

연관 마이닝을 이용한 고객 관계 관리 적용

Customer Relation Management Application using Associative Mining

정경용*, 김종훈**, 류중경**, 임기욱***, 이정현****
상지대학교 컴퓨터정보공학부*, 대림대학 컴퓨터정보계열**,
선문대학교 컴퓨터정보학부***, 인하대학교 컴퓨터정보공학부****

Kyung-Yong Chung(kyjung@sangji.ac.kr)*, Jong-Hun Kim(ddcome@korea.com)**,
Joong-Kyung Ryu(jkryu@daelim.ac.kr)***, Kee-Wook Rim(rim@sunmoon.ac.kr)***,
Jung-Hyun Lee(jhlee@inha.ac.kr)****

요약

유비쿼터스 상거래에서 회사가 정보를 효율적으로 이용할 수 있도록 제어하고 필터링하는 일을 도와주는 고객 관계 마케팅이 등장하였으며, 더 나아가 고객이 원하는 제품을 예측하고 추천해주고 있으며 이를 위해 데이터 마이닝 기술을 적용하고 있다. 본 논문에서는 연관 마이닝을 이용한 고객 관계 관리 적용 사례 연구를 제안하였다. 제안된 방법으로는 연관 마이닝을 이용하여 후보 고객 집합으로 빈발 고객을 구성하고 연관 고객 규칙을 생성한다. 생성된 연관 규칙의 향상도에 따라서 하이퍼 그래프 분할을 이용하여 효율적인 구매고객들의 특성을 분석한다. 그러므로 기존 고객에 대한 교차 판매와 격상 판매의 전략들을 도출하게 된다. 성능평가를 위해 설문조사 데이터 집합에서 기존의 방법과 비교 평가하였다. 실험 결과 제안한 방법이 기존의 다른 방법들보다 정확도면에서 우수함을 확인하였다.

■ 중심어 : | 고객 관계 관리 | 데이터 마이닝 | 연관 규칙 |

Abstract

The customer relation marketing in which companies can utilize to control and to get the filtered information efficiently has appeared in ubiquitous commerce. It is applying data mining technique to build the management that can even predict and recommend products to customers. In this paper, we proposed the case of customer relation management application using the associative mining. The proposed method uses the associative mining composes frequent customers with occurrence of candidate customer-set creates the association rules. We analyzed the efficient the feature of purchase customers using the hypergraph partition according to the lift of creative association rules. Therefore, we discovered strategies of the cross-selling and the up-selling. To estimate the performance, the suggested method is compared with the existing methods in the questionnaire dataset. The results have shown that the proposed method significantly outperforms the accuracy than the previous methods.

■ keyword : | Customer Relation Management | Data Mining | Association Rule |

* "본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음"
(IITA-2008-C1090-0801-0020)

접수번호 : #080521-002

접수일자 : 2008년 05월 21일

심사완료일 : 2008년 06월 23일

교신저자 : 정경용, e-mail : kyjung@sangji.ac.kr

1. 서론

시장이 국제화되고 기업들 간의 경쟁이 치열해짐에 따라 방대한 양의 데이터베이스와 유비쿼터스 컴퓨팅을 기반으로 한 유비쿼터스 상거래의 발전으로 인하여 적응적 고객 관계 관리의 필요성이 더욱 부각되었다. 이러한 적응적 고객 관계 관리는 기업의 데이터를 분석하여 고객들에게 시간의 절약과 동시에 유용한 정보를 제공해 주고 비용 절감과 생산성 향상의 효과를 가져와야 한다. 데이터의 분석 기술 중에서 연관 마이닝은 고객 관계 관리를 하는데 가장 성공적인 기술이라 할 수 있으며 또한 현재 많은 기술들이 개발되고 연구되어지고 있다[1][2]. 본 논문에서는 연관 마이닝을 이용한 고객 관계 관리를 제안하였다. 제안된 방법은 보다 폭 넓은 고객들에게 서비스의 기여도 및 적극적인 고객 마케팅이 가능하며 새로운 비즈니스 기회와 교차 판매와 격상 판매를 통한 성과를 높임과 동시에 고객 만족도를 증진시킬 수 있다.

본 논문의 구성은 2장에서는 관련연구를 통해 고객 관계 관리와 연관 마이닝에 대해 살펴보고 3장에서는 제안한 연관 마이닝을 이용한 고객 관계 관리 방법을 기술하고 적응적 고객 관계 관리 전략을 소개한다. 4장, 5장에서는 본 제안된 방법을 T-test를 사용하여 다른 시스템과 성능을 평가하며 결론을 도출한다.

II. 관련연구

2.1 고객 관계 관리

적응적 고객 관계 관리는 기업이 지속적이고 안정된 성장을 유지하기 위하여 고객 관계 관련 데이터를 이용하여 고객의 가치를 파악하고 기존 고객에 대한 교차 판매와 격상 판매를 위한 일련의 활동으로 고객 마케팅, 고객 서비스 등이 포함되어 있다. 기업이 보유하고 있는 고객 데이터를 수집, 추출, 정제, 변형, 적재하여 고객 개개의 특성에 맞게 개인화 마케팅 전략을 계획하고 실시, 재평가하기 위한 과정이다. 특히, 유비쿼터스 컴퓨팅 기술의 발전으로 유비쿼터스 인터넷 시

대의 도래와 함께 온·오프 연동 및 유·무선 연동의 비즈니스 모델이 반영된 유비쿼터스 고객 관계 관리가 필요하다. 유비쿼터스 적응적 고객 관계 관리는 분석, 운영, 협업으로 나눈다. [그림 1]은 유비쿼터스 적응적 고객 관계 관리의 구성도이다. [그림 1]에서 데이터 분석에서 가장 중요한 정보를 추출하기 위해 효과적으로 구현하기 위해서는 사용자 모델링, 모바일 디바이스 서비스, 고객 데이터 웨어하우스, 다차원 분석 도구 및 데이터 마이닝, 캠페인 시스템, 채널 시스템, 자동화 시스템, 고객 지원 및 서비스 시스템이 있어야 한다.

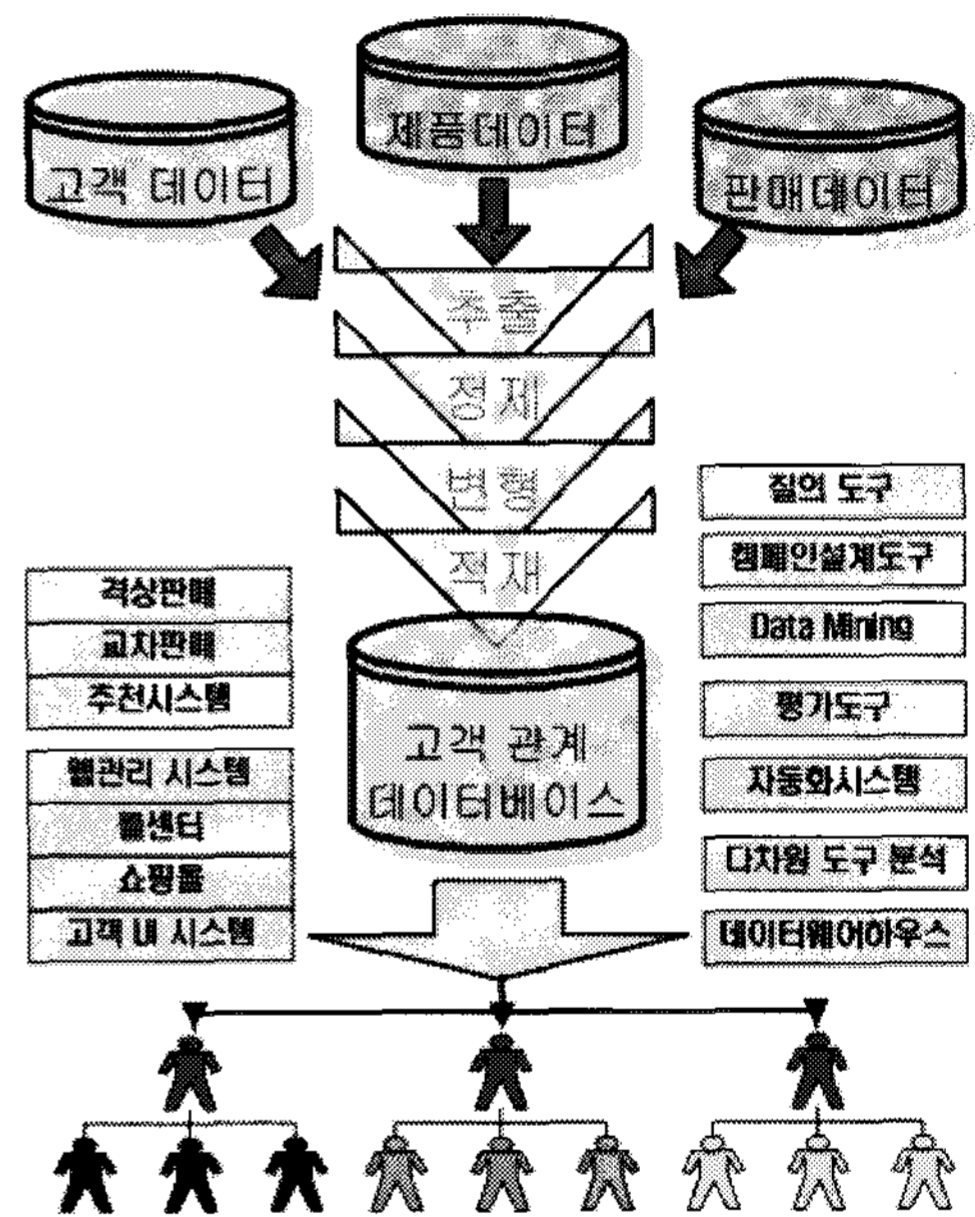


그림 1. 고객 관계 관리 구성도

유비쿼터스 적응적 고객 관계 관리의 목적은 고객에 대한 정보를 분석하여 보다 질 높은 개인화 서비스를 제공하고 고객과 적극적인 관계를 구축할 수 있는 수단으로 활용함으로써 고객의 가치를 극대화하는데 있다.

2.2 연관 마이닝

연관 규칙은 한 항목들의 집합과 다른 항목들의 집합 간에 강한 연관성이 있음을 나타내준다. 고객이 구매하는 물품들의 집합을 트랜잭션이라고 하고 이러한

트랜잭션들을 일정한 기간 동안 저장한 것을 데이터베이스라고 하면 제품 P₁을 구매한 고객은 제품 P₂도 구매한다는 것을 규칙으로 표현하면 “제품 P₁→제품 P₂, 지지도 15%, 신뢰도 84%”와 같이 표현할 수 있다 [3][5]. 여기서 15%의 지지도라는 것은 주어진 데이터베이스의 트랜잭션 중에서 15%가 제품 P₁과 제품 P₂를 동시에 구매했다는 것이고, 84%의 신뢰도라는 것은 제품 P₁을 구매한 고객들 중에서 84%가 제품 P₂를 구매했다는 것을 의미한다. 마이닝에서 연관 규칙 탐사는 발생한 트랜잭션들에서 제품들 상호간의 연관성을 발견해낼 수 있다.

하이퍼 그래프 분할은 다단계 과정을 걸쳐서 진행된다. 이를 멀티 하이퍼 그래프 분할이라고 하는데, 하이퍼 그래프 축소하는 단계, 축소된 하이퍼 그래프에 대하여 분할하는 단계, 분할된 하이퍼 간선을 원래 크기의 하이퍼 그래프로 확대하는 단계를 반복하면서 이루어진다[6].

III. 연관 마이닝을 이용한 고객 관계 관리 적용

고객이 자신의 선호도, 관심도, 구매 항목과 같은 정보를 기업에 제공하면 기업은 고객이 제공한 자료를 기반으로 가장 알맞은 정보를 제공한다. 적응적 고객 관계 관리를 통해서 고객에 관한 자료를 얻고 고객의 지속적인 이용이나 구매를 얻어낼 수 있게 되며 고객은 자신에게 맞는 적절한 정보를 편리한 방법으로 얻을 수 있다. 이러한 적응적 고객 관계 관리를 하기 위해 유비쿼터스 상거래에서 자주 사용되는 기법 중의 하나가 연관성 정보를 이용하는 것이다. 이는 데이터 마이닝을 이용하여 제품에 대해서 공동 구매한 고객의 집합 사이에서 연관성을 찾아내는 것이다. 특정 제품에 대해서 구매한 고객이 존재하고 동시에 같은 트랜잭션에 다른 고객이 존재할 때 이러한 고객들 간의 연관성을 발견하여 고객 관계 관리를 적용하는 방법에 대해서 기술한다.

3.1 고객 관계 트랜잭션 변환

온라인 기업의 데이터베이스는 제품 정보 테이블,

고객 테이블, 구매 테이블, 주문 테이블 등으로 구성되어 있다. 이러한 데이터를 기반으로 고객이 구매한 제품 트랜잭션으로 고객 관계 트랜잭션을 재구성한다. 적응적 고객 관계를 만들기 위하여 Apriori 알고리즘 [5]을 적용하기 전에 반드시 고객 관계 트랜잭션으로 데이터베이스로 전환시킬 필요가 있다. [그림 2]는 기업이 제공해준 데이터베이스에서 고객이 구매한 제품 트랜잭션이다. 구매한 제품 트랜잭션은 P₁, P₂,..., P_i(1 ≤ i ≤ 8)으로 표현하며 8개 구매한 제품들로 이루어져 있다. [그림 3]은 고객 관계 트랜잭션이다. 이는 C₁, C₂,..., C_j(1 ≤ j ≤ 20)으로 표현하며 20명의 고객들로 구성되어 있다.

[그림 2]에서 고객이 구매한 제품 트랜잭션을 [그림 3]의 고객 관계 트랜잭션으로 재구성한다. 고객 관계 트랜잭션에서 트랜잭션 번호는 고객이 구매한 제품을 의미하며 추출된 고객들은 후보 고객 집합과 고빈도 고객 집합을 구성하기 위한 것이다.

TID	Product set	TID	Product set
C ₁	P ₄ , P ₅	C ₁₁	P ₅ , P ₆ , P ₈
C ₂	P ₂ , P ₃ , P ₆ , P ₈	C ₁₂	P ₈
C ₃	P ₃ , P ₈	C ₁₃	P ₂
C ₄	P ₁	C ₁₄	P ₅ , P ₇ , P ₈
C ₅	P ₂ , P ₃ , P ₅ , P ₆	C ₁₅	P ₄ , P ₅
C ₆	P ₁	C ₁₆	P ₇
C ₇	P ₂ , P ₅ , P ₆ , P ₈	C ₁₇	P ₄
C ₈	P ₄	C ₁₈	P ₃
C ₉	P ₄	C ₁₉	P ₇
C ₁₀	P ₁	C ₂₀	P ₇

그림 2. 고객이 구매한 제품 트랜잭션

TID	Customer set
P ₁	C ₄ , C ₆ , C ₁₀
P ₂	C ₂ , C ₅ , C ₇ , C ₁₃
P ₃	C ₂ , C ₃ , C ₅ , C ₁₈
P ₄	C ₁ , C ₈ , C ₉ , C ₁₅ , C ₁₇
P ₅	C ₁ , C ₅ , C ₇ , C ₁₁ , C ₁₄ , C ₁₅
P ₆	C ₂ , C ₃ , C ₅ , C ₇ , C ₁₁
P ₇	C ₁₄ , C ₁₆ , C ₁₉ , C ₂₀
P ₈	C ₂ , C ₇ , C ₁₁ , C ₁₂ , C ₁₄

그림 3. 고객 관계 트랜잭션

3.2 연관 마이닝을 이용한 고객 관계 관리

연관 마이닝을 이용하여 고객의 구매 데이터베이스로부터 고객관계를 형성하는 과정을 기술한다. 고객의 구매 데이터베이스에서 고객이 구매한 제품들로 고객 트랜잭션으로 재구성한다. 고빈도 고객 집합과 후보 고객 집합은 고객 트랜잭션에 나타나는 고객들이다. 이러한 경우 n번의 검색을 통해 n개의 고객들로 구성된 연관 고객들을 마이닝한다. 각 단계에서 저빈도의 연관 고객은 연관 고객 마이닝에서 제외된다. 마이닝한 결과 고객들은 연관 고객들의 집합으로 나타내어진다. [그림 4]는 [그림 3]의 고객 관계 트랜잭션에서 추출된 고객들을 데이터 마이닝의 Apriori 알고리즘에 적용하여 연관 고객들을 유추하는 과정이다.

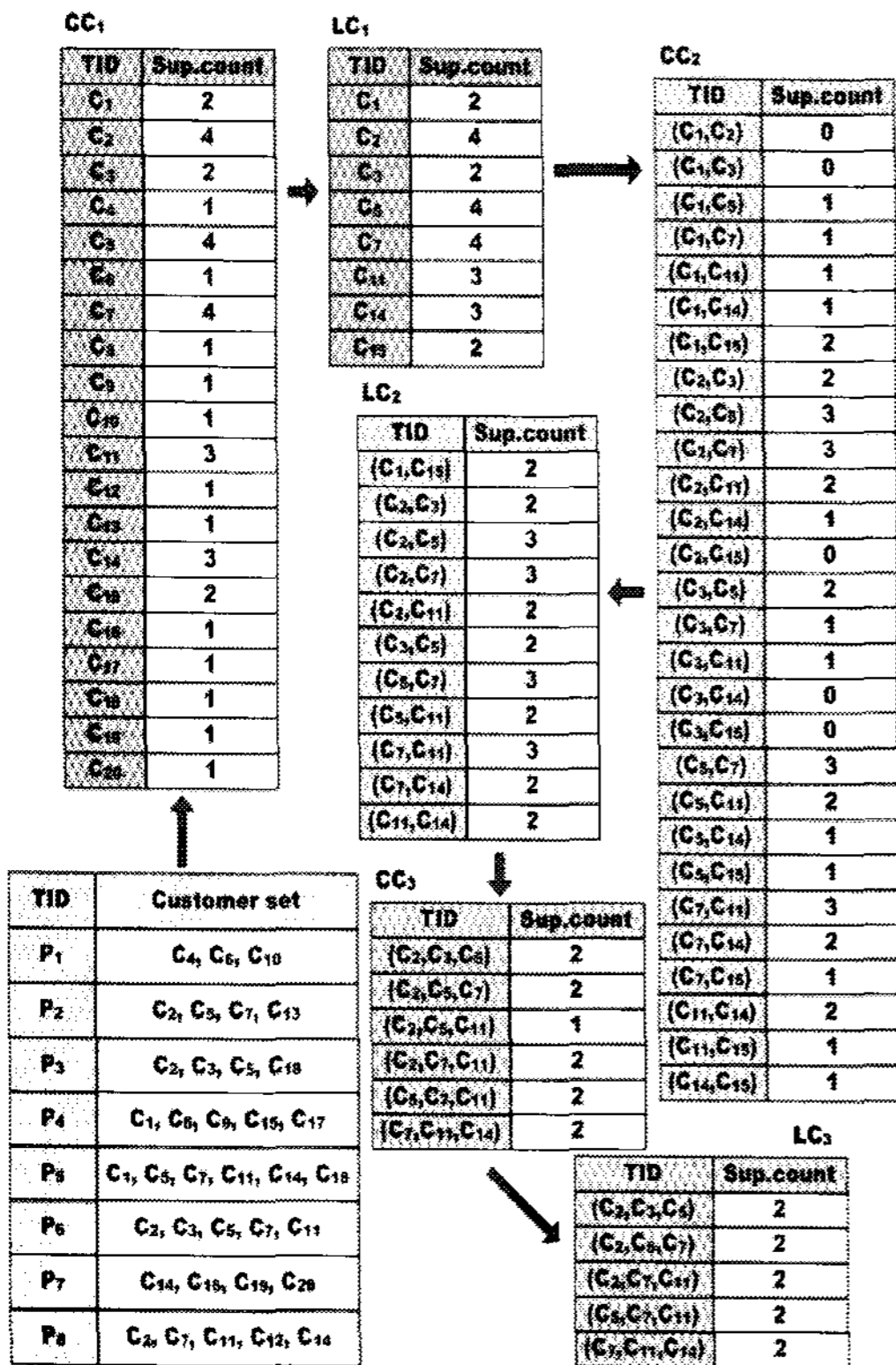


그림 4. 마이닝을 이용한 연관 고객 유추 과정

연관 마이닝은 첫 단계에서 후보 고객 집합(CC1)을 구성하며 이들의 지지도를 확인하기 위해 데이터베이스를 검색하고 고빈도 고객 집합(LC1)을 구성할 수 있

다[3][9]. 이와 같은 방법으로 Apriori 알고리즘의 두 번째 단계에서는 CC2, LC2를 구성하며 세 번째 단계에서는 CC3, LC3를 구성한다. LC3의 연관 고객 집합 {C2, C3, C5}, {C2, C5, C7}, {C2, C7, C11}, {C5, C7, C11}, {C7, C11, C14}로 추출된다.

3.3 하이퍼 그래프 분할을 이용한 고객관계 형성

연관 규칙과 하이퍼 그래프 분할을 이용하여 트랜잭션 기반의 데이터베이스에서 연관된 고객들 간의 관계를 형성한다. 연관 규칙을 사용하여 서로 관련있는 고객들을 군집하고, 고객들의 클러스터를 이용하여 서로 관련있는 트랜잭션을 군집한다. 하이퍼 그래프 H=(V, E)는 고객들로 구성된 정점들의 집합 V와 빈번한 고객 집합들을 나타내는 하이퍼 간선들의 집합 E로 구성된다[4][5]. 이는 항목간의 거리가 아닌 가중치를 이용하기 때문에 항목간의 거리 계산이 어려운 데이터 집합에 대한 군집에 유용하다[6]. 연관 규칙을 유도하기 위해 사용되는 고빈도 고객 집합은 고객들을 군집하여 하이퍼 간선으로 변환시키는데 사용되고 하이퍼 그래프 분할을 위한 가중치로는 연관 규칙의 향상도를 사용한다.

향상도 $lift(C_1 \rightarrow C_2)$ 를 결정하기 위한 식(1)은 다음과 같이 구해진다. 고객 트랜잭션의 수에서 고객 C2를 포함하고 있는 트랜잭션의 비율보다는 고객 C1을 포함하고 트랜잭션에서 고객 C2를 포함하고 있는 트랜잭션의 비율이 더 클 것이다. 따라서 고객 C1과 C2를 포함하고 있는 트랜잭션이 서로 상호 관련이 없다면, $Pr(C_2|C_1)$ 는 $Pr(C_2)$ 과 같게 된다.

$$lift(C_1 \rightarrow C_2) = \frac{Pr(C_2|C_1)}{Pr(C_2)} = \frac{Pr(C_1 \cap C_2)}{Pr(C_1)Pr(C_2)} \quad (1)$$

식(1)에서 $Pr(C_2|C_1)$ 의 값은 $Pr(C_2)$ 의 값보다 향상도가 배수만큼 크다. 그러므로 향상도는 1보다 크면 연관 고객들이 양의 연관 관계, 1보다 작으면 연관 고객들이 음의 연관 관계를 의미한다. 본 논문에서는 의미 있는 연관 고객의 규칙이 되려면 항상 향상도의 값이 1이상이 되어야 한다.

연관 고객 집합에서 모든 연관 규칙과 향상도를 구한 후, 하이퍼 그래프 분할을 이용하여 연관 고객 관계를 형성한다. 여기서 [그림 4]의 LC₃의 고빈도 고객 집합에서 하이퍼 그래프의 분할을 위한 가중치는 연관 규칙의 향상도를 사용한다.

예를들어 LC₃의 고빈도 고객 집합이 {C₂, C₃, C₅}라면 하이퍼 그래프는 고객들로 구성된 정점들의 집합 {C₂, C₃, C₅}과 연관 규칙으로 연결된 하이퍼 간선들의 집합으로 구성된다. 하이퍼 그래프 분할을 위한 가중치는 하이퍼 간선의 모든 고객들을 포함하는 연관 규칙의 향상도를 사용한다. 연관 고객 집합 {C₂, C₃, C₅}에서 동시에 일어날 수 없는 경우의 연관 규칙을 구한 후 각각의 연관 규칙의 향상도를 하이퍼 그래프 분할의 가중치로 부여한다. [표 1]은 향상도를 이용한 가중치를 나타낸다.

표 1. 향상도를 이용한 가중치 부여

연관규칙	향상도	연관규칙	향상도
{C ₂ }→{C ₃ , C ₅ }	1.34	{C ₃ }→{C ₂ , C ₅ }	0.99
{C ₂ , C ₃ }→{C ₅ }	1.52	{C ₃ , C ₅ }→{C ₂ }	0.89
{C ₂ , C ₅ }→{C ₃ }	1.35	{C ₅ }→{C ₂ , C ₃ }	1.21

[표 1]에서 {C₂, C₃, C₅}의 향상도가 하이퍼 그래프 분할을 위한 가중치이다. 연관 고객 집합에서 하이퍼 그래프를 이용해서 하이퍼 그래프 분할을 한 군집 결과는 [그림 5]와 같다[5]. 결과적으로 하이퍼 그래프 분할은 [그림 3]의 고객 관계 트랜잭션에 나타난 20명의 고객들을 [그림 5]와 같이 3개로 분할된 연관 고객 관계를 형성할 수 있다.

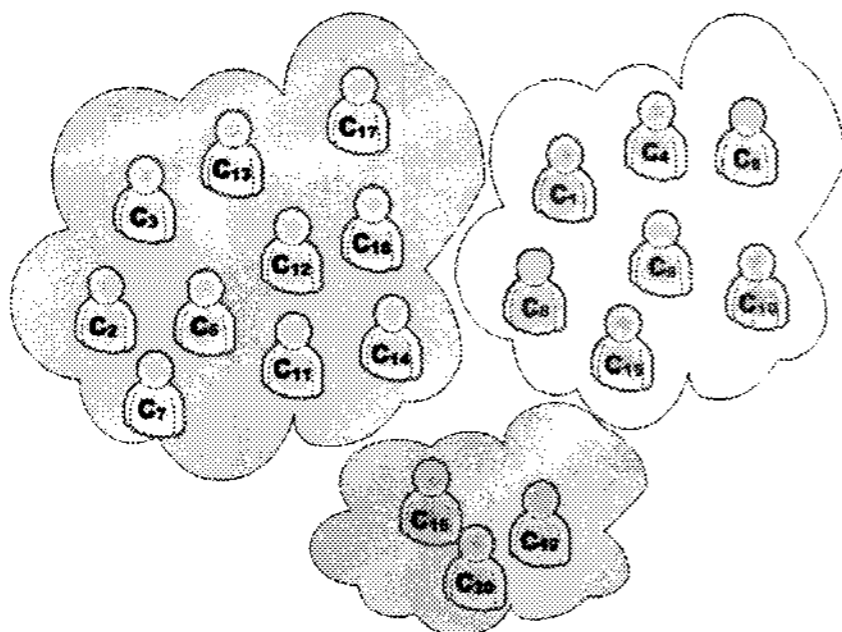


그림 5. 연관 고객 집합에서 하이퍼 그래프 분할

3.4 적응적 고객 관계 관리 전략

고객 관계 관리를 적용하는 온라인 기업의 경우 수천 명의 고객들에게 수천 가지의 서로 다른 제품을 판매하고 있다. 이러한 경우 차원 수는 매우 크게 된다. {연관 고객-제품} 행렬에서 고객 간의 연관 관계를 유지하면서 차원 수를 감소시키기 위해서 하이퍼 그래프 분할을 이용하여 연관 고객을 형성하게 된다. 따라서 고객의 성향과 유사한 성향을 갖는 고객 집단을 찾아 그 유사한 집단의 패턴 정보를 근거로 하여 서비스를 제공하게 된다. 예를 들어 동일한 패턴을 가진 고객 집단이 어떤 제품을 구매했다면 이와 유사한 성향을 나타내는 또 다른 고객에게 제품을 추천하는 것이다. 교차 판매와 격상 판매를 할 수 있는 기회를 보다 효과적으로 창출하고 신제품 홍보를 보다 빠르게 수행하게 함으로써 고객 마케팅을 효과적으로 수행할 수 있게 한다.

적응적인 고객 관계 관리는 고객으로부터 제품에 대한 피드백을 얻을 수 있고 이러한 고객의 피드백을 추적하고 분석함으로써 가장 적절하고 효율적인 고객 관계에서 캠페인 전략이 무엇인지 알 수 있다. 고객의 충성도를 높여 고객의 이탈 가능성을 줄이고 고객과 기업이라는 유기체를 통해서 효율적인 비용으로 성공적인 마케팅을 구현할 수 있다. 마지막으로 서로 다른 관심을 가진 고객들에게 적응적 관리를 통해 새로운 분야를 역동적으로 제시함으로써 확장성과 유연성이 있는 서비스를 제공할 수 있다[7][8].

IV. 실험 및 성능평가

4.1 실험 환경 및 평가 데이터

본 연구는 연관 마이닝을 이용한 고객 관계 관리 적용 사례 연구를 제안하였고, 제안한 방법의 적용 기업은 홈 네트워크 제조, 생산, 판매, 솔루션 개발하는 업체이다. 고객 만족도 평가를 위한 요소는 교차 판매, 격상 판매의 서비스 만족도이다. 평가를 위한 제품으로는 컴퓨터 부품, 홈 네트워크 구축 장비이다[9].

제안한 방법의 유효성과 타당성을 검증하기 위해 상

지대학교 취업 및 창업동아리인 상지벤처클럽 학우들의 도움으로 실제 현장의 오프라인 데이터를 수집하고 온라인 설문조사의 결과를 기반으로 평가 데이터를 구축하였다. 평가 데이터는 100명의 고객이 참여하여 제안한 방법에 의해 교차, 격상 판매 서비스된 제품과 일반적인 사용자간의 서비스된 제품에 대하여 1에서부터 5까지 1간격으로 명시적인 만족도 평가를 진행하였다. 만족도 평가에서 100명의 고객은 업체에서 파는 제품을 한번이라도 구매한 고객들로 남 65명, 여 45명으로 구성되었다. 평가 데이터에서 1은 설문한 요소에 대해서 매우 부정적인 평가를 의미하고, 5는 매우 긍정적인 평가를 의미한다. 온라인과 오프라인 설문조사는 15일 동안 진행하여 총 3,000개의 평가 데이터를 수집하였다.

본 논문에서 제안한 방법은 MS-Visual Studio C++ 6.0으로 구현되었으며 온라인 평가 데이터는 상지대학교 컴퓨터정보공학부 지능시스템 연구실 서버에 구성하였다.

4.2 분석 및 성능평가

성능평가는 연관 마이닝을 이용한 연관 고객 집합간의 교차, 격상 판매와 연관 마이닝 방법을 사용하지 않은 일반 고객 집합간의 교차, 격상 판매의 서비스 평가 데이터의 분포를 분석하고 T-test를 사용하여 연관 마이닝을 이용한 고객 관계 관리의 유용성을 평가한다.

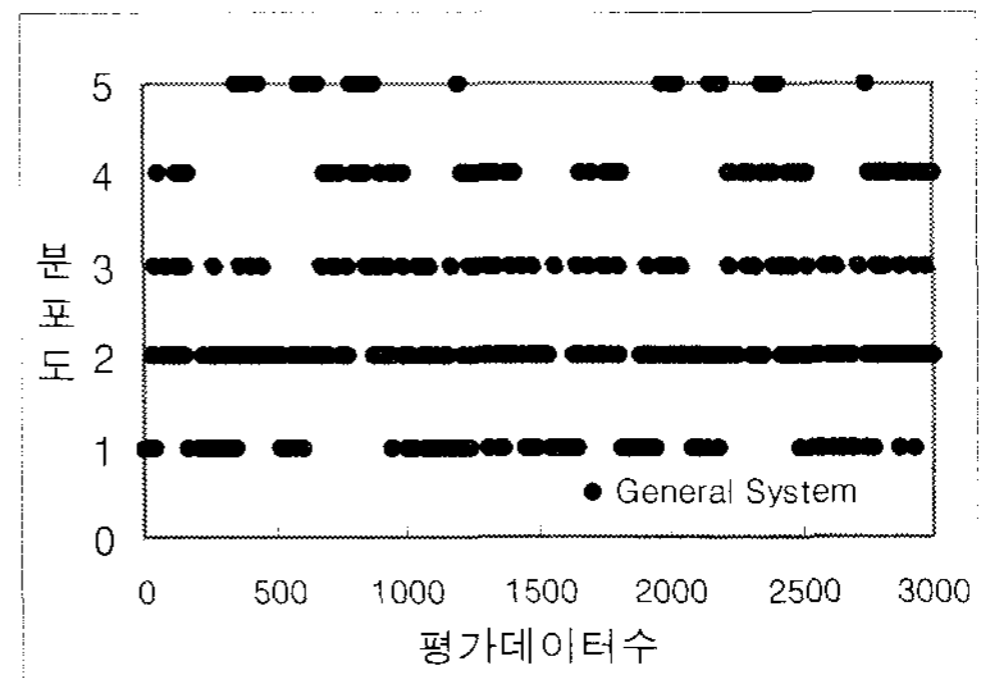
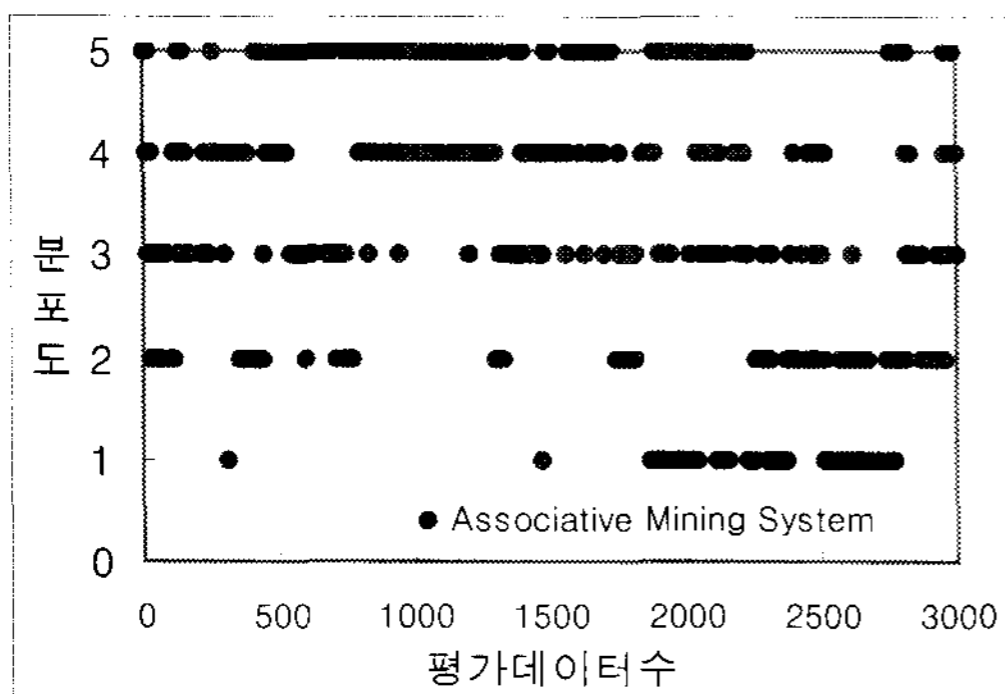


그림 6. 서비스에 대한 평가 데이터의 분포

[그림 6]은 서비스 만족도에 대한 평가 데이터의 분포를 나타낸 것이다. 평가 데이터에서 제안한 방법에 의해 서비스된 만족도의 분포가 만족도 측면에서 많은 고객들이 3점 이상으로 평가한 것을 볼 수 있다. 이는 고객들이 제안한 방법에 의해 서비스된 제품에 대해서 연관 마이닝을 적용하지 않은 방법보다 더 긍정적으로 만족하고 있음을 보여준다.

T-test는 통계적 검증을 함에 있어서 t 분포를 활용하는 기법이다. t 분포는 정상분포와 비슷하게 좌우대칭이며, 사례수의 크기에 따라 분포의 첨도가 달라지는 분포이다. T-test는 주로 두 집단 간의 평균에 차이가 있는지 여부를 검증하기 위해 사용된다. T-test는 크게 비교하고자 하는 두 집단이 독립표집인 경우와 두 집단이 종속표집인 경우로 구분한다[10].

본 논문에서는 성능 평가를 위해 100명을 대상으로 GS(General System)과 AMS(Association Mining System) 간의 서비스 평가 데이터의 차이를 통계적으로 검증하기 위해서 종속표집 T-test를 사용한다. 가설 H_0 는 “GS와 AMS의 서비스 만족도에는 통계적으로 차이가 없다.” 이고 H_a 는 “GS와 AMS의 서비스 만족도에는 통계적 차이가 있다.” 이다. [그림 7]은 AMS와 GS의 서비스 만족도 평가 데이터의 평균과 표준편차를 보여준다. 평균의 차이는 1.3593이다.

		평균	N	표준편차	평균의 표준오차
대응1	AMS	4.2733	3000	.78666	.01436
	GS	2.9140	3000	1.23657	.02258

그림 7. AMS-GS 대응표본 통계량

대응책					t	자유도
평균	표준편차	평균의 표준오차	사이의 95%신뢰구간			
			아만	상만		
1.35933	1.54327	.02818	1.30409	1.41458	48.244	2999

그림 8. AMS-GS 대응표본 검정

[그림 8]은 AMS와 GS의 대응표본 T-test 결과이다. 유의수준 α 가 0.05일 때 의사결정을 위한 값은 $t < 1.30409$ 이거나 $t > 1.41458$ 이다. 본 논문의 평가 데이터의 T-test에서는 t 가 $48.244 > 1.41458$ 이므로, H_0 를 기각하고 H_a 를 수용하게 된다. 따라서 GS와 AMS의 서비스 만족도에는 통계적인 차이가 있으며 평가 데이터의 평균의 차이인 1.3593 만큼 AMS가 GS보다 만족도가 높음을 확인할 수 있다.

V. 결론

기업은 지속적이고 안정된 성장을 유지하기 위하여 고객 관계 관련 데이터를 유비쿼터스 컴퓨팅 기술을 이용하여 고객의 가치를 파악하고 기존 고객에 대한 마케팅과 개인화 서비스를 해야 한다. 더 나아가 고객이 원하는 제품을 예측하고 추천해주고 있으며 이를 위해 데이터 마이닝을 고객 관계 관리에 적용하고 있다. 본 논문에서는 연관 규칙을 이용한 적응적 고객 관계 관리 전략을 제안하였다. 연관 마이닝을 이용하여 후보 고객 집합으로 빈발 고객을 구성하고 연관 고객 규칙을 생성한다. 생성된 연관 규칙의 향상도에 따라서 하이퍼 그래프 분할을 이용하여 적응적인 고객 관계를 형성한다. 따라서 고객의 성향과 유사한 성향을 갖는 고객 집단을 찾아 그 유사한 집단의 패턴 정보를 근거로 하여 서비스를 제공하게 된다. 이는 고객의 성향과 형태별로 세분화하여 각각의 고객 관계별로 적합한 마케팅 전략을 하거나 차별화된 서비스를 제공하는 것을 의미한다. 일반적인 상품 판매 서비스 시스템과의 성능 평가 결과, 상품 서비스에 대한 만족도의 차이가 통계적으로 의미가 있음을 증명하였고 본 논문의 연관 마이닝 고객 관계 관리 시스템이 일반적인 상품 서비스 시스템보다 더 높은 만족도를 보임을 확인하였

다. 따라서 본 논문의 방법을 활용하면 고객의 충성도를 높여줄 뿐 아니라 타겟 마케팅과 원투원 마케팅을 가능하게 해주고 기존 고객에 대한 교차 판매와 격상 판매의 전략들을 도출하게 된다.

참고 문헌

- [1] K. Y. Jung, H. J. Hwang, and U. G. Kang, "Optimal Associative Neighbor Mining using Attributes for Ubiquitous Recommendation Systems," LNAI 4027, pp.458-469, Springer-Verlag, 2006.
- [2] 조남재, 이용범, 정창미, "OLAP기반의 CRM 개발 전략 및 효과에 관한 연구: 유통산업의 사례를 중심으로", 한국데이터베이스학회 국제학술대회, pp.251-260, 2001.
- [3] 조동주, 정경용, 박양재, "연관관계 군집 분할 방법을 이용한 아이템 필터링 시스템", 한국콘텐츠학회논문지, 제7권, 제6호, pp.1-8, 2007.
- [4] M. O. Connor and J. Herlocker, "Clustering Items for Collaborative Filtering," In Proc. of the ACM SIGIR Workshop on Recommender Systems, Berkeley, CA, 1999.
- [5] H. Jiwei and K. Micheline, "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers, 2000.
- [6] E. H. Han, G. Karypis, and V. Kumar, "Clustering based on Association Rule Hypergraphs," In Proc. of the SIGMOD'97 Workshop on Research Issues in Data Mining and Knowledge Discovery, pp.9-13, 1997.
- [7] 이장희, "기업 대 기업 환경하의 지능적 CRM 활용", 한국지능정보시스템학회, 제1권, pp.226-229, 2003.
- [8] 이준규, 인터넷 개인화 아이템 추천 알고리즘에 대한 연구, 연세대학교 대학원, 석사학위논문, 2000.
- [9] 한기태, 정경용, 김종훈, 백준호, 류중경, 이정현, "연관 규칙을 이용한 적응적 고객 관계 관리 전

략”, 한국콘텐츠학회 2008 춘계종합학술대회 발표논문집, pp.84-86, 2008(5).

[10] 백순근, *교육연구 및 통계분석*, 교육과학사, 2007.

저자 소개

정 경 용(Kyung-Yong Chung) 정회원



- 2000년 2월 : 인하대학교 전자계산공학과(공학사)
- 2002년 2월 : 인하대학교 컴퓨터정보공학과(공학석사)
- 2005년 8월 : 인하대학교 컴퓨터정보공학과(공학박사)

- 2005년 9월 ~ 2006년 2월 : 한세대학교 IT학부 교수
- 2006년 3월 ~ 현재 : 상지대학교 컴퓨터정보공학부 교수

<관심분야> : 유비쿼터스 컴퓨팅, 인공지능시스템, 데이터마이닝, U-CRM

김 중 훈(Jong-Hun Kim) 정회원



- 2001년 2월 : 인천대학교 물리학과(학사)
- 2003년 2월 : 인하대학교 전자계산공학과(공학석사)
- 2007년 8월 : 인하대학교 컴퓨터정보공학과 박사수료

- 2008년 3월 : 대림대학 컴퓨터정보계열 전임강사
- <관심분야> : 유비쿼터스 컴퓨팅, 인공지능시스템, 데이터마이닝, U-CRM

류 중 경(Joong-Kyung Ryu) 정회원



- 1988년 2월 : 한국방송통신대학교 전자계산학과(이학사)
- 1991년 2월 : 인하대학교 산업대학원 정보공학과(공학석사)
- 1983년 ~ 1991년 : 대림산업 정보시스템실 대리

- 2003년 2월 : 인하대학교 컴퓨터정보공학과 박사수료
 - 1992년 3월 ~ 현재 : 대림대학 컴퓨터정보계열 부교수
- <관심분야> : 소프트웨어공학, HCI, ERP, CRM

임 기 욱(Kee-Wook Rim) 정회원



- 1977년 : 인하대학교 전자공학과(공학사)
- 1987년 : 한양대학교 전자계산학(공학석사)
- 1994년 8월 : 인하대학교 전자계산학(공학박사)

- 1977년 ~ 1983년 : 한국전자기술연구소 선임연구원
 - 1983년 ~ 1988년 : 한국전자통신연구소 시스템소프트웨어 연구실장
 - 1989년 ~ 1996년 : 한국전자통신연구원 시스템연구부장, 주전산기(타이컴)Ⅲ,Ⅳ 개발사업 책임자
 - 1997년 ~ 1999년 : 정보통신연구진흥원 정보기술전문위원
 - 2000년 ~ 현재 : 선문대학교 컴퓨터정보학부 교수
- <관심분야> : 실시간데이터베이스시스템, 운영체제, 시스템구조

이 정 현(Jung-Hyun Lee) 정회원



- 1977년 2월 : 인하대학교 전자공학과(공학사)
- 1980년 9월 : 인하대학교 전자공학과(공학석사)
- 1988년 2월 : 인하대학교 전자공학과(공학박사)

- 1979년 ~ 1981년 : 한국전자기술 연구소 시스템 연구원
 - 1984년 ~ 1989년 : 경기대학교 전자계산학과 교수
 - 1989년 1월 ~ 현재 : 인하대학교 컴퓨터공학부 교수
- <관심분야> : 자연어처리, HCI, 음성인식, 정보검색, 고성능 컴퓨터구조