

## 데이터마이닝을 이용한 e-CRM System 연구

최성, 김승찬, 유정근, 한정란\*

남서울대학교 컴퓨터학과, 협성대학교 경영정보학과\*

### Study on the e-CRM by Using DataMining System

Choi Sung, Yu Jeoung Geun, Kim Seoung Chan, Han Jung Lan\*

Dept. of Computer Science, Namseoul University

Dept. of Management Information System, Hyupsung University\*

#### 요 약

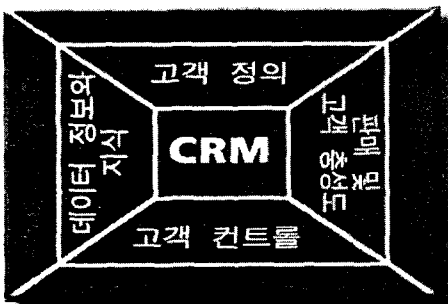
고객 중심의 기업활동의 중요성이 인식되면서 e-Business에서도 지속적인 고객의 가치를 증대하기 위해 eCRM을 실현하고 있다. eCRM은 비즈니스 모델에서의 다양한 프로세스를 지원할 수 있는 여러 시스템으로 구성된다. 데이터마이닝은 eBusiness 데이터로부터 유용한 지식을 추출할 수 있는 기술로서 eCRM에서 핵심 요소로 인식된다. 본 논문에서는 eCRM의 개념적 정의와 그 구성 프로세스를 살펴보고 eCRM에서의 데이터마이닝의 역할을 연구하였다.

#### 1. 서론

최근 급속도로 빠른 인터넷과 정보기술의 발달에 따라 CRM(Customer Relationship Management)은 기업 활동 기반의 필수요소로서 자리잡고 있다. CRM은 고객의 중요성을 인식하고 고객 중심적인 가치기준을 기업활동에 적용하는 것을 기초로 한다. 이로써 기업은 수익의 원천인 고객에 대하여 신규고객 유치, 우수고객 유지, 고객의 가치증진, 잠재고객의 활성화, 평생고객 확보등을 목표로 고객의 분석을 통해 고객을 이해하고 이를 통해 고객과의 지속적인 관계를 유지함으로써 고객 가치를 극대화하게 되는데 이러한 일련의 프로세스들을 CRM이라고 한다.

CRM은 지난 수십년간 꾸준히 발전되어 정착된 OLTP시스템, 최근 수년간 각광받는 기술로서의 데이터웨어하우스 및 OLAP 시스템 그리고 데이터마이닝 기술 등을 배경으로 하여 ERP, 워크플로우 시스템 등과 함께 기업 내의 정보시스템을 구성하게 되었다.

eCRM(Electronic CRM)은 CRM의 기술적 기반위에 진보된 형태의 e-Business에서의 고객관계 관리 요소이다. 기업수익 중 e-Business 채널을 통해 발생하는 비율이 증대되면서 기업의 상당수는 전통적 사업방식에서 웹을 근간으로 하는 방식을 선호하게 되었다. 따라서 인터넷을 통한 고객과의 가치 창출 및 증대는 기업의 중대 성공요소가 되었다. 인터넷을 통해 고객이 인식하지 못하는 차원의 데이터까지도 수집하여 고객의 모든 정보와 성향을 실시간으로 분석하고 마케팅활동으로 바로 연결이 가능한 솔루션이 eCRM이다. 정보기술(IT) 전문 시장 조사분석 기관인 IDC는 최근 발표한 시장 조사 및 예측 보고서에서 데이터마이닝 시장이 향후 고성장을 지속, 2004년이면 19억 달러 규모를 형성할 것이라고 전망했다. 특히 CRM시장은 매년 평균 22%의 지속적인 성장을 예측했다. 또한 데이터마이닝이 고객 행동 및 위험을 예측, 기업이 고객을 이해하는데 결정적인 역할을 담당함에 따라 고객 데이터를 기반으로 수익성 높은 고객별 특성 파악,



행동 예측을 필요로 하는 통신사업자나 카드, 보험, 은행 등에서 수요가 급증할 것으로 예측했다.

## 2. e-CRM 시스템 구성

### 2.1 e-CRM 정의

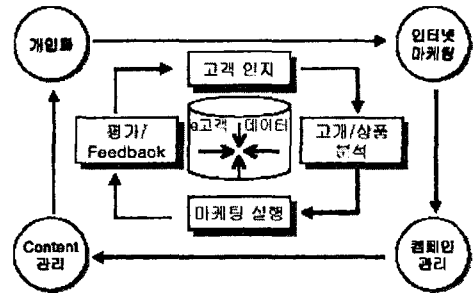
인터넷과 정보기술이 발달함에 따라 고객의 데이터 확보 및 추적, 관리가 용이해 지고 있다. 즉 이제는 개별 고객의 정보를 인터넷을 통해 손쉽게 획득할 수 있게 된 것이다. Click stream 데이터 분석 같은 기법을 이용하면, 개별 고객이 웹사이트에서 실행하는 모든 클릭 내용을 추적하여 데이터로 저장할 수 있다. 또한 고객 개인 개인이 회원을 가입하면서 입력하게 되는 모든 데이터 역시 고객정보로 저장이 되게 된다. 또한 인터넷의 발전과 함께, 기존에는 크게 중요시되지 않았던 무형자산들이 큰 가치를 나타내게 되었다. 즉 기업이 지적자산과 고객과의 관계, 고객에 관련된 정보 등의 무형자산은 기업의 경쟁력을 위해서는 없어서는 안될 중요한 자산으로 인식되고 있다. 이러한 변화 속에서 고객은 더욱 더 높은 수준의 서비스를 기대하게 되고, 요구와 취향 역시 복잡하고 다양해지고 있다. 이러한 상황속에서 다른 기업보다 경쟁력을 가지고 고객의 요구에 대응해 나아가려면 고객에 대한 이해가 매우 중요하며, 고객관리관계에서 인터넷을 이용해서 기존의 고객관리 시스템을 재구성하는 것을 말한다. e-CRM이란 고객관리관계에서 인터넷을 이용해서 기존의 고객관리 시스템을 재구성하는 것을 말한다. e-CRM은 고객만족을 극대화 하면서 동시에 관련 비용을 감소시킬 수 있는 새로운 고객관리 개념이라고 할 수 있다.

	CRM	e-CRM
고객접촉 채널	영업원방문, TM, DM 등 복수의 분산된 채널	인터넷을 활용한 단일 통합채널
고객요청 처리 과정	복잡하고 처리과정에 오류 기입 가능성	On-demand access로 단순한 절차와 실시간 처리
비용	신규 고객 유치와 관리 비용이 상대적으로 높음	초기 Set-up 비용이 높은 반면 유지, 관리 비용은 낮음
Online 서비스대응	주로 TM을 이용한 단순한 질의 응답	음성, 동화성 Multimedia, FAQ 기타 고객 관심 분야 정보제공
시간,공간적 범위	제한된 영업시간, 지역적 한계 존재	하루 24시간 전세계를 대상으로 가능

e-CRM이 되면, 고객의 주문의 처리속도도 빨라지고 주문 절차도 단순, 명확해지며, 고객의 불만이나 추가적인 서비스의 요구에도 신속하게 대응할 수 있어 고객의 만족도가 높아짐은 물론, 정보의 수집이 수월해짐은 물론, 통합된 정보의 공유로 처리과정이 단순해짐으로써 오류가 줄어들어 비용의 감소를 얻을 수 있다. 즉 고객의 만족도를 높여 이탈 고객을 감소시키고, 비용의 감소를 동시에 얻을 수 있으므로, 궁극적으로는 기업의 영업수익을 기대할 수 있게 된다.

### 2.2 e-CRM 프로세스

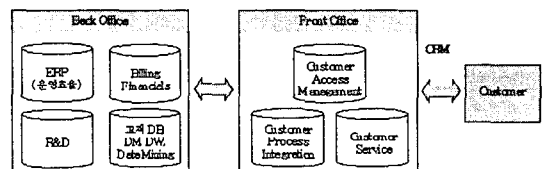
eCRM 프로세스 모델에 대한 관점은 연구기관, 컨설팅 및 솔루션을 제공하는 기업에 따라 다양하게 소개되고 있다. 즉, 이러한 모델은 제품 및 콘텐츠, 서비스 및 비즈니스, 정보획득 및 분석 등의 요소들을 통해 실질적으로 다양한 프로세스가 도출될 수 있다.



[그림 1] eCRM의 프로세스

### 2.3 e-CRM의 기능

eCRM 시스템은 온라인 상에서 하루 24시간 내내 신속하고 자동화된 고객 서비스를 제공하여야 한다. 따라서 eCRM 시스템은 Front-Office와 Back-Office의 연결과 고객을 위한 지원 Tool로서 기능에 초점이 맞추어 진다. [그림 2]에는 eCRM 업무 프로세스에 해당하는 Front-Office와 Back-Office의 기능을 나타내고 있다.



[그림 2] Back&Front Office로서의 CRM 모델의 위치

### 3. 데이터마이닝 구성

#### 3.1 데이터마이닝 방법론

데이터마이닝이란 대량의 데이터로부터 쉽게 드러나지 않는 유용한 정보들을 추출하는 과정을 말한다. 여기서 정보는 묵시적이고 잘 알려져 있지 않지만 잠재적으로 활용가치가 있는 정보를 의미한다.

다시말해 데이터마이닝이란 기업이 보유하고 있는 일일 거래자료, 고객자료, 상품자료, 마케팅 활동의 피드백 자료와 기타 외부자료를 포함하여 사용가능한 데이터를 기반으로 숨겨진 지식, 기대하지 못했던 패턴, 새로운 법칙과 관계를 발견하고 이를 실제 경영의 의사결정 등을 위한 정보로 활용하고자 하는 것이다.

정보를 찾아내는 방법은 어떤 특정 기법과 그 기술 자체만을 의미하는 것이 아니고, 비즈니스 문제를 이해하고 이러한 문제를 해결하기 위하여 정보기술을 적용하는 포괄적인 과정을 의미한다. 즉 유용한 정보의 추출을 위한 방법론이라고 할 수 있다. 따라서 데이터마이닝을 효율적으로 수행하기 위하여 시계열분석 등 각종 통계기법과 데이터베이스 기술 뿐만 아니라 산업공학, 신경망, 인공지능, 전문가시스템, 퍼지논리, 패턴인식, 기계적 학습(Machine Learning), 불확실성 추론(Reasoning with Uncertainty), 정보검색에 이르기까지 각종 정보기술과 기법들을 사용하게 된다. 또한 경영전략, 마케팅 기법 등의 최신 경영기법들의 이용도 필요하다. 이러한 데이터마이닝을 통하여 거대한 데이터베이스에 숨어있는 전략적인 정보를 발견할 수 있으며 이러한 정보는 주요 비즈니스 프로세스 개선의 가장 원초적인 단계에서 사용될 수 있다.

데이터마이닝은 흔히 Knowledge Discovery in Database(정보발견)라고도 불리우며 그외에 Knowledge Extraction(지식추출), Information Harvesting(정보추수), Data Archaeology(정보고고학), Data Pattern Processing(자료패턴처리) 등으로도 불리운다. 데이터마이닝이란 말은 대부분 통계학자들과 데이터베이스 학자들이 최근 MIS 분야와 경영분야에서 사용하는 말이다. 관련 용어로 데이터웨어하우스, 의사결정지원시스템, OLAP, 지식관리 등을 더 살펴보면 이해하는데 도움이 될 것이다.

#### 3.2 데이터마이닝의 단계

표준화된 단계는 없지만 일반적으로 데이터 준비, 데이터 분석과 분류, 지식 획득, 그리고 예측 등의 단계로 진행된다.

◎ 데이터 준비 : 데이터마이닝에 사용될 데이터 집합을 파악하고 오류가 포함된 데이터는 삭제한다. 데이터 마이닝에 있어서 핵심 과제 중의 하나는

마이닝을 실시하기에 적절한 데이터 집합을 찾는 일이다. 데이터 마이닝을 실시하기에 적절한 데이터 집합을 찾는 일이다. 데이터 마이닝을 위해서는 잘 정제되고, 통합되고, 일관성이 있는 단일 원천의 데이터가 요구된다. 데이터 웨어하우스에 저장된 데이터는 미리 통합되고 여과된 것이기 때문에 데이터 마이닝의 적절한 대상이 된다.

◎ 데이터 분석과 분류 : 이 단계에서는 데이터를 분석하여 공통 속성이나 패턴을 파악한다. 이 단계에서 데이터 마이닝 도구는 특정 알고리즘을 이용하여 데이터를 집단화하고 분류하거나, 데이터간의 의존성이나 관계를 찾아내고, 데이터의 패턴과 추세를 파악한다.

◎ 지식 획득 : 지식 획득 단계에서 데이터 마이닝 도구는 데이터 분석과 분류 단계에서 얻어진 결과를 이용하여 적절한 모델링 기법이나 지식 포착 알고리즘을 선택한다. 많이 사용되는 알고리즘으로는 인공 지능, 신경망, 귀납 규칙, 술어 로직, 의사결정 트리, 그리고 유전자 알고리즘 등이 있다.

◎ 예측 : 데이터 포착 단계에서 파악된 결과를 이용하여 향후 행동을 예측하고 사업의 성과를 예측한다. 예를 들어, 데이터 마이닝 결과물을 토대로 고객 중 지난 1개월간 우리 이동전화를 사용하지 않은 고객의 62%는 다른 이동전화 회사로 옮길 가능성이 83%이다 등의 사업 예측을 하고, 그에 대한 대책을 마련한다.

데이터 마이닝은 아직 초기 단계이며, 경우에 따라 경영자에게 별 의미 없는 결과를 제시할 수도 있다. 데이터 마이닝의 결과가 모두 사업적으로 의미있는 것일 수는 없지만, 여러분야에서 성공 사례가 속속 발표되고 있다. 특히 고객의 구매 패턴 파악, 보험 사기 색출, 도난 카드 사용 발견, 주식 시장 분석등의 분야에서 큰 위력을 발휘하는 것으로 알려져 있다.

#### 3.3 데이터마이닝의 문제점

◎ 잡음과 결손 자료

데이터베이스는 일반적으로 오류가 존재할 수 있기 때문에 자료가 완전히 정확하다고 단정할 수는 없다. 측정과 판단에 의존하는 속성들은 오류의 가능성이 있으므로 경우에 따라서는 잘못 분류될 가능성이 있다. 일반적으로 속성 값 또는 클래스 정보의 오류는 잡음으로 알려져 있으며, 이러한 잡음은 규칙 전체의 정확성에 영향을 미치기 때문에 반드시 제거되어야 한다.

◎ 데이터에 대한 준비 부족

대부분의 데이터들은 잘 정리되고 통합된 하나의 데

이터베이스로부터가 아니라 다양한 원천(source)로부터 얻어진다. 또한 사전에 준비되고 계획적으로 수집된 데이터를 다루는 실험연구에서와는 다르게, 데이터 마이닝에서는 비계획적으로 쌓여져 있는 과거의 데이터를 다루는 경우가 대부분이다. 따라서 분석의 목적에 맞는 데이터를 준비하기 위해 구체적이고 충분한 시간과 기술적 검토를 가져야 한다. 실제로 프로젝트의 지연 원인이 되는 대부분의 문제들은 데이터 준비 단계에서 발생한다.

◎ 확실성

대부분의 분석 도구가 휴리스틱한 방법이므로 데이터 마이닝 방법을 이용한 해답에 대해 확신을 갖기에는 무리가 있다. 따라서 데이터 마이닝 방법은 보관되어 있는 자료에 기초하여 정확한 추론을 할 수 있는가에 대한 확신을 가질 수 없는 문제가 발생된다.

◎ 시간차이 문제

데이터베이스에서 데이터를 처리하여 저장하는 데는 시간적 지연이 발생하므로, 현재 분석하고 있는 데이터는 이미 과거 시점의 데이터일 가능성이 많다. 그러나 분석으로부터 얻어진 결과는 미래에 적용된다. 따라서 과거와 현재 그리고 미래에 있어서의 시장과 기업의 환경변화를 충분히 고려하여 프로젝트를 계획해야 한다. 때로는 데이터마이닝으로부터의 결과가 과거의 기업전략과 고객행동에 대한 재확인에 불과한 경우가 있다.

◎ 적용상의 문제

데이터마이닝으로부터 얻어진 결과를 다양한 기업환경에서 어떻게 적용할 것인지를 고려해야 한다. 가치 있는 발견이 효율적으로 사용될 수 있도록 조직적이고 전략적인 계획이 수립되어야 한다.

◎ 크기, 갱신과 관련없는 분야

데이터베이스는 규모가 방대해지고 내용이 동적으로 변화하는 경향이 있다. 정보가 추가되거나 수정 또는 제거되면서 내용이 계속 변화되고 있다. 데이터 마이닝의 관점에서 데이터베이스의 문제점은 어떻게 하면 항상 정확한 규칙이 되도록 규칙을 갱신시킬 수 있고, 가장 최근의 정보와 일치되게 보장하는가 하는 문제이다.

◎ 장기적이고 구체적인 계획의 부족

데이터마이닝 프로젝트는 대규모의 데이터를 다루어야 하며 여러 분야의 연구자들이 함께 참여하기 때문에 때로는 지루하고 시간이 오래 걸리는 작업이 될 수 있다. 따라서 프로젝트의 최종적인 목표와 작은 목표들에 대해서 명확하게 정의하고, 실행상의 계획 등에 대해서 구체적으로 논의하고 결정해야 한다.

◎ 부서 및 프로젝트들 간의 비협조체제

분석, 전산, 적용을 담당하는 부서들은 다른 관심과 문제들이 있고, 데이터마이닝 프로젝트에 서로 다른 기대를 가지고 있을 지도 모른다. 또한 한 기업 내의 프로젝트들은 밀접한 연관을 가지고 있을 수도 있다. 관련 부서나 프로젝트들이 유기적으로 협조할 수 있도록 하는 것은 프로젝트의 효율적이고 성공적인 수행을 위해서 매우 중요하다.

3. 결론.

CRM에 대한 밝은 전망에도 불구하고, CRM 시스템을 이미 구축한 기업들을 보면 구축되는 CRM 시스템의 약 70%가 긍정적인 효과를 가져오지 못하고 있다는 연구결과를 내놓고 있다. 이는 많은 시간과 비용이 투입된 대부분의 CRM 구축 프로젝트가 기업경영 전반의 목표에 부합되는 비즈니스 프로세스 중심이 아닌 단순한 고객 프로세스의 자동화에 초점을 맞추는 기술 중심(Technology-centric)으로 이루어지기 때문인 것으로 분석하고 있다.

따라서, 향후 e-CRM에 대한 접근 방식은 단순한 업무 자동화 및 최적화에 의한 효율성 향상이 아니라, eBusiness의 특성을 감안하여 효과적인 타이밍 기법을 접목하는 동시에 기업 수익성 향상을 위해서 기존 데이터마이닝 기법이 안고 있는 문제점 해결과 더불어 경영 전략과 어떻게 통합, 활용되어야 하는가의 측면으로 변화될 필요성이 있다.

[참고문헌 및 사이트]

- [1] R. Agrawal, Tomasz Imielinski, and Arun Swami, Database mining; Aperformance perspective, IEEE Transactions on Knowledge and Data Engineering
- [2] CRM을 위한 데이터마이닝 : 스테판 스미스, 알렉스 버슨, 커트 티어링 대청(대청미디어)
- [3] 김영균, 서성보, 이준옥, 류근호, 전자상거래에 적용 가능한 데이터 마이닝 시스템 구현 및 적용, 한국정보과학회 충청지부 학술발표 논문집, 2000
- [4] 이용준, "시간 데이터 마이닝 프레임 워크", 충북대학교 대학원 전자계산학과 이학 박사 학위논문 제안 2001
- [5] www.itwarehouse.co.kr
- [6] www.inew24.com