on Christaller's central place theory. They used a digital elevation model (DEM) which inverted the kernel density of retail businesses, which was measured by using bandwidths of pre-determined 500, 1000 and 5000 m, respectively. In fact, their method is not a fully data-based approach in that they used pre-determined kernel bandwidths, however, this paper has been planned to improve Lee and Lee's method by using a kind of data-based approach of the L-index that describes clustering level of point feature distribution. The case study is implemented to automobile-related retail businesses in Seoul, Korea with selected Kernel bandwidths, 1211.5, 2120.2 and 7067.2 m from L-index analysis. Subsequently, the kernel density is measured, the density DEM is created by inverting it, and boundaries of market areas are extracted. Following the study, analysis results are summarized as follows. Firstly, the L-index can be a useful tool to complement the Lee and Lee's market area analysis method. At next, the kernel bandwidths, pre-determined by Lee and Lee, cannot be uniformly applied to all kinds of retail businesses. Lastly, the L-index method can be useful for analyzing the space structure of market areas of retail businesses, based on Christaller's central place theory.

Keywords : Market Area, Hydrological Modeling, Kernel Density, L-index

1. Introduction

A market area analysis has become an interesting topic in marketing field and therefore popular methods such as Huff's (1963) model have been developed. However, most of the previous methods have limitation in that they were not based on geographic information system (GIS) and thus, did not use effectively location and density information. Although various methods, such as Thiessen polygon for utilizing GIS function, have been developed, all of them are not based on market area theories. In that sense, Lee and Lee (2014) is worthy of notice because they developed a new method using hydrological analysis of GIS based on a famous market area theory, Christaller's central place theory.

Lee and Lee's method does not require any customer information but only the location information of point features to measure the density of retail businesses. They measure density by a kernel density function and invert it for a digital elevation model (DEM). In this process, the most important thing is that kernel bandwidths have to be decided in advance to measure the density of retail businesses. To testify usefulness of their method, Lee and Lee analyzed market areas of retail businesses in Seoul, Korea and used three kernel bandwidths: 500, 1000 and 5000 m. These bandwidths were determined by considering a consumer's behavior. They correspond to the size of neighborhood unit, neighborhood

district and sub-center's hinterland, respectively. In spite of prudent determination, however, these kernel bandwidths were pre-determined regardless of the spatial distribution pattern of retail businesses. In other words, it means that Lee and Lee (2014) do not use a fully data-based approach, which has advantage in reflecting reality.

In an attempt to decide a kernel bandwidth based on the distribution pattern of point features, the L-index was suggested by Borruso (2003), Borruso and Schoier (2004), Goodwin and Unwin (2000), Hwang (2004), and Lee (2008). The L-index is an indicator showing the clustering level of point feature distribution and it has been known that there is the strongest clustering at a distance with the maximum point of the L-index. In general, the distribution pattern of point features is explained by density and therefore, it is addressed best when using a distance with maximum point of the L-index as a kernel bandwidth. In this sense, the L-index can be useful tool to improve Lee and Lee's (2014) method. In this paper, we aim at improving the Lee and Lee's (2014) market area analysis method by using the L-index, which reflects the spatial distribution pattern of retail businesses. While Lee and Lee (2014) used total retail business data to find urban space structure, but we only use automobile related retail business data because we are not interested in urban space structure but market area structure of individual business.

2. Literature Review

Most market area analysis methods are originated from Christaller's central place theory and Reilly's law of the retail gravitation. The central place theory was developed by W. Christaller, a German geographer in 1933 (Pirounakis, 2013). A central place is a concentrated area which provides goods and services to a hinterland. A central place has an order that is determined by the level of goods and services. The higher central places nest lower central places, so the hierarchical structure appeared. A central place exerts influence in a hexagon shape called the market area, which differs in size depending on the order of a central place (Christaller, 1966; Lee and Lee, 2014). Christaller's main assumption is an isotropic plane which means identical in all directions (Pirounakis, 2013). For an urban space, the isotropic plane

assumption cannot be adequate and therefore, a hierarchical structure with hexagon market areas does not appear (Clark and Rushton, 1970; Lee, 1992; Davies *et al.*, 1999; Lee and Lee, 2014). Despite this limitation, the concept of hierarchical distribution of the central places is still valuable when it is used to explain market structure of an urban space. (Lee and Lee, 2014)

The Reilly's (1931) law of retail gravitation was inspired by Newton's law of universal gravitation. It proposes a way to estimate the proportional ability of two cities to attract trade from a location. The proportional ability is related to the ratio of the two populations and inversely related to the square of the ratio of the distances from the two cities to the location (Pirounakis, 2013). Reilly's law has inspired Converse's (1949) break-point model, Huff's (1963) model, and other gravity models. In gravity models, a considerable amount of research has undergone for delineating market areas, which has been an important issue (Bucklin, 1971; Peterson, 1974; Huff and Batsell, 1977; Huff and Rust, 1984; Current and Storbeck, 1994). Unfortunately, these models have been limitedly used in actual businesses fields because they need plenty of manpower and time.

Recently, this limitation is being improved by the business GIS, which is interested in dynamic data that is geo-demographics, or customer data. Dynamic data are mostly used as point features with the coordinates (Lee, 2006). The buffer analysis, range analysis and the Thiessen polygon are mainly used in the business GIS. The buffer analysis uses a certain distance and the range analysis uses a certain travel time through road networks. The Thiessen polygon is a polygon formed from the seed points and the boundary of the Thiessen polygon is set up so that the distance to the seed point on the polygon is closer than any other seed points on nearby polygons (Lee and Lee, 2014). These methods mainly use distances between point features but seldom use density. In Korea, some studies developed new methods based on raster density that applies only location information of retail businesses (Shin *et al.*, 2002; Lee and Kim, 2008). They cannot analyze entire urban space but only local area because they are not based on market area theory which explains urban space structure. Lee and Lee (2014) are noteworthy because they use the density of retail businesses and analyze urban

structure, based on the Christaller's central place theory. For the procedures, the pre-determined bandwidths are applied to measure density instead of a data-based approach. In general, a data-based approach has an advantage in reflecting data state, but has a disadvantage in required the sufficient data for preventing the biased results analysis.

3. Methodology

3.1 The analogy method of hydrological modelling

Lee and Lee (2014) suggested the algorithm dividing market areas using an analogy of hydrological modeling. A catchment area, which corresponds to the market area in a methodology of Lee and Lee (2014), means a basin collecting water in hydrology. The boundary of a catchment area was manually delineated in the past, yet it is now automatically extracted by GIS analysis function and DEM. The boundary extracted by DEM will be consistent regardless of users. As an advantage, only the accuracy and resolution of DEM will affect the accuracy of catchment boundaries. DEM is a digital record set of terrain elevations for ground locations of grids, which are divided pieces of ground with cells of equal size. The general procedure of delineating catchment areas are comprised of 4 steps as shown in Fig. 1.
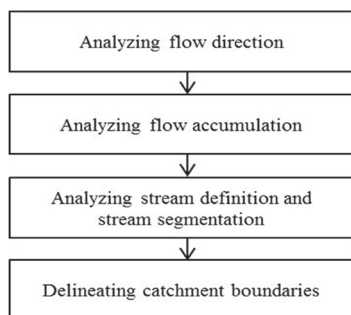

Fig.1. The general procedure of delineating catchment areas

In hydrological analysis, the rainwater flows from the higher point to lower one. This is similar to the traffic concentration to a central place with high density from nearby areas. The gradient of elevation differences becomes flow direction for the hydrological analysis, but gradient of business density differences becomes traffic direction in market area analysis (Lee and Lee, 2014). The first process to

divide the market areas following the analogy, Kernel density needs to be computed by store locations, since it has been known that kernel density is the most appropriate analysis method to describe a point feature distribution (Lee and Kim, 2008). The kernel density is calculated by the number of point features included within a search radius of every raster grid cell. Here, points near the cell appear to be high weight values, as points far from the cell have low weight values. The kernel density estimation has an advantage because only search radius and weighting function will affect the result. A large search radius helps in revealing the global trend, and the details can be showed using a short search radius. An inverse quadratic function is generally used as weighting function (Borruso, 2003; Goodwin and Unwin, 2000).

In market area analysis, the traffic flow begins from points of lower density to higher ones. On the other hand, in hydrological analysis, water flows from higher points to lower ones. Therefore, in order to make density DEM, kernel density should be inverted as shown on Fig. 2. By applying a general process of hydrological analysis to density DEM, the boundary of a catchment area can be extracted to the region with low kernel density as a boundary.
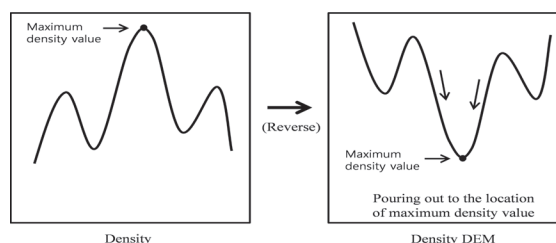

Fig.2. The concept of density DEM (Source: Lee and Lee, 2014)

3.2 Bandwidth setting by the L − index

The L-index, which has been suggested as a bandwidth setting method of kernel density analysis in several previous studies, is induced from the K-index. The K-index is calculated using Eq. (1) and decides whether point distribution is random with comparing the number of points existed in a certain distance from a specific point and theoretically expected.

$$\hat{K}(\tau) = \frac{R}{n^2} \sum \sum \frac{I_h(d_{ij})}{w_{ij}} = \frac{1}{\lambda^2 R} \sum \sum \frac{I_h(d_{ij})}{w_{ij}} \quad (1)$$

here, R is area of a target space, *n* is the number of events,

$\lambda(= \frac{n}{R})$ is the average density of events, d_{ij} is a distance between s_i and s_j, $I_h()$ is an indicator function (1 in case of $d_{ij}<\tau$, and 0 for the rest cases), and w_{ij} is a weighting to remove a boundary line effect.

In a random pattern, the probability of points existing in all locations is same and independent. Therefore, $\lambda\pi\tau^2$ is the number of average points, expected to be found in a certain distance from the specific point. A dense distribution appears in τ, whose $\widehat{K}(\tau)$ is bigger than $\lambda\pi\tau^2$.

However, the L-index, likely in Eq. (2), which is square root of the K-index, is practically used because a graph of K-index is exponentially increasing.

$$L(\tau) = \sqrt{\frac{K(\tau)}{\pi}} - \tau \qquad (2)$$

Eq. (2) is the L-index, originally suggested by Ripley (1976), but redeemed by Cressie (1991). The L-index on complete spatial randomness (CSR) distribution has the merit of coinciding with the x-axis of a graph due to a deducted distance. Diggle (1985) and Hwang (2004) reported that a distance with the maximum point of the L-index is suitable for a bandwidth of the kernel density.

4. Data

The spatial scope of this research is Seoul, Korea. Automobile-related retail businesses under the retail business categories from February 2009 phone book was used. Retail business category in a phone book is composed of 26 categories, including automobile-related retail businesses, which category is composed of 11 sub-categories as shown on Table 1.

Consequently, 3541 phone numbers are used among 61,292 for retail business category. Geo-coding is performed in order to assign coordinates for phone number using the address information included in the phone book. As shown on Fig. 3, they are widely dispersed in many places in Seoul.
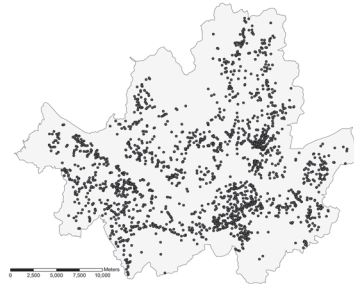

Fig.3. Automobile-related retail businesses in Seoul

5. Results of Analysis

To find distances applying as kernel bandwidths, $L(\tau)$ was analyzed using the CrimeStat III program, which was specialized in point process analysis. Fig. 4 shows a graph of $L(\tau)$ against distance for automobile-related retail businesses in Seoul. As seen the graph in Fig. 4, $L(\tau)$ has a global maximum point at distance 1211.5 m and local maximum points at 2120.2 and 7067.2 m, respectively. This reveals that there is strong clustering at 3 distances; the strongest clustering is at a distance of 1211.5 m, the second and the third locate at 2120.2 and 7067.2 m, respectively. Consumer goods can be categorized into convenience, shopping, and specialty goods. The classification scheme is based on ways of purchase. Businesses can be also categorized into convenience store, shopping store, and specialty store. Automobile-related retail businesses can be regarded as specialty store and they tend to be clustered in particular places in Seoul; for example, Janganpyeong market is famous for used car trade and there are many retail businesses. Therefore, the analysis result of L-index can be considered to be corresponding with characteristics of automobile-related retail businesses.

Three distances of L-index method, 1211.5, 2120.2 and 7067.2 m are used as bandwidths to measure kernel density of

Table 1. Automobile-related retail business data in Seoul

Category	Sub-category	Count
D23 Automobile-related retail business	A01 Automobile Sales	1272
	A02 Used Car	570
	A03 Car Audio	63
	A04 Tire	303
	A05 Navigation	19
	A06 Automobile Parts	844
	A07 Accessories	126
	A08 Used Tire	3
	A09 Imported Automobile	248
	A10 Used Imported Automobile	11
	A11 Car Glass	82
	Sum	3541

automobile-related retail businesses. When they were pairing off with pre-determined bandwidths of Lee and Lee's method of 500 m, 1000 m and 5000 m, they were larger than Lee and Lee's method. Kernel density was measured by the Spatial Analyst Tools of ArcGIS 10 program and the pixel size was set at 100m. A density DEM was created inverting kernel density as shown on Fig. 2. Catchment areas, market areas in this study, were extracted by applying arcHydro Tool to a density DEM. The fig. 5, 6 and 7 display boundaries of market areas when using 1211.5, 2120.2 and 7067.2 m from the L-index method as bandwidths, and also 500, 1000 and 5000 m of the Lee and Lee's method (2014) as bandwidths.
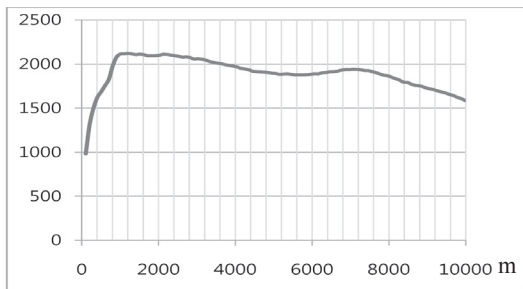

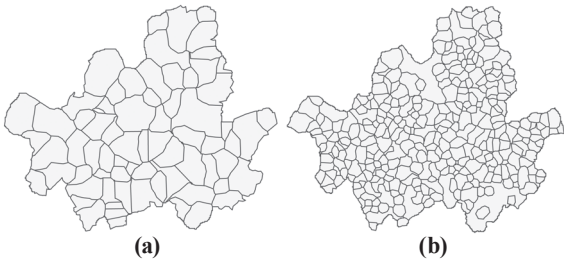**Fig. 4 . L(τ) change against distances for automobile-related retail businesses**

(a)   (b)

Fig. 5. Boundaries of market areas by 1211.5 m (a) and 500 m (b) bandwidth

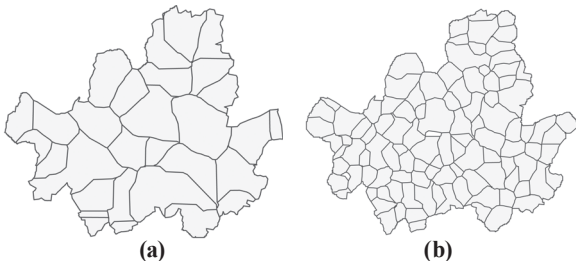(a)   (b)

Fig. 6. Boundaries of market areas by 2120.2 m (a) and 1000 m (b) bandwidth
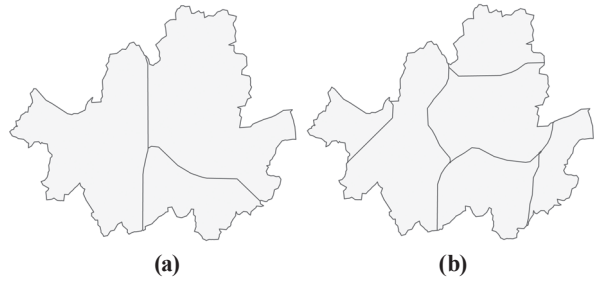

(a)   (b)

Fig. 7. Boundaries of market areas by 7067.2 m (a) and 5000 m (b) bandwidth

Table 2. The analysis result of market areas

	L-index Bandwidth			Lee and Lee's (2014) Bandwidth		
	1211.5 m	2120.2 m	7067.2 m	500 m	1000 m	5000 m
Number of market areas	70	30	3	324	106	6
Range of market areas (km²)	1.7 ~ 25.7	3.0 ~ 47.1	95.2 ~ 278.3	0.4 ~ 21.6	1.5 ~ 15.5	46.5 ~ 174.3
Average size of market areas (km²)	8.2	19.9	202.2	1.7	5.3	101.1

As shown on the Fig. 5, 6 and 7, results from the L-index method shows overall larger market areas than those from the Lee and Lee's method. According to Table 3, the average size of 1211.5 m market areas which are the smallest in the L-index method is larger than 1000 m market areas, which are the 2nd largest in Lee and Lee's method. On the contrary, numbers of market areas of the L-index method is much smaller than numbers of the Lee and Lee's method. Considering the meaning of L-index, this means that kernel bandwidths of Lee and Lee's method do not properly reflect the distribution pattern of automobile-related retail businesses. This can be easily agreed in that Lee and Lee's method is not a data-based approach but a pre-determined approach.

According to the Christaller's central place theory, the lower ordered market areas shall be embraced to those of a higher ordered central place (Christaller, 1966). Lee and Lee (2014) showed such phenomenon by overlaying market areas. The fig. 8 (a) and (b) show embraced market areas for automobile-related retail businesses by overlaying market areas of the Fig. 5, 6 and 7. This means that L-index method can be usefully applied to the Christaller's central place theory.

(a)	**(b)**

Fig. 8. Overlapping market areas of Fig. 5, 6 and 7

6. Conclusion

Lee and Lee (2014) suggested the analogy method of hydrological modelling for delimiting boundaries of market areas of retail businesses, based on the Christaller's central place theory. The method used the raster-based density DEM, which inverts the kernel density of retail businesses. The kernel density is measured by using 500, 1000 and 5000 m bandwidths for considering the size of the neighborhood unit, neighborhood district and sub-center hinterland, respectively. However, these bandwidths have limitation in that they are not a fully data-based but pre-determined. To overcome this limitation, this paper suggested a method to improve the market area analysis algorithm of Lee and Lee (2014) by using kernel bandwidths which are extracted from the L-index.

A case study was implemented to automobile-related retail businesses in Seoul, Korea. The kernel bandwidths, 1211.5, 2120.2 and 7067.2 m were extracted from L-index analysis. Subsequently, kernel density was measured, the density DEM was created, and market areas were delineated. Analysis results are summarized as follows. First, L-index method can be a method to complement Lee and Lee's method since the spatial distribution pattern of retail businesses is well reflected. Second, Lee and Lee's pre-determined bandwidths cannot be applied to all kinds of retail businesses uniformly. Finally, the L-index method can be utilized for delineating the space structure of market areas of retail business, based on the Christaller's central place theory.

In fact, there is an advantage in data-based approach that it reflects retail business distribution but also has a disadvantage in that it requires the sufficient data for non-biased of results

analysis. Although the results contribute to developments of the market area analysis method by applying the L-index application as a data-based approach, we cannot confirm that the L-index approach is superior to the Lee and Lee's approach. One reason is that we did not accomplish sufficient tests for various kinds of retail businesses, and the other is that we did not evaluate results analysis. Lee and Lee (2014) evaluated analysis results by internal traffic data of market areas, but this study cannot evaluate analysis results because there is not suitable traffic data for the automobile-related retail businesses.

References

Borruso, G. and Schoier, G. (2004), Density analysis on large geographical databases, Search for an Index of Centrality of Services at Urban Scale, *ICCSA 2004*, LNCS 3044, pp. 1009-1015.

Brunsdon, C. (1995), Estimating probability surfaces for geographical point data: an adaptive kernel algorithm, *Computers and Geosciences*, Vol.21, No.7, pp. 877-894.

Bucklin, L.P. (1971). Trade area boundaries: some issues in theory and methodology. *Journal of Marketing Research*, Vol.8, No.1, pp. 30-37.

Christaller, W. (1966), *Central places in Southern Germany* (C. W. Baskin, Trans.), Prentice Hall, Upper Saddle River.

Clark, W. A. V., and Rushton, G. (1970). Models of intra-urban consumer behavior and their implications for central place theory, *Economic Geography*, Vol.46, NO.3, pp. 486-497

Converse, P.D. (1949). New laws of retail gravitation, *Journal of Marketing*, Vol.14, No.3, pp. 379-384.

Cressie, N. (1991), *Statistics for Spatial Data*, John Wiley and Sons, New York.

Current, J.R. and Storbeck, J.E. (1994). A multiobjective approach to design franchise outlet networks. *The Journal of the Operation Research Society*, Vol.45, No.1, pp.71-81.

Davies, C. S., Holz, R.K., and Robertus, D. (1999). A test of central place theory using shuttle imaging radar (SIR-A) of China's north central plain. *Geocarto International*, Vol.4, No.1, pp.14-23.

Diggle, P. J. (1985), A kernel method for smoothing point

process data, *Applied Statistics*, Vol.34, No.2, pp. 138-147.

Goodwin, M. T. and Unwin, D. (2000), Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations, *Transactions in GIS*, Vol.4, No.4, pp. 305-317.

Hwang, S. (2004), Temporal extensions of K function, *UCGIS Fall 2004*, pp. 1-18.

Huff, D. L. (1963), A probabilistic analysis of shopping center trade areas, *Land Economics*, Vol.39, No. February, pp. 81-90.

Huff, D. L. and Batsell, R. R. (1977), Delimiting the areal extent of a market area. *Journal of Marketing Research*, Vol.14, No.4, pp. 581-585.

Huff, D.L. and Rust, R.T. (1984), Measuring the congruence of market areas, *Journal of Marketing*, Vol.48, No.Winter, pp. 68-74.

Jin, C.-J., Park, H.-S., and Kang, J.-M. (2012), An empirical analysis of locational tendency of coffee shops around Hongik university, *Journal of the Urban Design Institute of Korea*, Vol.13, No.5, pp. 71-82. (in Korean with English abstract)

Lee, B. (2008), Applying the L-index for analyzing the density of point features, *The Journal of GIS Association of Korea*, Vol.16, No.2, pp. 237-247. (in Korean with English abstract)

Lee, B. and Lee, S.-K. (2014), Delimiting boundaries of market areas of central places using the density of retail facilities in an urban space, *International Journal of Urban Sciences*, Vol.18, No.1, pp. 90-102.

Lee, B.-K. and Kim, J.-Y. (2008), Employment center

identification by density-based clustering method, *Journal of the Korean Regional Development Association*, Vol.20, No.3, pp. 1-20. (in Korean with English abstract)

Lee, C.-S. (1992), *A Study on the Hierarchy and Typology Analysis of the Commercial Zone in Seoul City* (Unpublished doctoral dissertation). Seoul National University, Seoul.

Lee, S.-K. and Lee, B. (2013), Assessing the appropriateness of the spatial distribution of Standard Lots using the L-index, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol.31, No.6-2, pp. 601-609. (in Korean with English abstract)

Lee, Y.-I., Hong, S.-E., and Park, S. (2010), Methodology of the spatial boundary determination for trade area of large-scale discount stores using GIS and road connectivity, *Journal of Korea Spatial Information Society*, Vol.8, No.1, pp. 11-17. (in Korean with English abstract)

Pirounakis, N. G. (2013), *Real Estate Economics*, Routledge, London.

Reilly, W. J. (1931), *The Law of Retail Gravitation*, Knickerbocker Press, NY.

Ripley, B. D. (1976), The second-order analysis of stationary point processes, *Journal of Applied Probability*, Vol.13, pp. 255-266.

Shin, W.-J., Jung, C.-M., and Lee, S.-K. (2002), A study on the changing patterns of the commercial power based on retail trade area analysis, *Journal of the Korea Real Estate Analysts Association*, Vol.8, No.1, pp. 1-11. (in Korean with English abstract).