

Print ISSN: 1738-3110 / Online ISSN 2093-7717
<http://dx.doi.org/10.15722/jds.14.1.201601.67>

Evaluation on Development Performances of E-Commerce for 50 Major Cities in China

중국 주요 50개 도시의 전자상거래 발전성과에 대한 평가

Dong-Bin Jeong(정동빈)*, Qiang Wang(왕강)**

Received: December 8, 2015. Revised: January 11, 2016. Accepted: January 15, 2016.

Abstract

Purpose – In this paper, the degree of similarity and dissimilarity between pairs of 50 major cities in China can be shown on the basis of three evaluation variables(internet businessman index, internet shopping index and e-commerce development index). Dissimilarity distance matrix is used to analyze both similarity and dissimilarity between each fifty city in China by calculating dissimilarity as distance. Higher value signifies higher degree of dissimilarity between two cities. Cluster analysis is exploited to classify 50 cities into a number of different groups such that similar cities are placed in the same group. In addition, multidimensional scaling(MDS) technique can obtain visual representation for exploring the pattern of proximities among 50 major cities in China based on three development performance attributes.

Research design, data, and methodology – This research is performed by the 2013 report provided with AliResearch in China(1/1/2013~11/30/2013) and utilized multivariate methods such as dissimilarity distance matrix, cluster analysis and MDS by using CLUSTER, KMEANS, PROXIMITIES and ALSCAL procedures in SPSS 21.0.

Results – This research applies two types of cluster analysis and MDS on three development performances based on the 2013 report of Aliresearch. As a result, it is confirmed that grouping is possible by categorizing the types into four clusters which share similar characteristics. MDS is exploited to carry out positioning of both grouped locations of cluster and 50 major cities belonging to each cluster.

Since all the values corresponding to Shenzhen, Guangzhou

and Hangzhou(which belong to cluster 1 among 50 major cities) are very large, these cities are superior to other cities in all three evaluation attributes.

Twelve cities(Beijing, ShangHai, Jinghua, ZhuHai, XiaMen, SuZhou, NanJing, DongWan, ZhangShan, JiaXing, NingBo and FoShan), which belong to cluster 3, are inferior to those of cluster 1 in terms of all three attributes, but they can be expected to be the next e-commerce revolution.

The rest of major cities, in particular, which belong to cluster 4 are relatively inferior in all three attributes, so that this automatically evokes creative innovation, which leads to e-commerce development as a whole in China.

In terms of internet businessman index, on the other hand, Tainan, Taizhong, and Gaoxiong(which belong to cluster 2) are situated superior to others. However, these three cities are inferior to others in an internet shopping index sense. The rest of major cities, in particular, which belong to cluster 4 are relatively inferior in all three evaluation attributes, so that this automatically evokes innovation and entrepreneurship, which leads to e-commerce development as a whole in China.

Conclusions – This study suggests the implications to help e-governmental officers and companies make strategies in both Korea and China. This is expected to give some useful information in understanding the recent situation of e-commerce in China, by looking over development performances of 50 major cities. Therefore, we should develop marketing, branding and communication relevant to online Chinese consumers. One of these efforts will be incentives like loyalty points and coupons that can encourage consumers and building in-house logistics networks.

Keywords: Development Performances, Dissimilarity Distance Matrix, Multidimensional Scaling, Cluster Analysis, STRESS.

JEL Classifications: C02, C18, M21.

* First and corresponding author, Professor, Dept. of Information Statistics, Gangneung-Wonju National University, Korea. Tel: +82-33-640-2274. E-mail: dj@gwnu.ac.kr.

** Graduate student, Dept. of Statistics, Gangneung-Wonju National University, Korea. Tel: +82-33-640-2274. E-mail: wq890730@gmail.com.

1. 서론

중국의 전자상거래는 1997년 인터넷 도입 이후 3단계(초창기, 발전기, 성숙기) 과정을 단계적으로 밟고 있으며, 2014년 현재 중국 전자상거래 시장거래규모는 12조 3천억 위안으로 전년대비 21.3%로 안정적으로 증가하였고, 특히 인터넷쇼핑은 48.7% 증가하여 전자상거래 시장을 이끄는 견인차 역할을 하고 있다. 인터넷 쇼핑은 B2B에 비해 비중이 작지만 그 성장의 폭이 폭발적으로 확대되고 있어 전자상거래를 주도하고 있는 양상을 보이고 있다. 또한 C2C에서 B2C로 변화, B2B와 B2C의 융합 및 정보플랫폼에서 거래플랫폼으로 전환하는 질적 고도화가 진행되고 있으며 Alibaba와 같은 거대 전자상거래 기업들이 인수 및 합병으로 해외진출이 활발하게 진행되고 있다. 중국의 전자상거래는 인터넷 및 모바일 기기 보급확대, 정부의 적극적인 정책지원, 기업의 구매 및 판매방식 효율화, 물류비용의 저하 등의 요인으로 그 원인을 찾을 수 있다. 따라서 중국의 전자상거래 시장은 중국경제의 성장세 지속과 인터넷 사용확대를 배경으로 증가세가 유지되며, 여러 기업이 미국증시에 상당한 상태로 인터넷 기업의 해외진출이 더 가속화될 것으로 전망할 수 있다(Jhun, 2014).

모바일 시장의 빠른 발전에 따라 전자상거래 기업이 중국의 3급 및 4급 도시와 심지어 농촌시장까지 확대하며 블루오션을 찾아 국제화전략을 추진하고 있다. 특히 상무부는 전자상거래를 통한 소매수출 지원정책을 발표하고, 2012년 하반기에는 상해, 충칭, 항주, 영파, 정주를 전자상거래 통관 서비스 시범도시로 지정하였으며, 2013년부터 전국적으로 조건을 만족하는 지역부터 지원하기로 하였다. 전자상거래를 통한 소매 수출을 위한 새로운 세관 감독관리 방식 및 전문 통계방식 설립, 새로운 검사검역 방식의 출현, 기업의 정상적인 수입, 결제, 외환업무지원, 금융기관 및 지불기관의 지불서비스 장려, 전자상거래 수출신용 시스템 건립 등이 주요골자이다.

Kim & Seo(2012)는 급변하는 스마트폰 시장의 성장을 따라 스마트폰 이용자들의 주요 구매동기의 물입을 근거로 충성도와 구매의도에 영향을 미치는지 여부를 유의성 검정하였다. 여기에서 구매동기란 개인 각각의 욕구를 충족하기 위한 특정 상품이나 서비스를 구매하는 동기를 의미한다. 구매동기와 연관되는 이유는 특정목표로 행동을 유발하는 지속적인 경향이며, 특히 외부적 환경안에 내재하는 목표를 대해서 선별적으로 생체에너지를 활성화시키는 내적인 현상으로 정의할 수 있다(Lee, 2011; Lee, 2013).

Jeong(2013)은 국내 학생수 1만 명 이하인 14개 국립대학교를 대상으로 교육과학기술부에서 제안한 주요 교육지표를 이용하여 포지셔닝 분석한 반면에, Jeong(2014)는 국내 28개 국공립대학교의 연구성과지표에 대해 다변량분석을 적용시켰다. Lee & Kim(2013)은 집단분류방법으로 군집분석을 이용하여 의도된 소비습관에 대한 저축률 및 소비습관에 대한 교육의 저축률의 차이에 대한 유의성 영향력을 살펴보았다. Jeong(2012)은 대학을 대상으로 미국의 말콤볼드리지(MBNQA) 국가품질상 교육의 기준을 사용하여 지도력, 전략기획, 고객 및 인적자원 중요시, 프로세스관리, 측정분석 및 지식경영 간의 인과관계 및 매개효과를 검정하였다.

Jeong(2015)은 국내 16개 행정구역과 13개 업종을 대상으로 전자상거래 구매동기에 대해 행정구역 및 업종을 유사한 속성을 지닌 군집끼리 분류 및 포지셔닝하였다. 그 결과 구매프로세스 단순화, 구매비용절감 및 공급자와 접촉 용이성 변수에 대해서 도소매업, 제조업, 서울과 경기도가 우위에, 제주는 열세함을 알 수 있었다. 반면에 업무추진속도향상 변수에 대해서는 도소매업과 숙박음식업, 대구 및 경기도가 열등한 위치에, 단체수리 및 개인서비스

업, 충북과 경남이 우위에 있음을 입증하였다.

상기 논의에 의해 구체적인 연구의 목적은 다음과 같다.

첫째, 중국은 2009년부터 전자상거래를 활성화를 위한 노력으로 선전시를 시범도시로 지정한 바 있다. 이에 관련하여 중국의 주요 50개 도시를 대상으로 발전성과면에서 각 도시들의 유사성 및 상이성 정도를 비교함으로써 특성을 파악하고자 한다. 여기에서 전자상거래 발전성과의 지표로서 세 개의 평가속성(인터넷상인지수, 인터넷쇼핑지수, 전자상거래발전지수)을 고려하였다(Sheng, Chen, & Zhang, 2014).

둘째, 발전성과 속성에 근거하여 50개 주요도시 중에서 유사성이 있는 도시들끼리 그룹화하며, 각 집단에 속한 도시들이 갖는 속성의 우위성과 열성을 다시 집단과 함께 상대적으로 비교 분류한다.

셋째, 주요도시 간 유사성 또는 비유사성을 근거로 도시들의 상대적 거리에 근거하여 2차원평면에 도시들을 포지셔닝 함으로써 도시들끼리 비교하여 어떤 이미지 속성을 가지는지 탐색 및 파악한다. 그 분석결과를 바탕으로 중국 주요도시들의 전자상거래 발전성과의 현황과 향후 방향성을 위한 바람직한 시사점을 도출하고자 한다.

본 연구의 목적을 달성하기 위하여 IBM SPSS Statistics 21.0을 이용하여 상이성 거리행렬분석, 계층적 군집분석, K-평균 군집분석, 다차원척도법을 실시하였다.

2. 연구방법

2.1. 분석자료

본 연구에서 고려하는 자료는 기준시점 2013년 1월 1일에서 11월 30일까지 취합하여 전자상거래와 관련하여 발전도시를 나타낸 것으로 아리연구원(Sheng et al., 2014)에서 발췌하였다. 중국의 주요 50개 발전도시의 전자상거래 발전성과를 평가할 세 개의 평가속성은 다음과 같이 측정하여 계산되었다.

$$1) \text{ 인터넷상인지수} = (\text{인터넷상인수 밀도지수} + \text{인터넷기업평균 교역지수}) \times 0.5$$

$$\text{여기에서, 인터넷상인 밀도지수} = (\text{B2B 인터넷상인수} + \text{소매인터넷상인수}) / \text{인구수}$$

$$\text{인터넷기업평균 교역지수} = (\text{인터넷상인수교역액} / \text{소매인터넷상인수})$$

참고로, B2B 인터넷상인수는 2013년 11월 30일 기준으로 송신통(誠信通)신청 회원수와 중국 납품업자 회원수, 소매인터넷상인수는 2013년 11월 30일 기준 Taobao와 Tian mao 인터넷 쇼핑사이트수, 인구수는 제6차 인구조사에 근거한 내륙인구와 대만인구와 홍콩인구를 합한다.

$$2) \text{ 인터넷쇼핑지수} = (\text{인터넷쇼핑밀도지수} + 1 \text{인당 평균인터넷쇼핑지수}) \times 0.5$$

$$\text{여기에서, (인터넷쇼핑밀도지수} = \text{인터넷쇼핑소비지수} / \text{인구수},$$

$$1 \text{인당 평균인터넷쇼핑지수} = \text{인터넷소비액} / \text{인터넷소비지수}.$$

참고로, 인터넷소비자수는 2013년 1월부터 11월까지 Taobao와 Tian mao에서 최소 한 번이라도 인터넷 쇼핑한 사람수를 의미한다.

$$3) \text{ 전자상거래발전지수} = (\text{인터넷상인지수} + \text{인터넷쇼핑지수}) \times 0.5$$

2.2. 군집분석

군집분석이란 고려한 개체 중에서 유사성이 있는 것들끼리 집 단화하여, 각 집단의 특성을 파악함으로써 자료 안에 숨어있는 구조를 이해할 수 있는 탐색적 다변량분석 기법이다.

다음과 같이 비유사성(또는 유사성)거리를 측정하는 방법이 있다. 고려하는 대상 i 의 관측값을 $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$, $i = 1, 2, \dots, n$ 이라고 할 때, 개체 간 유사성을 정의하는 제곱 유클리드거리는 $\sum_{j=1}^n (x_{ij} - x_{kj})^2$, 유클리드 거리는 $\sqrt{\sum_{j=1}^n (x_{ij} - x_{kj})^2}$, 맨

해튼 거리는 $\sum_{j=1}^n |x_{ij} - x_{kj}|$, 체비쇼프 거리는 $\max_j |x_{ij} - x_{kj}|$, 민코프스키 거리는 $\left(\sum_{j=1}^n |x_{ij} - x_{kj}|^p\right)^{1/p}$ 이다(Yang, 2013).

계층적(hierarchical) 군집방법은 한 개의 대상으로 출발하여 군집이 타 군집을 포함시키거나 군집 간의 중복이 인정되지 않고 주위의 대상을 군집화하여 최종으로 한 개의 군집을 형성시켜 나뉘어가지 형태를 취한다. 군집 내 가까우면 유사성이 크며, 군집 내 멀면 비 유사성을 갖는 개념에서 다음과 같이 사용되는 대상들을 군집화하는 방법이 있다(Kaufman & Rousseeuw, 2005; Aldenderfer & Blashfield, 1985).

1) 최단연결법(single linkage method): 두 군집 간 최단거리를 군집 간 거리로 삼으며, 그 결과 군집의 형태가 가늘고 길게 되는 현상이 발생한다.
 2) 최장연결법(complete linkage method): 두 군집 간 가장 먼 거리를 군집 간 거리로 삼으며, 군집 간의 내부적 응집력에 중점을 두며 가장 유사성이 큰 군집을 묶는다.
 3) 중심연결법(centroid linkage method): 군집전체를 군집의 중심으로 대변하여 군집 간 거리를 계산하며 새로운 개체를 포함시키는 경우 군집의 중심과 가장 가까운 군집을 개체에 포함시킨다.
 4) 평균연결법(between- or within-groups linkage): 두 군집의 모든 개체 간 거리들의 평균을 군집 간 거리로 하거나 새로운 개체를 기존 군집에 포함시킬 때 새로운 개체를 포함한 모든 개체 간의 평균거리라 최소로 하는 군집을 포함시키는 방법이다.

반면에 K-평균 군집분석법은 계층적 군집방법에서 결정된 군집 수(K)에 근거하여 K개의 각 군집에 한 개씩 개체를 심고, 모든 개체를 각각 가장 가까운 군집중심을 찾아 배속시킨다. 그리고 새로운 군집중심을 계산한 후 변화가 없을 때까지 앞 단계를 반복하여 진행한다(Borg & Groenen, 2005; Savaresi & Boley, 2004).

2.3. 다차원분석

군집분석과는 달리 다차원척도법은 개체 간 상대적 유사성(또는 비유사성)거리를 근거로 하여 2차원평면상에 개체들을 포지셔닝 분석방법으로, 개체들의 물리적 거리뿐만 아니라 응답자들의 심리적 거리를 계산하여 주관적 등급을 적용시킬 수 있다(Jeong, 2013).

Torgerson(1952)은 처음으로 다차원척도법을 착안하여 2차원 공간에 개체 간의 상대적 거리를 나타내었고, Shepard(1962)는 비계량형 다차원척도법을 제안하였고, Kruskal(1964)은 이 방법을 정밀하게 보완시킨 비계량형 알고리즘을 제시하였다. Carroll & Chang(1970)은 가중 유클리드 거리를 사용한 알고리즘인 INDSCAL(individual differences scaling)을 제공하였으며, Takane,

Young & Lewycky(1977)는 INDSCAL 알고리즘을 수정 보완한 ALSCAL(alternating least squares scaling)을 제안하였다.

다차원척도법은 계량심리학분야에서 태동하였지만, 현재는 심리학, 마케팅을 비롯한 다양한 사회과학 분야에서 활용하고 있다. 세분화된 고객 사이의 회사에 대한 이미지 측정, 후보자들의 이미지 분석, 소비자들이 인식하고 있는 브랜드 및 제품 속성에 대한 분석, 소비자들이 중요시 인식하는 브랜드 및 특성파악, 광고효과 측정 등 경영학분야에서 그 응용도를 찾아볼 수 있다.

일반적으로 사용되는 측도의 유형은 다음과 같다.

- 1) 유사성(similarity) 측도 : 개체 간 유사성을 표현하는 측도로 값이 클수록 유사성이 강하며, 대표적으로 피어슨의 상관계수, 캔달의 타우, 스피어만의 상관계수 등이 해당된다.
- 2) 상이성(dissimilarity) 측도 : 개체 간 상이성을 나타내는 측도로, 값이 클수록 상이성이 크며, 대표적으로 유클리드 거리, 체비쇼프 거리, 민코프스키 거리 등을 이에 해당된다.
- 3) 근접성(proximity) : 개체 간 가까움을 나타내는 측도로, 유사성 측도 및 상이성 측도로 이용하기도 한다.

다차원척도법에서 이용되는 비유사성(dissimilarity)측도는 개체 간 상이성을 나타내는 것으로, 값이 클수록 상이함이 강하며, 대표적으로 유클리드 거리, 체비쇼프 거리, 민코프스키 거리 등을 들 수 있다.

개체 간 상이성 거리를 나타내는 유클리드 거리(d_{ij})는 각 개체에 대한 관측 $(x_{i1}, x_{i2}, \dots, x_{ip})$, $i, j = 1, \dots, n$ 으로부터 개체 i 와 j 사이의 거리로 아래와 같이 측정된다.

$$d(x_i, x_j) = \left\{ \sum_{k=1}^n (x_{ik} - x_{jk})^2 \right\}^{1/2}, \quad i, j = 1, \dots, n$$

이러한 상이성 거리의 개념을 착안하여, Kruskal은 부적합도(badness-of-fit)를 나타내는 측도로 스트레스(STRESS; standardized residual sum of squares)를 다음과 같이 제안하였다(Kruskal, 1964; Kruskal & Wish, 1977).

$$STRESS = \sqrt{\sum_{i < j} \{d_{ij} - \hat{d}_{ij}\}^2 / \sum_{i < j} d_{ij}^2}$$

여기에서, \hat{d}_{ij} 는 개체 i 와 j 간 유사성의 값인 d_{ij} 의 적합값(또는 추정된 거리)이다. STRESS 값은 개체들 간의 적합된 상대적 거리와 실제거리 간의 오차의 정도를 표현한다. 만일 고려한 자료의 정확도 및 일관성이 높을수록 두 값의 거리가 일치하게 되므로 0에 가까워질 것이다. 따라서 STRESS 값이 작을수록 좋으며, 5% 이내이면 대체로 좋은 것으로 20% 이상이면 나쁜 것으로 평가하는 것이 통론이다.

46	3.1	1.2	0.7	0.7	0.6	0.9	1.0	0.6	0.6	0.9	0.4	0.9	0.4
47	3.0	1.3	0.8	0.7	0.7	1.1	1.2	0.7	0.6	1.1	0.4	1.0	0.5
48	3.2	1.2	0.8	0.8	0.7	0.9	1.0	0.7	0.7	0.9	0.5	0.9	0.4
49	2.9	1.5	1.0	0.8	0.8	1.2	1.3	0.9	0.6	1.2	0.5	1.2	0.6
50	3.5	1.0	0.9	1.1	0.9	0.8	0.9	0.7	1.0	0.8	0.7	0.7	0.5
27	28	29	30	31	32	33	34	35	36	37	38	39	
27	0.0												
28	0.4	0.0											
29	3.2	0.7	0.0										
30	2.7	1.2	0.5	0.0									
31	2.9	1.0	0.3	0.1	0.0								
32	3.7	0.2	0.5	1.0	0.8	0.0							
33	3.9	0.1	0.6	1.1	0.9	0.1	0.0						
34	3.3	0.6	0.1	0.6	0.4	0.4	0.5	0.0					
35	2.6	1.4	0.7	0.2	0.3	1.1	1.3	0.7	0.0				
36	3.8	0.2	0.5	1.1	0.9	0.9	0.0	0.5	1.2	0.0			
37	3.0	1.0	0.4	0.4	0.2	0.7	0.9	0.3	0.4	0.8	0.0		

The numerical values above stand for the following:
 1: Shenzhen 2: Guangzhou 3: Hangzhou 4: Beijing 5: Shanghai 6: Jinhua 7: Zhuhai 8: Xiamen 9: Suzhou 10: Nanjing 11: Dongwan 12: Zhongshan 13: Jiaxing 14: Ningbo 15: Foshan 16: Wuhan 17: Wenzhou 18: Putian 19: Wuxi 20: Tainan 21: Fuzhou 22: Chengdu 23: Sanya 24: Taizhou 25: Changzhou 26: Changsha 27: Taizhong 28: Zhoushan 29: Zhengzhou 30: Huzhou 31: Shaoxing 32: Xian 33: Taiyuan 34: Hefei 35: Quanzhou 36: Haikou 37: Qingdao 38: Kelamayi 39: Jinan 40: Dalian 41: Tianjin 42: Nanchang 43: Lishui 44: Gaoxiang 45: Zhengjiang 46: Huzhou 47: Yangzhou 48: Huizhou 49: Nantong 50: Kunming

<Table 1>에 주어진 값은 전자상거래 발전성과를 나타내는 세 개의 평가변수에 근거하여 주요도시 간의 상이성 정도를 거리로 계산하여 행렬로 표현한 것으로, 값이 커질수록 두 도시 간의 상이성이 크다는 것을 의미한다.

첫 세 개의 도시인 선전(1), 광주(2), 항주(3)끼리는 서로 다른 도시들에 비해 아주 작은 값을 가지며, 순위가 낮은 도시와는 달리 월등히 큰 값을 가진다. 이들은 서로 상대적으로 동일한 특성을 가지는 반면에 나머지 도시들과는 상이성이 두드러지게 나타남을 알 수 있다.

특히, 대남(20), 대중(27), 고웅(44)끼리, 그리고 대중(27)은 주산(28)과 서로 작은 값을 가지므로 이들끼리는 상대적으로 동일한 성격을 가진 집단으로 간주할 수 있으며, 반면에 이들은 나머지 도시들과는 월등히 큰 값을 가지기 때문에 그만큼 상이성이 크다고 간주할 수 있다. 또한 앞에서 언급한 도시들을 제외한 나머지 도시들은 서로 비슷한 작은 값을 가지기 때문에 그들 간의 상이성 정도가 작다고 판단할 수 있다.

종합적으로 상이성 정도를 단계적으로 살펴보면, 첫 번째 집단은 선전(1), 광주(2), 항주(3)로 현저하게 큰 값들을 가지며, 두 번째 집단으로는 대략 북경(4), 상해(5), 금화(6), 주해(7), 하문(8)으로, 나머지 22개 도시는 세 번째 집단으로 50개 주요 도시를 크게 세 개의 집단으로 구분할 수도 있을 것이다.

3.2. 군집분석

3.2.1. 계층적 군집방법

계층적 군집방법을 이용하여 군집화 단계표에서 계수가 급속하

게 변하는 시점에서 적절한 군집의 수를 결정할 수 있다. 다시 말해서, <Table 2>에서 이러한 상대적으로 큰 변화가 감지되는 시점을 네 시점(즉, 단계 45에서 46, 단계 46에서 47, 단계 47에서 48, 단계 48에서 49)을 발견할 수 있기 때문에 네 개 군집으로 유형화할 수 있다.

3.2.2. K-평균 군집방법

1) 소속군집과 최종군집중심

앞에서 다룬 상이성 거리행렬에서 지적인 바와 같이 군집 1에 속한 세 개의 도시들은 선전(1), 광주(2), 항주(3)이다. 이들은 다른 군집에 비해 세 개의 모든 발전성과속성에 대해 월등한 위치에 있는 반면에, 대다수 32개 도시를 포함한 군집 4는 다른 군집에 비해 열등한 위치에 있다. 특히 인터넷상인지수와 전자상거래발전지수는 다른 군집에 비해 매우 저조하다고 판단된다.

<Table 2> Agglomeration schedule

stage	cluster combined		coefficients	stage cluster first appears		next stage
	cluster 1	cluster 2		cluster 1	cluster 2	
1	41	42	0.002	0	0	13
2	21	22	0.003	0	0	9
3	33	36	0.009	0	0	6
4	17	18	0.009	0	0	20
5	2	3	0.013	0	0	43
6	32	33	0.014	0	3	14
7	46	47	0.014	0	0	10
8	12	13	0.016	0	0	33
9	21	26	0.017	2	0	16
10	46	48	0.024	7	0	21
11	37	39	0.028	0	0	22
12	29	34	0.029	0	0	28
13	40	41	0.033	0	1	27
14	32	38	0.036	6	0	23
15	30	31	0.037	0	0	24
16	19	21	0.041	0	9	25
17	43	45	0.046	0	0	22
18	9	10	0.055	0	0	30
19	11	14	0.056	0	0	30
20	15	17	0.065	0	4	35
21	46	49	0.072	10	0	31
22	37	43	0.072	11	17	27
23	28	32	0.078	0	14	34
24	30	35	0.094	15	0	26
25	19	23	0.096	16	0	29
26	24	30	0.131	0	24	38
27	37	40	0.141	22	13	31
28	25	29	0.145	0	12	29
29	19	25	0.174	25	28	34
30	9	11	0.178	18	19	33
31	37	46	0.251	27	21	32
32	37	50	0.291	31	0	38
33	9	12	0.397	30	8	39
34	19	28	0.442	29	23	41
35	15	16	0.542	20	0	39
36	7	8	0.595	0	0	42
37	4	5	0.599	0	0	42
38	24	37	0.617	26	32	41
39	9	15	0.709	33	35	45
40	20	27	0.735	0	0	44
41	19	24	0.853	34	38	45

42	4	7	1.143	37	36	46
43	1	2	1.444	0	5	47
44	20	44	2.161	40	0	48
45	9	19	2.746	39	41	48
46	4	6	4.146	42	0	47
47	1	4	6.793	43	46	49
48	9	20	11.45	45	44	49
49	1	9	14.335	47	48	0

군집 2에 속한 도시는 대남(20), 대중(27), 고웅(44)이며, 이미 <Table 2>의 상이성 거리행렬에서 이들 간의 유사성을 확인한 바 있다. 군집 2에 속한 도시들은 다른 군집에 속한 도시들에 비해서 인터넷상인지수가 가장 우월한 위치에 있는 반면에 인터넷쇼핑지수는 가장 열등한 위치에 있다.

마지막으로, 군집 3에 속한 12개의 도시들(북경(4), 상해(5), 금화(6), 주해(7), 하문(8), 소주(9), 남경(10), 둥관(11), 중산(12), 가흥(13), 영파(14), 불산(15))은 군집 1에 속한 도시들에 비해 세 개의 발전성과 속성 모두에 대해 열세에 있지만, 모두 작은 양 값을 가지므로 군집 1의 차세대 도시로 기대할 수 있을 것이다. 다시 말해서 군집 2에 속한 도시들은 인터넷쇼핑지수를, 군집 4는 모든 평가속성을 높임으로써 군집 1과 같은 도시로 전자상거래 발전성 과를 향상할 수 있을 것으로 예상된다.

<Table 3> Final clustering centers

variables to be evaluated	cluster			
	1	2	3	4
internet businessman index	2.06	2.10	0.51	-0.58
internet shopping index	2.36	-2.16	0.75	-0.30
e-commerce development index	2.87	-0.48	0.84	-0.54

2) 군집중심간 거리와 분산분석

각 군집 간 평균차이를 조사하면, (군집 1과 군집 2)와 (군집 1과 군집 4)는 차이가 크며, (군집 3과 군집 4)는 다른 집단들에 비해서 상대적으로 유사한 집단임을 알 수 있다. <Table 3>에서 전반적으로 양의 점수가 가장 큰 군집 1은 군집 3을 제외한 다른 군집보다 거리가 훨씬 멀리 떨어져 있다(Table 4 참고).

<Table 4> Distance between final cluster centers

cluster	1	2	3	4
1				
2	5.622			
3	3.011	3.564		
4	5.053	3.260	2.043	

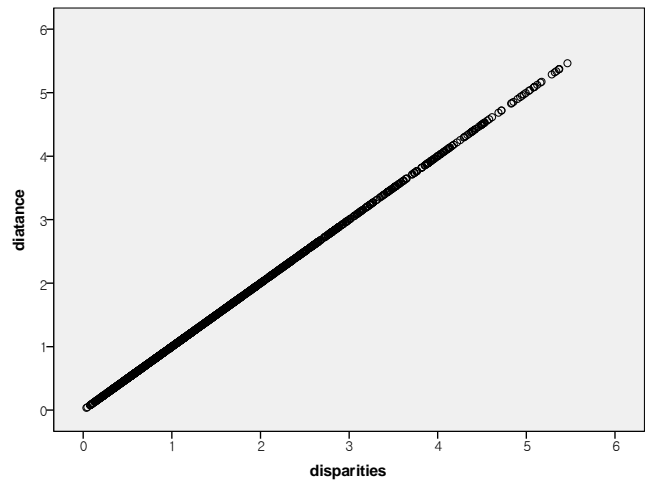
<Table 5> Analysis of variance

variables to be evaluated	cluster		error		F	Sig.
	MS	df	MS	df		
internet businessman index	13.156	3	.207	46	63.496	.000
internet shopping index	13.423	3	.190	46	70.727	.000
e-commerce development index	14.366	3	.128	46	111.972	.000

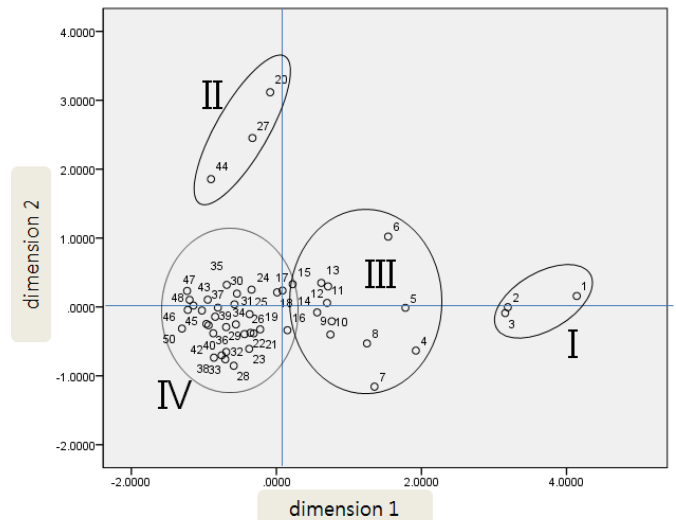
<Table 5>는 세 개의 평가변수에 대해 네 개의 군집별 평균차이를 검정한 것이며, 세 개의 평가속성 모두 네 개의 군집 분류에 유의한 변수임을 재확인할 수 있다($p < 0.001$). 따라서 세 개의 속성 모두 네 개의 집단에 대한 차이를 크게 식별하고 있음을 결론지을 수 있다.

3.3. 다차원척도법

<Figure 1>은 주요도시 간 근접성 변환값과 2차원 상대적 공간 상에 타점된 주요도시 간의 거리를 플롯한 것으로, 대응점들이 직선으로 위치하고 있는 것을 시각적으로 확인할 수 있다. 이는 <Figure 2>의 다차원척도 맵이 도시들을 매우 정확히 표현하고 있음을 의미한다.



<Figure 1> Scatter plot of linear fit



<Figure 2> Derived stimulus configuration

<Figure 2>는 2차원 다차원척도에서 중국 주요 50개의 도시들이 위치할 상대적 좌표를 계산한 후 2차원 산점도로 나타낸 것이

다. 다차원공간상의 도시별 좌표를 구한 후 계산된 STRESS는 '0.00001'이므로 채택할 수 있다고 판정할 수 있다.

각 속성들은 원점에서 가로축 또는 세로축 방향으로 멀어질수록 각 속성의 특성이 강하게 나타난다. 또한 발전성과 평가항목속성과 도시간의 관계는 수직 또는 수평으로 이동하였을 때 가깝게 접근하는 도시일수록 발전성과 평가속성의 속성을 많이 가진 도시로 판정할 수 있을 것이다.

<Figure 2>의 2차원평면에 표시된 I, II, III, IV는 각각 3.2절에서 분류한 네 개의 군집(군집 1, 군집 2, 군집 3, 군집 4)를 표현한 것이며, 네 개의 타원은 각 군집을 시각적으로 나타낸 것이다.

세 개의 발전성과 속성 중 인터넷상인지수는 원점을 중심으로 세로축의 위로 향할수록, 인터넷쇼핑지수는 세로축의 아래로 향할수록, 전자상거래발전지수는 가로축의 오른쪽으로 갈수록 발전성과 속성이 강하게 나타나며, 각 속성들이 그 반대방향으로 향할수록 발전성과 속성이 약해진다.

이러한 사실에 근거하여 대중(27), 고웅(44)과 함께 세로축 맨 끝에 놓여있는 대남(20)은 인터넷상인지수가 다른 도시에 비해 월등히 큰 반면에, 주산(28), 주해(7), 삼야(23), 서안(32), 극람마의(38) 등은 상대적으로 작은 인터넷상인지수를 갖는 것으로 판단할 수 있다. 또한 광주(2), 항주(3)와 더불어 가로축의 맨 오른쪽에 위치한 선전(1)은 무호(46), 양주(47), 해주(48), 곤명(50) 등과 상반되게 큰 전자상거래발전지수값을 가지는 것으로 추측할 수 있다. 마찬가지로 군집 2에 속하는 도시들(대남(20), 대중(27), 고웅(44))을 제외한 나머지 도시들은 세로축을 아래에 위치하여 상대적으로 이들에 비해 인터넷쇼핑지수가 높다는 것을 알 수 있다.

이러한 점을 고려하면 각각 다른 도시들의 위치를 판단하여 세 개의 발전성과지표를 그 상대적인 우위 및 향후 방향성을 전략적으로 활용할 수 있을 것이다(Jeong, 2015).

4. 결론

본 연구에서는 아리연구원(Sheng et al., 2014)의 연구자료에 근거하여 국가정보센터의 '중국정보사회지수(ISI)', '아리바바 전자상거래발전지수(AEDI)'에서 언급한 세 종류의 전자상거래 발전성과 속성을 기준으로 두 종류의 군집분석을 단계적으로 적용시켜, 중국의 주요 50개 주요도시를 유사성 있는 네 개 군집으로 분류하였다. 그리고 다차원척도법을 사용하여 각 50개 주요도시와 서로 비교하여 어떠한 이미지 속성을 가지는지 살펴보았다.

본 연구에서 상이성거리행렬, 군집분석 및 다차원척도법을 실시한 분석결과에 근거하여 요약하면 다음과 같다.

첫째, 50개의 주요도시를 네 개의 군집으로 뚜렷하게 유형할 수 있었으며, 그 중 맨 앞에 있는 선전은 광주, 항주와 더불어 세 가지 평가속성인 인터넷상인지수, 인터넷쇼핑지수, 전자상거래발전지수 모두 우위에 있음을 알 수 있었다.

특히 선전의 경우를 살펴보면, 2009년 중국의 국가발전개혁위원회와 상무부는 선전시를 국가 전자상거래모범도시로 선포하고 노력을 기울인 결과, 다음과 같은 특징을 갖게 되었다.

- 1) 2013년 현재 전자상거래 교역액이 약 50%, 전국평균 보다 속도가 33.5% 증가하여 높은 성장의 특징이 나타났다.
- 2) 높은 침투성이다. 2013년 선전시 경제무역 및 정보위원회의 자료에 따르면, 인터넷쇼핑인구수가 500만 명을 초과하고 인터넷 교역액이 889억 위안되어, 선전시의 소비총액의 20.1% 전국의 7.8%를 차지하였다.
- 3) 높은 혁신성이다. 선전시의 전자상거래발전을 위해서 정책, 상업, 기술 등 다방면을 혁신하였으며, 한 예로, 선전시장감

독관리국에 전자상거래회사 등록할 때 특정 사무실 없어도 허가를 내주었다. 강도 깊은 혁신을 추진한 3년 만에 "중국 일반무역수출사업백위" 중에 전국 94위에서부터 전국 5위로 급부상하게 되었다(Sheng et al., 2014).

둘째, 대남(20), 대중(27), 고웅(44)은 다른 도시들에 비해서 인터넷상인지수가 가장 우월한 위치에 있는 반면에 가장 열등한 위치인 인터넷쇼핑지수를 향상하기 위해서는 선전과 같이 다각적인 혁신성이 활성화되어야 할 것이다.

셋째, 12개의 도시들(북경(4), 상해(5), 금화(6), 주해(7), 하문(8), 소주(9), 남경(10), 등관(11), 중산(12), 가흥(13), 영파(14), 불산(15))은 앞에서 선전이 보여준 세 가지 특징을 갖도록 총력을 기울여 차세대 전자상거래 발전도시로 도약해야 할 것이다.

최근 중국에서 온라인을 통한 공동구매거래가 증가함에 따라 B2C 뿐만 아니라 B2T(온라인공동구매/소셜커머스)가 대도시 청년층을 중심으로 새로운 구매형태로 지속적으로 증가하는 추세에 있다. 수출에 의존하는 한국입장에서는 중국 소비자의 소비성향을 파악하여 맞춤형 서비스를 제공하면서 온라인 쇼핑시장의 시장점유권을 넓혀 나가야 할 것이다. 이와 더불어 온라인 쇼핑의 기반이자 최우선적으로 개선해야 할 물류인프라를 확충시켜야 할 것이다. 이를 위해서는 스마트 물류 네트워크 시스템을 보다 정밀도 있게 구축한 후 제조, 택배, 유통업체 간 공급망 사슬(supply chain)의 공동발전을 기대해야 할 것이다(Han et al, 2013). 또한 전 세계적으로 해외직접구매가 급증하면서 유통의 패러다임이 무너지고 있다. 따라서 중국 소비자의 기호와 구매형태를 파악하여 중국의 해외직접구매를 견양하여 온라인 쇼핑몰을 다각화로 브랜드화해야 할 것이다(KPMG, 2014).

중국내 제한된 전자상거래 자료로 인하여, 본 논문에서는 전자상거래 발전성과를 나타내는 세 개의 평가속성만을 고려하였다. 향후에는 더욱 연관성 있는 속성들을 추가로 포함시킴으로써 보다 설명력있는 유용한 분석결과를 도출할 수 있을 것이다.

Reference

- Aldenderfer, M. S., & Blashfield, R. K. (1985). *Cluster analysis*. Los Angeles, CA: Sage Publications.
- Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling*(2nd ed.). New York, NY: Springer-Verlag.
- Carroll, J. D., & Chang, J. J. (1970). Generalization of the singular value (Eckart-Young) decomposition to N-way tables. *Psychometrika*, 35, 238-319.
- Han, J. P., Park, J. H., & Kim, Y. S. (2013). A study on the problem and improvements of the B2C E-commerce logistics services in China. *Electronic trading Reviews*, 11(1), 1-25.
- Jeong, D. B. (2013). A study on universities positioning using multidimensional scaling. *Korean Business Education Review*, 28(2), 1-15.
- Jeong, D. B. (2014). Evaluation of research performances for 28 national universities. *Journal of Korean Data & Information Science Society*, 25(6), 1241-1251.
- Jeong, D. B. (2015). A study on cluster and positioning of domestic electronic commerce based on purchasing motivation. *Journal of Korean Data & Information Science Society*, 29(4), 841-856.

- Jeong, I. H. (2012). A study on the causal relationship of educational performance factors for university educational competence - focus on Malcolm Baldrige national quality award model-. *Korean Business Education Review*, 27, 207-239.
- Jhun, I. H. (2014). The Recent development status and prospective of E-commerce in China. *Oversea Economic Focus*, 2014(30), 1-14.
- Kaufman, L., & Rousseeuw, P. (2005). *Finding groups in data – An introduction to cluster analysis* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.
- Kim, H., & Seo, S. S. (2012). A study on the performance of online community of purchasing motivation by smart phone. *Journal of Korean Electronic Commerce Research Association*, 13, 41-57.
- Kruskal, J. B. (1964). Major MDS based on a firm numerical analysis foundation. *Psychometrika*, 29, 1-27.
- Kruskal, J. B., & Wish, M. (1977). *Multidimensional Scaling*. Beverly Hills, CA: Sage Publications.
- Lee, G. H., & Kim, P. S. (2013). Effect of savings rate by adolescent consumption habit and business/economic education. *Korean Business Education Review*, 28, 363-377.
- Lee, K. Y., Lee, H. J., Park, D. H., & Kim, S. K. (2014). Online shopping patterns and new change in China. *KPMG International*, April, 1-36.
- Lee, L. J. (2011). A study on the purchase Intent of Internet Shopping Mall. *Journal of Korean Electronic Commerce Research Association*, 12, 93-106.
- Lee, J. Y. (2013). The current on-line shopping market and its prospect based on internet and mobile. *KISDI*, 25, 96-108.
- Savarese, S. M., & Boley, D. (2004). A comparative analysis on the bisecting k-means and PDDP clustering algorithm. *Intelligent Data Analysis*, 8, 345-362.
- Shanghai Assetplus (2015). *E-commerce industry in China* (pp.1-10). Retrieved March 10, 2015, from <https://www.chinawindow.co.kr/download.php?category=spt6...pd>
- Sheng, Z., Chen, L., & Zhang, R. (2014). *Report on 2013 development performances of e-commerce for 50 major cities in China* (pp. 25-26). Retrieved June 4, 2014, from <http://www.aliresearch.com>
- Shepard, R. N. (1962). Nonmetric algorithm. *Psychometrika*, 27, 219-246.
- Takane, Y., Young, F. W., & DeLeeuw, J. (1977). Combined all previous major MDS developments into a single unified algorithm. *Psychometrika*, 42, 7-67.
- Torgerson, W. S. (1952). Multidimensional scaling: 1. Theory and method. *Psychometrika*, 17, 401-419.
- Yang, B. H. (2013). *Understanding multivariate analysis*. Seoul, Korea : Communication books.