

모바일 환경하에 RFM 기법을 이용한 개인화된 추천 시스템 개발

조영성*, 허문행**, 류근호***

Implementation of Personalized Recommendation System using RFM method in Mobile Internet Environment

Young Sung Cho *, Moon Haeng Huh **, Keun Ho Ryu ***

요 약

모바일 환경하에 RFM 기법을 이용한 개인화된 추천 시스템을 제안한다. 사용자의 평가 자료에 의존하지 않고 사용자에게 번거로운 질의 응답 과정이 없이 묵시적인(Implicity) 방법을 이용하여 고객정보와 구매이력정보를 기반으로 RFM 기법을 이용하여 고객 세분화와 아이템 세분화 통해서 대상 사용자에게 구매 가능성이 높은 아이템을 추천한다. 또한 기존의 추천시스템의 문제점의 해결 방안으로 신규 고객이나 신규 아이템 추천을 고려하여 적용한다. 추천 아이템과 사용자가 구매한 아이템 이력 데이터를 비교하여 추천된 아이템이 중복 추천을 제거하였고 현업에서 사용하는 데이터 셋을 구성하여 실험을 통해서 효용성과 타당성을 입증 및 평가하여 개인화된 일대일 웹 마케팅을 실현하였다.

Abstract

This paper proposes the recommendation system which is a new method using RFM method in mobile internet environment. Using a implicit method which is not used user's profile for rating, is not used complicated query processing of the request and the response for rating, it is necessary for user to keep the RFM score about users and items based on the whole purchased data in order to recommend the items. As there are some problems which didn't exactly recommend the items with high purchasability for new customer and new item that do not have the purchase history data, in existing recommendation systems, this proposing system is possible to solve existing problems, and also this system can avoid the duplicated recommendation by the cross comparison with the purchase history data. It can be improved and evaluated according to the criteria of logicity through the experiment with dataset, collected in a cosmetic cyber shopping mall. Finally, it is able to realize the personalized recommendation system with high purchasability for one to one web marketing through the mobile internet.

▶ Keyword : 모바일(mobile internet), RFM기법(RFM Method), 추천시스템(recommendation system)

• 제1저자 : 조영성 교신저자 : 류근호

• 접수일 : 2008. 1. 25, 심사일 : 2008. 2. 15, 심사완료일 : 2008. 2. 20.

* 동양공업전문대학 전산정보학부 산업체겸임 부교수 ** 안양대학교 디지털미디어학과 부교수

***충북대학교 전기전자컴퓨터공학부 교수

※ 이 논문은 2008년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원(지방연구중심대학육성사업/충북BIT 연구중심대학육성사업단)과 건설교통부 첨단도시기술개발사업-지능형국토정보기술혁신사업과제의 연구비지원(과제 번호:07국토정보C02)에 의해 수행되었습

I. 서론

유비쿼터스(ubiquitous) 네트워크 환경이 도래하여 언제(anytime), 어디서나(anywhere) 시간과 장소에 제약받지 않고 고성능·초소형 지능형 이동 단말기로 인터넷을 즐기 위한 수요가 더욱 증대되고 있다. 이러한 시점에서 전자상거래에서 취급하는 아이템의 종류가 다양해져서 유무선 정보통신망을 통해 컴퓨터나 휴대단말기(mobile device)로 아이템을 보다 효과적으로 추천이 가능하고 사용자를 만족시킬 수 있는 추천시스템(recommendation system) 프레임워크에 대한 연구[1, 2, 3]가 활발히 진행되고 있다. 오프라인과 달리 온라인 쇼핑물에서는 사용자 스스로가 자신이 원하는 아이템들을 구매하기 위해서 많은 아이템 정보를 처리해야 하는 부담이 발생하여 정보 과부하(information overload) 현상으로 사용자의 구매 의욕까지 상실시킬 수 있는 요인이 되고 있으며, 데이터의 널(null) 값으로 인한 데이터 희소성(sparsity) 문제, 입력정보가 주관적이고 서술적 정보이기 때문에 부정확성의 문제가 있다. 아울러 사용자 수의 증가로 인한 확장성 문제, 수행 속도가 저하로 인한 효율성과 실시간성(real time accessiblity) 및 민첩성(agility) 등에 많은 문제가 존재한다. 모바일을 이용한 m-CRM(Mobile Customer Relationship Management) 기반 고객 서비스의 경우 사용자 프로파일 기반의 많은 평가 자료에 대한 질의 및 응답처리에 따른 과도한 시간의 지체와 무선 과금 문제로 현실적으로 많은 어려움이 있다. 이러한 문제점을 해결하기 위해서 기존 추천 시스템에서 많이 사용하고 있는 사용자 프로파일 기반의 사용자의 평가 자료에 의존하지 않고, 사용자에게 번거로운 질의 응답 과정이 없이 묵시적인 방법을 이용하여 정보의 처리 시간을 감소시키고, 고객 회원정보와 고객 구매정보를 기반으로 RFM 기법을 이용하여 고객 세분화와 아이템 세분화 통해서 대상 사용자에게 고객이 원하는 구매 가능성이 높은 아이템을 효과적으로 추천할 수 있는 효율적인 모바일 추천 시스템을 제안한다. 아울러 신규 고객이나 신규 아이템 추천에서의 기존 추천 시스템에서의 문제점의 해결 방안을 제시한다. 본 논문의 구성은 다음과 같다. 제 I 장은 서론, 제 II 장은 관련 연구를 다루었으며 제 III 장에서는 제안한 추천 시스템에 대한 프로토타입 설계 및 구현에 대한 상세를 설명하고 제 IV 장에서는 시스템 구현에 따른 환경 구성 및 구현을 다루며, 실험을 통해서 제안한 추천기법의 적용과정과 시스템의 효용성과 타당성을 입증 및 평가한다. 마지막으로, 제 V 장에서는 본 논문의 결론 및 향후 연구 과제에 대하여 기술한다.

II. 관련연구

2.1 WAP 프로토콜

Web 접근의 위한 표준 프로토콜인 WAP(Wireless Application Protocol)은 무선망과 인터넷을 연결하기 위한 WAP 포럼에서 제정한 표준 규약이다. WAP에서 동작하는 무선 프로토콜 마크업 언어인 WML(Wireless Markup Language)은 XML을 기반으로 하는 이기종간의 플랫폼에서도 구현과 활용이 가능하다. WAP은 이동전화나 PDA 등 소형 무선 단말과 네트워크상의 서버 사이의 통신을 가능케 하는 표준규격을 갖춘 요소들의 집합으로서, 유선 망 위주의 Web을 고려하여 만들어졌기에 전체적인 구조에서 많은 유사점을 가지고 있다. 물론 기술적으로는 HTTP와 별도의 WAP 프로토콜이 기존기술과의 호환성을 제공하고 응용 애플리케이션의 개발이 가능하며 많은 유연성을 제공하며 지원한다[4,5].

2.2 데이터마이닝

데이터마이닝이란 지식 발견(knowledge discovery), 정보 발견(information discovery), 정보 수확(information harvesting) 등으로 알려진 자동화되고 지능을 갖춘 데이터 분석 기법이다[6]. 데이터마이닝은 대용량의 데이터들로부터 이들 데이터 내에 존재하는 관계, 패턴, 규칙 등을 탐색하고 찾아내어 모형화 함으로써 유용한 지식을 추출하는 일련의 과정들이다[7]. 데이터마이닝이란 일반적으로 데이터에서 필요한 부분의 정보를 구하거나, 의사결정을 하게 해주는 지식을 제공하는데 사용되는 수많은 기술 또는 기법이라고 할 수 있다[8]. 데이터 마이닝은 대량의 데이터로부터 패턴 인식, 통계적 기법, 인공지능 기법 등을 이용하여 데이터간의 상호 관련성, 패턴, 추세 등 의사결정에 유용한 정보를 추출해 내는 과정으로서 지식탐사의 핵심적인 역할을 담당한다. 데이터마이닝은 아이템 추천시스템에서도 많이 활용되고 있다. 최근에는 효율적인 고객관리 및 마케팅전략을 수립하기 위해 고객을 세분화하거나 성향을 분석하여 그룹별 혹은 개인별로 차별화된 마케팅을 지원하는 고객관계관리(Customer Relationship Management, CRM) 등에서 데이터 마이닝이 광범위하게 사용되고 있다. 추천시스템은 준비단계에서 수집된 고객 프로파일과 고객 행위 정보를 기반으로 데이터마이닝 기법에 의해 분석된 결과를 토대로 고객에게 맞춤 서비스를 제공하는 방법을 사용한다. 이러한 데이터마이닝 기법으로는 협업 필터링

(collaborative filtering), 내용 기반 필터링(contents based filtering) 등이 있으며, 이러한 기법들을 조합한 하이브리드(hybrid) 기법이나 연관 규칙(association rule) 등도 활용되고 있다.

추천시스템은 준비단계에서 수집된 고객 프로파일과 고객 행위 정보를 기반으로 데이터마이닝 기법에 의해 분석된 결과를 토대로 고객에게 맞춤 서비스를 제공하는 방법을 사용한다. 또한 정보 필터링(information filtering)이라고도 부르는 내용 기반 필터링 기법과 협업 필터링 기법은 모두가 고객 프로파일을 통해 과거 구매나 추천 결과를 쉽게 반영할 수 있는 있으며 추천속도가 빠른 장점이 있다. 그러나 유사한 아이템을 찾거나 유사한 고객을 탐색하기 위해서는 최소한 한 개 이상의 아이템을 구매한 고객에 대해서만 아이템 추천이 가능하다. 이러한 제한성 때문에 많은 종류의 상품을 취급하는 대형 온라인 쇼핑몰에서는 실질적인 적용이 불가능하다. 다음은 기존 연구와 본 제안과의 차이점을 비교한 표를 보여주고 있다.

표 1. 기존 연구와 본 제안과의 차이점
Table 1. Comparison existing study and new proposal

| 구분 | Profile 사용 | 피드백 사용 | Input | 대상고객 선정 | 추천고객 단위 |
|--------|------------|--------|--------------|--------------|---------|
| 협업적 | ○ | ○ | 각항목 점수 부여 | 전체고객 | 전체고객 |
| 내용기반 | ○ | ○ | 각항목 점수 부여 | 전체고객 | 전체고객 |
| 인구통계학적 | ○ | ○ | 인구정보 | 고객군 | 고객군 |
| 사례기반 | ○ | ○ | 고객선호 각 항목 사용 | 고객군/개인 | 개인 |
| 본 논문 | X | X | 고객/제품 점수 | 4가지 고객성향별 분류 | 개인 |

2.3 RFM

RFM(Recency, Frequency, Monetary)은 다이렉트 마케팅 실무자들 사이에 30여 년 동안 널리 사용된 모델이다. 여기서 R값은 최근 구매일자를 의미하며, F값은 일정기간 동안의 총 구매횟수를 그리고 M값은 일정기간 동안의 총 구매금액을 의미한다. 먼저 RFM에 대한 정의로는, RFM의 세가지 요소를 가산 평균하여 단일 지표로 나타내고 이를 고객의 가치를 평가하는 지표로 삼는 방식이 바로 RFM에 의한 고객 점수 부여 방법이다[9]. 즉 RFM 분석은 구매가능성이 높은 우수고객에게 집중한 마케팅 전략을 실행하고 불필요한 자원 낭비를 방지함으로써 근본적으로 매출액 증대보다는 수익을

창출하기 위한 모형이라고 할 수 있으며 우선 R값에 의해 5개의 세분화 세그먼트가 나누어지게 되고, 생성된 5개의 세그먼트는 F값에 의해 다시 5개로 나누어지며, 마지막으로 M값에 의해 각각의 세그먼트가 5 등분되어진다. 따라서 전체 고객데이터베이스는 결국 $5 \times 5 \times 5 = 125$ 개의 세그먼트로 분할되어 진다. 여기서 각각의 R값, F값, M값에 대한 계산 방식은 학자에 따라, 기업에 따라 다르며, 그 종류도 다양하므로 RFM을 이용하는데 있어서 무엇보다도 중요한 것은 자사에 목적에 적합한 RFM을 구현하는 것이다. 이렇게 구해진 RFM은 최근 구매시기(recency), 구매빈도(frequency), 구매금액(monetary)등 고객의 수익기여도를 나타내는 세 가지 지표들의 선형결합으로 구한 점수(score)로 표현할 수 있는데 이를 구하는 방법으로 가장 일반적인 RFM 모형은 식은 다음과 같다.

$$RFM = A \times Recency + B \times Frequency + C \times Monetary \dots (1)$$

여기서 A, B, C는 각 요인에 대한 가중치가 되고, RFM은 각 요인들과 가중치의 선형결합에 의한 점수(score)가 되는 것이다. 따라서 RFM 분석에서 가장 중요한 부분은 가중치(A, B, C)를 구하는 것이며, 이러한 가중치 또한 기업에 따라 다르므로 기업을 대표할 수 있는 가중치를 찾아내어 적용하는 것이 중요하다고 할 수 있다.

III. 제안 시스템

3.1 시스템 구조

전체 시스템을 구성하는 세부적인 서브 시스템인 에이전트 모듈은 다음과 같다. 무선 모바일 기기를 통한 입출력 접속을 담당하는 인터페이스 모듈, 분석 에이전트는 첫째 회원 가입을 통한 고객 집단을 고객 분류코드를 부여하여 고객 정보를 생성 및 관리한다. 로그인 고객에 맞는 고객 분류코드를 체계화하여 유사한 고객 그룹으로 분류하여 추천을 위해 활용한다. 둘째 로그인 고객의 고객 분류코드에 해당하는 유사 고객 그룹 내의 구매데이터를 기반으로 선호도 확률 산출과 그 선호도에 따른 추천에이전트는 분석에이전트가 제시한 선호도 확률을 기반으로 아이템 소분류에 해당하는 아이템 RFM점수가 높은 아이템을 TOP-4로 추천 목록을 생성하고 로그인 고객의 구매목록 즉 고객이 구매한 아이템(구매이력 상황) 데이터를 비교하여 추천된 아이템이 중복 추천되지 않도록 추천

목록을 작성한다. 다음 <그림 1>은 추천시스템의 전체 시스템 구성도를 나타낸 것이다.

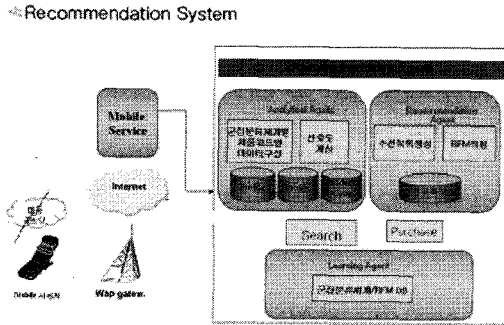


그림 1. 추천시스템의 전체 시스템 구성도
Fig. 1. system configuration for recommend system

3.2 시스템 주요 기능

<그림 1>과 같은 추천시스템의 전체 시스템 구성도로 운영 되는 본 시스템의 소프트웨어의 기능은 다음과 같다(1, 10).

(1) 서비스 접속 및 요청 기능

고객은 자신의 모바일 기기(이동통신 휴대폰 단말장치)를 이용하여 원하는 정보에 접근하기 위해 로그인을 한다. 고객이 로그인을 하기 위한 방법은 고객이 직접 해당사이트의 주소를 입력하여 접속할 수 있다. 또한 정보제공자로부터 제공 되는 개인화 서비스 메일, 즉 신상품 안내 메시지, 축하 메시지 등에 첨부된 해당 인터넷의 URL을 통해 접속 가능하다.

(2) 모바일을 통한 모바일 서버로의 접속 및 전송

고객은 로그인 인증과정을 거쳐서 서버에게 접속하여 고객에 대한 정보를 알리고 서버는 이 과정을 통하여 접속한 고객이 추천시스템에 접속을 신청한 고객임을 알 수 있다.

(3) 고객 회원정보의 분류코드에 근거한 정보 추출

고객이 인증이 완료되면, 모바일 추천시스템 서버에 유지된 고객에 대한 회원 정보를 중심으로 해당 웹 사이트의 정보를 추출하기 위한 작업이 요청된다. 즉, 고객의 개인 정보를 기준으로 웹 사이트의 구매이력 데이터를 추출하고, 이를 기반으로 추천 대상인 로그인 고객에게 추천에 대한 정보가 전달된다. 초기에 고객에게 전달되는 정보는 사전 정의된 인터페이스의 틀에 따라 제공된다.

(4) 아이템 추천 서비스 지원

모바일 단말기를 사용하는 고객의 입력된 ID를 식별하여

웹 서버에 접속되면 제공되는 추천 서비스는 현재 접속한 고객의 회원정보에 따른 구매이력을 기반으로 한 고객 RFM점수에 따라 그 내용이 고객마다 서로 다르게 보여 진다. 즉, 고객의 휴대폰에 나타나는 추천 정보는 고객의 성향에 따라 서버의 에이전트들을 통하여 추천결과가 서로 다르게 나타내게 된다. 고객이 접근하여 추천에 따른 구매 행위에 대한 사항은 즉시 구매데이터에 반영되어 활용되기 위하여 서버의 관련 DB에 반영된다.

3.3 추천 시스템 알고리즘

기존 추천 시스템에서 많이 사용하고 있는 고객 프로파일 기반의 고객의 평가 자료에 의존하지 않고, 고객에게 번거로운 질의 응답 과정이 없이 묵시적인방법을 이용하여 정보의 처리 시간을 감소시키고, 고객 회원정보와 고객 구매정보를 기반으로 RFM 기법을 이용하여 고객 세분화와 아이템 세분화 통해서 대상 고객에게 고객이 원하는 구매 가능성이 높은 아이템을 효과적으로 추천할 수 있는 효율적인 모바일 추천시스템을 제안한다. 우선 회원 가입 시 고객 회원정보에 고객 집단을 분류할 수 있는 분류코드를 신설하여 고객의 회원정보를 생성 및 관리한다. 로그인 고객은 고객 회원정보를 읽어 자신이 속한 분류코드를 이용하여 고객을 분류하고 이 분류된 고객을 기반으로 구매정보를 분류하여 아이템 분류코드 (대, 중, 소분류) 별로 데이터를 재구성하여 선호도를 구한다. 고객이 아이템 카테고리(아이템 소분류)에 대해 갖는 선호도의 임계값은 0.25이상의 값을 가지는 아이템 카테고리(아이템 소분류)가 된다. 로그인 고객이 속한 군집내의 고객의 선호도를 계산하고 유사한 고객을 선정한다. 로그인 고객과 유사한 고객 TOP-4의 아이템 목록을 생성한다. 추천 시 추천된 아이템을 로그인 고객의 구매 이력정보와 체크하여 중복 추천되지 않도록 한다. 신규고객이나 구매이력이 없는 고객에게 아이템 추천 시 군집화 된 고객 집단 중에서 RFM 점수가 가장 높은 고객이 구입한 아이템에서 가장 많이 팔린 아이템 4개 추천 아이템을 추천할 수 있게 한다. 소분류에 해당하는 단말노드인 브랜드화 된 아이템을 추천할 때 실 시간으로 평가 관리되는 RFM기법이 적용된 아이템 RFM점수 테이블을 적용하여 추천 리스트를 작성하여 추천하는 아이템 추천 시스템을 제안한다. 다음(표 2)과 (표 3)는 아이템 세분화의 RFM 기법을 적용한 개인화된 추천 시스템 절차 알고리즘과 고객 세분화의 RFM 기법을 적용한 개인화된 추천 시스템 절차 알고리즘을 각각 나타낸 것이다.

표 2. 아이템 세분화의 RFM 기법을 적용한 개인화된 추천 시스템 절차 알고리즘

Table 2. Procedure algorithm for recommendation system using item segmentaion

| |
|--|
| 입력 : 회원정보, 구매정보, 아이템 RFM 점수 정보 /*본 논문에서는 회원정보의 분류코드로 유사 고객을 사용하였음 */ 출력 : 유사고객, 선호도계산, 아이템 RFM점수가 높은 아이템 추천 조건 : 로그인 고객이 구매한 적이 있을 경우 (구매데이터가 있을 경우) |
| begin 1. 로그인 고객을 기준한 고객 분류코드를 기준으로 한 구매 정보를 발췌하여 분류한다. /* 고객 성향 분류코드를 이용하여 고객군집을 추출한다.*/ 1.1 신규고객시 회원 등록 및 수정 1.2 로그인 고객이 기존고객일 경우 해당 분류코드기준으로 유사 고객군을 생성한다. 1.3 유사 고객군에 해당하는 구매 데이터를 발췌하여 데이터를 생성한다. 2. for (고객-아이템 소분류코드) - 고객별 구매 아이템 소분류 건수집계한다. - 각 아이템 소코드별 각 아이템 소코드를 전체 고객의 아이템 소분류로 나누어 각 소코드별 확률은 구한다. (1식->P1) - 전체 구매 아이템 소코드별 나누어 각 소코드별 확률을 구한다. (2식->p2) $- 3식 = (2 * P1 * P2) / (P1 + P2)$ /* 고객별 구매 아이템 소분류코드별 선호도 확률을 계산한다. 3. 선호도가 높은 소분류코드(카테고리)에 해당하는 정보를 이용하여 아이템 점수가 로그인사용자와 유사한 RFM점수의 아이템을 추천한다. 3.1 로그인 고객이 구매한 아이템은 중복을 제거하여 추천한다. /* 아이템 RFM점수를 이용하여 아이템점수가 로그인사용자와 유사한 RFM점수의 아이템을 추천한다. */ end. |

표 3. 고객 세분화의 RFM 기법을 적용한 개인화된 추천 시스템 절차 알고리즘

Table 3. Procedure algorithm for recommendation system using customer segmentaion

| |
|--|
| 입력 : 회원정보, 구매정보, 고객 RFM 점수 정보 /*본 논문에서는 회원정보의 분류코드로 유사 고객을 사용하였음 */ 출력 : 유사고객, 선호도계산, 고객 RFM점수가 높은 아이템 추천 조건 : 로그인 고객이 구매한 적이 없을 경우 (신규고객 추천 경우) |
| begin 1. 로그인 고객을 기준한 고객 분류코드를 기준으로 한 구매 정보를 발췌하여 분류한다. /* 고객 성향 분류코드를 이용하여 고객군집을 추출한다.*/ 1.1 신규고객시 회원 등록 및 수정 1.2 로그인 고객이 신규고객일 경우 해당 분류코드기준으로 유사 고객군을 생성한다. 2. 유사 고객군에 해당하는 고객 RFM점수에서 가장높은 점수의 고객의 구매 아이템을 추천한다. /* 고객 RFM점수를 이용하여 고객점수가 가장 높은 고객이 구매한 아이템을 추천한다. */ end. |

IV. 시스템 구현

4.1 실험 및 분석

시스템 구현 및 실험 환경은 윈도우 운영체제하에서 Apache/PHP/MySQL 웹서버 환경하에서 사용하였고 버전은 다음과 같다.

- OS: Window XP
- Web Server :Apache HTTP Server Version 1.3.34 / WAP
- Database : MySQL Version 4.0.26
- Server-Side Application : PHP Version 4.4.3
- WML/WMLScript

실험데이터의 목록으로 로그인 고객은 고객 회원정보를 읽어 자신이 속한 분류코드를 이용하여 구매정보에서 군집화 된 집단(cluster)을 인지하고 해당 군집화에 속한 고객 249명의 고객 정보와 화장품 아이템 339개를 대상으로 총 구매건수 2262건의 구매데이터를 이용하여 로그인 고객의 군집성향에 맞게 구매데이터를 화장품 대분류, 스킨케어 중분류, 11개의 카테고리 소분류 89개를 아이템 군 별로 데이터를 취합하였다. 신규고객이나 구매이력이 없는 고객을 위한 유사 구매 아이템 추천을 고려한 RFM 세분화 기법을 적용하여 구매 가능성이 높은 개인화된 아이템 추천이 가능하다. 실험데이터의 구성은 현재화장품을 전문적으로 판매하는 인터넷 쇼핑몰을 화장품관련 쇼핑몰 P사의 아이템 분류에서 대,중,소 분류로 실 쇼핑몰에서 사용되는 실 데이터를 중심으로 실험데이터 셋과 테스트 데이터 셋을 구성하였다.

4.2 데이터베이스 구조

실험데이터구성을 위해서 데이터베이스 테이블을 6개의 테이블로 구성하였다. 다음 <그림 2>는 ERD로 나타난 데이터베이스 구조를 나타낸 것이다.

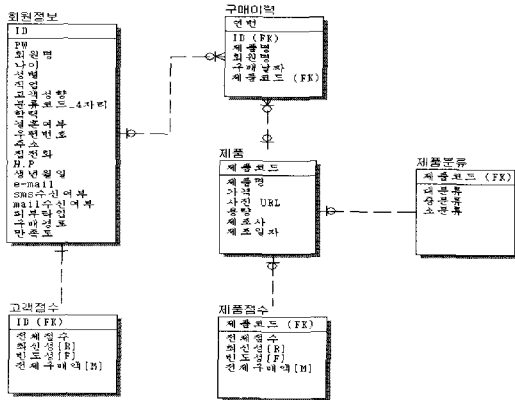


그림 2. ERD로 나타난 데이터베이스 구조
Fig. 2. ERD of database structure

다음은 ERD(Entity Relationship Diagram)으로 나타낸 테이블간의 관계 설명이다.

회원정보(member), 제품(items), 구매이력(sale), 제품분류(gp), 고객점수(point), 제품점수(s_point)

- member & sale table : 다 대 다
- gp & items table : 일 대 일
- items & sale table : 다 대 다
- point & member table : 일 대 일
- s_point & items table : 일 대 일

4.3 시스템 구현 및 테스트

시스템 구현 측면에서 본 시스템에 대한 효율성을 위해서 다음과 같은 표준 플랫폼 지원성을 분석하여 다음과 같은 주요한 2가지측면에서 구현 및 테스트를 수행하였다.

다음은 구현된 과정에 대한 개략적인 결과화면을 나타내 주고 있다. 로그인 사용자의 고객 성향을 바탕으로 분류코드를 얻어내어 분류코드의 사용자의 가장 많이 구입한 제품중 RFM 점수를 토대로 로그인 사용자 점수와 유사한 점수이상인 화장품을 추천한다. WML의 흐름은, 최초 Startq.wml 파일이 실행되며 1번 째 화면이 출력되고 이후 menu.wml이 실행되어 대분류 제품목록을 보여준다. 이후 menu.wml에 다른 카드아이디로 중분류로 이동하여 중분류를 클릭하면 다시 card id를 찾아서 그 해당 분류에 해당하는 제품을 보여주고 제품을 클릭하면 구매나 추천을 원하는 것으로 간주하여 로그인창이 뜨고 로그인이 되면 RFM점수와 분류코드를 기반으로한 추천을 한다. 다음 <그림 3>은 구현 과정에 따른 모

바일 인터넷 처리 결과를 통한 시스템 구현화면을 나타낸 것이다.



그림 3. 시스템 구현화면
Fig. 3. The screen of each step for result of system implementation

4.4 실험적 데이터분석

본 논문에서는 실험 데이터를 통해서 추천을 위해서 고객 정보 및 구매정보를 이용하여 데이터 분석과 추천을 위한 RFM 점수를 분석 및 실험하였다. 다음은 구매이력데이터를 기반으로 한 RFM 점수를 분석하여 현황을 나타낸 것이다. 상품(아이템) 및 고객의 RFM이 점수들은 각각 R은 0.15, F는 0.35, M은 0.5로 100점을 만점으로 안분하였다. 추천 시스템 설계에서의 점수 산출 식은 마켓터가 쇼핑물의 상황에 따라 결정하며 본 제안시스템의 점수 계수는 해당 물에서 구매 액에 비중을 두었다. 아이디마다 부여된 필드의 종합 점수(score)에 입력된다. 최고 점수는 100점, 최하 점수는 0점이다. 다음은 고객별 RFM 점수 판단 방식과 아이템(상품)별 RFM 점수 판단방식을 나타내어 RFM 점수에 따른 점수현황을 분석한 것을 나타낸 것이다.

표 4. 고객별 RFM 점수 판단 방식
Table 4. decision method for RFM score of customer

| | R | F | M | 총합 |
|------|--|---|--|------------------------------------|
| 점수 | 15점 | 35점 | 50점 | 100점 |
| 산출방식 | 현재날짜- 최근 구매 일자 27월 이내 = 5점 47월 이내 = 4점 67월 이내 = 3점 87월 이내 = 2점 87월 이상 = 1점 | 구매수량 60개 이상 = 5점 50개 이상 = 4점 30개 이상 = 3점 20개 이상 = 2점 20개 미만 = 1점 | 상품 총 구입 가격 / 100 7000점 이상 = 5점 5000점 이상 = 4점 3000점 이상 = 3점 1000점 이상 = 2점 1000점 미만 = 1점 | R x 3 + F x 7 + M x 10 |

표 5. RFM점수에 의한 전체 인원분포(80점이상)
Table 5. The statistic of total counts for member by RFM score(80 point grade more)

| | 90점 이상 | 80점 이상 | 70점 이상 | 60점 이상 | 50점 이상 | 50점 이하 | 총인원 |
|--------|--------|--------|--------|--------|--------|--------|-----|
| 수량 (개) | 12 | 26 | 39 | 57 | 46 | 69 | 249 |

표 6. 아이템 RFM 점수 판단 방식
Table 6. decision method for RFM score of customer

| | R | F | M | 총합 |
|------|--|---|---|-----------------------|
| 점수 | 15점 | 35점 | 50점 | 100점 |
| 산출방식 | 현재날짜- 상품 등록 일자 27월 이내 = 15점 47월 이내 = 13점 67월 이내 = 11점 87월 이내 = 8점 127월 이내 = 5점 12개월 이상 = 2점 | 구매빈도수 180개 이상 = 35점 160개 이상 = 33점 140개 이상 = 31점 120개 이상 = 29점 80개 이상 = 25점 50개 이상 = 22점 30개 이상 = 19점 30개 미만 = 15점 | 상품 가격 * 판매수량 / 100 80000점 이상 = 50점 60000점 이상 = 47점 50000점 이상 = 44점 40000점 이상 = 40점 35000점 이상 = 36점 30000점 이상 = 32점 20000점 이상 = 26점 10000점 이상 = 20점 10000점 미만 = 15점 | R + F + M |

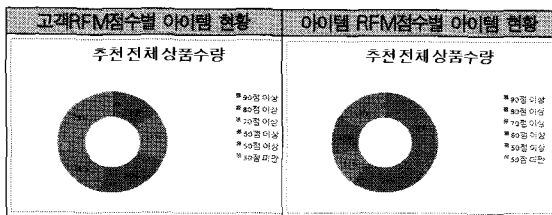


그림 4. 추천 전체 상품 수량 현황
Fig. 4. The statistic of total counts for item of purchase by RFM score

구매이력데이터를 기반으로 각종 구매 현황을 참조로 RFM 점수별 분석이 가능하다. 고객별 RFM 점수와 상품별 RFM 점수는 전체고객 및 전체 아이템에 대한 RFM 점수가 실시간으로 계산 및 관리되며 RFM 별 구매현황은 RFM 점수 가중치에 기반이 된다고 하겠다. 로그인 사용자와 같은 분류 코드 내에서 로그인 사용자와 고객RFM 점수가 유사한 고객이 구매한 아이템들을 Top-N으로 추천이 가능하며 로그인 사용자와 유사한 고객이 구매한 아이템을 대상으로 같은 분류 코드내의 신규 아이템 추천도 가능하며 아이템별 RFM 점수 판단방식을 이용하여 아이템이 갖고 있는 아이템 소분류코드의 신규 아이템에 대한 구매 이력데이터를 기반으로 한 RFM 분석을 통하여 신규아이템 추천이 가능하다.

4.5 제안 시스템 평가

제안 시스템에 대한 평가는 쇼핑몰을 이용해 본 경험이 있는 고객 249명의 고객 정보와, 그리고 현업에서 사용하는 화장품 아이템 339개를 대상으로 그들의 추천 2262건의 구매 데이터를 이용하였다. 실험은 특정 로그인 사용자와 성향이 같은 분류코드가 적용된 군집데이터 즉 social data에 기반한 화장품 아이템 추천에서의 각종 RFM 점수 유형에 따른 추천의 유효성과 추천시스템의 전체적인 성능 평가의 두 방향으로 진행되었다. 우선 초기 화장품 아이템 추천의 유효성을 실험에 참가한 고객들의 구매데이터와 제시되는 화장품 아이템의 비교를 통해 이루어졌으며, 평가 방법으로는 정보검색 분야에서 보편적으로 사용되는 평가척도인 정확도 (precision)을 응용하여 사용하였다. 정확도는 전체 추천된 상품 개수 중에서 사용자가 선호하는 아이템 개수의 비율로, 여기에서는 구매데이터에 기록된 아이템을 사용자가 선호하는 아이템으로 간주하였다.

precision =

$$\frac{\text{추천조건에 따른 추천된아이템과 일치하는 구매 데이터아이템수}}{\text{Social data를 기반으로한 초기 추천아이템수}} \dots\dots\dots (2)$$

추천받는 대상이 되는 특정 로그인 사용자와 성향이 같은 인구통계학적 변수를 갖는 분류코드(분류코드 : 2211)가 동일한 군집(clustering)내의 구매데이터를 이용한다. 로그인 사용자와 성향이 같은 분류코드의 Social data에 기반한 초기 추천 아이템을 분모로 하여 7가지 유형의 추천을 제시하여 평가 하였다.

조건 : 특정 로그인 사용자와 같은 분류코드 내에서의 실험을 대상으로 함

추천 조건은 다음과 같다.

- ①임의 추천
- ②최고 RFM점수를 가지는 고객들의 추천
- ③80점 이상의 RFM점수를 가지는 고객들의 추천
- ④60점 이상의 RFM점수를 가지는 고객들의 추천
- ⑤40점 이상의 RFM점수를 가지는 고객들의 추천
- ⑥20점 이상의 RFM점수를 가지는 고객들의 추천
- ⑦20점 이하의 RFM점수를 가지는 고객들의 추천

표 7. 추천조건별 추천 평가결과
Table 7. The result of recommending ratio for recommendation each condition type

| 구분 | 추천① | 추천② | 추천③ | 추천④ | 추천⑤ | 추천⑥ | 추천⑦ |
|-------|-----|-----|-----|-----|-----|-----|-----|
| 분류코드내 | 42 | 26 | 74 | 263 | 238 | 0 | 0 |
| 전체 | 575 | 575 | 575 | 575 | 575 | 575 | 575 |

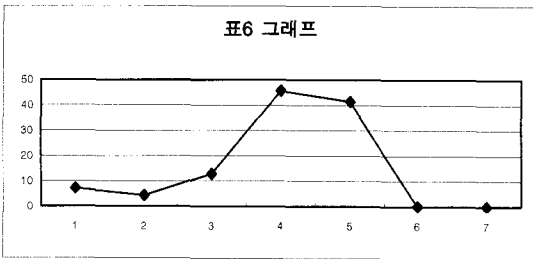


그림 5. 추천조건별 추천 평가결과
Fig. 5. The result of recommending ratio for recommendation each condition type(graph)

표 8. 추천조건별 전체 건수별 추천 평가결과
Table 8. The result of recommending ratio of total count for recommendation each condition type

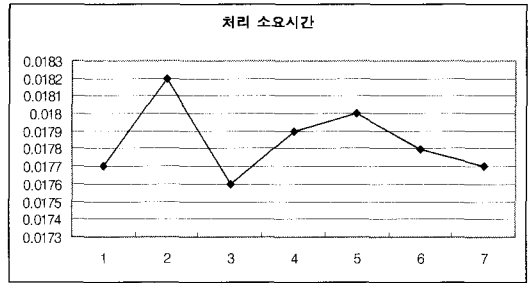
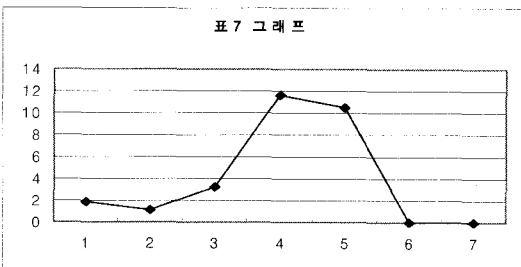


그림 6. 추천조건별 전체 건수별 추천 처리 시간 결과
Fig. 6. The result of processing time of total count for recommendation each condition type (unit: sec)

실험결과 추천조건에 따른 Precision 측정결과 추천 받는 대상에게 임의로 아이টে를 제시하는 것보다 추천 받는 대상과 성향이 같은 고객 RFM 점수가 비슷한 고객이 구매했던 상품을 추천하는 것이 훨씬 높은 결과를 보였다. 또한 추천에 걸리는 시간은 고객 RFM 점수를 이용하는 측면에서는 거의 차이가 없음을 알 수 있었다. 즉 실험 결과에서 알 수 있었던 사항은 아이টে를 추천받는 대상이 되는 로그인 사용자와 인구 통계학적 변수가 같더라도 다른 추천 결과가 나왔으며 추천받는 대상이 되는 로그인 사용자에게 RFM 점수가 80점에 맞는 고객의 구매했던 아이টে를 추천해도 정확도 면에서 추천 받는 대상이 되는 로그인 사용자와 비슷한 RFM 점수대의 추천방식④에서 높은 결과를 보였다. 이는 개인화된 추천이 추천의 결과를 높일 수 있다는 평가임이 입증된 결과라 하겠다.

V. 결론 및 향후 과제

유비쿼터스 네트워크 환경이 도래하여 언제, 어디서나 시간과 장소에 제약 없이 받고 지능형 이동 단말기로 무선 인터넷을 즐기려는 수요가 더욱 증대되고 있다. 이러한 시점에서 전자상거래 기업들의 마케팅 전략은 고객에 대한 많은 가치를 부여하는 개인화(personalization) 전략으로 발전되고 있다. 본 논문은 모바일 환경하에 RFM 기법을 이용한 개인화된 추천 시스템을 제안하였다. 기존 추천시스템에 많이 사용하는 사용자 프로파일 기반의 평가 자료에 의존하지 않고 사용자에게 번거로운 질의 응답 과정 없이 묵시적인 방법을 이용하여 고객 구매이력정보를 기반으로 RFM 기법을 이용하여 구매 효율성과 추천 효율성이 높은 추천 알고리즘을 제시하기 위하여 고객 세분화와 아이টে 세분화 통해서 대상 사용자에게 아이টে를 추천하는 새로운 방식을 제안하였다. 기업에서는 고객서비스의 중요성이 더욱 증대되어 갈 것이다. 극심한 경쟁

속에서 전자상거래 기업들이 살아남기 위해서 더욱 다양한 마케팅 전략을 활용이 시급하다. 본 논문에서는 핸드폰 요금 정액제 활성화를 기반으로 m-Commerce가 증가되는 시점에 시간과 장소에 제약받지 않는 모바일을 이용한 추천시스템에 RFM 기법을 이용한 개인화된 추천알고리즘을 제안함에 큰 의미가 있으며 중소규모의 유선 인터넷 전자쇼핑몰을 모바일을 이용한 추천시스템으로 전환하는 데 있어서 현실적이고 구매 촉진 방안으로 고객 개인화를 바탕으로 구매 효율성과 추천의 성능이 좋은 추천 시스템의 프레임워크를 제시함으로써 방향성 있는 대안이 될 것으로 생각된다. 아울러 기존 추천시스템의 문제점인 신규고객이나 신규아이템 추천에 대한 해결방안 제시하여 원투원 마케팅의 실현을 더욱 가능하게 하였다. 향후과제로는 추천의 효율성을 극대화 할 수 있는 방안으로 시스템 튜닝작업과 유무선을 통합한 환경 하에서 제공될 수 있는 추천시스템의 통합 솔루션 작업이 남아 있다.

참고문헌

- [1] 홍장의, 김종필, "모바일 환경에서 고객 맞춤형 정보서비스를 위한 MoSA-CRM 시스템", 한국데이터베이스학회, 2006.
- [2] 박정석, 정영진, 신문선, 류근호, "모바일 추천 서비스를 위한 이동 객체 질의 처리 시스템", 한국정보처리학회 논문지 제14-D권, 제7호, pp 707-718, 12월, 2007.
- [3] 구미숙, 최남규, 황정희, 류근호, "유비쿼터스 상거래 환경의 컨텍스트 기반", 제11-D권, 제7호, pp 1417-1426, 12월, 2004.
- [4] 조영성, "유무선을 통합한 웹프로그래밍실습", 기한재, 2007.
- [5] Evas, H., and Ashworth, P., "Getting Started with WAP and WML", Sybex, 2001.
- [6] 강현철 등저. "SAS Enterprise Miner 4.0을 이용한 데이터마이닝 : 방법론 및 활용", 자유아카데미, 2001.
- [7] 안광호, 김상용, 김주영. "인터넷마케팅원론", 법문사, 2001.
- [8] 허준, 최병주. "클레멘타인을 이용한 데이터마이닝", 고려정보산업, 2001.
- [9] 박찬욱. "데이터베이스 마케팅-고객정보의 활용을 통한 기업경쟁력의 강화", 연암사, 1996.
- [10] 최중민, "에이전트의 개요와 연구방향", 정보과학회지, pp.7-16, 1997.
- [11] 문영수, "데이터마이닝기법에 관한 연구", 석사학위논문, 서경대학교 대학원 인터넷정보학과, 2005.
- [12] Alper Caglayan and Colin Harrison, Agent Sourcebook, the Wiley press, 1997.
- [13] 하창승, "연관규칙 탐사와 사례기반 추천 기법에 의한 유무선 통합 지능형 검색 에이전트 시스템의 구현에 관한 연구", 박사학위논문, 한국해양대학교, 2004.
- [14] Elsenpeter, R. C. and Velte, T. J., "e-bussiness A Beginner's Guide", McGraw-Hill, Berkeley 2001.
- [15] 장형욱 한창엽 "CRM의최근 연구동향 및 향후과제", 경성대학교 경연구논문지 제21권, 제1호, pp. 83-101, 2005.
- [16] 허경희, "CRM-신경영패러다임", 정보과학회지, 2001
- [17] 장창복, "모바일 환경에서의 정확도와 처리 속도 개선을 위한 상황 적응적 서비스추천", 박사학위논문, 한남대학교 대학원 컴퓨터공학과, 2007.

저 자 소개



조 영 성

1989년 연세대학교 전산학과(공학석사)
 2007년 충북대학교 전산학과 박사과
 정수료
 1982~2001년 국제청 전산실, 미국
 콘트롤레이터 한국지사(과장),
 시나이미디어(부장), 유경제어
 기술연구소(수석), 네오아이엔
 씨(CEO)
 2002년~현재, 가나소프트(대표), 동
 양공진 전산정보학부 산업체
 겸임/부교수
 관심분야 : 시공간 데이터베이스, 데이
 터 마이닝, 기계학습, 웹서
 비스, 웹 2.0, ebXML



허 문 행

1979년2월 숭실대학교 전산학과 (학사)
 1989년2월 연세대학교 전산학과 (석사)
 2003년8월 충북대학교 전산학과 (박사)
 2003년12월 멀티미디어 기술사
 1980년 4월~1984년 1월 한국전자
 통신 연구원 선임연구원
 1983년 8월~1984년 2월 스웨덴
 ERICSOM 과건연구원
 1984년 1월~2000년12월 KT 연구
 개발원 책임연구원
 2001년 1월~2004년 2월 한국소프
 트웨어진흥원 단장
 2004년 3월 ~ 현재 안양대학교 디
 지털미디어학과 부
 교수 산학협력단장
 관심분야: 디지털콘텐츠, 미디어컨버
 전스, 유비쿼터스



류 근 호

1976년 숭실대학교 전산학과(이학사)
 1980년 연세대학교 공학대학원 전산
 전공(공학석사)
 1988년 연세대학교 대학원 전산전공
 (공학박사)
 1976~1986년 육군군수 지원사 전산
 실(ROTC 장교), 한국전자통
 신연구원(연구원), 한국방송통
 신대 전산학과 (조교수) 근무
 1989년~1991년 Univ. of Arizona
 Research Staff(TempIS
 연구원, Temporal DB)
 1986년~현재 충북대학교 전기전자
 컴퓨터공학부 교수
 관심분야 : 시간 데이터베이스, 시공간
 데이터베이스, Temporal
 GIS, 지식기반 정보검색
 시스템, 유비쿼터스 컴퓨
 팅 및 스트림데이터처리,
 데이터 마이닝, 데이터베
 이스 보안, 바이오 인포메
 틱스