

공간적 패턴클러스터링을 위한 새로운 접근방법의 제안 : 슈퍼마켓고객의 동선분석

정인철 · 권영식[†]

동국대학교 산업시스템공학과

A New Approach to Spatial Pattern Clustering based on Longest Common Subsequence with application to a Grocery

In-Chul Jung · Young S. Kwon

Department of Industrial and Systems Engineering, Dongguk University-Seoul

Identifying the major moving patterns of shoppers' movements in the selling floor has been a long-standing issue in the retailing industry. With the advent of RFID technology, it has been easier to collect the moving data for a individual shopper's movement. Most of the previous studies used the traditional clustering technique to identify the major moving pattern of customers. However, in using clustering technique, due to the spatial constraint (aisle layout or other physical obstructions in the store), standard clustering methods are not feasible for moving data like shopping path should be adjusted for the analysis in advance, which is time-consuming and causes data distortion. To alleviate this problems, we propose a new approach to spatial pattern clustering based on longest common subsequence (LCSS). Experimental results using the real data obtained from a grocery in Seoul show that the proposed method performs well in finding the hot spot and dead spot as well as in finding the major path patterns of customer movements.

Keyword: customer path, shopping behavior, exploratory analysis, LCSS, RFID

1. 서론

대형 유통 매장은 판매를 통한 매출 이익 증대를 가장 큰 목표로 한다. 그러나 최근 대형 유통매장 시장은 지속적인 가격 경쟁 심화와 신규 점포의 경쟁적 확장 그리고 경기침체로 인한 소비 감소 현상이 겹치면서 점포 수 증가에 비해 성장률은 미진한 상황이다. 이는 더 이상 규모의 경제를 통한 판매 가격 인하만을 추구해서는 경쟁력이 없음을 의미함과 동시에 대형 유통 매장의 해외 진출과 내실 강화에 주력해야하는 시대적 상황을 보여 주고 있다. 따라서 비용을 줄이며 운영효율을 극대화 하고 고객 서비스 고도화를 통한 구매 유도는 매우 중요한 과제이다.

이를 위해 매장 관리자들은 세밀한 판매 전략을 세우고 고객에게 적극적인 판매 활동을 전개하고 있는데, 대표적으로 구매 이력이나 인구통계학적 정보를 활용한 장바구니 분석이나 상권분석 등이 이러한 시도들이다. 예로써, 이들은 고객의 구매 이력 정보를 기반으로 고객의 관심 상품 정보를 분석하여, 고객 세분화를 통한 개별 맞춤 상품 추천뿐만 아니라 상권 지역의 인구 통계학적 정보와 융합하여 수익성 높은 점포의 위치를 파악하기 위해 노력하고 있다. 그러나 고객들에 대한 이해를 높일 수 있는 쇼핑 시의 고유한 행동 패턴이나 습관을 파악하기 위해서는 기존의 구매 이력과 인구통계학적 정보만을 가지고는 부족한 면이 있다. 특히 고객의 소비 의사결정에

본 연구는 지식경제부 산업원천기술개발사업(10035352) 수행의 일환으로 이루어진 것임.

[†]연락처 : 권영식, 100-715 서울특별시 중구 필동 3가 26번지 산업시스템공학과,

Fax : 02-2269-2212, E-mail : yskwon@dgu.edu

투고일(2011년 07월 11일), 심사일(1차 : 2011년 10월 19일, 2차 : 2011년 10월 30일), 게재확정일(2011년 10월 31일).

관한 연구에 따르면 고객 소비 행동이 매장의 환경(매대배치, 상품배치 등)에 밀접한 영향을 받는 것으로 알려져 있기 때문에 고객 행동 분석을 통한 매장 환경 개선이나 매출 증대 방안 등에 대한 관심이 높아지고 있는 상황이다(Newman *et al.*, 2002). 예를 들어, 매장 환경 개선을 위해 현재 매장 내의 어느 지역이 판매가 활발히 이루어지고 있으며, 고객들이 붐비는 매대는 어느 지점인지, 그리고 고객들이 오랫동안 구경하는 매대는 어느 곳인가와 같은 정보를 근거로 매대 배치를 변경하거나 상품 배치 위치를 변경함으로써 소비 촉진을 유발 할 수 있다.

이와 같은 상황에서 고객 소비 행동 가운데 하나인 쇼핑 동선을 분석하기 위해, 매장의 일부 고객들을 대상으로 연구자가 따라다니며 기록하여 그 데이터를 바탕으로 매장 내의 동선 분석을 시도한 연구가 이루어졌다(Farley and Ring, 1996). 그러나, 최근의 매장 환경은 일일 이용 고객 수가 너무 많기 때문에 연구자가 따라다니며 개개인별 소비 동선을 파악하는 것은 현실적으로 가능하지 않다. 이와 같은 문제를 해결하기 위해 RFID 기술과 클러스터링 기법을 이용하여 고객의 쇼핑 동선을 분석하는 연구들(Larson *et al.*, 2005; Hui *et al.*, 2009; Gil *et al.*, 2009)이 진행 되었으나, 매장 내의 고객 동선 분석에 클러스터링을 적용하는 데는 다음과 같은 문제점이 있다. 우선, 고객의 쇼핑 동선을 클러스터링 하는 과정 중에 클러스터가 매대나 상품 가판대와 같은 장애물의 위치에 생성되거나, 장애물을 중심으로 나누어져 있는 동선들이 같은 클러스터로 군집되는 문제점이 있다. 고객들은 매대 사이의 통로로만 이동하지 매대나 장애물을 넘어서 이동할 수는 없기 때문에, 매장 내의 물리적인 환경이나 장애물(매대, 상품 가판대 등)과 같은 제약조건을 고려해야 한다. 또한 클러스터링 적용을 위해서는 모든 고객의 이동거리가 동일해야 하는 문제점이 있다. 그러나 실제 매장 내의 고객들의 이동 거리는 매우 다양하다. 예를 들어, A라는 고객은 5분 만에 쇼핑을 끝내고 나갈 수 있는 반면, B라는 고객은 30분 동안 천천히 구경을 하면서 매장을 구경할 경우 A 고객과 B 고객의 이동 경로는 서로 매우 상이하기 때문에, 이를 클러스터링 알고리즘에 적용을 하기 위해서 같은 이동 길이로 변환해야 하는데, 이러한 과정에서 정보의 손실 또는 왜곡이 발생할 수 있다.

따라서, 본 연구에서는 클러스터링 기법의 적용상의 문제점을 해결하기 위해 이동 경로 그래프로 동선 정보를 변경함으로써 물리적 제약조건들을 해결하는 동시에 고객 이동 경로의 길이가 상이한 경우에도 데이터 변형 없이 고객 이동 정보간의 유사성을 측정할 수 있는 공간적 패턴 발견 기법을 제안한다. 이를 위해 최장 공통 부분열(Longest Common Subsequence : 이하 LCSS) 알고리즘에 기반을 둔 주동선 발견 알고리즘을 개발하고 서울의 한 대형 슈퍼마켓에 적용하여 결과를 분석하고자 한다.

2. 관련 연구

대형 유통 매장은 고객들에게 보다 많은 상품을 싸게 제공함

으로써 수익을 창출해 왔고 최근 들어서는 타 업체와 경쟁 우위를 위해 다양한 마케팅 전략을 구사해왔다. 이를 위해 고객에 대한 이해는 필수적인 것이며, 주로 인구 통계학적 정보나 구매 정보를 활용해 진행 되어 왔다. 그러나 매장 내에서 발생하는 고객들의 행동 패턴을 발견하기 위해서는 거래 내역 정보만을 가지고는 알 수가 없는데, 이는 고객 행동에 대한 데이터 수집 기술의 한계에 의한 것으로 고객 정보와 고객 거래 데이터를 통해서 고객 행동과 관련하여 특정 고객이 해당 매장에 몇 번 방문하였는지, 또는 어느 시점에 방문하였는지 등의 단순한 정보만을 알 수 있기 때문이다.

고객 행동 정보를 얻기 위해서 주로 직접관찰이나 설문조사 방식을 통하여 상품 디스플레이에 따른 매출액이나 브랜드 변경(Harris, 1958; McClure and West, 1969), 비계획적인 구매에 대한 요인의 탐색(Cox, 1964), 구매 빈도의 정도와 고객의 인구 통계학적 특성(David and Ronald, 1967), 매장 내 정보 프로세싱과의사결정(Dickson and Sawyer, 1986), 매장에 대한 지식과 쇼핑가용 시간이 계획된 구매의 실패가 일어나는 정도, 비계획적인 구매의 수준, 브랜드나 제품의 전환, 구매량의 확장과 같은 매장 내 고객의 쇼핑 행동에 미치는 영향(Hoyer, 1984)을 파악하는 연구가 진행 되었다. 그러나 매장 전체 이용 고객들을 대상으로 할 수는 없었고, 일부 고객들을 선별하여 실시한다 하더라도 많은 비용이 드는 문제점이 있었다. 최근에는 RFID나 영상 카메라 등의 기술 발전으로 인해 보다 적은 비용으로 많은 고객을 분석할 수 있게 되어 고객 행동 분석에 관한 연구들이 활발히 진행되고 있다(Larson *et al.*, 2005; Uotila and Skogster, 2007; Gil *et al.*, 2009; Hou and Chen, 2011). 이러한 연구를 통해 보다 쉽게 고객의 특성에 따른 매장 내 행동의 관계에 대한 기초적인 이해를 할 수 있게 되었으며, 이를 매장 운영이나 마케팅에 활용 가능성을 제시 하고 있다. 특히, Larson *et al.*(2005)와 Hui *et al.*(2009)는 RFID 기술과 유클리드 거리 기반의 K-medoids 기법을 사용하여 주요 동선을 파악하고자 하였다(Larson *et al.*, 2005; Hui *et al.*, 2009). 그러나 유클리드 거리 기반의 통계적 클러스터링의 경우 클러스터 과정 중 클러스터가 매대나 상품 가판대와 같은 장애물의 위치에 생성되거나, 장애물을 중심으로 나누어져 있는 동선들이 가까운 거리에 위치했을 경우에도 같은 클러스터로 군집되는 경향이 있기 때문에, 동선 군집 정확도를 높이기 위해서는 실제 매장 내의 장애물(매대, 가판대 등) 환경에 대한 추가적인 고려가 필요하다. 예를 들어, <그림 1> 에서와 같이 a 위치에서 b와 c의 위치 중에서 유클리드 거리의 거리 측정을 하게 되면, c가 b보다 a에 보다 가까운 위치가 된다. 하지만, 실제 매장에서 고객이 지나가는 동선을 기준으로 본다면, 실제로는 a에서 c는 매대를 건너가야 하기 때문에 b에 비해 더욱 먼 거리를 걸어야 도달할 수 있는 위치에 있기 때문이다.

또한 유클리드 거리 기반 클러스터링 기법을 사용할 경우 알고리즘 특성상 클러스터 입력 변수의 개수가 동일해야 하는데, 쇼핑 동선의 경우 <그림 2>의 좌측 그림과 같이 고객에 따

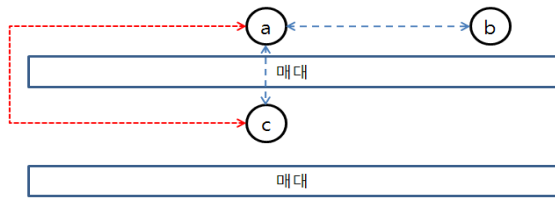


그림 1. 매장 환경에서의 유클리드 거리 사용 문제점

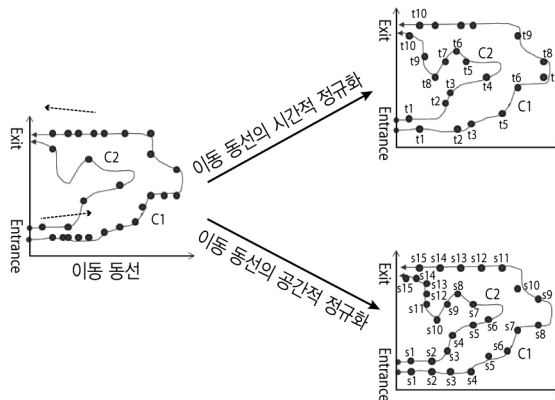


그림 2. 동선의 정규화 기법

라 여러 동선 형태(C1, C2)로 나타나 방문 지점의 수나 이동 거리가 각각 다르다. 이를 위해, 동선 정규화(normalization)기법이 적용되었으며(Yanagisawa et al., 2003), <그림 2>의 우측 그림과 같이 시간적 정규화(temporal normalization) 기법이나 공간적 정규화(spatial normalization)기법으로 동선의 길이를 똑같은 수로 맞추어 주어야 한다. 시간적 정규화란 고객의 이동 전체 시간을 기준으로 일정 방문 지점 수로 정규화 하는 방법이며, 공간적정규화란 전체 이동 거리를 기준으로 같은 방문 지점 수로 정규화 하여 입력 값의 수를 같게 하는 방법이다. 이러한 정규화의 문제점은 실제 이동한 원래의 동선과는 달라진 동선을 입력 값으로 취하게 되기 때문에 원래에 정보가 왜곡되거나 소실될 수 있는 문제점이 있다.

따라서 본 연구에서는 이러한 문제점을 보완할 수 있는 새로운 접근방법을 제안하고자 한다.

3. 고객 쇼핑 동선 패턴 분석

3.1 RFID를 이용한 데이터의 획득

고객의 구매행동분석을 위해서는 고객의 이동 동선 정보와 구매이력 정보가 필요하기 때문에, 구매 현황 및 물품에 대한 정보는 기존의 POS 시스템과 주문 DB를 통해서 수집하고, 고객 개인마다 매장에서 이루어지는 이동 동선 정보를 수집하기 위해서 매대 곳곳에 RFID 리피터(repeater)와 매장 천정에 리더기 역할을 하는 싱크 노드(sink node)를 설치하고 고객의 쇼핑 카트에 RFID 태그를 설치하였다(<그림 3> 참조).

동선 정보 획득을 위해 2.4GHz 주파수 대역을 사용하는 Ultra-low Power Wireless System One-Chip을 능동형 태그 형태로 개발하였다. 능동형 태그는 쇼핑 카트에 부착하여 일정한 주기로 데이터를 송신하며, 매장 내부의 주요 위치마다 리피터를 설치하여 수신 가능 영역에 쇼핑 카트가 지나갈 때 태그로부터 데이터를 수신하여 싱크노드로 전송을 하고 다시 정보를 동선 저장 서버에서 처리할 수 있도록 데이터를 가공하여 서버로 전송하도록 RFID 시스템 설계를 하였다(<그림 4> 참조).

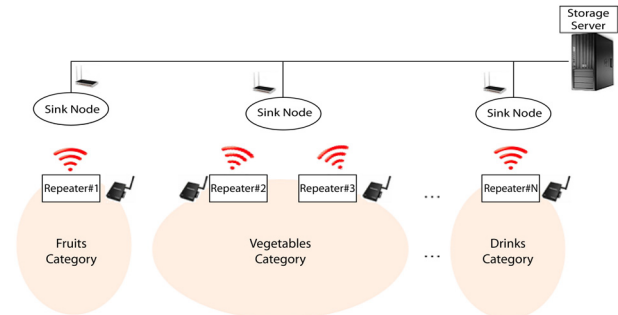


그림 4. 리피터와 RFID 리더기



그림 3. 매장과 쇼핑 가트에 설치된 RFID 장비

이와 같이 리피터를 통해 RFID 신호를 중계함으로 해서 모든 위치마다 RFID 리더기 장비를 설치하는 비용보다 적은 비용으로 시스템을 구성할 수 있는 장점이 있다. 리피터의 설치 위치는 진열 상품의 매장 분류기준(대, 중, 소, 세 등급)을 근거로 하여 설치 지점을 정하였고, 이때 RFID의 전파는 설치 지역의 위치와 설치 구조물 모양, 그리고 구조물의 종류에 따라 수신률이 달라질 수 있기 때문에, 위치별로 리피터의 수신 감도 조절을 하였고, 능동형 태그의 부착위치는 카트의 전면부 하단(<그림 3>)에 설치 구성하였다. 이를 통해서 고객의 쇼핑 카트에 장착된 능동형 RFID 태그를 통해 고객의 이동 위치 및 시간 데이터를 수집하게 된다. 쇼핑 카트에 장착된 태그는 고유한 ID를 갖고 있기 때문에, 고객이 쇼핑을 끝낸 후 계산대에서 계산을 한 후, 고객 정보(예를 들어, Family Card ID)와 태그 ID를 매핑 하여 개별 고객의 이동정보와 시간 및 구매 상품 정보 등을 매핑하게 된다.

3.2 최장 공통 부분열(Longest Common Subsequence)을 이용한 고객 동선 패턴분석

최장 공통 부분열 문제는 주어진 여러 개의 수열 모두의 부분 수열이 되는 수열들 중에 가장 긴 수열을 찾는 문제로(Hirschberg, 1977), 예를 들어 두개의 수열 $X = (x_1, x_2, \dots, x_m)$, $Y = (y_1, y_2, \dots, y_n)$ 이 있을 때, 최장 공통 부분열은 다음과 같이 정의된다.

$$LCSS(X_{1 \dots i}, Y_{1 \dots j}) = \begin{cases} 0 & \text{if } i=0 \text{ or } j=0 \\ LCSS(X_{1 \dots i-1}, Y_{1 \dots j-1}), x_i & \text{if } x_i = y_i \\ \max(LCSS(X_{1 \dots i}, Y_{1 \dots j-1}), LCSS(X_{1 \dots i-1}, Y_{1 \dots j-1})) & \text{otherwise} \end{cases} \quad (1)$$

이 문제는 최적 부분 구조(optimal substructure) 성질을 갖기 때문에, 동적계획법(dynamic programming)으로 최적해를 구할 수 있다. 우선, 두 문자 수열의 마지막 문자가 같다면, 이 문자는 최장 공통 부분열의 일부에 해당되며, 이 공통 문자를 두 문자 수열에서 삭제한다. 이로 인해 길이가 줄어든 문자 수열을 이용하여 다시 공통 부분열의 여부를 결정한다. 즉, $X_i = \{x_1, x_2, \dots, x_i\}$ 와 $Y_j = \{y_1, y_2, \dots, y_j\}$ 와 같은 문자 수열이 있고, 마지막 문자 x_i 와 y_j 가 똑같다면, $LCSS(X_i, Y_j) = \{LCSS(X_{i-1}, Y_{j-1}), x_i\}$ 처럼 x_i 는 최장 공통 부분열에 해당되며, 이를 제외한 나머지 문자열들로 작은 문제로 변환하여 다시 재귀적으로 최장 공통 부분열을 찾는다. 만일 마지막 문자 x_i 와 y_j 가 같지 않다면, $LCSS(X_{i-1}, Y_j)$ 와 $LCSS(X_i, Y_{j-1})$ 중에 보다 긴 길이의 공통 부분열을 도출하는 식을 선택하여 해를 찾게 된다. 예를 들어, 문자 수열 ACD와 AGD의 최장 공통 부분열을 찾기 위하여 마지막 문자부터 시작하여 계속해서 작은 길이의 문자 수열로 줄여가며 공통 부분열을 구하게 되는데, 이를 설명하기 위하여 <표 1>과 같이 ACD와 AGD를 테이블 형식으로 표현을 하였다. 재귀적으로 문자열 ACD와 AGD의 마지막 문자 D부터 최종적으로 A인 첫 번째 문자까지 공통 부

분열 여부를 확인해야하기 때문에, 문제를 풀기 위해 모든 열과 행을 처음부터 하나씩 증가시켜 가며 공통 부분열 여부를 찾는다. <표 1>의 행을 기준으로 문자열을 증가시키며 공통 부분열을 찾아보면, 가장 작은 단위의 문제로 ACD의 A와 AGD의 A인 $LCSS(A, A) = A$ 이며, 그다음의 $LCSS(A, AC)$ 나 $LCSS(A, ACD)$ 모두 A이다. 이와 같은 식으로 $LCSS(AG, A)$, $LCSS(AG, AC)$, $LCSS(AG, ACD) = A$ 이며, 마지막으로 $LCSS(AGD, A)$ 과 $LCSS(AGD, AC)$ 역시 A이며, $LCSS(AGD, ACD)$ 에서 D가 공통적으로 나타나며 앞의 A와 결합되어 AD가 최장 공통 부분열이 된다.

표 1. ACD와 AGD의 최장 공통 부분열 검색

	A	C	D
A	{A}	{A}	{A}
G	{A}	{A}	{A}
D	{A}	{A}	{AD}

이와 같이 공통적으로 중복되면서 가장 긴 길이를 갖는 부분열을 찾아내는 특성을 이용하여, 물체의 이동 지점의 나열을 문자 수열로 가정하고 동선 간에 이동 지점이 비슷한 동선을 찾아내는데 최장 공통 부분열을 활용할 수 있다.

최장 공통 부분열 알고리즘을 활용하여 공간상에서 이동하는 물체의 궤적간의 유사성을 파악하고자 하는 시도들이 있으며(Cao *et al.*, 2007; Kang *et al.*, 2008; Kim and Chang, 2008), 본 연구에서는 매장 내에서의 고객의 이동경로를 그래프로 표현하고, 최장 공통 부분열 알고리즘을 이용해 고객들의 동선을 클러스터링 하여 특징적인 동선패턴을 추출하고자 한다. <그림 5>는 매장의 통로에 설치된 RFID 리피터(repeater)를 노드로 하여 이동 경로를 그래프로 표현한 예이다. 고객의 쇼핑 동선은 방문한 노드의 순서에 따라 이동하는 위치에 있는 리피터 노드의 ID를 참조하여 표현된다.

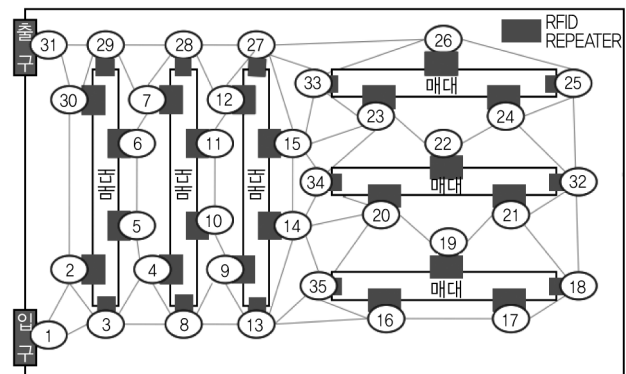


그림 5. 이동 경로 그래프 예

두 명의 고객이 이동 경로 그래프 상에서 이동한 노드 ID의 순서열의 유사성을 비교할 때, 두 고객의 최장 공통 부분열 길이가 크다는 것은 공통된 노드를 많이 지났다는 의미가 되고, 따라서 두 고객의 동선 패턴이 유사함을 나타내는 것이다. 그

러나 최장 공통 부분열 크기만을 고려하여 동선 유사도를 판단할 경우, 다음과 같은 문제점이 있을 수 있다. <그림 6>에서와 같이 고객 1의 동선이 A-B-C-D-E-F이고 나머지 고객 3명(고객 2, 고객 3, 고객 4)의 동선이 각각 A-B-E-Z-F, A-B-C-F, A-B-Y-Z-F와 같을 때, 최장 공통 부분열을 적용하면, 고객 1과 고객 2의 최장 공통 부분열은 ABEF, 고객 1과 고객 3의 최장 공통 부분열은 ABCF, 고객 1과 고객 4의 최장 공통 부분열은 ABF가 된다. 이때, 고객 2와 고객 3의 경우, 둘 다 공통 부분열 길이가 4로써 모두 같은 유사 정도를 보이기 때문에, 더 유사한 이동 동선을 갖는 고객을 변별해 내는 것이 어렵다.

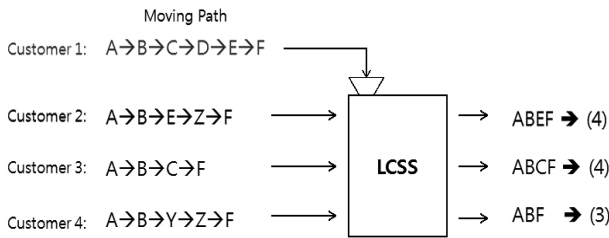


그림 6. 최장 공통 부분열(LCSS)의 예

변별력 있는 유사도 계산을 위해서는 최장 공통 부분열의 공통 부분 길이만을 기준으로 유사함을 판단하는 대신에 두 고객이 이동한 전체 이동 노드 길이 중에서 공통되는 최장 공통 부분열 노드의 길이 비율이 높을수록 보다 가까운 이동 패턴을 가지는 것으로 판단하는 것이 더 변별력이 있다고 볼 수 있다. 이를 위해 다음과 같은 최장 공통 부분열을 변형한 유사도 함수를 다음과 같이 제안하였다.

$$\text{이동패턴 유사도}(x, y) = \frac{LCSS(x, y)}{x\text{이동거리} + y\text{이동거리}} \quad (2)$$

제안한 이동 패턴 유사도함수는 보다 비슷한 이동 패턴을 가지는 동선들을 찾아내기 위하여, 두 고객의 전체 이동 거리 중에서 보다 많은 공통 노드를 지나는 동선들을 비슷한 동선 패턴으로 결정을 하는 것이다. 이 유사도 함수를 이용하여 위의 예를 재적용 해보면, <그림 7>에서와 같이 고객 1과 고객 3의 동선의 유사도는 $4/(6+4) = 0.4$ 로 고객 1과 고객 2의 동선의 유사도 $4/(6+5) = 0.36$ 보다 더 유사함을 알 수 있다. 고객 1과 고객 3은 최종적으로 각각 6개의 노드와 4개의 노드를 이동한 동선을 지났으며 최종적으로 4개의 공통 노드를 지나온 반면에, 고객 1과 고객 2의 경우는 각각 6개의 노드와 5개의 노드를 이동했으며 역시 공통적으로 4개의 노드를 지나왔다. 전체적으로 이동한 거리가 고객 1과 고객 3의 경우가 고객 1과 고객 2의 경우보다 짧음에도 불구하고, 공통된 이동 노드의 수는 4로 똑같기 때문에, 고객 1과 고객 3의 경우가 보다 유사한 이동 패턴을 가지고 있다고 판단할 수 있다. 이로써, 최장 공통 부분열 길이의 크기만을 비교하는 것에 비해, 전체 이동 길이 중에 공유되는 노드의 개수의 비율이 높을수록 보다 유사한 이동 패

턴으로 결정이 되도록 하는 것이다.

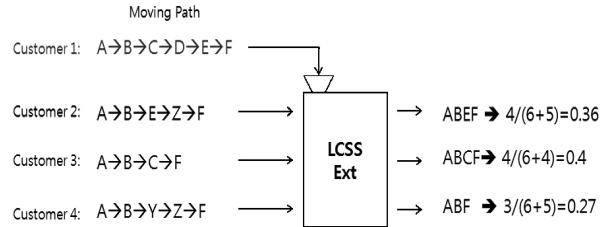


그림 7. 제안 이동 패턴 유사도

이와 같이 최장 공통 부분열을 확장한 유사도 측정 지표를 사용함으로써, 최장 공통 부분열이 가지고 있는 순차적인 이동 순서뿐만이 아니라 제약 환경들이 많이 있는 매장환경에서의 유클리드 거리 기반의 클러스터링 기법들의 한계점을 극복할 수 있는데, 유클리드 거리 기반의 클러스터링을 이용한 쇼핑 동선 패턴을 파악하는 경우 클러스터 과정 중 클러스터가 매대나 상품 가판대와 같은 장애물의 위치에 생성되거나, 장애물을 중심으로 나누어져 있는 동선들이 가까운 거리에 위치했을 경우에도 같은 클러스터로 군집될 수 있으나, 최장 공통 부분열은 이동 동선을 이동 그래프 노드로써 표현을 하고, 쇼핑 동선간의 최장 공통 부분열을 찾아내어 쇼핑 패턴을 분석하기 때문에 이러한 단점을 해결할 수 있다. 또한 이동거리가 상이한 동선 데이터를 비교하기 위하여 같은 거리로 전처리 없이도 동선간의 유사도를 측정할 수 있다.

본 연구에서 제안하는 공간적 주동선 패턴 클러스터링 알고리즘은 N개의 고객 쇼핑 동선들 가운데 주요 동선을 파악하기 위하여 다음과 같은 순서로 쇼핑 동선 패턴을 파악 한다. 모든 고객 동선 중에서 임의의 $K_{init} = \{k_1, k_2, \dots, k_m\}$ 의 초기 동선들을 시작으로 각각의 k_i 동선과 유사한 동선 그룹을 식 (2)를 사용하여 유사한 동선 그룹을 만든다. 그리고 다시 각 k_i 의 그룹 중에서 새로운 기준 동선 시퀀스를 추출하는데, 이때는 초기 k_i 와 가장 가까운 최장 공통 부분열을 가지는 동선이 k_i 로 설정되게 된다. 이와 같은 방식으로 반복하는 가운데, 각 집합이 변화가 생기지 않거나 사용자 정의한 수까지 반복을 하고나면 최종적으로 K 집합의 LCS 시퀀스가 도출되게 된다. <그림 8>은 제안한 동선 패턴 클러스터링 기법에 대한 pseudo 코드이며, <그림 9>는 알고리즘 동작 방식에 대한 표현이다.

1	Main-Shopping-Path-Pattern-Clustering(input : moving-paths, iterate-count)
2	randomly selecting K-initial Trajectory in moving-paths
3	
4	loop (stable? or user determined iterate-count)
5	for-each k in K
6	if (find the closest moving-path with k by using(2))
7	insert LCSS(closest moving-path, k) to k group
8	redefine K-initial Trajectory with first LCSS in each k groups
9	
10	return k

그림 8. 제안한 최장 공통 부분열 Pseudo code

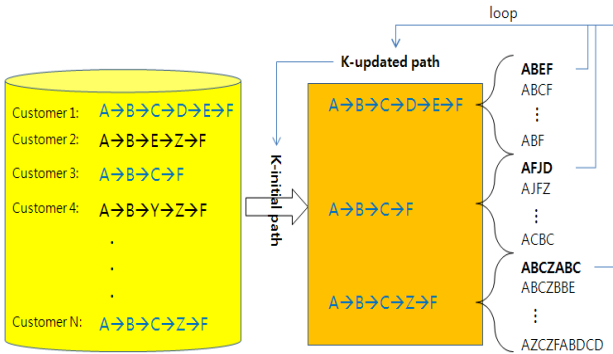


그림 9. 제안 알고리즘 구성

제안하는 주동선 발견 알고리즘의 특징으로는 매장 내 고객의 쇼핑을 위한 동선 시퀀스뿐만 아니라 이동 방문이 집중되는 부분과 이동 방문이 이루어지지 않는 지역을 알 수가 있는데, 매장 내에서 이러한 방문 빈도가 높은 주요 방문 지역(hot spot)을 인지하거나 반대로 상대적으로 방문이 적게 이루어지는 지역(dead spot)을 알아내는 것은 매장 관리자 입장에서 중요한 정보이다. 본 논문에서는 이러한 중요지점을 hot spot이라는 용어를 사용하여 표현을 하겠으며, 반대로 중복 방문이 발생되지 않는 지역을 dead spot이라는 용어를 사용하겠다. 제안 알고리즘은 최장 공통 부분열의 특성으로 인해 이동 순서를 나타내는 주요 공통 방문 시퀀스 집합이 도출되면서 고객들의 방문이 중복되는 노드들이 함께 나타나기 때문에, 최종적으로 도출된 동선패턴이 겹치는 지역을 hop spot이라 판단할 수 있으며, 그 반대로 동선 패턴이 나타나지 않은 지역을 dead spot이라 판단할 수 있다.

4. 실험 결과와 분석

4.1 적용 매장 환경

본 연구에서 제안한 고객 쇼핑 동선 패턴을 발견하기 위하여, 서울에서 운영되고 있는 대형 할인 매장의 발생 데이터를 이용하여 분석을 진행하였다. 대상 할인 매장은 전체 면적이 약 4,290m²의 단층 구조이며, 월 평균 고객 수는 약 218,794명 정도로 하루 평균 약 554명의 이용 고객을 보유하고 있으며 고객 당 평균 쇼핑 시간은 약 31분 정도이다. <그림 10>은 적용 매장의 매대 구조와 RFID 리피터가 설치된 현황을 보여 주고 있다. 능동형 RFID 태그 300여 개가 쇼핑 카트에 부착되고 200여개의 리피터가 매장에 설치되었으며 25개의 리더가 매장 천장에 설치하였다.

<그림 10>에서 점으로 표시되어 있는 지점이 RFID 리피터가 설치된 지점이며 동시에 경로 그래프의 한 노드가 된다. 따라서 본 연구 환경에서는 200여 개의 노드를 갖는 이동 경로 그래프가 사용되게 된다. RFID 리피터는 1부터 200까지 고유의 ID를 갖으며 이를 참조하여 고객의 이동 순서를 추출하게 된다.



그림 10. 매장 RFID 시스템 장비 설치 영역

데이터 수집기간은 2011년 2월 중 일주일 동안의 데이터를 대상으로 하여 추출한 쇼핑 동선 데이터를 <표 2>에서 보는 바같이 평일 월요일, 수요일, 금요일의 각각 439개, 434개, 595개의 동선 데이터를 획득하였다. 분석을 위해 전체 데이터 중에서 평일의 쇼핑 동선 데이터를 분석 대상으로 한정 하였는데, 이는 실험 대상인 매장의 위치 특성상 도심가에 위치하여 주말 방문 고객 수와 평일 방문의 고객의 수가 비슷하고, 주말 바로 전날인 금요일에 방문 고객의 수가 가장 많은 특징을 갖고 있기 때문이다.

실험 기간 동안의 매장의 평균 구매 실적은 각각 약 21,928원, 약 27,460원, 약 30,680원이었다. 이를 통해서 대상 매장의 경우 주로 금요일에 구매가 비교적 많이 이루어지고 있음을 알 수 있는데, 이는 언급한 바와 같이 매장의 특성상 주로 주말 이전에 미리 쇼핑을 마치는 성향의 고객들이 주를 이루고 있으며 상대적으로 주중의 경우 구매 및 방문횟수가 떨어짐을 알 수 있다.

표 2. 대상 매장 고객 쇼핑 이용 현황

순서	분석 대상	1인 소비자 평균 구매액(원)	동선 추출 수
1	월	21,928.79	439
2	수	27,460.15	434
3	금	30,680.01	595

<표 3>는 일별 판매 순위를 판매 수량을 기준으로 상위 10위까지 정리한 자료이다. 채소와 테일러리 상품, 과자류 등이 상위 3위까지 있지만, 전체적으로는 비슷한 판매 속성을 보여주고 있다. 이를 통해서 대상 매장은 주로 가정 식가공류 위주의 상품 판매가 주를 이루고 있음을 알 수가 있다.

표 3. 일별 판매 순위 분석

(단위 : 개수)

Rank	월		Rank	수		Rank	금	
1	채소	381	1	채소	267	1	채소	534
2	데일리	199	2	데일리	228	2	데일리	300
3	과자	148	3	과자	183	3	과자	224
4	냉장냉동	102	4	과일	95	4	냉장냉동	151
5	조리식품	99	5	인스턴트	92	5	인스턴트	143
6	인스턴트	86	6	주방용품	83	6	조미식품	132
7	닭고기	73	7	생선	81	7	생선	126
8	생선	69	8	조미식품	81	8	조리식품	121
9	과일	63	9	조리식품	80	9	과일	121
10	조미식품	56	10	닭고기	69	10	닭고기	107

그룹	동선 시퀀스(월)	
1	채소, 채소 2, 닭고기, 채소 3, 양념육, 양념육 2, 돈육, 델리카, 액상 요거트, 대용량, 생수팔레트, 과자, 세제, 세제 2, 수예, 인테리어, 애완원예용품	
2	채소, 채소 2, 닭고기, 채소 3, 튜브/바, 대용량, 옛날과자, 위생용품, 델리카, 싱글존/대용량존, 미역/다시마, 맛살/유부/ 어묵/단무지, 육포/취포, 채소4	
3	채소, 채소 2, 닭고기, 양념육, 양념육 2, 돈육, 델리카, 튜브/바, 델리카 2, 냉장기타반찬류, 떡갈비, 냉장음료, 영유아두유/축산물통조림, 포장김치, 햄소세지, 채소 3, 청과	
4	채소, 채소 2, 닭고기, 닭고기, 채소 4, 채소 3, 채소, 채소, 양념육 2, 식용류/믹스, 영유아두유/축산물통조림, 1L물, 컵라면, 국수류, 청소욕실, 청소욕실 2	
5	채소, 채소 2, 닭고기, 채소, 수산, 채소 3, 가루식품, 케찹/마요네즈, 농산통조림, 국수류, 커피믹스, 주방용품, 생수팔레트, 채소, 닭고기, 채소 2	

그룹	9일 동선 시퀀스(수)	
1	닭고기, 채소, 양념육 2, 델리카, 물만두, 델리카 2, 싱글존/대용량존, 물만두, 자연/가공조미료/소금, 청소욕실, 청소욕실 2	
2	자연/가공조미료/소금, 1L물, 과자, 세제, 세제 2, 수예, 인테리어, 애완원예용품, 자동차	
3	채소, 닭고기, 채소, 양념육, 햄소세지, 자연/가공조미료/소금, 영유아두유/축산물통조림, 1L물, 주방용품	
4	채소, 채소 4, 채소, 가루식품, 국수류, 커피믹스, 용기, 커피믹스, 녹차, 식수용차, 국산차, 식수용차, 녹차, 세제, 주방용품, 간장/고추장, 설탕/밀가루/농산가루, 간장/고추장, 가루식품, 채소 4, 청소욕실, 과자, 세제, 수예, 식수용차	
5	채소, 채소, 양념육 2, 돈육, 영유아두유/축산물통조림, 식용류/ 믹스, 햄소세지, 양념육, 수산, 닭고기, 채소, 채소 2, 청과	

그룹	11일 동선 시퀀스(금)	
1	애완원예용품, 인테리어, 과자, 세제, 과자, 옛날과자, 대용량, 호상 요거트, 생수팔레트, 옛날과자, 위생용품, 과자, 세제, 수예, 인테리어, 자동차	
2	닭고기, 채소, 양념육 2, 돈육, 델리카, 액상요거트, 호상요거트, 대용량, 국수류, 농산통조림, 주방용품	
3	채소, 채소, 채소 3, 닭고기, 양념육 2, 돈육, 델리카, 델리카 2, 완자, 액상요거트, 옛날과자, 과자, 애완원예용품	
4	채소, 채소, 수산, 양념육 2, 델리카 2, 과자, 세제, 과자, 녹차, 주방용품	
5	닭고기, 채소, 식용류/믹스, 아임리얼/자연(크림/피자치즈, 세제, 수예, 인테리어, 애완원예용품, 자동차	

그림 11. 분석 동선 시퀀스

4.2 고객 쇼핑 동선 분석

고객 쇼핑 동선 분석을 위해, <표 2>와 같이 추출된 동선 수인 439개, 434개, 595개에 대하여 본 연구에서 제안한 고객 쇼핑 동선 패턴 분석 알고리즘을 적용하였다(<그림 8> 참조). 동선 패턴의 발견을 위해서 알고리즘의 입력 파라미터인 군집군을 본 실험에서는 5로 하여 5개의 주요 쇼핑 동선 패턴 분석하였다.

고객들의 최장 공통 부분열의 주요동선 패턴은 <그림 11>과 같으며, 이 정보를 바탕으로 매장 레이아웃에 시각화한 결과는 <그림 12>와 같다. <그림 12>의 좌측 상단은 매장 내의 판매 순위 10위의 제품들(<표 3> 참조)이 진열되어 있는 레이아웃(Layout of Product)을 보여주고 있으며, 우측 상단은 월요일, 좌측 하단은 수요일, 우측 하단은 금요일의 일별 주요 고객

쇼핑 동선 패턴을 나타내고 있다.

고객 쇼핑 동선 패턴 분석을 통해 나타난 고객 이동의 특징으로 출입구를 시작으로 시계 반대방향으로 쇼핑객 동선이 주로 발견이 되었다. 이는 출입문의 위치가 <그림 12>에서 보는 바와 같이 매장 구조도의 좌측 하단에 위치하고 있으며, 계산대가 좌측 상단에 위치되어 운영되고 있기 때문에 시계반대 방향으로 자연스럽게 고객 이동이 발생하는 것으로 판단된다. 또한 <표 3>의 일별 판매 랭킹 10위 제품 중 상당수가 매장의 하단 측면에 배치되어 있고 이로 인해 자연스럽게 구매 및 방문이 가장 많이 발생 되고 있는 것을 알 수 있다. 특히, <그림 12>의 원 모양으로 표시를 한 지점은 입구에서 가장 가까이 위치해 있거나, 출구와 계산대와 근접해 있기 때문에 모든 동선이 집중되는 지점들으로써 매장에서 hot spot에 속함을 알 수가 있다.

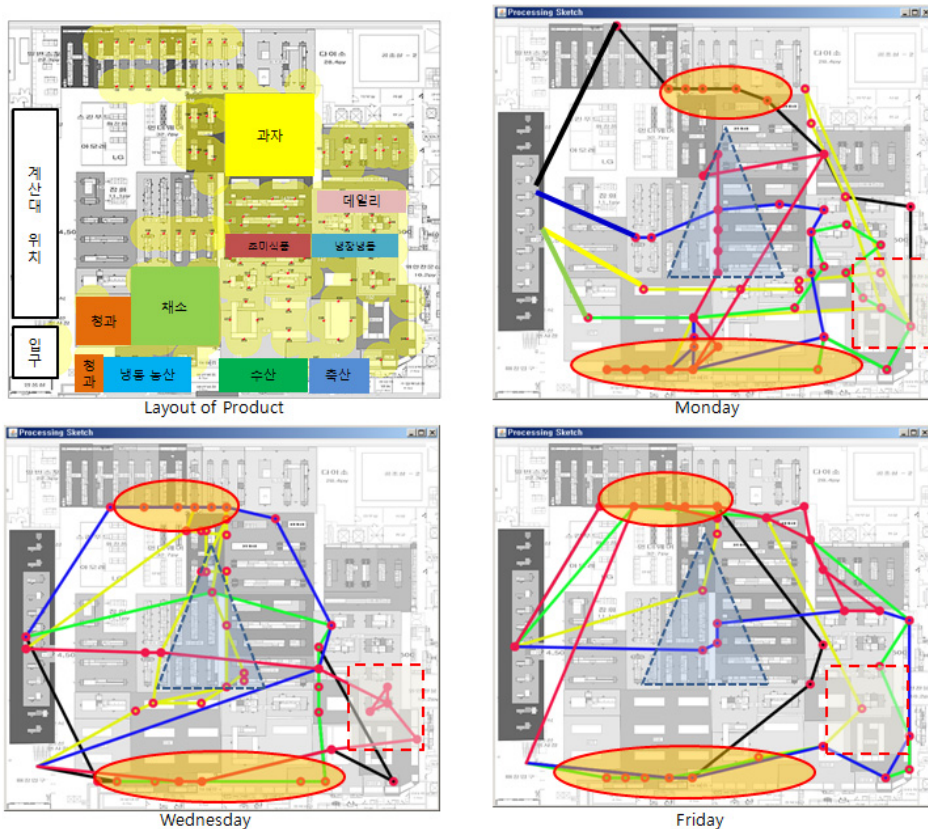


그림 12. 고객 주요 쇼핑 동선 추출 결과

이에 비해 상대적으로 최상단 쪽의 진열 제품들(인테리어, 취미용품, 자동차 제품들)은 실제 판매 실적도 저조하며 상위 순위(<표 3> 참조)에서 찾을 수 없으며, 동시에 쇼핑 동선 패턴 역시 거의 나타나지 않음을 알 수 있다. 그러나 이 지점들과 바로 접하고 있는 통로 복도는 쇼핑을 마무리 하고 계산대까지의 이동을 위한 동선 패턴들은 겹쳐지고 있기 때문에, 비록 이 부분은 판매는 저조하지만 매장 내의 두 번째 hot spot으로 판단 할 수 있다. 두 번째 hotspot은 첫 번째 hot spot에 비해 판매가 많이 발생하지는 않지만, 고객들이 쇼핑을 마치고 계산대로 이동하기 위해서 이동하는 통로로써 상품에 대한 전시 효과가 있기 때문에, 고객에게 소비 충동을 자극할 수 있는 프로모션 전략이나 상품을 진열하는 등의 전략을 수립하여야 할 것이다.

그런데, 매장의 중앙 위치에 삼각형 표시 영역은 원모양의 첫 번째 hotspot과 두 번째 hotspot 사이에 연결 다리 역할을 하고 있는데, 조미식품이나 주방용품의 판매가 이루어지는 지점이다. 이에 대비해서 같은 다리 역할을 하는 매장 우측의 사각형 모양 영역은 동선이 발생함에도 불구하고 많은 구매가 이루어지지 못하고 있다. 이는 판매 순위 상위의 제품이 주위에 배치되어 있어, 고객들의 이동이 자주 발생되지만 판매는 활발히 이루어지지 않는 지점으로, 이 지점에 대해 매장 관리자의 입장에서 상품배치의 변경이나 상품 변경을 통해서 판매를 유도시켜야한다는 것을 인지를 할 수 있다.

이러한 정보는 상품 장바구니 분석과 비교했을 때, 많은 시사점을 제공해 주고 있다. 장바구니 분석 기법은 실제 매장 데이터와 같이 대량의 상품군 가운데 특정 상품군에 대한 판매 트랜잭션이 주를 차지하는 경우에는 장바구니 순차 패턴의 지지율이 낮은 경향을 보이기 때문에 현업에서 사용하기에는 부적합하기 쉽다. 실제로 본 연구의 실험 데이터를 장바구니 분석을 시행해본 결과 <표 4>와 같은 결과를 보였는데, 역시 전체 상품군에 비해 특정 상품군에 대한 트랜잭션이 주를 차지하고 있기 때문에, 전반적으로 지지율이 높은 상품군은 나타나지 않았다. 지지율은 5% 이상의 데이터 결과만으로 한정을 하였다. <표 4>의 구매 순차 분석의 결과는 <표 3>의 일별 판매 순위 분석 결과를 통해서 미리 충분히 예상할 수 있는 상위 랭크 상품군들의 구매 시퀀스를 보여주고 있고 다른 유의한 의미를 제공해주지는 못하고 있다. 이와 같이 기존의 분석 기법들은 매장 내에서 고객들이 어떤 식의 흐름으로 쇼핑을 하는지 행동 패턴에 대해서는 알 수가 없으며, 고객들의 인구 통계학적 정보와 결합하여 더 많은 고객 분석이 가능하더라도 실제 구매하지 않은 경우에 대하여 분석하지 못한다.

이처럼, 동선을 기준으로 분석할 수 있는 정보들은 기존의 판매 상품의 트랜잭션 정보만을 활용한 분석에서는 알 수 없는 사실들이 분석이 가능하며, 본 연구의 제안 기법을 통해서 고객의 쇼핑 동선 패턴 도출 결과를 통해 새로운 매장 개선의 기회를 제공해 줄 수 있다.

표 4. 일별 구매 순차 분석

월요일				수요일				금요일			
support %	confidence %	Item 1	Item 2	support %	confidence %	Item 1	Item 2	support %	confidence %	Item 1	Item 2
19.82	32.1	채소	데일리	17.82	35.98	채소	데일리	27.24	43.16	채소	채소
13.21	21.4	채소	닭고기	12.27	24.77	채소	채소	20.3	32.17	채소	데일리
12.98	21.03	채소	조리식품	11.34	28.99	데일리	데일리	17.6	27.88	채소	생선
11.62	18.82	채소	냉장냉동	11.11	22.43	채소	생선	12.86	20.38	채소	냉장냉동
11.62	18.82	채소	생선	9.72	24.85	데일리	과자	12.69	20.11	채소	닭고기
9.57	15.5	채소	과자	8.56	17.29	채소	조리식품	12.01	19.03	채소	조리식품
7.52	12.18	채소	인스턴트	8.56	28.46	과자	과자	11.34	17.96	채소	조미식품
7.06	31.31	조리식품	데일리	8.33	16.82	채소	과자	11.34	31.75	데일리	데일리
6.61	18.24	데일리	과자	7.87	35.79	과일	채소	10.83	17.16	채소	과자
6.38	10.33	채소	과일	7.18	14.49	채소	닭고기	10.32	16.35	채소	인스턴트
5.92	35.62	닭고기	데일리	6.71	13.55	채소	과일	9.81	15.55	채소	과일
5.24	14.47	데일리	인스턴트	6.71	42.03	닭고기	데일리	8.12	44.86	닭고기	채소
				6.48	40.58	닭고기	채소	7.61	12.06	채소	주방용품
				6.25	35.53	인스턴트	과자	7.45	20.85	데일리	과자
				6.25	12.62	채소	냉장냉동	7.28	11.53	채소	소고기
				6.25	15.98	데일리	인스턴트	7.11	34.71	과일	채소
				6.02	12.15	채소	인스턴트	6.94	24.55	과자	과자
				5.79	11.68	채소	조미식품	6.26	34.58	닭고기	데일리
				5.56	14.2	데일리	채소	6.26	30.58	조리식품	데일리
				5.56	30	조리식품	데일리	6.26	17.54	데일리	인스턴트
				5.32	13.61	데일리	주방용품	6.26	29.37	생선	데일리
				5.32	10.75	채소	주방용품	6.09	33.64	닭고기	생선
				5.32	17.69	과자	주방용품	5.75	16.11	데일리	채소
				5.09	10.28	채소	소고기	5.58	15.64	데일리	냉장냉동
								5.41	8.58	채소	건해산물
								5.08	14.22	데일리	주방용품

5. 결론 및 향후 연구

기존의 매장 내 고객 분석은 장바구니 분석이나 통계 기법 등이 주로 사용되어 왔고 매장 고객 소비 행동 패턴 분석 등은 널리 시행되지 못했다. 그러나 매장 환경에서의 고객 쇼핑 동선 패턴을 분석함으로써 기존의 트랜잭션 정보와 인구 통계학적 정보를 활용한 분석 기법들로는 알 수 없었던 고객 쇼핑 동선의 공간적 패턴 클러스터링이나 중요 지점의 위치 등을 파악할 수 있고, 이러한 결과물을 통해 매장 환경 개선이나 상품 배치 등에 활용되어 소비 증진에 기여할 수 있다.

이를 위해 본 연구에서는 기존 고객 동선 분석 기법들에서의 문제점인 유클리드 거리 지표와 데이터 전처리 문제점들을 해결하기 위하여 최장 공통 부분열 기법을 기반으로 하여 새롭게 고객 쇼핑 동선 공간적 패턴 클러스터링 분석 기법을 제안하였다. 유클리드 거리 기반의 쇼핑 동선 패턴 분석 기법은

클러스터 과정 중 클러스터가 매대나 상품 가판대와 같은 장애물의 위치에 생성되거나, 장애물을 중심으로 나누어져 있는 동선들이 가까운 거리에 위치했을 경우에도 같은 클러스터로 군집될 수 있으나, 최장 공통 부분열은 이동 동선을 이동 그래프 노드로써 표현을 하고, 쇼핑 동선간의 최장 공통 부분열을 찾아내어 쇼핑 패턴을 분석할 수 있기 때문에 이러한 단점을 해결할 수 있었다. 또한 쇼핑 동선 패턴뿐만 아니라 이동 방문이 집중되는 부분과 이동 방문이 이루어지지 않는 지역을 구분할 수 있었다. 이러한 정보는 매장 관리자 입장에서 중요한 정보로써, 분석 결과를 바탕으로 매장 내의 제품 진열이나 레이아웃 변경할 때, 고객의 소비 행동의 이해를 바탕으로 배치나 진열에 적용할 수 있을 것이다. 이는 기존의 판매 수량 통계 정보나 매장 직원의 경험에 의한 정성적인 측면을 정량화시키고, 객관화된 수치의 개발을 통해서 매장은 고객에게 질 좋은 서비스를 제공할 수 있으며, 동시에 매출 상승까지 기대할 수

있을 것이다.

향후에는 고객 이력 정보와 같은 레거시 시스템 정보들과 융합하여 매장 운영에 효율화와 매출 향상에 도움이 되는 지능형 매장 분석 시스템 개발이 필요하며, 보다 다양한 고객의 쇼핑 행동 분석을 위한 분석 모형의 개발, 동시에 매장 환경 분석을 위한 다양한 측정 지표들에 대한 연구가 필요하다. 또한 다차원적인 분석 기법과 결합하여 고객 행동과 매장 환경의 영향을 정량화 시킬 수 있으며, 객관적인 측정 지표를 제안할 수 있을 것이다. 이는 매장에 있어서 꼭 필수 요소로 더 다양한 분석이 가능할 것으로 기대되며, 고객서비스 향상을 통한 제조, 유통업체, 소비자 모두의 이익증대를 실현시킬 수 있는 고객 서비스 기술이 될 것이다.

참고문헌

- Cao, H., Mamoulis, N., and Cheung, D. W. (2007), Discovery of Periodic Patterns in Spatiotemporal Sequences, *Knowledge and Data Engineering, IEEE Transactions on*, 19(4), 453-467.
- Cox, K. (1964), The Responsiveness of Food Sales to Shelf Space Changes in Supermarkets, *Journal of Marketing Research*, 1(2), 63-67.
- Dickson, P. R. and Sawyer, A. G. (1986), Point-of-Purchase Behavior and Price Perceptions of Supermarket Shoppers, *Working Paper*, 86-102, Marketing Science Institute, 1000 Massachusetts Ave., Cambridge, MA 02138.
- Farley, J. U. and Ring, L. W. (1996), A Stochastic Model of Supermarket Traffic Flow, *OPERATIONS RESEARCH*, 14(4), 555-567.
- Gil, J., Tobar, E., Lemlij, M., Rose, A., and Penn, A. (2009), The Differentiating Behaviour of Shoppers : Clustering of Individual Movement Traces in a Supermarket, *Proceedings of the 7th International Space Syntax Symposium*.
- Harris, D. H. (1958), The effect of display width in merchandising soap, *Journal of Applied Psychology*, 42(4), 283-284.
- Hirschberg, D. S. (1977), Algorithms for the longest common subsequence problem, *Journal of ACM*, 24(4), 664-675.
- Hou, J.-L. and Chen, T.-G. (2011), An RFID-based Shopping Service System for retailers, *Advanced Engineering Informatics*, 25(1), 103-115.
- Hoyer, W. D. (1984), An Examination of Consumer Decision Making for a Common Repeat Purchase Product, *Journal of Consumer Research*, 11(3), 822- 829.
- Hui, S. K., Bradlow, E. T., and Fader, P. S. (2009), Testing Behavioral Hypotheses Using an Integrated Model of Grocery Store Shopping path and purchase Behavior, *Journal of consumer research*, 36, 478-493.
- Hui, S. K., Fader, P. S., and Bradlow, E. T. (2009), Path Data in Marketing : An Integrative Framework and Prospectus for Model Building, *Marketing Science*, 28(2), 320-335.
- Kang, H.-Y., Kim, J.-S., Hwang, J.-R., and Li K.-J. (2008), Similarity measures for trajectories of moving objects in cellular space, *GIS Association of Korea*, 16(4), 291-301.
- Kim, Y.-C. and Chang, J.-W. (2008), Grid-based Similar Trajectory Search for Moving Objects on Road Network, *Journal of Korea Spatial Information System Society*, 10(1), 29-40.
- Larson, J. S., Bradlow, E. T., and Fader, P. S. (2005), An exploratory look at supermarket shopping paths, *International Journal of Research in Marketing*, 22(4), 395-414.
- McClure, P. J. and West, E. J. (1969), Sales Effects of a New Counter Display, *Journal of Advertising Research*, 9, 29-34.
- Newman, A. J., Yu, D. K. C., and Oulton, D. P. (2002), New insights into retail space and format planning from customer-tracking data, *Journal of Retailing and Consumer Services*, 9(5), 253-258.
- Uotila, V. and Skogster, P. (2007), Space management in a DIY store analysing consumer shopping paths with data-tracking devices, *Facilities*, 25(9), 363-374.
- Yanagisawa, Y., Akahani, J.-I., and Satoh, T. (2003), Shape-based Similarity Query for Trajectory of Mobile Objects, *Proceedings of MDM*, 63-77.



정인철

시립인천대학교 인문대 학사
 동국대학교 산업시스템공학과 석사
 현재 : 동국대학교 산업시스템 박사과정
 관심분야 : 기계학습, 데이터 마이닝, 에이전트,
 지능형 정보시스템



권영식

서울대학교 산업공학 학사
 한국과학기술원 산업공학 석사
 한국과학기술원 공학박사
 현재 : 동국대학교 산업시스템 공학 교수
 관심분야 : 데이터 마이닝, 지능형 정보 시스템