

# 고객 쇼핑 동선 분석시스템의 개발

## Development of Shopping Path Analysis System(SPAS)

정인철(In-Chul Jung)\*, 권영식(Young S. Kwon)\*\*, 이용한(Yong-Han Lee)\*\*\*

### 초 록

최근의 RFID와 모바일 기술을 포함한 정보 기술의 진보로 인해 매장 내에서 고객 쇼핑 동선을 추적하는 것이 가능해지고 있다. 특히, 고객 쇼핑 동선은 매장 내에서 발생하는 고객 쇼핑 행동을 이해하는데 중요한 단서를 제공한다. 따라서 본 연구에서는 고객의 동선 분석을 위하여 동선 정보를 획득하고 분석하는 고객 쇼핑 동선 분석 시스템을 개발한다. 시스템 개발을 위해 RFID를 사용하였고, 대형 유통 매장의 현업 전문가와 함께 다양한 매장 측정 변수를 정의하여 혼잡도, 방문비율등과 같은 매장 프로파일 분석과 고객 쇼핑 패턴 식별을 위해 동선 클러스터링 분석 기법을 개발하여 적용하였다. 최종적으로는 개발한 시스템의 실효성을 평가해보기 위해 실제 서울의 대형 유통 매장에 적용하여 사례분석까지 실시해보았다.

### ABSTRACT

Technological advancements in information technology including RFID and mobile technologies have made it feasible to track the customers travel path in a store. The customer travel paths provide valuable implications to understanding the customer behaviors in a store. In our research, we develop a shopping path analysis system to track and analyze the customer travel path. The proposed system consists of RFID systems for collecting the customer paths and analysis system. The analysis system conducts clustering for identifying the distinctive shopping patterns, and analyzes the profile of a grocery, such as congestion rate, visiting rate, and staying time, etc. We show the applicability of our proposed system using the actual data obtained at a grocery in Seoul as a case study.

**키워드 :** RFID 동선, 쇼핑 동선, 고객 행동, 실험 분석, 분석 시스템

RFID, Customer Path, Shopping Behavior, Exploratory Analysis, Analysis System

---

\* First author, Department of Industrial & Systems Engineering, Dongguk University-Seoul

\*\* Corresponding Author, Department of Industrial & Systems Engineering, Dongguk University-Seoul  
(E-mail : yskwon@dgu.edu)

\*\*\* Co-author, Department of Industrial & Systems Engineering, Dongguk University-Seoul  
2012년 08월 16일 접수, 2012년 09월 14일 심사완료 후 2012년 10월 05일 게재확정.

## 1. 서 론

유통업계에서는 지난 수년간의 IT 인프라 기술의 발달로 인하여 업무 효율화와 비용 절감 그리고 편리한 서비스 제공을 하고 있다. 대표적 유통업체인 월마트(Walmart)나 메트로(Metro)는 이미 RFID 기술을 활용해 물류와 창고 관리에서 프로세스를 효율화하는 노력을 기울여 왔다. 예를 들어 유통망을 구축하여 협력 업체에서부터 판매점까지의 전체 공급망 가시성을 확보하는 노력을 들 수 있다. 협력 업체의 재고 및 판매 현황을 실시간으로 한 눈에 알 수 있고, 그 결과 남아도는 상품을 최소화하여 영업이익의 신장에 노력을 기울이는 것이다. 그 외에도 새로운 기술을 활용한 다양한 활동들을 시도하고 있는데, 메트로 그룹은 Future Store 프로젝트에서 남성 의류 매장 현장에 RFID 기술을 적용하여 판매 의류에 대한 추가 정보를 고객에게 제공하는 서비스와 더불어 판매된 상품들을 판단하여 실시간으로 해당 상품을 자동 발주하는 매장 시스템을 구축하였고, 특별한 고객 경험을 제공하기 위하여 스마트 거울과 스마트 드레싱 룸(Smart Dressing Rooms)을 개발하여 구매 선택에 도움을 주는 서비스를 개발하였다. 또한 My GROCER 프로젝트에서는 슈퍼마켓을 이용하는 고객 가정의 식료품 정보를 이용하여 부족한 상품을 자동으로 보충해주는 유비쿼터스 시스템을 개발하였다[4]. 그러나 물류나 새로운 서비스 제공 이외에 고객 이해를 위해 유비쿼터스 기술을 활용하는 연구는 많이 진행되지 않고 있는 실정이다.

과거에는 고객 쇼핑 행동을 이해하기 위하여 구매 이력을 바탕으로 장바구니 분석이나

인구 통계학적 정보를 활용하여 상권분석이나 고객 세분화 분석 등을 실행하여 고객 성향을 파악해 왔다. 그러나 앞으로 빠르게 변화하는 소비 환경과 늘어나는 경쟁의 심화로 인해 유통 매장을 이용하는 고객 쇼핑 행동에 대한 세밀한 분석 능력은 중요한 자산이 될 수밖에 없다. 특히 쇼핑 행동의 중요한 요소인 쇼핑 동선 분석은 매장 환경 개선이나 프로모션 전략 수립에 사용할 수도 있고, 유통 분야뿐만이 아니라 교통, 날씨, 동물, 통신과 같이 다양한 분야에서 활용이 될 수 있다 [7, 10].

그러나 현재까지는 고객 동선 파악을 위해 연구자가 직접 고객과 함께 동행하며 일일이 기록을 하는 방법밖에 없었다. 이와 같은 방식은 매장 방문 고객이 하루에도 수백 명 이상인 매장에서는 관찰 고객 수가 너무 작아 객관적인 수치로 사용하기가 어렵고 비용 역시 많이 든다. 객관적인 근거를 사용하기 위해서는 자동화된 동선 데이터 수집 시스템 개발이 필요하다. 따라서 본 논문에서는 쇼핑 동선 수집 및 분석을 위한 시스템을 개발하여 이용 고객 쇼핑 동선 분석뿐만이 아니라 매장 환경 분석을 하고자 한다. 이를 위해 문제가 되는 이슈들을 정리해보고 각각의 해결 방안들을 제시한다. 개발한 동선 분석 시스템은 크게 동선 수집 모듈과 분석 모듈로 나누어지며, 동선 수집을 위하여 RFID 기술을 사용하였고, 데이터 분석을 위하여 통계 분석과 데이터마이닝 알고리즘을 사용하였다. 또한 본 연구의 동선 분석을 위하여 동선 클러스터링 알고리즘을 새롭게 개발하였다[2]. 그리고 최종적으로는 구현한 쇼핑 동선 분석 시스템을 실제 유통매장에 적용해본 결과, 실제

유통 매장에서 과거에는 알 수 없었던 정보를 제공해줄 수 있었다.

## 2. 관련 연구

쇼핑 동선은 쇼핑물에서 고객이 입장하여 퇴장할 때까지 매장 안에서 상품의 구매를 위해 쇼핑하는 이동 경로를 말하며, 이러한 이동 경로를 도식화하여 고객이 쇼핑을 하면서 방문한 장소, 머무른 장소, 머무른 시간, 지나간 통로들을 파악할 수 있고, 파악된 정보를 기반으로 매장의 통로 구조를 변경 하거나 상품을 재배치하여 최적의 고객 쇼핑 동선을 유도할 수 있다[1]. 기존 동선 수집 기법은 쇼핑 매장에 들어온 고객을 관찰자가 직접 따라다니며 고객이 이동한 경로를 평면도에 그려 넣어 분석하였다. 그러나 인력 투입에 대한 부담으로 인해 적극적인 분석을 실시하고 있지는 못하고 있다. 그러나 최근의 IT 기술의 발달로 인하여 RFID, WLAN, CCTV 기술 등을 이용하여 고객의 동선을 추적하고 분석하는 시스템이 제안되어 왔다. 유통 매장 분야에서는 기술과 가격 면에서 점점 안정 되어가고 있고 RFID가 태그를 읽은 시간과 위치를 파악할 수 있기 때문에 주로 RFID 태그를 쇼핑 카트와 장바구니에 부착하여 매장의 고객 위치를 검출하고 쇼핑 동선을 수집하여 고객의 쇼핑형태를 분석하는 연구들이 진행되어 왔다[8, 9, 11, 12]. 이와 같은 동선 정보들을 이용하여 안재명은 동선분석을 통한 실시간 선호상품 추천에 대하여 연구하였다. 오프라인 쇼핑물에서의 상품추천의 경우, 회원 고객들을 대상으로 프로파일과 구매정보를 확인

하여 관련된 상품이나 행사 상품 등을 고객에게 안내 할 수 있지만, 고객의 현재 주변 상황을 고려하지 않은 추천방식이므로 스팸 광고로 여겨질 여지가 있다. 따라서 이러한 단점을 보완하여 오프라인 쇼핑물 환경에 적합한 선호 상품 추천이 가능하도록 실시간으로 고객의 위치를 파악하여 현재 주변 상품 선호도를 계산한 후 추천하는 실시간 상품 추천 기법을 제안하고 있다[1]. 다른 연구로는 매장 내에서 일어나는 고객들 간에 영향이 어떻게 미칠 수 있는지에 대한 연구 시도가 있었다. 고객들이 다른 고객들이 구입하는 상품을 따라서 구매를 한다거나 봄비는 지역을 피해가는 경향성이나 그 반대등과 같은 행동 분석을 위해 이동 데이터 정보를 활용하고자 하는 연구가 진행되었다[6]. 동선 클러스터링 기법 연구로는 k-medoids 클러스터링 알고리즘을 이용하여 동선 패턴 분석 한 연구가 있었으나, 이를 위해서는 다양한 동선 길이를 똑같이 전처리 한 후에 적용해야 한다. 이때 정보의 손실이나 손상이 될 수도 있기 때문에 전처리 작업이 매우 중요하다[8, 13]. 이러한 문제를 해결하기 위하여 Jung[2]은 동선의 길이와 상관없이 클러스터링을 할 수 있도록 LCSS를 기반한 클러스터링 알고리즘을 제안하였다. 그러나 이러한 동선 클러스터링 연구들은 모두 초기값 선정에 대한 기준이 없기 때문에 주로 임의로 무작위 추출하여 초기값을 선정하였다. 초기값 선정은 알고리즘 성능에 중요한 영향을 미치기 때문에 보다 적절한 선정방법이 필요하다.

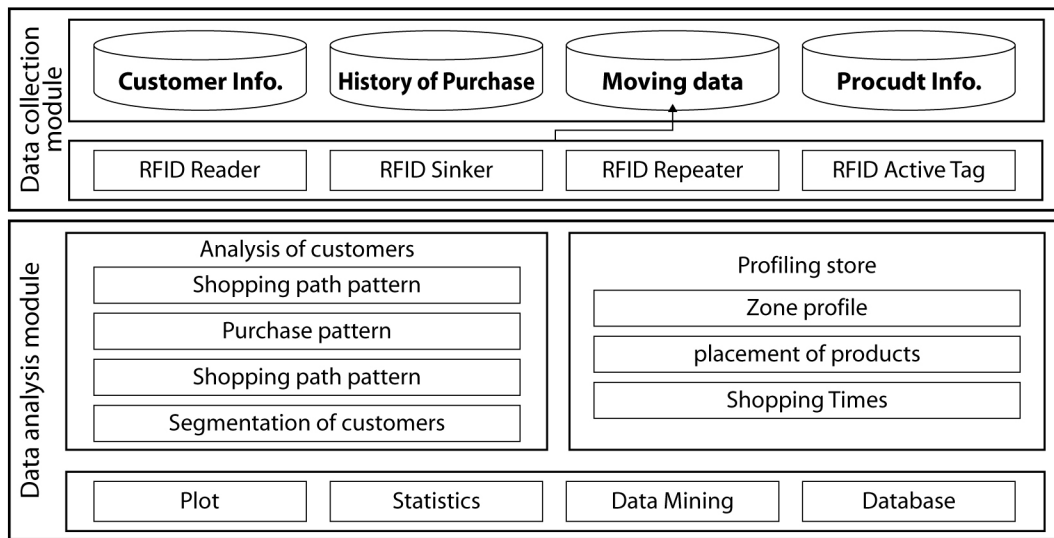
따라서, 본 연구에서는 동선 분석 시스템을 개발하기 위하여 기존 연구의 문제점을 정리하고 개선된 시스템을 개발하고자 한다.

### 3. 시스템 설계 및 구현

본 장에는 <Figure 1>과 같이 쇼핑 동선 분석을 위한 시스템(SPAS)을 개발하였다. 이를 위해 해결해야 할 문제점들을 정리하고 해결 방안들을 제시한다. 시스템 개발을 위하여 현업 전문가와 기존 연구를 분석하여 시스템의

요구사항을 <Table 1>과 같이 정리하였고 이에 대응하여 본 연구의 시스템을 개발하였다.

개발한 동선 분석 시스템은 동선 수집 모듈과 분석 모듈로 나누어지며, 동선 수집을 위하여 RFID 기술을 사용하였고, 데이터 분석을 위하여 통계 분석과 데이터마이닝 알고리즘을 사용하였다.



<Figure 1> Shopping Path Analysis System(SPAS)

<Table 1> System Requirement and Related Developed Module

|   | System Requirement                                                           | Solution                                              | Related modules                                                                                                                  |
|---|------------------------------------------------------------------------------|-------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| 1 | Labor-intensive<br>Small sample count                                        | - Automation tech. using RFID                         | 3.1 moving data collection module<br>- RFID Tag and Repeater                                                                     |
| 2 | Correlation between purchasing<br>and shopping path                          | - mapping with customer info and<br>shopping path     | 3.1 moving data collection module<br>- Integration total information<br>(customer, product, moving data,<br>history of purchase) |
| 3 | Need of measures for<br>congestion in store                                  | - development of measures of<br>congestion            | 3.2 analysis module<br>- definition of zone profile                                                                              |
| 4 | Recognition of behavior of<br>customers in front of<br>non-purchase products | - analysis shopping time of non-<br>purchase products | 3.2 analysis module<br>- analysis of times                                                                                       |
| 5 | Recognition of dead zone                                                     | - development new shopping<br>pattern algorithm       | 3.2 analysis module<br>- moving data clustering algorithm                                                                        |

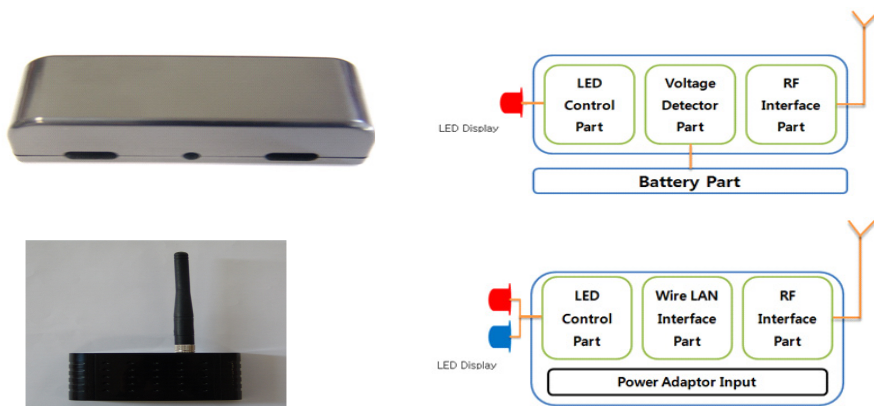
### 3.1 동선 수집 모듈

본 연구에서는 정확한 쇼핑 동선 정보의 수집을 위하여 Active Tag, Repeater, Sink와 같은 RFID 시스템을 개발하였다. <Figure 2>는 개발된 RFID 장비로써 실제 매장에 설치되어 활용하게 된다. Active Tag는 2개의 출력 채널로 RF 신호 전달이 가능하기 때문에 출력 채널마다 신호의 세기를 변화해서 보낼 수 있도록 개발되었다. 이러한 특징을 이용하여 2채널의 상이한 출력신호로 인식 거리를 조정할 수 있으며, 일정 시간간격으로 RF 신호를 전송한다. Active Tag의 설치위치는 쇼핑 카트의 전면 부분이 된다. Repeater는 쇼핑 카트에 설치되어 있는 Active Tag에서 송출하는 ID 정보를 받아 수집하며, 수집된 정보는 Repeater의 고유 식별 ID를 추가하여 Sink 노드로 전송을 한다. Active Tag와 Sink의 중간에 위치하여 데이터를 전달하는 역할을 한다. 이때 Repeater의 개별 ID를 근거로 매장에서의 쇼핑카트의 이동 위치를 파악하게 된다. Repeater가 설치 되어있는 지역으로 Active Tag가 장착되어 있는 쇼핑 카트가 근접하게

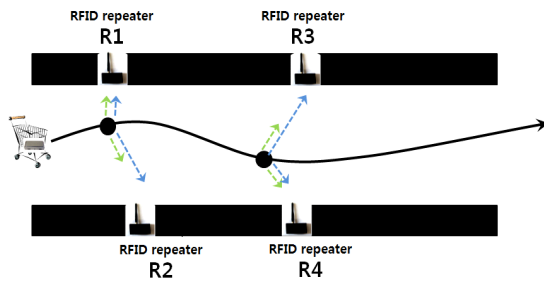
되면 고유의 ID가 Active Tag로부터 Repeater로 전송이 되며, Repeater는 자신의 고유개별 ID와 함께 추가하여 Sink로 정보를 전달한다. 이때 Active Tag에서는 2채널로 각기 다른 신호의 세기를 일정하게 전송한다. Sink는 Repeater에서 전송하는 데이터를 RF로 수신하여 Wire LAN을 통하여 동선 데이터 저장서버로 전송을 하는 기능을 한다. Sink의 설치 위치는 Repeater의 데이터를 모두 수신하기 위하여 수신 지역이 중첩이 되도록 설치하게 된다.

이처럼 RFID 동선 수집시스템은 설치된 모든 Repeater의 고유 아이디를 통하여 고객의 이동 동선 정보를 도출하기 때문에 다음과 같은 과정이 필요하다. 쇼핑 카트에 설치한 Tag의 고유의 ID가 Repeater에 수집이 되고 수집된 Tag의 ID와 함께 Repeater의 ID 정보를 조합하여 최종적인 쇼핑 동선 정보가 생성된다.

예를 들어 <Figure 3>과 같이 Repeater R1~Repeater R4까지 설치되어 있는 매대가 있고, 쇼핑 카트는 그림과 같이 R1 지점으로 이동을 한 후에 R4 지점으로 이동을 하고 있다고 가정을 하자. RFID 동선 수집 시스템에서는 R1~R4의 모든 ID가 수집이 될 수도



<Figure 2> RFID System(Active Tag and Repeater)



〈Figure 3〉 Moving Path Extraction from RFID Data

|   | Time     | Repeater id |
|---|----------|-------------|
| 1 | 09:54:38 | R01         |
| 2 | 09:54:45 | R05         |
| 3 | 09:54:50 | R22         |
| 4 | 09:54:55 | R19         |
| 5 | 09:54:57 | R21         |
| 6 | 09:55:02 | R27         |
| 7 | 09:55:09 | R16         |

있지만, 개발한 Active Tag는 2채널의 각기 다른 출력으로 Tag ID를 전송하기 때문에, R2보다는 R1으로 이동을 하였고, R3보다는 R4로 이동을 하고 있는 이동 동선 정보를 도출 할 수가 있다.

이와 같은 과정을 통해서 최종적으로 쇼핑 동선 데이터를 추출하게 된다. 향후에 이 정보를 바탕으로 최종적으로는 구매 이력 데이터와 고객 정보 데이터와 연결을 하고 최종적으로 분석데이터를 만들어 낼 수 있다.

### 3.2 데이터 분석 모듈

본 논문의 데이터 분석 모듈은 고객의 구입 목록 정보를 기반한 고객군 분석이나 연관 구매 상품 분석뿐만 아니라 다양한 구매 관련 정보와 동선 정보를 다차원적으로 융합하여 분석을 함으로써 매장과 고객에 대한 더 높은 이해를 할 수 있는 시스템을 디자인 하였다.

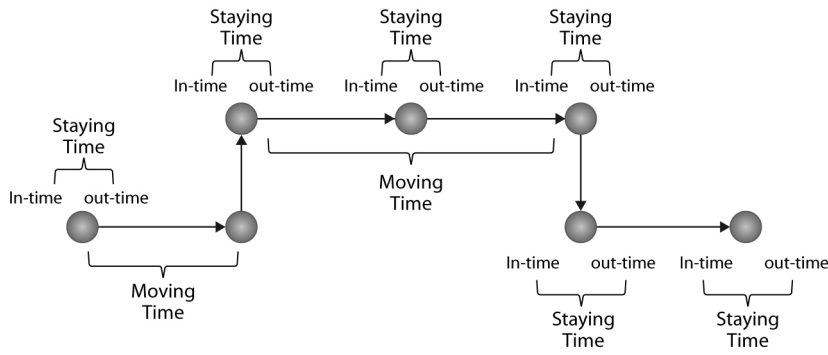
RFID 동선 수집 시스템에서 나오는 정보는 시간, 장소 등의 단순한 형태의 데이터이지만 매장의 POS 시스템과 연계하여 데이터를 가공 처리하여 고객 동선과 매장 환경이 고객 구매에 미치는 영향을 분석하기 위하여 고객의 쇼핑 시간 분석, 매장 내의 존(Zone)별 프로파일링(profiling), 동선 분석 등을 할 수 있다.

#### 3.2.1 쇼핑 시간 분석

고객의 쇼핑 시간에 대한 분석을 통해, 매장에 방문한 고객들의 평균 이동시간, 구매 물건 중 가장 시간이 오래 걸리는 상품 리스트 분석, 비구매한 물건 중 가장 시간이 오래 걸리는 상품 분석을 통해서 고객의 구매 성향에 대해 보다 정확하게 파악할 수 있다. 이러한 분석을 하기 위해서는 RFID 동선 정보를 이용하여 고객이 이용하는 시간과 POS 데이터를 이용하여 구입 상품 리스트를 결합하여, 구입한 상품의 소요 시간과 쇼핑 시간 등을 추출하는 전처리 과정이 필요하다.

〈Figure 4〉은 RFID에서 쇼핑 시간을 추출하기 위한 정보를 획득하는 방법을 보여주고 있다. 이를 통해 특정 지역에서 쇼핑한 시간(존 체류시간)과 이동 시간은 추출할 수가 있다.

이러한 쇼핑 시간 분석을 통해 상품의 재배치 필요 여부 및 혼잡으로 인한 서비스 불편의 영역을 파악할 수가 있다. 예를 들어, 상품 자체의 이유로 구매시간이 소요되는 것인지 혼잡구간이기 때문에 오래 소요되는 것인지를 구분하여 상품이나 배치의 변경을 통해 구매 소요시간을 줄일 수 있을 것이다. 또한 구매가 이루어지지 않은 상품 중에서 쇼핑 시간이 오래 소요가 되는 상품의 경우, 관심 상품으로 관리하여 추가적인 프로모션 전략을 통해 구매 촉진을 유발할 수 있다.



〈Figure 4〉 Shopping Time and Moving Time Transformation from RFID Data

### 3.2.2 매장에서 존(Zone)별 프로파일 분석

매대별 고객 방문횟수, 체류시간, 구매액뿐만 아니라 이를 조합한 파생변수를 생성시켜 각 매대의 혼잡도를 매출정보와 연계하는 분석을 할 수 있다. 제안한 시스템에서는 매대별 유효방문 고객수, 유효고객 체류시간, 고객구매액 비율, 유효방문당 구매액, 유효 체류시간당 구매액 비율 등 분석할 수 있다. 이처럼 혼잡도 분석을 위한 측정 지표 정의는 <Table 2>와 같다.

이러한 분석을 통해, 매대별 각 혼잡도 지수를 통해 측정된 값을 순위화하여 평가하며, 혼잡정도를 “High/Moderate/Low”로 등급화할 수 있고, 이 때 “High”인 경우를 “Hot spot”, “Low”

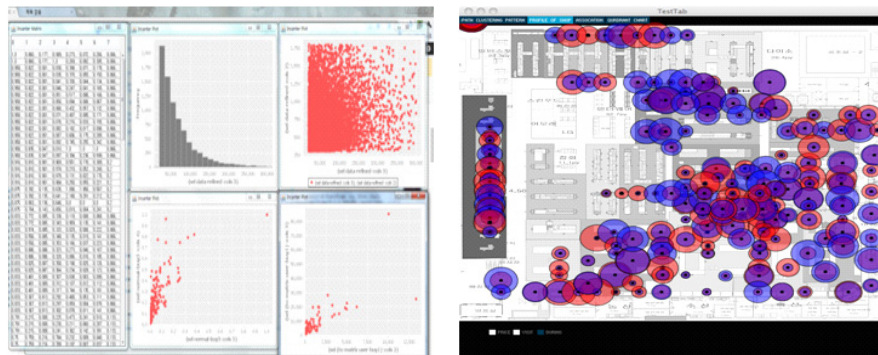
인 경우를 “Dead zone”으로 구분하며, 사용자가 이해하기 쉽도록 혼잡도 결과를 시각화하여 보여줄 수 있다. <Figure 5>은 이러한 분석을 위해 개발한 분석 시스템의 UI를 보여주고 있다.

### 3.2.3 쇼핑 동선 분석

쇼핑 동선 분석을 위해서 본 연구에서는 동선을 클러스터링 하여 쇼핑 동선 패턴을 추출하였다. 그러나 일반적으로 클러스터링 알고리즘은 벡터 기반의 유클리디안 거리 지표를 사용하지만 동선은 고객 마다 이동 동선 길이가 제각각 이기 때문에 적용을 할 수 없다. 본 연구에서는 이와 같은 문제점을 해결하기 위하여

〈Table 2〉 Measurement Definition

| Measures                              | description                                                           |
|---------------------------------------|-----------------------------------------------------------------------|
| Ratio of visiting in each product     | Visiting customers in each product/Entire customers                   |
| Ratio of staying in each product      | Staying time in each product/total staying time                       |
| Ratio of purchasing in each product   | Purchase price in each product/entire purchased price                 |
| Ratio of purchasing in visiting count | Ratio of purchasing in each product/Ratio of visiting in each product |
| Ratio of purchasing in staying time   | Ratio of purchasing in each product/Ratio of staying in each product  |



〈Figure 5〉 UI Examples of Shopping Path Analysis System

LCSS[5]를 확장한 LCSS\_SIM 유사도 거리 함수를 제안하고 이를 기반으로 하여 새롭게 동선 패턴 군집화 알고리즘을 제안 개발하였다[3]. LCSS\_SIM은 아래 식과 같이 2명의 x와 y의 쇼핑 경로의 유사성을 측정할 수 있다.

$$LCSS \sim (x, y) = \frac{lcss(x, y)}{xlength + ylength - lcss(x, y)}$$

또한 군집화를 위한 초기값 선정에서 빈발 패턴을 사용하기 때문에 주요한 동선들을 우선적으로 군집화 할 수 있었다. 제안한 군집화 알고리즘의 성능 평가를 위하여 가상의 데이터를 생성하여 실험한 결과 우수한 성능을 보였다[3]. 〈Figure 6〉은 제안한 클러스터링 알고리즘의 동작 순서를 보여주고 있다.

|            |                                                                                    |
|------------|------------------------------------------------------------------------------------|
| Function : | INIT_CLUSTER_CADIDATE                                                              |
| Input :    | Moving Path Trajectories (T), Support (S)                                          |
| Output :   | Initial Cluster of Moving Path (CT)                                                |
| 1          | L = Apriori(T, S) // Frequent Pattern over Support(S)                              |
| 2          | FOR (list in L) {                                                                  |
| 3          | FP= max(L)                                                                         |
| 4          | }                                                                                  |
| 5          | FOR (traj in T) {                                                                  |
| 6          | CT = max(LCSS_SIM (FP, traj)) //selecting initial cluster moving path              |
| 7          | }                                                                                  |
| 8          |                                                                                    |
| 9          | RETURN CT                                                                          |
| Function : | Path_CLUSTERING                                                                    |
| Input :    | Moving Path Trajectories T = {x1, x2, ..., xn}, Support (S), Sim_LCS Threshold (D) |
| Output :   | Clusters (Cn)                                                                      |
| 1          | WHILE (CT = INIT_CLUSTER_CADIDATE(T, S)) {                                         |
| 2          | for (traj in T) {                                                                  |
| 3          | If (LCSS_SIM (traj, CT) > D) {                                                     |
| 4          | Cn <= traj // Insert Cluster group                                                 |
| 5          | }                                                                                  |
| 6          | }                                                                                  |
|            | RETURN Cn                                                                          |

〈Figure 6〉 Pseudo Code of Moving Data Clustering

〈Table 3〉 Purchased Product Ranking

| rank | product       | ratio(%) | rank                 | product    | ratio(%) |
|------|---------------|----------|----------------------|------------|----------|
| 1    | vegetables    | 22.6     | 14                   | beverage   | 2.0      |
| 2    | snack         | 12.6     | 15                   | detergent  | 1.6      |
| 3    | dairy         | 12.0     | 16                   | cleaning   | 1.6      |
| 4    | fruit         | 6.0      | 17                   | pork       | 1.1      |
| 5    | seasoning     | 5.6      | 18                   | health-tea | 1.0      |
| 6    | cold/frozen   | 5.3      | 19                   | hygiene    | 0.9      |
| 7    | finfish       | 4.8      | 20                   | interior   | 0.5      |
| 8    | cooking       | 4.6      | 21                   | pet-turf   | 0.2      |
| 9    | fast-food     | 4.1      | 22                   | vehicle    | 0.1      |
| 10   | chicken       | 3.9      | 23                   | handicraft | 0.05     |
| 11   | kitchen       | 3.1      | 24                   | alcohol    | 0.05     |
| 12   | beef          | 2.6      | 25                   | toy        | 0.03     |
| 13   | dried-seafood | 2.6      | 25 category products |            |          |

#### 4. 제안 시스템의 사례 적용

##### 4.1 고객 동선 데이터 수집

매장 고객 쇼핑 동선을 수집을 위해 대상 매장에 RFID 태그 약 300개를 쇼핑 카트에 부착하고 약 200여 개의 리피터가 매장에 설치하였고 27개의 싱크 노드를 매장 천장에 설치하였다. RFID 리피터는 1부터 200까지 고유 ID를 갖고 있으며 이를 참조하여 고객의 이동 동선을 추출하게 된다. 사례 분석 대상 매장은 채소류, 과자류, 유가공품 순으로 판매 순위를 보이고 있으며 〈Table 3〉은 25개의 대분류 상품의 판매 순위를 나타내고 있다. 동선 정보는 약 한달 정도의 쇼핑 동선 수집을 통해 총 16,640명의 고객의 동선 데이터를 축적하였다. 이 데이터 중 전처리 과정을 거쳐 최종적으로 총 13,334명의 고객 쇼핑 동선 데이터를 수집하였다.

##### 4.2 동선 분석 결과

###### 4.2.1 매대별 쇼핑 시간

〈Figure 7〉과 같이 쇼핑 시간이 오래 걸리는

매대로는 완구, 위생용품, 주류 매대로 볼 수가 있는데, 방문 수가 많지는 않지만 완구, 위생용품, 주류 같은 상품은 다른 일상품과는 달리 구경을 하는데 오랜 시간이 걸리는 것을 알 수 있다. 이를 통해 일상 용품인 채소, 냉장냉동, 조미식품, 인스턴트 같은 경우는 쇼핑시간이 오래 걸리지 않는 대신 자주 고객들이 지나치면서 소비를 하는 성향이 있지만, 일상품이 아닌 주류, 완구, 위생용품은 자주 방문을 하지는 소비를 위해 고객들이 많은 고민을 하고 있음을 알 수 있다.

이러한 쇼핑 시간 분석을 위하여 〈Table 4〉는 매대 별 고객 쇼핑 시간을 보여주고 있으며 〈Table 5〉은 특정 상품을 구매하기 까지 평균 쇼핑 시간과 비 구매 상품들의 평균 쇼핑 시간을 보여주고 있다. 〈Figure 8〉~〈Figure 9〉과 같이 완구, 자동차, 주류, 위생용품들은 구매까지 매우 오랜 시간이 소요되고 있고 비 구매 경우에도 완구, 자동차, 주류, 위생용품은 오랜 시간 동안 시간을 소비하는 것을 알 수 있다. 소고기와 채소의 경우는 구매를 위한 소요 시간이 길다

는 사실을 보여주고 있다. 이는 소고기와 채소의 경우 구매를 위해 가격과 품질의 비교를 위한 시간이 소요되는 것으로 추측해 볼 수 있다.

또한 분석 대상 매장의 시간대별 특성을 분석하면 <Figure 10>과 같이 주로 고객들이 오전 11시부터 오후 21시까지 방문이 주로 하고 있고 오후 17시에 가장 많은 방문 인원이 판매를 올리는 것을 볼 수 있다. 또한 <Table 6>과 같이 매장에서는 출근 시간을 넘긴 주부

들이 방문을 하여 퇴근 전까지 매장의 구매가 가장 활발히 이루어지고 있음을 알 수 있다. 이는 해당 매장이 서울 시내에 위치해 있고 판매 물품이 주로 일일 생활 용품을 주로 다루고 있기 때문에 생필품 위주로 매일 매일 판매가 되고 있는 것을 알 수 있다. 또한 매장 직원의 스케줄을 새로 작성할 때 매장의 피크(peak) 시간을 고려하여 새롭게 마련을 할 필요가 있다는 것을 보여주고 있다.

<Table 4> Summary of Average Shopping Time of Selves

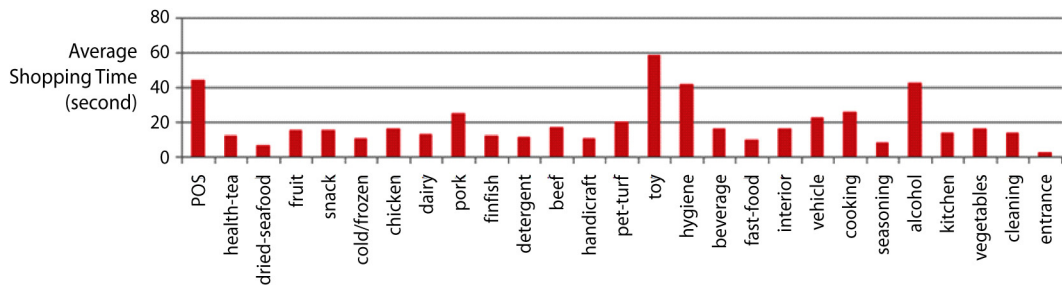
| product       | Entire Shopping time(hour) | Average Shopping time(second) |
|---------------|----------------------------|-------------------------------|
| POS           | 204.8                      | 44.4                          |
| health-tea    | 94.6                       | 13.1                          |
| dried-seafood | 62.0                       | 6.9                           |
| fruit         | 86.1                       | 15.6                          |
| snack         | 207.3                      | 15.6                          |
| cold/frozen   | 239.1                      | 11.3                          |
| chicken       | 70.0                       | 16.8                          |
| dairy         | 203.1                      | 13.3                          |
| pork          | 44.6                       | 25.1                          |
| finfish       | 79.9                       | 12.8                          |
| detergent     | 65.4                       | 12.1                          |
| beef          | 136.6                      | 17.6                          |
| handicraft    | 28.0                       | 11.2                          |
| pet-turf      | 46.6                       | 20.2                          |
| toy           | 24.5                       | 59.0                          |
| hygiene       | 75.7                       | 42.4                          |
| beverage      | 54.5                       | 17.0                          |
| fast-food     | 162.1                      | 10.5                          |
| interior      | 38.2                       | 16.4                          |
| vehicle       | 56.4                       | 22.9                          |
| cooking       | 138.8                      | 26.0                          |
| seasoning     | 161.9                      | 8.6                           |
| alcohol       | 4.4                        | 42.7                          |
| kitchen       | 82.2                       | 14.1                          |
| vegetables    | 503.0                      | 16.4                          |
| cleaning      | 39.8                       | 14.4                          |
| entrance      | 15.1                       | 3.1                           |

〈Table 5〉 Summary of Average Shopping Time of Selves for Non-Purchase

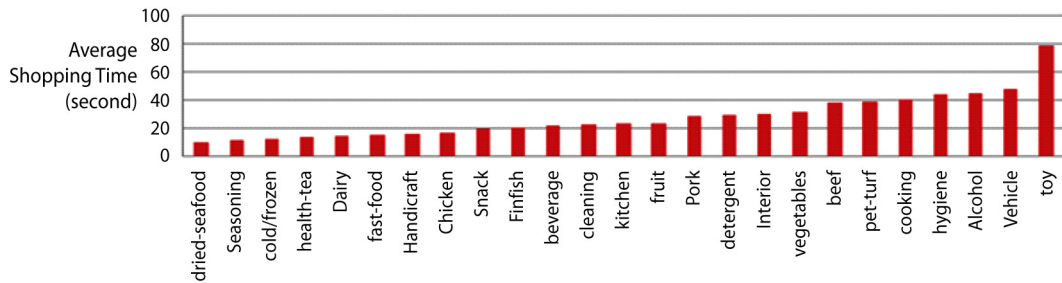
|               | Non-purchase average shopping time |                                | purchase average shopping time |                                |
|---------------|------------------------------------|--------------------------------|--------------------------------|--------------------------------|
|               | Total shopping time (hour)         | Average shopping time (second) | Total shopping time (hour)     | Average shopping time (second) |
| entrance      | 15.1                               | 3.1                            | 0.0                            | 0.0                            |
| dried-seafood | 54.6                               | 6.7                            | 7.4                            | 10.0                           |
| seasoning     | 143.5                              | 8.3                            | 18.7                           | 11.5                           |
| fast-food     | 144.0                              | 10.1                           | 18.5                           | 15.7                           |
| detergent     | 52.0                               | 10.5                           | 13.6                           | 29.2                           |
| finfish       | 51.3                               | 10.6                           | 28.7                           | 20.7                           |
| handicraft    | 27.8                               | 11.2                           | 0.3                            | 16.2                           |
| cold/frozen   | 220.6                              | 11.2                           | 18.9                           | 12.3                           |
| fruit         | 45.0                               | 11.9                           | 41.1                           | 23.9                           |
| vegetables    | 299.5                              | 12.3                           | 204.5                          | 31.5                           |
| kitchen       | 61.6                               | 12.5                           | 21.0                           | 23.3                           |
| cleaning      | 29.2                               | 12.7                           | 10.6                           | 22.8                           |
| dairy         | 152.2                              | 12.8                           | 51.1                           | 14.8                           |
| health-tea    | 90.5                               | 13.0                           | 4.2                            | 14.1                           |
| snack         | 136.5                              | 14.1                           | 71.3                           | 19.7                           |
| interior      | 33.4                               | 15.3                           | 4.8                            | 30.1                           |
| beef          | 108.3                              | 15.4                           | 28.5                           | 37.9                           |
| beverage      | 41.8                               | 15.8                           | 12.9                           | 22.4                           |
| chicken       | 50.9                               | 16.7                           | 19.4                           | 17.2                           |
| pet-turf      | 43.8                               | 19.6                           | 2.9                            | 39.0                           |
| cooking       | 85.1                               | 21.2                           | 53.7                           | 40.4                           |
| vehicle       | 53.8                               | 22.3                           | 2.6                            | 47.9                           |
| pork          | 35.0                               | 24.2                           | 9.7                            | 28.8                           |
| hygiene       | 63.1                               | 41.9                           | 12.6                           | 44.4                           |
| alcohol       | 3.7                                | 42.3                           | 0.7                            | 45.2                           |
| POS           | 207.1                              | 44.8                           | 0.0                            | 0.0                            |
| toy           | 23.9                               | 58.5                           | 0.7                            | 78.9                           |

〈Table 6〉 Average Purchase Price in a Time Slot

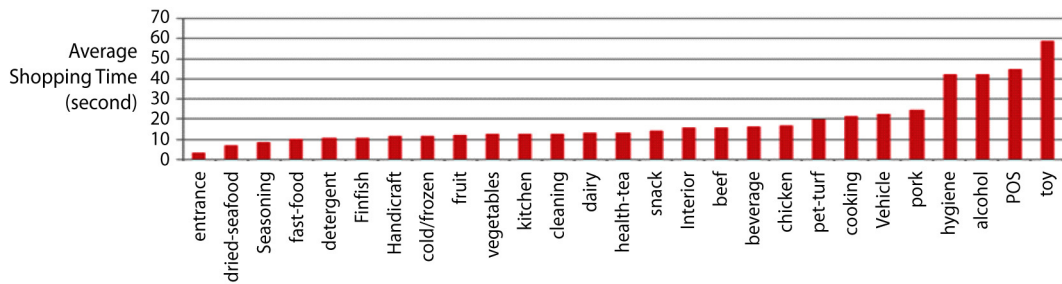
| Hour | count | Total price | Average price |
|------|-------|-------------|---------------|
| 9    | 54    | 2,566,655.  | 47,530.6481   |
| 10   | 304   | 18,905,658. | 62,189.6645   |
| 11   | 650   | 42,795,791. | 65,839.6785   |
| 12   | 764   | 50,799,448. | 66,665.9423   |
| 13   | 894   | 56,656,056. | 63,373.6644   |
| 14   | 1012  | 61,656,025. | 60,924.9259   |
| 15   | 1264  | 78,594,196. | 62,178.9525   |
| 16   | 1551  | 91,591,054. | 59,052.9039   |
| 17   | 1536  | 96,373,378. | 62,743.0846   |
| 18   | 1428  | 86,153,057. | 60,331.2724   |
| 19   | 1232  | 73,147,227. | 59,372.7492   |
| 20   | 1108  | 63,088,826. | 56,939.3736   |
| 21   | 859   | 52,577,888. | 61,208.2515   |
| 22   | 558   | 33,107,041. | 59,331.6147   |
| 23   | 120   | 5,999,588.  | 49,996.5667   |



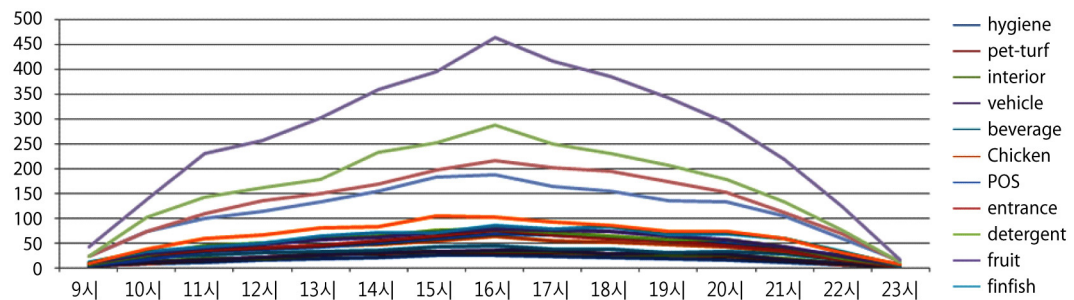
〈Figure 7〉 Average Shopping Time of Selves(second)



〈Figure 8〉 Average Shopping Time of Selves for Purchase(second)



〈Figure 9〉 Average Shopping Time of Selves for Non-Purchase(second)



〈Figure 10〉 Average Visiting Customers of Selves in a Time Slot

#### 4.2.2 매대 프로파일링 분석

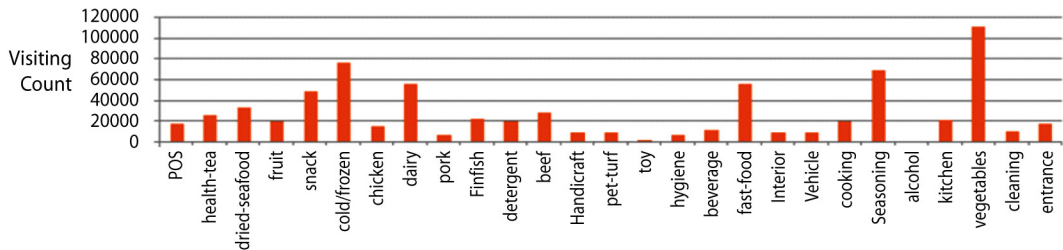
매대(Zone)마다 고객 방문 횟수, 체류시간, 구매액 뿐만 아니라 이들을 조합한 새로운 파생 변수를 생성시켜 다양한 조합으로 매대별 방문 고객수 비율, 매대별 고객 구매액 비율, 매대별 방문당 구매액 비율의 분석 지표를 개발하여 분석하였다. 이러한 분석을 통해 매장의 매대 별로 도출된 지표 값들을 순위화하여 매대 별 특징을 파악할 수 있다. 이것은 판매 이력 정보에만 초점을 두는 것이 아니라 매장에서 발생할 수 있는 여러 요인들과 결합하여 새로운 관계나 특이 사항의 유무를 발견할 수 있기 때문에 매장 운영 및 고객 서비스 개발 측면에서 중요하게 활용할 수 있다. 예를 들어, 고객의 방문수가 적으면서 판매가 많이 발생되지 않는 상품의 경우, 매장 운영자는 그 매대 지역의 문제는 없

는지 혹은 상품에 문제가 없는지 등을 파악하는 판단 기준이 될 수 있다.

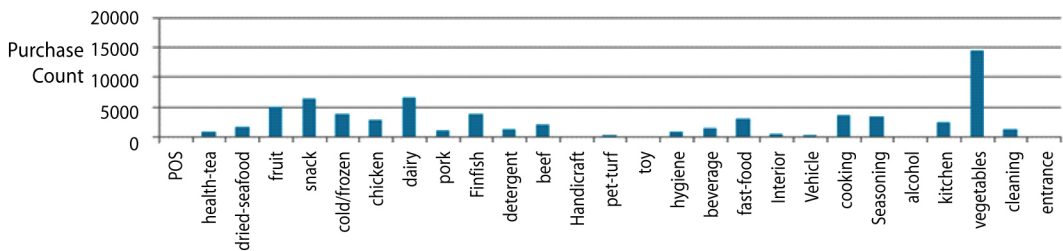
<Table 7>은 분석 대상 매장의 매대 별 특징을 보여주고 있다. 이러한 매장 매대의 프로파일 정보는 동선 쇼핑 패턴들과 함께 매장의 전반적인 현황 분석을 위한 자료로써 활용된다. 예를 들어, 매대 별 방문 수 분석을 한 결과 <Figure 11>와 같이 채소, 냉장냉동, 조미식품, 인스턴트 상품이 가장 활발히 고객 방문이 많은 판매 매대인 것을 알 수 있다. 이는 분석 대상 매장은 주로 일상 용품을 주로 판매하는 대형 슈퍼마켓이기 때문에 주로 소비되는 상품은 <Figure 12>에서 보는 바와 같이 식료품이 주를 이루고 있음을 알 수 있다. 또한 전반적으로 방문 수가 많은 상품들이 구매로 이루어지는 경향이 높다는 것을 <Figure 13>을 통해 판단 할 수 있다.

<Table 7> Profile Summary of Selves

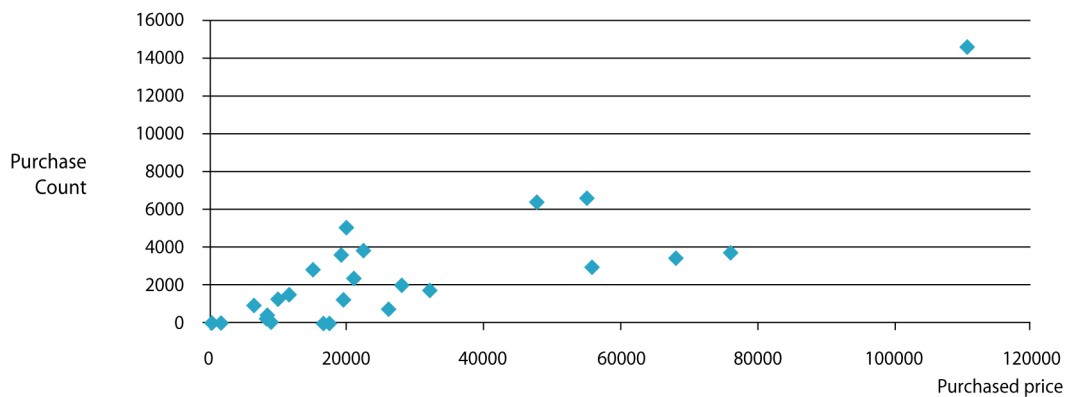
|               | Visiting count | Purchase count | Purchase count/visiting count |
|---------------|----------------|----------------|-------------------------------|
| POS           | 16600          | 0              | 0                             |
| health-tea    | 26081          | 732            | 2.806641                      |
| dried-seafood | 32156          | 1723           | 5.358254                      |
| fruit         | 19839          | 5066           | 25.53556                      |
| snack         | 47736          | 6398           | 13.40288                      |
| cold/frozen   | 76219          | 3751           | 4.921345                      |
| chicken       | 14971          | 2834           | 18.92993                      |
| dairy         | 55011          | 6607           | 12.01033                      |
| pork          | 6412           | 1012           | 15.78291                      |
| finfish       | 22401          | 3812           | 17.0171                       |
| detergent     | 19459          | 1256           | 6.454597                      |
| beef          | 28010          | 2018           | 7.20457                       |
| handicraft    | 9018           | 36             | 0.399202                      |
| pet-turf      | 8286           | 227            | 2.739561                      |
| toy           | 1495           | 25             | 1.672241                      |
| hygiene       | 6432           | 844            | 13.12189                      |
| beverage      | 11545          | 1523           | 13.19186                      |
| fast-food     | 55702          | 2979           | 5.348102                      |
| interior      | 8407           | 434            | 5.162365                      |
| vehicle       | 8851           | 152            | 1.71732                       |
| cooking       | 19218          | 3595           | 18.70642                      |
| seasoning     | 68048          | 3428           | 5.037621                      |
| alcohol       | 371            | 50             | 13.47709                      |
| kitchen       | 20964          | 2365           | 11.28124                      |
| vegetables    | 110538         | 14623          | 13.22893                      |
| cleaning      | 9951           | 1280           | 12.86303                      |
| entrance      | 17445          | 0              | 0                             |



〈Figure 11〉 Visiting Customers of Selves



〈Figure 12〉 Visiting Customers of Selves(Only Purchasing Customers)



〈Figure 13〉 Purchased Price vs. Visiting Customers

#### 4.2.3 동선 패턴 군집 분석

##### (1) 전체 이동 데이터 동선 군집 분석

고객 동선 분석을 위하여 본 연구에서 제안한 동선 군집화 알고리즘을 적용하여 고객 동선 분석을 실시하였다. 제안한 알고리즘의 임계값인 LCSS\_SIM을 0.17로 정하고 결과 분석한 결과 전체적으로 10개의 동선 패턴으로 군

집화 되었다. 각 군집의 특성은 <Table 8>와 같다. <Figure 14>은 도출된 군집들을 평균 이동 길이를 x축으로 매대별 방문 시간을 y축으로 원의 크기를 구매 금액을 기준으로 시각화한 결과이다. 이동 거리가 길어질수록 매장에서 쇼핑하는 시간이 늘어나고 있는 추세를 보여주고 있다. 또한 이동 거리가 길어질수록 구매액도 높아지고 있는 것을 보이고 있는데 마

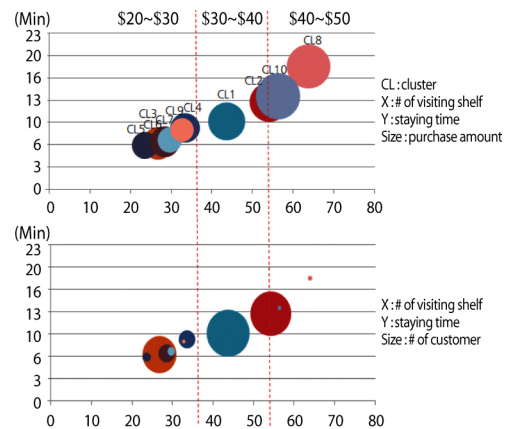
〈Table 8〉 Summary of Moving Clusters

|                       |                | C1     | c2     | c3     | c4     | c5     | c6     | c7     | c8     | c9     | c10    |
|-----------------------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Cluster Size          |                | 5640   | 4992   | 3567   | 801    | 221    | 788    | 188    | 67     | 53     | 54     |
| Average Visiting      |                | 43.7   | 54.2   | 26.8   | 33.5   | 23.6   | 28.5   | 29.6   | 63.8   | 32.7   | 56.3   |
| Average Shopping Time |                | 608.4  | 783.1  | 411.8  | 550.9  | 393.9  | 426.1  | 443.7  | 1102.5 | 534.9  | 832.7  |
| Sales                 | Entire         | 35,498 | 37,352 | 29,689 | 28,087 | 24,800 | 28,525 | 23,353 | 41,637 | 22,719 | 43,174 |
|                       | Below ₩150,000 | 32,526 | 33,750 | 26,757 | 27,072 | 24,021 | 26,731 | 23,353 | 29,955 | 16,466 | 35,527 |
|                       | Below ₩100,000 | 28,313 | 29,476 | 24,173 | 23,648 | 21,844 | 24,448 | 23,353 | 26,086 | 14,226 | 26,840 |

찬가지로 방문 시간이 늘어날수록 역시 구매액도 높아지는 경향을 볼 수 있다. 10개의 군집의 특성을 분석하기 위하여 평균 방문 매대 수, 평균 방문 시간, 평균 구매액, 평균 구매 수(대분류)를 분석하여 보았다. 구매액을 기준으로 2만원~3만원 사이의 구매 특징을 보내는 그룹으로 C3, C4, C5, C6, C7, C9 이 있으며, 3만원~4만원 사이의 구매 특징 그룹으로 C1, C2를 들 수 있고, 4만원~5만원 사이의 구매 특징으로는 C8, C10으로 나누어진다. 매대 방문 시간을 기준으로 분류를 해보면 400초~600초 사이의 그룹으로는 C3, C4, C5, C6, C7, C9이 있으며, 600초~800초 사이의 특징 그룹으로는 C1, C2가 있다. 마지막으로 800초 이상의 특성을 보이는 그룹으로는 C8과 C10으로 나눌 수 있다. 군집 집단의 고객 수를 기준으로 분석했을 때, 군집 C8과 C10은 개체 수가 상대적으로 작음에도 불구하고 구매액은 다른 군집들에 비해 높은 특징을 가지고 있고, C5, C7, C9는 개체수가 작지만 구매액은 C3, C4, C6과 비슷하다.

방문 특징을 기준으로 다시 10개의 군집들을 분석 해보면, C1, C2는 전체적으로 고르게 모든 매장을 방문하는 특징을 보이고 있으며, C5는 수예, 인테리어, 청소용품, 세제 매장을 상대적으로 방문을 자주하는 특징 있으며, C7은 완구, 애완원예, 자동차 매장을 상대적으로 방문을 자주하고 있다. C9 연기 완구, 애완원예, 자동차

코너를 상대적으로 방문을 많이 하고 특히 완구 코너에 방문율이 상당히 높다. 그에 비하여 C8은 닭고기, 과일, 채소, 생선, 견해산물 매장을 주로 방문을 하고 있으며, C10은 주류, 과자, 건강차 매장에 상대적 방문 비율이 높다.



〈Figure 14〉 Result of Moving Data Clustering

정리해 보면, C5, C7, C9는 주로 먹을거리 상품 보다는 공산품 위주로 관심을 많이 보이고 있으며 C8, C10은 먹을거리 위주의 상품에 주로 관심을 보이는 것으로 나타났다. 이를 통해서 C5, C7, C9에 속하는 고객들과 C8, C10에 속하는 고객들은 주로 관심 상품에 차이가 나는 것을 알 수 있게 되었으며 향후 프로모션에 다른 전략을 구사할 수 있을 것이다.

## 5. 결론 및 향후 연구

본 논문에서는 고객 구매 행동 정보를 이용하여 매장 운영에 효율화와 매출 향상에 도움이 되는 지능형 매장 분석 시스템을 제안하였다. 이를 통해서 이전의 단순한 구매 이력 데이터 분석에 그치지 않고 고객 행동과 매장 환경에 대하여 정량화하여 측정할 수 있었으며 이를 위해서 현업 전문가들과 함께 측정 지표를 개발하였다. 본 연구에서 제안한 시스템의 실제 적용 활용성을 평가하기 위하여 실제 대형 유통 매장을 선정하여 개발 시스템을 적용해 본 결과, 구매 이력에 대한 분석뿐만이 아니라 동선 분석을 통해 보다 다양한 분석이 가능해졌고 정량적으로 평가를 해볼 수 있었다.

그러나, 개인 신용정보 보호로 인해 보다 세세한 고객 분석이 부족하였다. 본 연구에서는 고객 정보로는 회원 아이디와 구매액과 구매 이력만을 사용할 수 있었기 때문에, 연령대별로 또는 직업별 또는 거주지별의 동선 특성을 파악할 수 없었다. 또한 시범 적용 기간인 한 달 동안의 동선 분석을 통해 새롭게 매장 배치를 바꾸어보고 이를 통해서 매출액이나 판매 추이의 변화 등까지는 분석하지 못하였다. 향후에는 보다 많은 고객 개인 정보와 충분한 기간의 운영을 하게 된다면 동선 분석 연구가 보다 활발히 진행될 것으로 기대되며, 보다 다양한 분석이 가능할 것으로 기대된다. 이를 통해서 유통 분야뿐만이 아니라 다양한 분야에 적용해 볼 수 있을 것으로 기대한다.

## References

- [1] An, J. M., A study on real-time preference goods recommendation using RFID-based shopping moving line analysis, Ph. D. Dissertation, Soongsil University, 2005.
- [2] Jung, I.-C. and Kwon, Y. S., "A New Approach to Spatial Pattern Clustering based on Longest Common Subsequence with application to a Grocery," IE Interface, Vol. 24, No. 4, pp. 447-456, 2011.
- [3] Jung, I.-C., A Study on Moving Object Trajectory Clustering(Case Study : Analysis of Shopper Movement in Grocery Supermarket), Ph. D. Dissertation, Dongguk University, 2012.
- [4] Kourouthanassis, P. and Roussos, G., "Developing consumer-friendly pervasive retail systems," Pervasive Computing, IEEE, Vol. 2, No. 2, pp. 32-39, 2003.
- [5] Hirschberg, D. S., "Algorithms for the longest common subsequence problem," Journal of ACM, Vol. 24, No. 4, pp. 664-675, 1977.
- [6] Hui, S. K., Bradlow, E. T., and Fader, P. S., "Testing Behavioral Hypotheses Using an Integrated Model of Grocery Store Shopping path and purchase Behavior," Journal of consumer research, Vol. 36, No. 3, pp. 478-493, 2009.
- [7] Hui, S. K., Fader, P. S., and Bradlow, E. T., "Path Data in Marketing : An Integrative Framework and Prospectus for Model

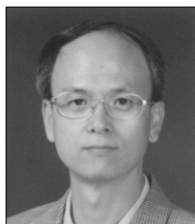
- Building,” *Marketing Science*, Vol. 28, No. 2, pp. 320-335, 2009.
- [8] Larson, J. S., Bradlow, E. T., and Fader, P. S., “An exploratory look at supermarket shopping paths,” *International Journal of Research in Marketing*, Vol. 22, No. 4, pp. 395-414, 2005.
- [9] Liao, I.-E. and Lin, W.-C., “Shopping Path Analysis and Transaction Mining Based on RFID Technology,” *RFID Eurasia*, 1st Annual, pp. 1-5, 2007.
- [10] Newman, A. J., Yu, D. K. C., and Oulton, D. P., “New insights into retail space and format planning from customer-tracking data,” *Journal of Retailing and Consumer Services*, Vol. 9, No. 5, pp. 253-258, 2002.
- [11] Uotila, V. and Skogster, P., “Space management in a DIY store analyzing consumer shopping paths with data-tracking devices,” *Facilities*, Vol. 25, No. 9, pp. 363-374, 2007.
- [12] Yada, K., “String analysis technique for shopping path in a supermarket,” *International Journal of Intelligent Systems*, Vol. 36, No. 3, pp. 385-402, 2011.
- [13] Yang, S. J., Shoppers’ shopping path pattern analysis using RFID data, M. E. Dissertation, Dongguk University, 2011.

## 저 자 소 개



정인철  
2003년  
2005년  
2012년  
관심분야

(E-mail : uhhaha@gmail.com)  
시립인천대학교 인문대 (학사)  
동국대학교 산업시스템공학 (석사)  
동국대학교 산업시스템공학 (박사)  
기계학습, 데이터 마이닝, Big Data 분석, 지능형 정보시스템,  
에이전트



권영식  
  
  
  
관심분야

(E-mail : yskwon@dgu.edu)  
서울대학교 산업공학 (학사)  
한국과학기술원 산업공학 (석사)  
한국과학기술원 산업공학 (박사)  
동국대학교 산업시스템 공학 교수  
데이터 마이닝, 지능형 정보 시스템



이용한  
1988년  
1990년  
1991년~1997년  
2002년  
2003년  
관심분야

(E-mail : yonghan@dgu.edu)  
서울대학교 산업공학과 (학사)  
한국과학기술원 산업공학과 (석사)  
대우자동차 기술연구소 (선임연구원)  
펜실베이니아 주립대학 산업공학과 (박사)  
동국대학교 산업시스템공학부 부교수  
RFID/USN, ABMS, BPM