J Korea Industr Inf Syst Res Volume 17 Number 5
http://dx.doi.org/10.9723/jksiis.2012.17.5.009

연관규칙 분석에 기초한 POS 데이터 분석 시스템†

(POS Data Analysis System based on Association Rule Analysis)

안 경 찬*, 문 창 배**, 김 병 만***, 신 윤 식****, 김 현 수**
(Kyung-Chan Ahn, Chang Bae Moon, Byeong Man Kim, Yoon Sik Shin, and HyunSoo Kim)

요 약 현재 전자상거래를 통한 상품추천 서비스가 많이 연구 되고 서비스되어지고 있다. 정보기술의 발달로 소규모 상점에서도 POS가 많이 보급되어 있지만 전자상거래에 비해 상품추천 서비스가 많이 이뤄지고 있지 않는 실정이다. 이러한 맥락에서 본 연구에서는 데이터마이닝 알고리즘을 POS 판매데이터에 접목하여 연관분석을 이용한 상품추천서비스 시스템을 구현하였다. 또한, 본 연구에서는 기존에 없는 서비스인 소멸규칙 및 새로운규칙, 상승 및 하향규칙을 제안하였다. 상품판매데이터를 이용하여 연관 분석한 결과는 고객에게 적용하여 상품추천서비스를 가능하게 하고 이와 더불어 소멸규칙 및 새로운규칙, 상승 및 하향규칙을 파악하여 경영자에게 경영 의사결정 정보로 제공해 주어 고객의 요구 변화에 신속한 대응이 가능하도록 하였다.

핵심주제어 : Apriori 알고리즘, 데이터 마이닝, 상품추천, POS, 연관규칙 분석

Abstract Merchandise recommendations service based on electronic commerce has been actively studied and on service these days. By virtue of progress in IT industry, POS has been widely used even in small shops, but the merchandise recommendations service using POS has not been much facilitated compared with that of using electronic commerce. This paper proposes a merchandise recommendations service system using association analysis by applying data mining algorithm to POS sales data. This paper, also, suggests novel services such as annihilation rule and new rule, and ascending and descending rules. The analysis results are applied to the customers enabling to offer merchandise recommendations service. In addition, prompt responses against the changes in demands from customers are possible by identifying the annihilation rule and new rule, and ascending and descending rules, and providing the management with the rules as managerial decision making information.

Key Words : Apriori Algorithm, Data Mining, Merchandise recommendations service, POS, Association Analysis

1. 서 론

정보기술의 발달로 소규모 상점에도 판매시스템(POS)을 사용하고 있으며 판매된 데이터가 컴퓨터에 실시간으로 저장되며 규모가 큰 프랜차이즈 업체는 각 지점에 쌓인 판매데이터가 수 분 안에 본사 서버

† 본 연구는 금오공과대학교 학술연구비에 의하여 연구되었음
* (주) 제니스정보기술 연구원
** 금오공과대학교 컴퓨터소프트웨어공학과
*** 금오공과대학교 컴퓨터소프트웨어공학과
**** 금오공과대학교 컴퓨터소프트웨어공학과(ysshin@kumoh.ac.kr)

에 옮겨진다[1, 2]. 이러한 판매데이터를 이용하여 고객에게 양질의 서비스를 제공하는 것은 고객확보의 중요한 요인이고 고객확보로 인하여 매출 역시 증대된다고 할 수 있다.

기존 판매데이터는 손익분석, 재고관리, 고객관리 등 제한적인 목적으로만 사용이 되었다. 이런 데이터는 단순 고객관리용으로만 사용되고, 고객의 구매습관을 파악하는데 거의 사용되지 않았다. 이러한 이유로 고객에게 제공되는 서비스는 판매자의 know-how에 의존할 수밖에 없었다. 만약 고객이 A라는 상품을 구입하고 A와 가장 어울리는 상품을 구입하고자 할 때 기존 상품 추천 방식은 판매자의 경험적 지식으로 상품을 추천하는 방식이다. 경험적 지식을 이용하여 상품을 추천하는 방법은 판매자의 경력에 영향을 받는다. 즉, 각 판매자들은 상이한 결과를 고객에게 제공할 수 있고 신입직원의 경우 고객에게 지식서비스를 제공하지 못하는 문제가 발생할 수 있다. 고객에게 양질의 서비스를 제공하기 위해 경험적 지식보다는 정형화된 서비스 제공 방법이 필요하다.

본 논문에서는 상품추천서비스에 더하여 기존에 없는 새로운 서비스 2가지를 제안한다. 첫 번째는 소멸규칙 및 새로운 규칙이며 두 번째는 상승규칙 및 하향규칙이다. 생성규칙이란 이전에 없던 규칙이 새롭게 발생되는 것을 말하고, 소멸규칙이란 이전에 존재하던 규칙이 소멸하는 것을 말한다. 이러한 정보들을 자동으로 분석해서 경영자에게 제공해 준다면 해당문제에 대한 신속한 대처가 가능할 것이다. 또한 경영자는 해당상품들에 대한 교차판매전략 구성, 기획 상품의 결정 등에 도움을 줄 수 있다.

소멸규칙과 새로운 규칙 외에 경영자에게 유용한 정보를 제공해 줄 수 있는 규칙이 하향규칙과 상승규칙이다. 만약 특정규칙을 분기별로 분석하였더니 서서히 상승하는 곡선을 나타내거나 이와 반대로 서서히 하강하는 곡선을 나타낸다면 이를 자동으로 분석하여 사용자에게 유용한 판매 분석 정보로 나타내 줄 수 있다. 경영자는 규칙들을 분석하여 해당 규칙이 나타내는 문제를 파악하여 빠른 대처가 가능할 것이다.

얼핏 보면 소멸 및 새로운 규칙과 상승 및 하향규칙을 비교해보면 비슷해 보이지만 경영자 입장에서는 전혀 다른 결과가 될 것이며 거기에 대처하는 정책도 달라질 것이다. 그 이유는 소멸 규칙이나 새로운 규칙은 해당규칙이 없어지거나 새롭게 나타나는 등 극단적인 형태로 발생되지만 하향규칙과 상승규칙은 그 규칙이 없어지거나 새롭게 발생되지 않고 서서히 변화되는 분석 결과이기 때문이다. 이런 특성으로 인해 경영자도 분석결과가 주는 특성에 맞는 대처를 할 수 있게 된다.

본 논문에서는 연관분석 방식을 이용하여 치킨 POS 데이터를 분석하였다. 연관분석의 대표적 알고리즘인 Apriori 알고리즘[3]을 사용하여 판매데이터를 분석하였고 이런 결과로 얻어진 분석 규칙을 이용하여 고객에게 추천하는 시스템을 구현 하였다. 이와 더불어 소멸규칙 및 새로운 규칙과 상승규칙 및 하향규칙을 발견하여 기업 경영자에게 제공함으로 빠른 의사결정에 도움을 줄 수 있도록 하였다. 실제 업무에 적용 가능성을 높이기 위해 약 2년 동안의 실제 판매데이터 2만 건을 사용하여 전처리 과정을 거쳐 분석하였다.

2. 관련연구

본 논문에서는 서론에 서술한 것과 같이 POS 판매데이터를 데이터마이닝 기법을 이용하여 분석하여 관련 규칙들을 찾아내고 이를 근거로 상승 및 하향 규칙과 소멸 및 새로운 규칙을 발견한다. 따라서 본 장에서는 먼저 상품추천시스템에 대한 관련 연구를 살펴보고 본 연구에서 데이터마이닝 분석 기법 중에서도 연관분석 기법에 대해서 살펴본다.

상품추천과 관련된 연구로 [1]은 상품추천 방식은 내용기반 상품추천방식과 협업필터링 상품추천방식이 있으며 두 방식 각각 장단점이 있으나 신규고객에 대한 추천 어려움과 판매자의 의도를 반영치 못하는 문제점이 있기 때문에 이를 보완하고 추천 정확도를 높이기 위해 온라인 전자상거래 데이터를 이용한 다중모형조합기법을 제시하였다.

[4]는 상품추천시스템을 위해 고객로그 기록 분석, 고객 구매패턴 분석, 웹 로그 분석을 수행하여 고객의 성향 및 패턴을 이용한 개인화된 상품을 추천하는 방법을 제시하였다.

[5]은 협업적 필터링을 통한 추천방식에서 이미 분류된 그룹에 특정 사용자가 분류되는 것이 아니라, 특정 사용자를 기준으로 그룹을 분류하는 방법을 사용하는 추천 시스템 기법을 제안 하였다.

[6, 7]에서는 내용기반 추천방식과 협업 필터링추천

방식의 문제점을 해결하기 위해 이 두 방법을 결합한 혼합기법들을 제공하였다.

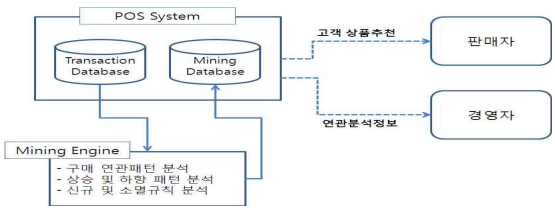
상기의 연구들은 상품의 내용을 분석하여 고객의 관심사항과 일치하는 상품들을 추천하거나 아니면 성향이 비슷한 다른 사용자의 구매이력을 이용하여 추천하는 방법들이다. 이러한 방법 외에 여러 사용자의 구매 자료를 분석하여 일반적인 구매패턴을 찾고 이를 바탕으로 추천하는 방법이 있다. 본 논문에서 사용하는 방법도 이 방법이며 구매 패턴을 파악하기 위해 데이터 마이닝 분야에서 많이 사용하는 연관분석 방법을 사용하였다.

Apriori 알고리즘은 연관 규칙을 찾아주는 알고리즘 중에서 가장 먼저 개발됐고, 또한 가장 많이 쓰이는 알고리즘이다. Apriori 알고리즘은 두 가지 단계로 구성된다. 우선 첫 번째 단계에서는 최소 지지도 설정 값에 따라 빈도수가 높은 항목들의 집합을 찾아내고, 집합들로부터 최소신뢰도에 보다 큰 값에 만족하는 연관 규칙을 모두 찾아낸다. Apriori 알고리즘에서 사용하는 중요한 법칙은 빈도수가 높은 항목 집합의 모든 부분 집합도 다 빈도수가 높다는 사실이다. 예를 들어 데이터에 {A, B, C}가 최소 지지도에 의해 빈도수가 높다면 당연히 {A, B}만을 봐도 빈도수가 높고, {B, C}을 봐도 빈도수가 높다. 다시 말해 어떤 집합이 주어졌을 때 새로운 항목을 더해주면 지지도는 절대로 전보다 증가할 수 없다[8].

3. 상품추천 및 의사결정 시스템

3.1 시스템 구성

POS 상품 추천 시스템의 주요 구성도는 <그림 1>과 같다.



<Fig 1> 시스템 구성도

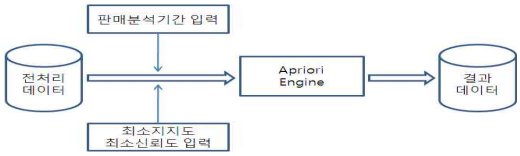
먼저 POS 시스템에 있는 POS Database에서 Transaction 정보를 추출하고, 데이터마이닝 엔진을 이용하여 구매 연관패턴을 분석한다. 연관 분석한 결과를 이용하여 다시 상승 및 하향 패턴을 분석하고 신규 및 소멸규칙을 분석한 후 분석한 결과를 Mining Database에 저장한다. 이렇게 추출된 패턴과 구매연관 패턴 분석 데이터를 이용하여 판매자에게는 고객 상품추천서비스에 이용하고, 상승 및 하향 패턴과 신규 및 소멸규칙 패턴은 경영자에게 경영에 도움이 되는 정보를 제공한다.

POS 상품추천 시스템의 분석 순서는 <그림 2>와 같다. 판매데이터 Database에서 기간별 판매정보를 가져와서 전처리 작업과 구매패턴분석을 한 후 이 결과를 이용하여 규칙들을 분석한다. 마지막으로 분석한 결과를 데이터마이닝 결과 Database에 저장 후 사용자의 요청이 있을 때 분석한 결과를 사용자에게 제공한다. <그림 2>의 전처리 과정은 일반 POS에 포함된 데이터를 바로 연관분석에 적용이 불가능하기 때문에 연관 분석에 적용하기 위한 전처리 작업이다.

<Fig 2> 상품추천 시스템 분석 순서

3.2 Apriori를 이용한 구매패턴 분석

구매패턴 분석절차는 <그림 3>과 같이 최소 지지도와 최소 신뢰도, 판매기간을 입력받은 후 연관분석을 실행하는 구조로 분석한 결과는 데이터마이닝 Database에 저장을 한다. 최종적으로 저장된 데이터를 사용하여 규칙들을 분석 한다.



<Fig 3> 구매패턴 분석 절차

3.3 상승 및 하향 규칙

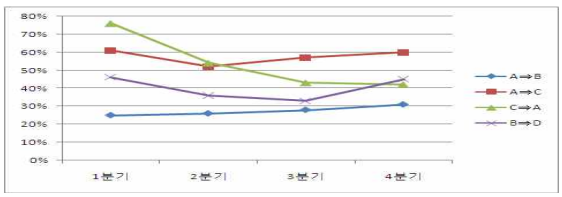
일반적으로 상품추천 서비스는 알고리즘에 의해 분석된 규칙을 그대로 사용자에게 적용하거나 경영자에게 제공하는 경우가 대부분이다. 하지만 분석된 규칙들 중에서도 특징적인 규칙을 자동으로 찾아서 경영자에게 제공한다면 소비자의 구매 변화에 대처가 쉬울 것이다. 이런 특징적인 규칙들 중의 하나가 바로 상승 및 하향규칙이다.

연관분석에서 사용되는 기준은 지지도와 신뢰도이지만 본 논문에서는 상승 및 하향규칙을 판단하는 기준으로 신뢰도를 사용하였다. 지지도는 전체거래에서 상품X와 Y가 판매되는 경우의 확률이며 신뢰도는 X라는 상품을 사면 Y라는 상품을 사는 경우의 확률이다. 즉, 지지도는 전체 거래수에 따라 수치가 변하지만 신뢰도는 오직 X와 Y의 판매만으로 구해지기 때문에 상승 및 하향규칙을 판단하는 기준으로 신뢰도를 사용하였다.

상승 및 하향규칙을 판별하기 위해서는 판매데이터에 대한 분석 작업이 선행되어야 하며 1차로 데이터마이닝을 통해 분석해낸 규칙을 이용하여 상승 및 하향규칙을 찾아낸다. 상승규칙은 분기별 판매규칙을 기준으로 꾸준히 상승하는 규칙을 말하며 하향규칙은 상승규칙과 반대로 꾸준히 하향곡선을 그리는 규칙이다. 상승규칙 및 하향규칙을 찾아내는 방법을 <표 1>에 예시하였고 이를 다시 <그림 4>에 그래프로 나타내었다.

<Table 1> 분기별 신뢰도 결과 및 판정 예

규칙	1분기	2분기	3분기	4분기	판정
치킨1⇒치킨2	25%	26%	28%	31%	상승 규칙
치킨1⇒찜닭	61%	52%	57%	60%	
찜닭⇒치킨1	76%	54%	43%	42%	하향 규칙
치킨2⇒소주	46%	36%	33%	45%	



<Fig 4> 간단한 분기별 신뢰도 그래프 예 (A : 치킨1, B : 치킨2, C : 찜닭, D : 소주)

<그림 4>에서 보는 바와 같이 “찜닭⇒치킨1”는 하향규칙이며 “치킨2⇒소주”는 상승규칙이다. 상승 및 하향규칙은 <표 1>에서 보면 쉽게 구분이 가지 않지만 그림 4에서 보이는 것처럼 꺽은선 그래프로 표시하면 쉽게 확인 가능하기 때문에 본 논문의 시스템에서도 그래프로 제공하는 방식을 사용한다.

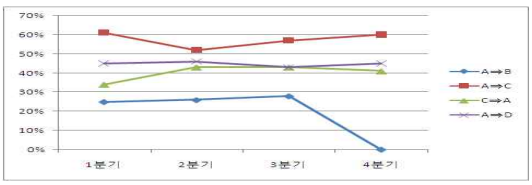
3.4 소멸규칙 및 새로운 규칙 파악 방법

소멸 및 새로운 규칙은 상승 및 하향규칙과 마찬가지로 신뢰도를 기준으로 분석하지만 상승 및 하향규칙이 분기별로 분석하는 것과 달리 기준월과 비교대상 월을 정하여 이 둘을 비교하여 소멸규칙 및 새로운 규칙을 찾아낸다. 기준과 비교대상은 날짜 범위를 서로 달리 할 수 있으며 기준을 1개월이나 특정분기로 정하고 비교대상기간을 1개월, 특정분기 혹은 1년을 정하여 서로 비교할 수 있다. 소멸규칙이란 기준을 3월로 하고 비교대상을 2월로 가정 했을 때 2월에 발생하였던 규칙이 3월에 발생하지 않았다면 이를 소멸규칙으로 판별한다. 생성규칙은 이와 반대로 2월에 없던 규칙이 3월에 발생하였을 경우이다.

소멸규칙의 예는 <표 2>와 같고, <표 2>에서 “치킨1⇒치킨2” 규칙은 1분기에서 3분기사이에는 규칙이 계속 발생되었지만 4분기에는 규칙이 나타나지 않았다. 이는 분기별로 신뢰도 수치가 변화되는 규칙과는 다른 각도로 해석이 필요하며 이런 소멸규칙만을 자동으로 분석하여 경영자에게 제공한다면 경영의사결정에 도움이 될 것이다. <그림 5>는 <표 2>를 그래프로 나타내었다.

<Table 2> 소멸규칙 발생 예

규칙	1분기	2분기	3분기	4분기	판정
치킨1⇒치킨2	25%	26%	28%	0%	소멸 규칙
치킨1⇒찜닭	61%	52%	57%	60%	
찜닭⇒치킨1	34%	43%	43%	41%	
치킨1⇒소주	45%	46%	43%	45%	

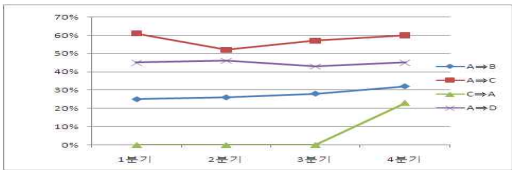


<Fig 5> 소멸규칙 발생 예 (A : 치킨1, B : 치킨2, C : 찜닭, D : 소주)

새로운 규칙의 예는 <표 3>과 같고, <표 3>에서 보는 바와 같이 규칙 “찜닭⇒치킨1”는 1분기에서 3분기까지는 규칙이 발생하지 않고 4분기에 새롭게 규칙이 발생하였다. 이런 규칙은 해당 규칙이 어떤 요인에 의해 판매가 갑자기 늘어났다는 의미이며 경영자에게 이런 자료가 제공된다면 세트 상품을 기획하거나 상품전시 위치의 변경 등을 위한 중요한 정보로 제공될 수 있다. <그림 6>은 <표 3>을 그래프로 나타내었다.

<Tabel 3> 새로운 규칙 발생 예

규칙	1분기	2분기	3분기	4분기	판정
치킨1⇒치킨2	25%	26%	28%	32%	
치킨1⇒찜닭	61%	52%	57%	60%	
찜닭⇒치킨1	0%	0%	0%	23%	새로운 규칙
치킨1⇒소주	45%	46%	43%	45%	



<그림 6> 새로운 규칙 발생 예(A : 치킨1, B : 치킨2, C : 찜닭, D : 소주)

4. 구현 및 실험결과

시스템 분석에 사용된 데이터는 국내 유명 치킨 프렌차이즈 업체 중 하나이며 대구 지역 판매점 1곳의 2008년부터 2009년까지 2년간 판매데이터이다. 본 논문에서는 전처리 과정을 거친 23,197건의 데이터를 사용하여 분석하였다.

4.1 구현

본 논문에서는 윈도우7 환경에서 Visual Studio 2008 C#을 사용하여 구현하였고, MS SQL 2005를 사용하여 데이터베이스를 구축하였다.

분석시스템의 메인화면은 <그림 7>과 같다. 사용자가 기간을 선택 한 후 조회를 하면 조건에 맞는 연관규칙이 조회 되며 정열 순서는 지지도가 높은 순으로 (descending) 조회되며 조회되는 항목은 규칙, 지지도(%), 신뢰도(%) 이다.

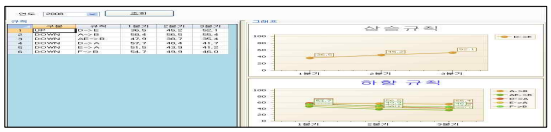
상세한 정보를 확인하기 위해 규칙을 클릭하면 선택한 규칙의 정보가 그래프형태로 우측에 제공된다. 우측 상단에 표시되는 그래프는 선택한 규칙이 월단위로 1월부터 12월까지의 지지도와 신뢰도 정보가 나타난다. 지지도는 Line 형태로 제공되고 신뢰도는 Bar 형태로 제공하여 구분을 쉽게 하였다.

우측 하단에는 선택한 규칙에 관련된 상품 각각 X와 Y의 월별판매량을 표시하였다. 하단에 각 상품별로 판매 현황을 보여준 이유는 규칙이 나타내는 지지도와 신뢰도의 수치와 상품 단독으로 판매되는 수치와의 관계가 비례하지 않을 수도 있기 때문에 사용자로 하여금 비교분석을 가능하도록 하였다.

<Fig 7> 분석 시스템 메인화면

규칙명은 특정업체의 상품명을 노출시키지 않기 위하여 본 논문에서는 수정된 이름으로 표시하였다.

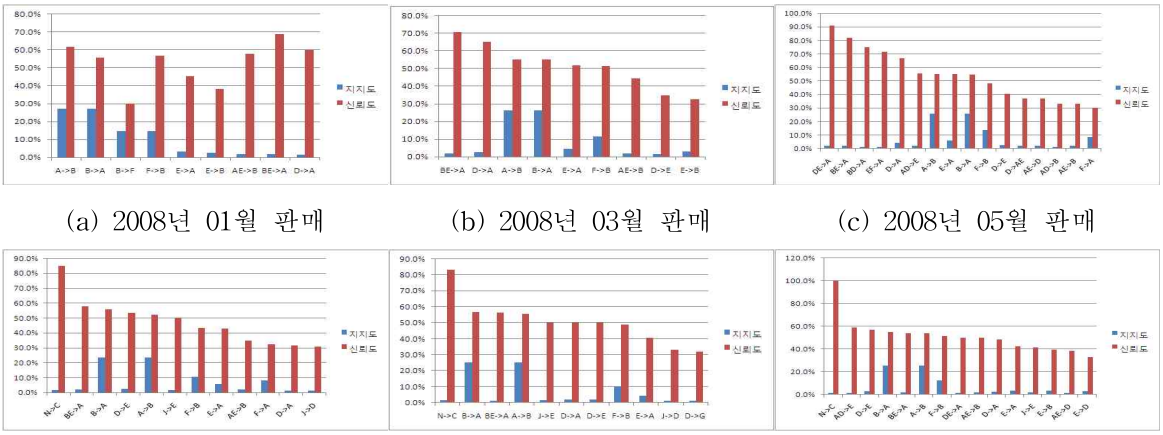
상승 및 하향규칙을 조회하는 화면은 <그림 8 a>와 같고, 새로운규칙 및 소멸규칙을 조회하는 화면은 <그림 8 b>와 같다.



(a) 상승규칙 및 하향규칙 조회화면

(b) 새로운 규칙 및 소멸규칙 조회 화면

<Fig 8> 규칙 조회화면



(a) 2008년 01월 판매 (b) 2008년 03월 판매 (c) 2008년 05월 판매

(d) 2008년 07월 판매 (e) 2008년 09월 판매 (f) 2008년 11월 판매

<Fig 9> 연관규칙 조회결과 (A : 치킨1, B : 치킨2, C : 찜닭, D : 소주, E : 생맥주, F : 치킨3, G : 날개, H : 콜라, I : 계란, J : 똥집, K : 순살치킨1, L : 순살치킨2, M : 강정치킨, N : 공기밥, O : 통구이, P : 바베큐, Q : 치킨4, R : 순살치킨3, S : 치킨5, T : 치킨6, U : 치킨7, V : 치킨8)

상승 및 하향규칙의 경우 연도를 선택한 후 조회를 하면 왼쪽에 상승 및 하항규칙이 분기별로 수치가 나타난다. 오른쪽 상단화면에는 상승규칙이 그래프로 표시되며 오른쪽 하단화면에는 하향규칙이 그래프로 표시된다. 새로운 규칙 및 소멸규칙의 경우 기간을 선택하고 조회하면 상단화면에는 연관규칙목록이 나타나며 왼쪽 하단에는 새로운 규칙목록이 나타난다. 오른쪽 하단에는 소멸규칙이 나타나는데 소멸된 규칙에 나타나는 지지도와 신뢰도는 전월 수치이다.

4.2 실험

4.2.1 연관분석 실험결과

Apriori 알고리즘을 이용하여 2008년 01월부터 12월까지의 홀수 월에 대하여 판매데이터를 최소 지지도 1% 최소 신뢰도 30% 기준으로 분석한 결과는 <그림 9>과 같다. 연관규칙 발생건수는 월별 최소 7개에서 최대 19개까지 발견되어 연관규칙건수가 월별로 최대 2배 이상 차이가 나는 것을 알 수 있다<그림 9>.

4.2.2 상승 및 하향 규칙 실험결과

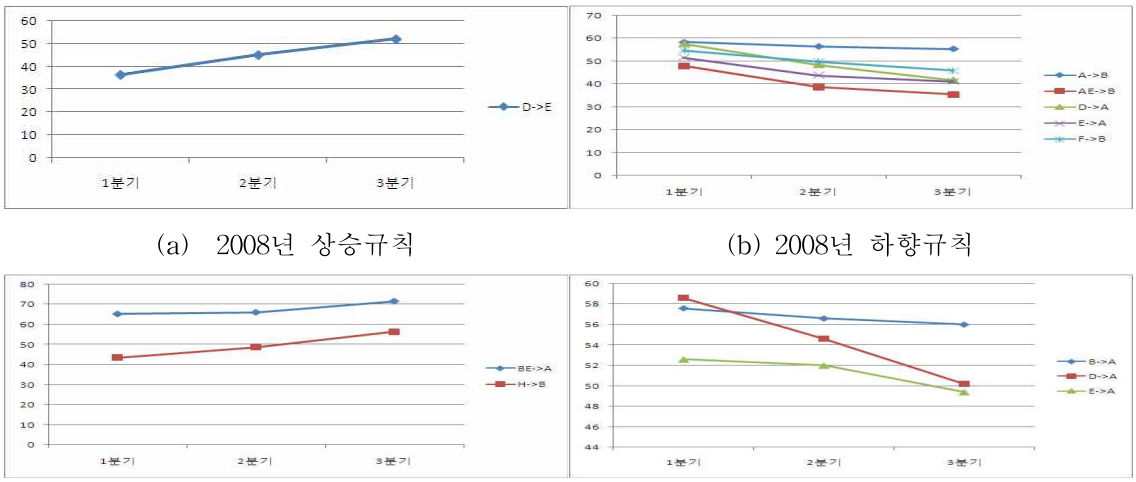
2008년 및 2009년 판매데이터를 분기별로 연관규칙을 찾아 1분기부터 4분기까지 상승규칙 및 하향규칙을 분석해 본 결과 해당되는 규칙을 찾지 못하였다.

본 논문에서는 상승규칙 및 하향규칙을 찾기 위해 1분기에서 3분기까지의 연관규칙을 기준으로 분석하였으며 결과는 <그림 10>과 같다.

적용결과 2008년 상승규칙은 “소주⇒생맥주”로 1건이 발생하였으며 3분기 까지 꾸준히 상승하였다. 하향규칙은 총 5건의 규칙이 발생되었으며 1분기에서 3분기까지 모두 하향곡선을 그린다. 2009년의 경우 상승규칙은 “치킨2,생맥주⇒치킨1”, “콜라⇒치킨2”등 총 2건이 발생되었고, 하향규칙은 “치킨2⇒치킨1”, “소주⇒치킨1”, “생맥주⇒치킨1”등 총 3건이 발생하였다.

4.2.3 소멸 및 새로운 규칙 실험결과

소멸규칙 및 새로운 규칙은 월단위로 분석이 가능하지만 본 논문에서는 2008년 07월부터 12월까지 하반기 규칙을 기준으로 2008년 01월부터 06월까지 상반기를 비교대상으로 하여 분석을 하였다. 분석한 결과는 <표 4, 5>와 같고, 2008년 하반기 연관규칙을 기준으로 No 9 ,10 규칙은 비교대상인 2008년 상반기에 없던 규칙이 새롭게 발생된 것이기 때문에 새로운 규칙이다. 반면, 2008년 상반기 연관규칙 목록 중에서 No 5는 2008년 하반기 기준으로 규칙이 발생되지 않아 2008년 하반기 기준으로 소멸규칙이다.



(a) 2008년 상승규칙 (b) 2008년 하향규칙

(c) 2009년 상승규칙 (d) 2009년 하향규칙

<Fig 10> 상승 및 하향규칙 조회결과 (A : 치킨1, B : 치킨2, C : 찜닭, D : 소주, E : 생맥주, F : 치킨3, G : 날개, H : 콜라, I : 계란, J : 똥집, K : 순살치킨1, L : 순살치킨2, M : 강정치킨, N : 공기밥, O : 통구이, P : 바베큐, Q : 치킨4, R : 순살치킨3, S : 치킨5, T : 치킨6, U : 치킨7, V : 치킨8)

<Table 4> 2008년 하반기 기준 2008년 상반기 대비 새로운 규칙

No	규칙	지지도	신뢰도	비고
1	치킨1->치킨2	25.80%	55.20%	
2	치킨2->치킨1	25.80%	56.90%	
3	치킨3->치킨2	11.40%	46.50%	
4	생맥주->치킨1	4.40%	44.10%	
5	소주->치킨1	2.90%	49.40%	
6	소주->생맥주	2.10%	36.10%	
7	치킨1, 생맥주->치킨2	1.90%	43.10%	
8	치킨2, 생맥주->치킨1	1.90%	64%	
9	공기밥->찜닭	1.40%	91.30%	새로운 규칙
10	똥집->생맥주	1.20%	36.50%	새로운 규칙
11	치킨1, 소주->생맥주	1%	36.20%	
12	소주, 생맥주->치킨1	1%	49.50%	

<Table 5> 2008년 하반기 기준 2008년 상반기 대비 소멸 규칙

No	규칙	지지도	신뢰도	비고
1	치킨1->치킨2	26.40%	57.50%	
2	치킨2->치킨1	26.40%	55.80%	
3	치킨3->치킨2	12.90%	52.30%	
4	생맥주->치킨1	4.20%	47.30%	
5	생맥주->치킨2	2.90%	32.40%	소멸규칙
6	소주->치킨1	2.20%	51.70%	
7	치킨1, 생맥주->치킨2	1.80%	43.20%	
8	치킨2, 생맥주->치킨1	1.80%	63%	
9	소주->생맥주	1.80%	42.10%	
10	치킨1, 소주->생맥주	1.10%	49.30%	
11	소주, 생맥주->치킨1	1.10%	60.70%	

5. 결 론

본 논문에서는 POS 판매데이터의 연관분석을 통하여 사용자에게 상품추천 서비스를 하고 경영자에게 상승 및 하향규칙과 소멸 및 새로운 규칙 정보를 제공하도록 하였다.

일반적으로 상품추천 시스템은 전자상거래를 통한 온라인 추천방식이 주를 이루고 있어서 POS 시스템에 대한 상품추천서비스가 많지 않고 또한 분석결과를 재분석하여 경영자에게 제공해주는 서비스 또한 많지 않다. 본 시스템을 실제 POS 시스템에 적용한다면 상품 판매자는 상품구매자에게 추가적인 서비스를 제공해 줄 수 있으며, 경영자에게는 경영 의사결정에 도움이 되는 정보를 줄 수 있다. 또한 본 시스템의 구조를 특정 POS 시스템에만 적용할 수 있도록 만들지 않고 다른 POS 시스템에 적용시키기 쉽도록 일반화하여 기존 POS 시스템에 이식이 쉽도록 하였다.

본 시스템은 상품추천 서비스와 더불어 상승 및 하향규칙을 찾는 시스템이다. 하지만 실험에 사용된 치킨 판매데이터가 판매되는 상품의 수가 제한적이고 또한 몇 개의 상품에 집중적으로 판매가 발생되어서 분석된 규칙도 일반적으로 생각하는 수준을 크게 벗어나지 않았다.

향후 본 논문은 상품의 수가 많고 여러 상품에 골

고루 판매가 일어나는 데이터에 적용한다면 현 실험보다 더 다양한 결과들을 산출해 낼 수 있을 것이다.

또한 본 논문에서는 상품별 연관관계를 찾는 실험을 진행하였으나 향후에는 좀 더 복합적인 환경을 고려한 연관관계를 찾는 연구가 병행되어야 할 것이다. 소멸 규칙과 새로운 규칙의 경우 비교 단위에 따라 다양한 결과가 발생할 수 있어 이에 대한 실험도 추가적으로 진행할 필요가 있다.

참 고 문 헌

[1] 이연경, 김경재, “데이터마이닝과 다중모형조합기법을 이용한 온라인상점 상품추천시스템 개발”, 한국지능정보시스템학회 추계학술대회논문집, pp. 340-348, 2004.

[2] 한창엽, 장형욱, “기업의 재고 수준 효율화를 위한 POS 시스템 활용에 관한 연구”, 한국산업정보학회 논문지, v.10, no.1, pp.81-88, 2005.

[3] 사재학, 남인길, “연관마이닝에 의한 데이터베이스 캐시 설계”, 한국산업정보학회논문지, v.7, no.2, pp. 16-32, 2002.

[4] 김지혜, 이경호, 이혜정, 박두산, "전자상거래를 위한 데이터 마이닝 기반상품 추천 시스템 개발", 한국정보기술학회논문지 제2권 제1호, pp. 47-54, 2004.

[5] 이세일, 이상용, “상대적 분류 방법과 시간에 따른 평가값 보정을 적용한 협력적 필터링 기반 추천 시스템”, 한국지능시스템학회 논문지 Vol 20, No. 2, pp. 189-194, 2010.

[6] 김병만, 이경, 김시관, 임은기, 김주연, "추천시스템을 위한 내용기반 필터링과 협력필터링의 새로운 결합 기법", 한국정보과학회논문지 : 소프트웨어 및 응용, 31권 3호, pp. 332-342, 2004.

[7] Qing Li, Sung Hyon Myaeng, Byeong Man Kim, “A Probabilistic Music Recommender Considering User Opinions and Audio Features”, Journal of IP&M, Vol. 43, No. 2, pp. 473-487, 2007.

[8] 심규석, “데이타 마이닝, 걷히는 안개를 바라보며”, 마이크로소프트웨어, 2001

안 경 찬 (Kyung-Chan Ahn)

- 비회원
- 1998년 : 경북대학교 식품공학과 (학사)
- 2011년 : 금오공과대학교 소프트웨어공학과 (석사)
- 현재 : (주)제니스정보기술
- 관심분야 : 소프트웨어공학, 디자인 패턴, 데이터마이닝, 전사적자원관리

문 창 배 (Chang Bae Moon)

- 학생회원
- 2007년 : 금오공과대학교 컴퓨터공학부 소프트웨어공학과 (학사)
- 2010년 : 금오공과대학교 소프트웨어공학과 (공학석사)
- 2010년 - 현재 : 금오공과대학교 소프트웨어공학과 (박사과정)
- 관심분야 : 인공지능, 감성공학, 영상처리, 정보검색

김 병 만 (Byeong Man Kim)

- 정회원
- 1987년 : 서울대학교 컴퓨터공학과(학사)
- 1989년 : 한국과학기술원 전산학과(석사)
- 1992년 : 한국과학기술원 전산학과(박사)
- 1992년 - 현 재 국립금오공과대학교 컴퓨터공학부 교수
- 1998년 - 1999년 미국 UC, Irvine 대학 방문교수
- 2005년 - 2006년 미국 콜로라도 주립대학 방문교수
- 관심분야: 인공지능, 정보검색, 정보보안

신 윤 식 (Yoon Sik Shin)

- 정회원
- 1982년 경북대학교 (학사)
- 1985년 한국과학기술원 전산학과 (석사)
- 1988년 - 현 재 국립금오공과대학교 컴퓨터공학부 교수
- 관심분야: 소프트웨어공학, 데이터마이닝

김 현 수 (HyunSoo Kim)

- 비회원
- 2008년 : 금오공과대학교 컴퓨터공학부(학사)
- 2010년 : 금오공과대학교 대학원 소프트웨어공학과 (공학석사)
- 2010년 - 현재 금오공과대학교 대학원 소프트웨어공학과 재학 중(박사과정)
- 관심분야 : 인공지능, 소프트웨어공학, 디자인 패턴

논 문 접 수 일 : 2012년 05월 03일
1차수정완료일 : 2012년 06월 05일
2차수정완료일 : 2012년 09월 24일
게 재 확 정 일 : 2012년 10월 15일