

# eCRM을 위한 고객구매패턴 분석에 관한 연구

김남호\*, 이도현\*\*

\*호남대학교 정보기술원

\*\*전남대학교 컴퓨터정보학부

e-mail:nhkim@itc.honam.ac.kr

## A Study on the Analysis of Customer Purchasing Pattern for eCRM

Nam-Ho Kim\*, Do-Heon Lee\*\*

\*Information Technology Center, Honam University

\*\*Dep. of Computer Science & Information,  
Chonnam National University

### 요약

개별화 웹 마케팅은 본질적으로 고객지향의 패러다임이다. 즉, 개별 고객의 특수한 니즈를 개별적으로 파악해서 각각의 고객에게 차별화된 서비스를 제공하는 것이 그 핵심이다. 웹 서버의 고객이 접근한 상품의 로그파일에 데이터마이닝의 연관규칙 기술을 이용하게 되면 고객행동 패턴의 파악 및 예측을 위한 기법으로 활용할 수 있다. 본 연구에서는 웹 사용자의 교차 판매를 위한 원투원 마케팅에 필요한 접근패턴을 분석하고자 하며, 이는 고객의 상품에 대한 접근이 기록된 액세스 로그 데이터베이스의 분석을 통하여 이루어진다. 이들 데이터 처리와 교차판매 지원을 위한 데이터마이닝 모델링, 이를 통한 원투원 마케팅 모델 제시, 그리고 이의 활용이 고객관계관리(eCRM)에 미치는 효과를 제시한다.

### 1. 서론

최근에 기업의 마케팅 환경이 고객관계관리(CRM, Customer Relationship Management)로 대표되는 고객중심적 마케팅으로의 패러다임 변화가 일고있다.[5] 특히 인터넷은 회사의 상품과 서비스를 마케팅하기 위한 강력한 매체로 등장하였다. 종전의 대중매체를 통한 마케팅이 광고 주관업체가 불특정 다수의 잠재고객에게 마케팅하는 One-to-Many의 일방적 형태의 의사소통을 행한 것에 반하여 인터넷 마케팅은 가상의 공간에서 소비자와의 관계형성 및 실시간 상호작용이 가능한 쌍방향 커뮤니케이션을 통한 일대일 개별화 마케팅이 실현될 수 있는 최적의 환경을 제공하고 있다. 개별화 웹 마케팅은 본질적으로 고객지향의 패러다임이다. 즉, 개별 고객의 특수한 니즈를 개별적으로 파악해서 각각의 고객에게 차별화된 서비스를 제공하는 것이 그 핵심이다. 이를 통해서 기업은 개별 고객과의 관계증진을 통해 고객 유지율을 상승시키고 해당 회사에 대한 충성심

을 유지해서 결국은 이익을 증대시킬 수 있는 것이다.

지금까지 개별화 웹 마케팅을 위한 데이터는 사용자들의 회원정보와 통계 데이터로부터 추출되었다. 하지만 웹 서버의 로그파일에 데이터마이닝의 연관규칙 기술을 이용하게 되면 고객행동 패턴의 파악 및 예측을 위한 기법으로 활용할 수 있다.[1]

로그파일 분석의 데이터마이닝 결과는 eCRM 주제 중 하나인 교차판매(Cross Selling)에 적용된다. 즉, 한 제품을 구입한 고객이 다른 제품을 추가로 구입할 수 있도록 유도하는 것을 목적으로 한다. 이로부터 기존 고객이 보다 수익성 있는 제품을 구매하도록 하거나 고객의 거래 폭을 넓혀서 이탈을 어렵도록 하여 충성도 향상의 효과로도 연결된다.

본 연구에서는 웹 사용자의 교차 판매를 위한 원투원 마케팅에 필요한 접근패턴을 분석하고자 데이터마이닝의 연관규칙 기법을 사용하였으며, 이는 기존의 웹서버 로그파일 분석이 아니라 고객의 상품에

대한 접근이 기록된 상품 액세스 로그 데이터분석을 통하여 이루어진다. 이들 데이터 처리와 교차판매를 위한 데이터마이닝 모델링, 이를 통한 원투원 마케팅 모델 제시, 그리고 이의 활용이 고객관계관리(eCRM)에 미치는 효과를 분석하고자 한다.

논문구성은 1장에서는 연구배경과 목표에 대하여, 2장에서는 지금까지 수행되어온 관련연구, 3장에서는 데이터마이닝을 이용한 고객구매패턴 분석에 대해서, 4장에서는 시스템 설계와 평가에 대해서 그리고 5장 결론에서는 기대효과 및 향후 연구에 대하여 논하고자 한다.

## 2. eCRM과 데이터마이닝

eCRM과의 통합을 통해 인터넷 비즈니스는 단지 저렴한 매스 비즈니스 수준에서 정교하게 타겟팅 된 수익성 높은 비즈니스의 모습으로 변화될 수 있다. 이들간의 연계와 상호 보완적인 기능의 통합은 더 이상 새로운 시장이 생겨나기 어려울 정도로 포화상태에 이른 시장에서 기업의 생존과 지속적인 성장을 위한 통합된 고객중심 마케팅 전략 실행으로의 혁명을 실제로 가능하게 할 것으로 보인다.

효율적인 eCRM을 위해서 꼭 필요한 것 즉, 고객에 대한 특성 정보와 구매 의사결정 기준 그리고 구매 행위에 대한 정보를 획득하고 관리하는 과정에는 상당한 비용과 시간이 소요되며, 그 보다는 최초에 입력되는 고객 정보는 신뢰성이 부족하여, 지속적으로 갱신되지 않으면 eCRM을 통해 얻을 수 있는 이점은 거의 없다.

이렇게 수집된 데이터가 각종 통계, 데이터 마이닝 기법으로 분석되어 고객 개개인을 타겟으로 한 수준 높은 마케팅 조치가 설계될 수 있기 때문에 기존의 CRM과 그 격이 달라진다.

웹사이트에서 수집한 고객등록 정보들과 그 내용들에 대한 마이닝을 통해 기업은 인구통계학적 소비자 선호도를 발견하여 특정 광고나 배너를 포지셔닝 할 수 있도록 하는 기초자료를 추출할 수 있다. 새로운 데이터나 정보가 웹사이트를 통해 수집되면 이 정보들은 지속적으로 데이터웨어하우스로 통합되어 향후의 의사결정에 도움을 주는 분석결과를 제공하고, 데이터베이스 마케팅과 전략기획을 위한 자료로 활용되는 것이다. 또한 웹사이트 데이터 마이닝을 통해서 온라인 상에서 제공하는 서비스와 제품간의 연관관계를 밝혀내어, 적절한 제품이 적절한 서비스와 함께 판매가 되고 있는지의 여부를 밝혀내게 될 수도 있다.

데이터 마이닝을 통해서 기업은 웹사이트상의 패턴을 의미 있는 정보로 종합해내고, 인터넷 상의 고객들과 예상치들을 이해하고 연관시킬 수 있게 된다. 데이터와 웹이 제공하는 방대한 사업지식의 흐름에 근거한 웹 마이닝은 온라인 고객과의 관계를 생성하고 유지시키며 생산성 있는 온라인 상점의 최전선을 구축하는데 있어 결정적 열쇠가 되는 것이다.

## 3. 고객 구매패턴 분석

### 1) 정보수집

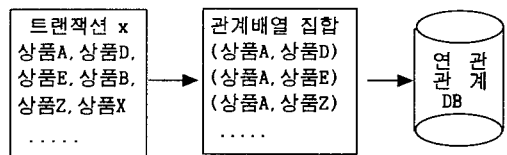
쇼핑몰 서버에서 고객의 상품에 대한 접근은 액세스 로그 데이터베이스에 기록되며, 로그는 고객이 접근한 상품의 웹 액세스 로그와 고객의 Session ID, 고객 ID를 기준으로 한다. 기존에는 웹서버가 생성하는 로그에서 필요한 정보를 수집하는데 이 경우에는 현재 데이터베이스와 연동되어 동작하는 CGI등을 이용한 시스템에서는 CGI파일의 로그만 생성되어 분석 시 가장 중요한 상품코드를 추출할 수 없다.

### 2) 전처리(preprocessing)

로그데이터에 대해 세션아이디를 사용하여 사용자별로 세션을 분류하고 세션 내에 유효한 회원 ID를 사용하여 사용자 별로 분류한다. 그리고 최대 시간 간격을 고려하여 Transaction단위로 구분한다.

### 3) 연관관계 분석

상품테이블의 상품코드로 전처리과정을 통해 정제된 데이터를 다음과 같은 방법으로 분석하여 상품간 연관관계를 계산한다. 상품테이블의 상품코드로 로그데이터의 한 트랜잭션 내에서 X라는 제품을 액세스한 고객이 Y라는 제품을 액세스한 연관관계를 검사한다. 각 트랜잭션단위에서 한 고객이 액세스한 상품들간의 관계를 관계배열 집합으로 생성하여 상품연관관계 테이블에 저장한다. 최종적으로 상품간의 연관관계 지지도와 신뢰도를 산정한다.



### ● 교차판매 방식에 의한 상품 프로모션

- 프로모션 대상 지지도 결정
- 상품간 교차판매 대상 확인
- 고객이 상품보기를 하였을 때 관련 상품리스트로

웹사이트에 프로모션

- 관련상품을 웹페이지에 리스팅할 때 연관관계에 의해 교환판매임을 명시하는 코드를 추가하여 실제 페이지뷰가 발생했을 때 로그에 남긴다.

● 프로모션 성공률 분석

일괄배치작업시 로그데이터로부터 프로모션된 상품의 로그만 추출하여 A라는 제품과 연관된 제품이 총 몇 회 프로모션 되었을 때, 고객들이 몇 번 액세스했는가를 검색하여 프로모션 성공률을 파악한다. 일정성공률이하인 연관정보는 프로모션 상품 대상 그룹에서 제외시킨다.

4) 제품지지도를 이용한 상품리스트 순위 변경

- 제품 지지도율이 높은 상품을 우선 순위로 상품테이블의 프로모션순번필드의 순위값을 변경
- 우선 순위에 의해 상품리스트를 소트하여 쇼핑물에 프로모션하여 상대적으로 우수한 상품을 고객들이 쉽게 액세스 할 수 있도록 한다.

5) 고객 유형에 의한 협업 필터링 (Collaborative Filtering)

고객분류테이블의 고객분류코드로 로그데이터를 다음과 같은 방법으로 분석 처리하여 고객분류그룹과 상품간의 연관관계를 파악한다.

- 고객분류테이블의 고객분류코드로 로그데이터 내에서 고객분류그룹이 각 제품을 액세스 카운트를 검사하여 고객 분류별 추천상품 테이블의 총 검색 카운트와 각 제품의 검출카운트를 로그
- 일정 지지도 이상이 되면, 프로모션을 시작
- 고객이 로그인하여 개인화 페이지에 액세스할 때 프로모션 회수를 증가
- 프로모션시 프로모션된 제품임을 코드값으로 명시하고 액세스 히스토리에 프로모션 코드를 저장
- 일정성공률이하인 상품은 프로모션을 정지

6) 고객 직중에 의한 협업 필터링

직중테이블의 직중코드로 로그데이터를 다음과 같은 방법으로 분석 처리하여 고객분류그룹과 직중간의 연관관계를 파악한다.

- 직중테이블의 직중코드로 로그데이터 내에서 직중 그룹이 각 제품을 액세스 카운트를 검사하여 직중별 추천상품 테이블의 총 검색 카운트와 각 제품의 검출카운트를 로그
- 일정 지지도 이상이 되면, 프로모션을 시작

- 고객이 로그인하여 개인화 페이지에 액세스할 때 프로모션 회수를 증가
- 프로모션시 프로모션된 제품임을 코드값으로 명시하고 액세스 히스토리에 프로모션 코드를 저장
- 일정성공률이하인 상품은 프로모션을 정지

7) 고객 취미에 의한 협업 필터링

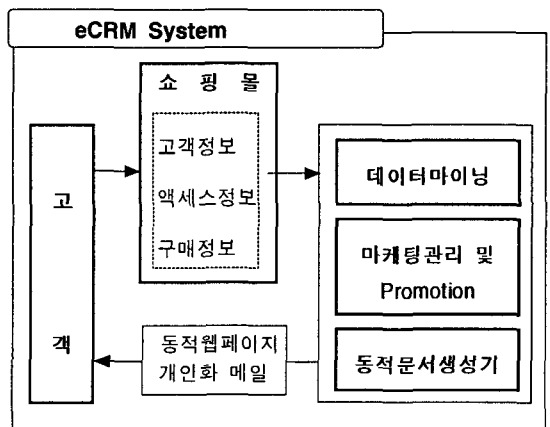
취미테이블의 취미코드로 로그데이터를 다음과 같은 방법으로 검색하여 고객분류그룹과 직중간의 연관관계를 파악한다.

- 취미테이블의 취미코드로 로그데이터 내에서 취미 그룹이 각 제품을 액세스 카운트를 검사하여 직중별 추천상품 테이블의 총 검색 카운트와 각 제품의 검출카운트를 로그
- 일정 지지도 이상이 되면, 프로모션을 시작
- 고객이 로그인 하여 개인화 페이지에 액세스할 때 프로모션 회수를 증가
- 프로모션시 프로모션된 제품임을 코드값으로 명시하고 액세스 히스토리에 프로모션 코드를 저장
- 일정성공률이하인 상품은 프로모션을 정지

4. 시스템 설계

1) 시스템 구성

본 시스템은 쇼핑몰 사이트 고객 개인의 구매성향을 파악하여 차별화된 서비스를 제공하는 것을 기본 목표 한다.



<그림1> 시스템 구조

CRM을 위한 분석용 데이터베이스를 구축하는 과정에서 첫번째 작업은 인터넷 사용자의 접점 데이터인 고객구매정보 항목들 중 분석에 필요한 항목들을 통합하는 모델을 작성하는 것이다. 이를 원시데이터의 전처리라고 하며 쇼핑몰 서버에서 고객의 상품에

대한 접근기록인 고객이 접근한 상품의 웹 액세스 로그와 고객의 Session ID, 고객 ID를 기준으로 한다.

데이터마이닝 모듈인 구매패턴분석기에는 전처리기의 트랜잭션에 포함된 접근한 상품정보가 항목으로 입력된다. 이 웹 상품정보에 연관규칙(Association Rule)을 적용하여 접근한 상품들간의 패턴지식을 찾아내고 패턴 데이터베이스에 저장한다.

마케팅관리 및 프로모션 시스템에서는 웹상에서 관리자가 조회 조건을 생성할 수 있는 인터페이스를 제공하여 고객의 성별, 연령, 관심도, 직업, 마일리지 등의 관련 조합에 의해 조회식을 용이하게 생성할 수 있도록 지원한다. 관리자는 생성된 조회식을 통해 마케팅 타겟을 결정할 수 있으며, 마케팅 정책에 의해 프로모션하는 상품을 정의하고 작성된 발송 메일 양식을 등록한다. 또한 발송일등을 스케줄링하여 자동발송될 수 있도록 한다.

개인화 메일 발송 시스템은 마케팅 타겟으로 결정된 고객 및 고객 집단에 대해 매핑된 특정 메일을 대상으로 발송할 수 있으며, 스케줄링을 통해 자동으로 발송 처리한다.

## 2) 시스템 평가

활용성을 중심으로 제안한 시스템을 평가하고자 한다. 웹사이트를 통한 소비자 구매 행동의 추적 및 이를 활용한 eCRM시스템 구축은 현 수준에서는 아마 남의 생각을 읽는 독심술 이상으로 막강한 마케팅 무기로 활용할 수 있을 것이다. 인터넷 사이트 운영자와 쇼핑물들이 일반적인 미디어와 차별될 수 있는 강력한 힘은 바로 소비자의 행동패턴을 통한 원하는 정보나 구매를 예측할 수 있다는 점이고 이는 지금까지의 인터넷 미디어들이 보완해야 할 부분이다.

구매행동 활용을 통한 타겟 마케팅이란 어떤 것인지 제안한 시스템의 예를 들어 살펴보자. 기념일을 맞이하여 쇼핑물의 꽃배달 서비스를 통해 장미패키지를 구매한 사람은 동일한 연인에게 보낼 샤넬 향수를 화장품코너에서 구매하고, 또 뮤직코너에서 좋아하는 가수의 CD를 구매했다고 한다. 이 사람에게 향후 구매했던 아이템을 근거로 일정 형태의 구매 패턴을 갖고있는 것이 확인되므로 그에 맞는 제안형 추천을 함으로써 직접 매출을 유발할 수 있을 것이다.

## 5. 결론

연관규칙을 이용한 데이터마이닝 기술을 소비자의 상품구매 데이터에 적용하여 인터넷 사용자의 구매 행동을 파악하였다. 데이터 처리를 위한 시스템과 생성된 규칙을 이용한 다양한 마케팅 기법들을 소개하였다.

새로운 방문자에 대한 예측 분류나 행동 양식에 대한 예측에 기초하여, 우리는 비로소 그 사람에게 가장 적절한 광고와 마케팅 메시지들을 전달할 수 있게 된다. 즉, 그 사람이 어떤 상품을 구매하고 싶어하는지를 먼저 예측하여 그 사람에게 적절한 상품 판촉 메일을 보낼 수 있게 되는 것이다. 이는 인터넷이 기업의 대고객 접점으로서 중요성이 날로 중요시 되는 현시점에서 쇼핑물이나 웹사이트 운영 기업의 고객관계관리(eCRM)를 위한 원투원 마케팅의 중요한 기초 데이터로 활용될 수 있다.

본 논문에서 제안한 시스템은 쇼핑물을 운영하는 기업의 입장에서 지속적인 고객관리에 목표를 둔 eCRM 시스템이다. 하지만 인터넷을 중심으로 한 디지털 경제가 이제는 무선인터넷 환경으로 변화하는 시점에 와있다. 무선인터넷 환경에 맞는 비즈니스 모델 개발과 여기에 맞는 eCRM에 관한 연구가 필요하다.

## 참고문헌

- [1] Dong-Ha Lee, Dong-Yal Seo, Nam-Ho Kim, and Jeon-Young Lee, Discovery and Application of User Access Patterns in the World Wide Web, 4th World Congress on Expert system 98, March 16-20, Mexico City.
- [2] Bamshad Mobasher, Robert Cooley and Jaideep Srivastava, "Automatic Personalization based on Web Usage Mining", Communications of the ACM, 43(8), pp.142-151, 2000.
- [3]이도헌, "데이터마이닝을 이용한 CRM", 정보과학회지, 18(11), pp4-11, 2000.11.
- [4]한국오라클, "Oracle CRM: Release11i Statement of Direction ", 오라클 백서, 2000.
- [5] 한재홍, "CRM특강", <http://www.dpc.or.kr/dbworld/document/9912/tech-1.html>
- [6] H. Edelstein, "Building Profitable Customer Relationships with Data Mining", SPSS White Paper, 2000.