

# 컨텍스트 인식 기반 상품 추천 시스템의 설계

이송희<sup>0</sup> 이근호 김정범 김태윤  
고려대학교 컴퓨터학과  
{pine<sup>0</sup>, root1004, qston, tykim}@netlab.korea.ac.kr

## A Design of Recommendation System based on Context-Awareness

Song-Hee Yi<sup>0</sup> Keun-Ho Lee Jeong-Beom Tai-Yun Kim  
Dept. of computer Science Engineering, Korea University

### 요 약

추천 시스템은 방문 고객 개개인의 취향이나 구매이력 등을 분석하여 고객이 필요로 하는 상품 또는 컨텍스트 정보의 서비스를 제공한다. 기존의 추천 시스템은 온라인에 초점을 맞추어 설계되었는데 본 논문에서는 무선 인터넷 서비스를 기반으로 무선 단말기(e.g. PDA, Cell Phone 등)를 통해 오프라인에서도 추천 정보를 제공하는 시스템을 제안한다. 사용자에게 제공이 되는 추천 정보는 상품이나, 컨텍스트 또는 이벤트 정보이며 제안된 시스템에서는 데이터 마이닝 기법을 통해 데이터를 분류, 측정 및 예측하고 지식 기반 방법과 collaborative filtering 방법을 혼합하여 양쪽의 장점만을 취하여 기존의 한정된 상품에 대한 정보와 웹상에서만 제공이 되는 서비스를 오프라인까지 통합한 추천 시스템을 제안한다.

## 1. 서 론

최근 무선 단말기를 이용한 무선인터넷의 발달로 이를 이용한 응용 서비스에 대한 관심이 고조 되고 있다. 무선 인터넷의 가장 큰 특징은 기존의 인터넷보다 더욱 개인화된 서비스와 사용자의 이동성에 기반한 고유의 정보를 제공하는 것이다[1].

본 논문에서는 사용자의 이동성을 통해 온라인과 오프라인을 통합한 더욱 특화되고 실시간적이며 개인화된 추천 시스템을 제안한다. 기존의 추천 시스템은 일반적으로 고정된 데스크 탑 환경에서 인터넷 쇼핑물이나 컨텍스트 사이트의 고객을 대상으로 사용자 프로필, 그리고 사이트 접속 및 활용도를 측정하거나 통계자료를 활용하여 주 고객의 성향 및 특성을 확인하였다. 제안된 시스템은 사용자가 가입한 사이트의 쇼핑센터를 오프라인상에서 실제적으로 방문했을때 사용자의 이동성을 이용한 컨텍스트 정보 인식을 통하여 온라인과 오프라인을 통합하여 실제적이고 개인화된 서비스를 웹과 무선 단말기로 제공한다.

컨텍스트(Context)는 컨텍스트 정보를 의미하며 “개체의 상황을 특성화하여 나타내는데 사용되어 질수 있는 어떠한 정보”로서 사용자 실제의 요구에 적용할 수 있게 하기 위해서 사용자의 정확한 기호와 성향 뿐만아니라 더 나아가 사용자의 주변 환경까지 고려해야한다[2].

본 논문에서 적용한 컨텍스트 요소는 무선 단말기를 통한 위치정보이다.

## 2. 관련연구

이 장에서는 위치기반 무선 인터넷 서비스와 일반적인 추천 시스템, 그리고 데이터 마이닝 기법에 대해 기술한다.

### 2.1 위치 기반 무선 인터넷 서비스

위치 기반 무선 인터넷 서비스는 그림 1과 같은 구성을 갖는다.

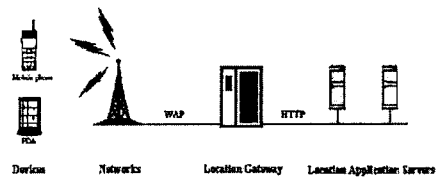


그림 1 위치 기반 무선 인터넷의 구성

그림 1 을 보면 단말기, 게이트웨이(즉위 서버), 그리고 응용 서버들이 각각 존재하게 되며 이는 HTTP와 WAP 프로토콜을 이용한다는 점에서 무선 인터넷 서비스를 위한 시스템의 구성과 같으며, 가장 큰 차이점은 WAP 게이트웨이와 함께 즉위 게이트웨이가 존재한다는 것이다. WAP은 제한된 이동통신환경에서 데이터 통신을 가능하도록 한 프로토콜이다[1][3].

### 2.2 추천 시스템(Recommendation System)

고객에게 개별 상품에 대한 선호도를 표시하도록 하여, 이를 근거로 한 고객 선호 벡터를 구성한 뒤, 이 선호 벡터와 비슷한 선호도를 가지는 다른 고객의 집단이 구매하거나 선호했던 상품을 추천하는 방법이다.

일반적으로 이러한 추천시스템을 구축하기 위한 기술로 Collaborative Filtering 기법을 사용한다. 추천시스템 기반기술의 개념적 배경은 일상생활에서 엿볼수 있는 구매패턴을 시스템화한 것이라 할 수 있다. 주로 물건을 사는 사람의 구매패턴을 예측(Prediction)하기 위한 방법은 구매패턴이 같은 다른 고객을 검색하여 그들의 구매

한 물건 중 가장 점수를 많이 얻은 고객에게 상품을 추천하는 것이 가장 적중률이 높기 때문이다[4].

### 2.3 데이터 마이닝(Data Mining)

데이터 마이닝이란 방대한 양의 데이터에 함축적으로 들어 있는 지식이나 패턴을 찾아내는 기술이라고 정의할 수 있다[5].

데이터 마이닝의 지식을 분류하면 친화도나 패턴을 찾아내는 연관규칙(association rule), 시간이라는 개념이 포함되어 동시에 발생할 가능성이 큰 데이터 집합을 찾아내는 순차 패턴(sequence pattern), 그룹화 되어있는 객체 집합에 대한 분류규칙(classification rule), 유사성이 높은 객체들을 그룹핑하는 클러스터링(clustering), 유사 패턴을 찾아내는 유사성 탐색(similarity search) 등이 있다.

### 3. 시스템 설계

본 논문은 사용자의 위치를 기반으로 온라인과 오프라인을 통합한 시스템으로서 컨텍스트 정보와 고객정보 구매정보 등을 데이터 마이닝 기법으로 측정, 분석하여 사용자의 구매성향이나 추천 리스트를 제공한다. 구체적으로, 우선 분류규칙과 클러스터링 기법을 사용하여 건물내 각 매장의 상품과 고객의 데이터를 분류하고 유사성이 높은 분류나 상품을 그룹핑하는데 사용하며 연관규칙 기법은 우선 단말기를 통한 사용자의 컨텍스트 정보와 건물내 공간DB를 통합하여 사용자의 이동 경로, 지체시간, 방문 우선순위를 추출하여 연관성을 측정하여 사용자의 구매 성향이나 구매패턴, 기호를 파악하여 추천리스트를 제공한다.

이 추천 리스트는 다시 추천 서버의 필터링 기법으로 최적의 추천 리스트를 생성한다.

상품 추천 시스템의 필터링은 상품 추천의 2가지 방법인 지식기반 상품 추천 방법과 collaborative filtering을 혼합하였으며 이 방법은 비슷한 취향을 가지는 고객들이 높이 평가하는 제품들을 최종적으로 고객들에게 추천하게 된다. 이 방법은 기존의 방법들이 모두 특정 한 상품 카테고리에 대해서만 효과적이었에 비하여 제안된 방법은 모든 상품 카테고리에 적용할 수 있으며 collaborative filtering 방법을 후보 추천 상품에 대해서만 적용시킴으로써 이 방법의 단점인 많은 계산량을 줄일 수 있다[4].

### 3.1 시스템 아키텍처

시스템의 구성은 그림 2와 같이 즉위 어플리케이션 서버(LAS : Location Application Server), 웹서버(Web server : Shopping mall), 메인프레임 데이터베이스(Mainframe Database), 마이닝 에이전트(DataMining Agent), 추천 서버(Recommendation Server)로 구성된다.

그림 3을 보면 WAP과 HTTP 사이의 즉위 게이트웨이 서비스의 흐름도가 있다. 선 단말기의 사용자가 서비스를 Request를 하면 즉위 게이트웨이를 통해 HTTP 프로토콜로 전달이 되어 즉위 어플리케이션 서버에 의해서 사용자의 위치정보가 제공이 되는 것을 볼 수 있다.

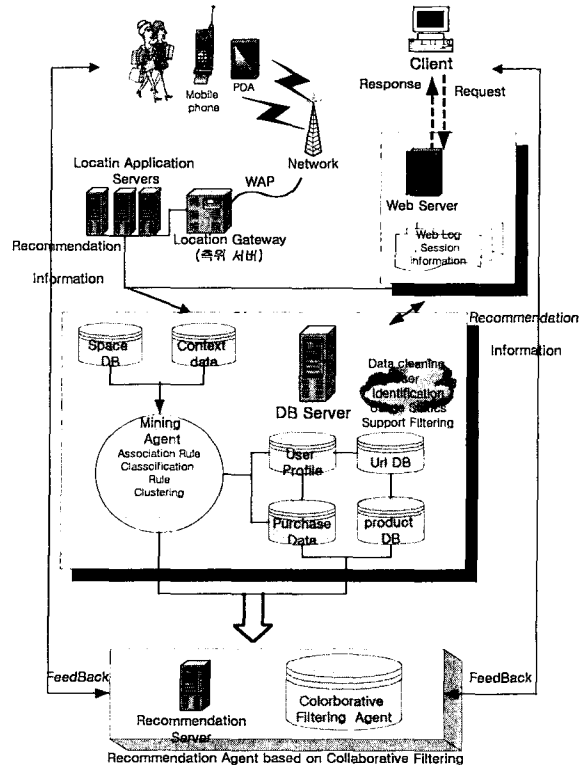


그림 2 System Architecture

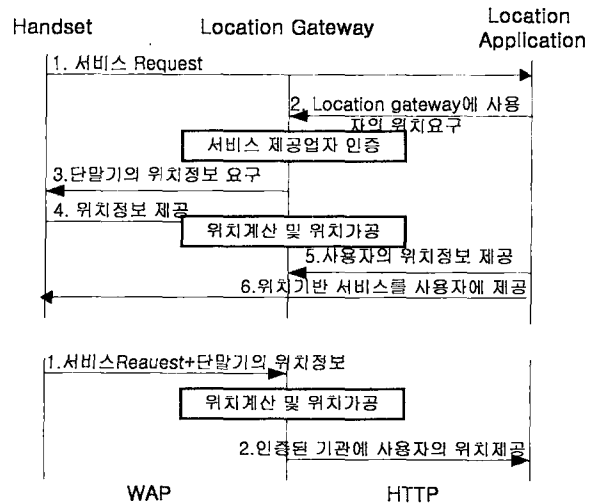


그림 3 즉위 게이트웨이 서비스의 흐름도

3.2 알고리즘

사용자에게 상품 및 이벤트 정보를 제공하기 위해서 건물안의 사용자의 위치와 각 매장 그리고 매장의 품목을 분류하여 표 1의 최적화된 DB 테이블을 생성하여 이들의 연관성 관계를 측정하기 위해서 연관규칙을 사용하였다[6][7].

표 1 최적화된 DB 테이블 생성

Name	Type	Describe
UserID	varchar(40)	사용자 ID
cornerID	varchar(40)	인접매장 ID
Item_1	varchar(40)	상품 대분류
Item_2	varchar(40)	상품 소분류
visit_Priority	varchar(30)	방문매방우선순위
visit_Count	varchar(30)	동일매장방문횟수
delay_Time	varchar(30)	매장 지체시간

$$S = P(X \cap Y) = \frac{\text{매장 } X \text{와 매장 } Y \text{의 공통품목수}}{\text{전체방문매장수 } N}$$

$$C = P(Y|X) = \frac{P(X \cap Y)}{P(X)} = \frac{\text{매장 } X \text{와 매장 } Y \text{에 포함하는 품목수}}{\text{매장 } X \text{에 포함하는 품목수}}$$

추천 시스템의 collaborative filtering 기법에서 항목 기반의 유사도 함수를 측정하기 위해서 Pearson correlation을 사용한다. 항목 매장 혹은 상품 x와 y의 상관 유사성 Corr(x,y)을 측정할 때,

$$\text{Corr}(x,y) = \frac{\sum_{u \in U} (R_{u,x} - \bar{R}_x)(R_{u,y} - \bar{R}_y)}{\sqrt{\sum_{u \in U} (R_{u,x} - \bar{R}_x)^2} \sqrt{\sum_{u \in U} (R_{u,y} - \bar{R}_y)^2}}$$

식을 사용하며 여기서 U는 사용자이고,  $R_{u,x}$ 와  $\bar{R}_x$ 은 사용자가 방문한 매장 또는 상품 벡터간의 유사도를 계산한다[6][7].

Collaborative filtering 모듈에서는 다음과 같은 식을 이용하여 사용자 m의 상품 n에 대한 추천 점수를 계산한다.

$$C_{mn} = \bar{C}_m + \kappa \sum_{i=1}^k w(m,n)(C_{in} - \bar{C}_i)$$

여기서  $C_{mn}$ 은 사용자 m의 매장 또는 상품 n에 대한 추정 점수이고  $\bar{C}_m$ 은 사용자 m의 평균 평가 점수,  $w(m,n)$ 은 매장 혹은 상품 m과 n의 거리, 연관도, 유사도 등을 의미한다. k는 해당 상품 평가에 참여한 고객의 숫자이다. 또한 k는  $w(m,n)$ 의 합의 값이 1이 되게 만드는 normalizing factor이며 이를 측정하는 방법은 GroupLens 프로젝트에서 쓰였던 Pearson correlation coefficient, cosine coefficient등 여러 가지가 있다[4]. 본 논문에서는 Pearson Correlation을 사용하였다.

4. 결론

본 논문에서는 사용자의 컨텍스트 정보를 활용하여 온·오프라인을 통합한 정보를 마이닝 기법과 추천 서버의 collaborative filtering과 Knowledge based 기법을 통해 추천 리스트를 제공하는 추천 시스템을 제안하였다.

사용된 컨텍스트 정보는 통신시 이용가능한 모든 개체들의 정보를 뜻하며 제안된 시스템에서는 실내에서의 사용자의 위치를 컨텍스트 정보로 사용하였다. 사용자의 위치정보는 무선 단말기를 통한 WAP과 HTTP, 그리고 측위 서버, 측위 게이트웨이 서버들을 통해서 측정된다.

이렇게 수집된 데이터들은 마이닝 기법을 통해 분류, 측정하여 예측된 추천 리스트가 생성이 되고 이 생성된 리스트는 추천 서버의 collaborative filtering 모듈로 연결되어 현재의 사용자와 비슷한 취향의 다른 사용자들로부터 높게 평가된 상품을 찾은 다음, 이 상품들을 최종 추천 리스트에 포함시킨다.

본 논문에서 제안된 시스템은 지식 기반 방법과 collaborative filtering 방법을 혼합 한 시스템으로서 양쪽의 장점만을 취하고자 했으며 기존의 추천 시스템들이 한정된 상품에 대해서만 효과적이거나 또는 오프라인에서는 제공이 되지 않았던 서비스에 반해 제안된 시스템은 오프라인에서도 실시간으로 제공이 되는 통합형 추천 시스템이다.

향후 본 시스템을 구현하고 좀더 개인화된 정보를 제공하는 추천 시스템에 대한 연구 및 개발이 필요하다.

참고문헌

- [1] 김욱, 지규인, 이장규, "위치 기반 무선 인터넷 서비스", Telecommunications review, 제10권 6호 2000. 11~12월.
- [2] David Mandato, Erno Kovacs, Fritz Hohl, and Hamid Amir-Alikhani, Son International (Europe) Gmb, "CAMP: A Context-Aware Mobile Portal", IEEE Communications Magazine, January 2002.
- [3] WAPFORUM : <http://www.wapforum.org>.
- [4] In-Gook Chun, "Knowledge-based Recommender System for Internet Shopping Mall"
- [5] 류근호, 안윤애, 이준욱, 이용준, "이동 객체 데이터베이스와 위치 기반 서비스의 적용", 한국정보과학회 Database research, 2001. 9월.
- [6] Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl, J., "Item-based Collaborative Filtering Recommender Algorithms", Accepted for publication at the WWW 10th Conference. May, 2001.
- [7] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Analysis of Recommendation Algorithms for E-Commerce", ACM Conference on Electronic Commerce, 2000.