

상품의 선호 특성 추출을 위한 고객 프로파일링 시스템

성경상*, 이종희, 김정재, 오해석
*숭실대학교 컴퓨터공학과
e-mail:actofgod@multi.ssu.ac.kr

A Customer Profiling System for Extraction of Preference Features of Goods

Kyung-Sang Sung*, Jong-Hee Lee*, Jung-Jae Kim*, Hae-Seok Oh*
*Dept of Computing, Soongsil University

요 약

최근에는 데이터마이닝(DataMining)이나 고객관계관리(Customer Relationship Management) 시스템을 이용하여 고객을 유치, 관리 및 서비스를 해오고 있으나, 이러한 시스템을 개발하고 관리하는데 있어서는 과도한 비용과 시간이 소요되는 점과 관리자가 시스템을 올바르게 이해하고 관리하는데 어려움이 따른다. 본 논문은 구매자와 유사한 신상정보를 지닌 사용자들로부터 선호하는 상품에 대해 추천을 받을 수 있으며, 하나의 상품에서도 여러 특성을 지니고 있다는 점을 고려하여 사용자의 구매 상품에 대한 특성을 파악하여 보다 정확하게 추천해 줄 수 있는 시스템을 개발하는데 그 목적이 있다.

1. 서론

최근 전자상거래의 발전은 인터넷의 대중화를 통하여 가속화되고 있으며 이로 인해 전자상거래 분야가 활발히 진행되면서 고객의 유치·관리 및 서비스에 대한 연구가 진행되고 있다.

쇼핑몰 등에서 자신에게 맞는 상품들을 추천 받는 서비스는 이미 미국 시장에서는 어느 정도 일반적인 일이 되었으며, 국내에서도 이미 그 자리가 다져지고 있는 실정이다. 철저하게 고객의 필요와 욕구에 밀착된 인터넷 서비스를 제공한다는 개인화(Personalize)의 취지에 비추어 볼 때, 이와 같은 추천 시스템은 필수적으로 입지를 굳힐 것으로 전망된다. 이런 상황이라면 고객에게 어느 정도의 개인화 서비스를 제공하느냐가 앞으로 쇼핑몰의 성패를 좌우하게 될 것임에 틀림없다.

개인화 서비스란, 무분별하고 과도하게 제공되는 정보들에 대해 여과 기술을 이용하여 무분별하게 제공되어지는 정보들을 정제(精製)함으로써 사용자 개인의 취향에 알맞은 정보를 제공하는 서비스를 말한다. 이러한 개인화 서비스는 최근에 정보의 추천을

통해 제공되고 있으나, 추천의 정확도에 있어서 한계를 보이고 있다는 문제점이 드러나고 있다.

또한 사용자들간의 유사도를 계산하는데 있어서 항목간의 중요도, 즉 가중치를 고려하지 못하는 문제점을 지니고 있을 뿐만 아니라 공통된 선호도 항목이 적을 경우 발생하는 희소성(sparsity)문제와 사용자의 수가 증가함에 따라 수행속도가 저하되는 범위성(slaability)문제로 인하여 효율성이 떨어지는 문제점들을 지니게 된다. 이러한 기존의 방식에 대한 문제점들을 어느 정도 해결하고자 새로운 기법을 활용하여 추천할 수 있는 방식을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 관련연구에 대해 살펴보고, 제 3장에서는 연구 목적 및 방법과 시스템의 전체적인 구조와 각 모듈들에 대해 다룬다. 제 4장에서는 제안하는 시스템의 주요 알고리즘에 대해 살펴보도록 하겠으며, 제 5장에서는 제안하는 시스템의 구현 환경과 추천되어져 가는 과정에 대해 자세히 살펴보도록 한다. 제 6장에서는 국내·외에서 현재 운용중인 개인화된 시스템을 기준으로 성능평가를 통해 제안하는 시스템의 우수성을

살펴본다. 제 7장에서는 결론과 향후 연구 과제를 통해 제안하는 시스템이 적용되어질 수 있는 방향을 제시하도록 기술한다.

2. 관련연구

2.1 WiseWire

WiseWire는 WiseWire사에서 만든 지능형 에이전트로, 인터넷에서 정보를 검색할 때 사용자의 기존 검색 패턴을 분석하여 얻은 프로파일을 이용하여 사용자들에게 각 개인이 원하는 정보를 보다 쉽게 검색할 수 있도록 하는 기능을 제공한다[4]. WiseWire는 관심정보와 비 관심정보를 기반으로 사용자의 관심(Profile)을 기계학습을 이용하여 학습하고, 이렇게 학습된 사용자의 관심을 이용하여 사용자 적용력 있는 검색을 수행한다. 또한, WiseWire는 각 사용자에 대한 관심 학습뿐 아니라, 일군의 그룹에 대한 학습(collaborative learning)을 수행한다.

2.2 오이 뮤직의 개인화 전략(http://www.oi.co.kr)

오이 뮤직에서는 회원 가입 과정을 거쳐 로그인을 하게 되면 인공지능 추천 센터를 통하여 사용자의 나이와 성별에 의해 일차적인 추천이 이루어진다. 또한 고객들의 구매 정보에 의한 조건에 따라 앨범에 대한 추천도 이루어지게 된다. 자신의 음악스타일에 따라 앨범 평가를 하는 과정을 거치게 되는 동안 진정한 개인화 서비스를 맞출 수 있게 된다. 또한 앨범을 구매한 다른 사람들의 구매 경향도 알 수 있어 구매에 도움을 받을 수 있다.

2.3 그룹화를 이용한 에이전트 기반의 전자도서관 시스템

에이전트 기반 전자도서관 시스템은 웹을 통해 도서정보를 검색하는 사용자들의 도서 정보 관심도를 모니터 에이전트에 의해 추출하게 되고 해당 정보를 대상으로 학습 과정을 거쳐 사용자별 프로파일을 구축 하는데 그 목적을 둔다. 이 과정에서 모니터 에이전트에서 추출되는 사용자 히스토리과 학습 예제 정보는 사용자가 행하는 다양한 작업의 결과를 이용한 차별화된 중요도가 적용된다. 학습 에이전트를 통해 구축된 사용자 관심 정보는 각 사용자에게 세부 DB 추천 모듈과 관심 정보를 푸쉬 해 주는 에이전트에 적용된다. 또한 전자 도서관에 가입하는 사용자들의 공통된 그룹 정보를 관리하여 이를 통해 동일 그룹내의 사용자들 사이의 정보 이용이 가능하게 한다.

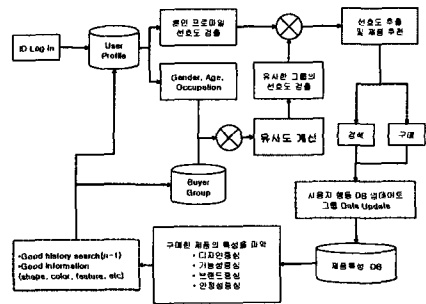
3. 프로파일링 에이전트의 설계

3.1 연구목적 및 방법

기존의 시스템에서는 사용자가 초기의 상황에 놓여 있을 때에 전혀 추천 해 줄 수 있는 여건이 마련되어 있지 못했다. 본 논문에서는 초기 상황 하에서도 사용자의 등록되어진 기본 정보와 유사한 등록정보를 지닌 사용자들을 그룹화하여 상품에 대한 특성들을 추출해서 사용자에게 아이템을 추천해 줄 수 있도록 하는데 그 취지가 있다.

3.2 제안하는 시스템의 구조 및 모듈

3.2.1 전체적인 시스템 구조

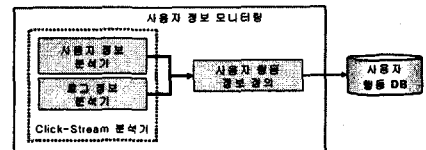


[그림 1] 전체적인 시스템 구성도

[그림 1]은 선호 특성 추출을 위한 고객 프로파일링 시스템의 전체적인 구성으로 학습 에이전트를 이용하여 사용자에 대한 선호도를 추출한다. 추출되어진 아이템에 대해 사용자의 구매 성향과 각 구매고객단의 구매 성향을 분석하여 선호도를 계산하게 되며, 이를 이용하여 추천해 줄 수 있도록 각 아이템에 대한 가중치를 고려하도록 한다.

3.2.2 구매 성향 파악을 위한 모니터 에이전트

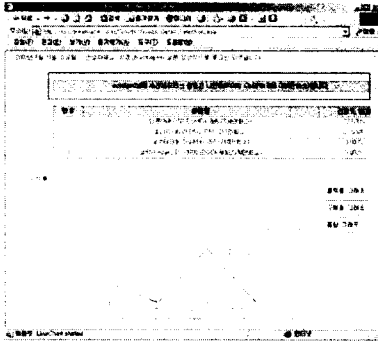
적용형 전자 상거래 시스템을 위해서 가장 먼저 수행 되어야 하는 것은 사용자가 취하는 행동들을 포괄적으로 탐지하는 것이다.



[그림 2] 모니터 에이전트

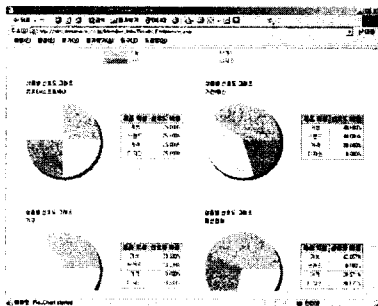
모니터 에이전트를 이용하여 사용자의 행위를 모니터하고, 각 사용자가 브라우징 한 상품, 구매한 상품 등에 대해 데이터베이스에 실시간적으로 적용되

프로파일링 과정을 거쳐 유사도를 추출하게 되고 중분류 카테고리 보이게 된다. 이러한 과정을 거쳐 최종 아이템의 선택에 이르기까지 사용자와 유사한 사용자 그룹은 축소되어지며, 상품 분류에 따른 카테고리의 특성을 이용한다할 수 있다.



[그림 5] 최종 아이템에 대한 추천 제품과 선호도 비율

[그림 5]에서 빨간 네모 박스 내부에 들어있는 내용이 사용자에게 제품 특성을 고려하여 추천하는 아이টে이 된다. 물론 사용자는 자신이 원하는 상품을 선택할 수 있다. 사용자가 물품을 구매하면 물품의 특성을 파악하게 되고, 파악되어진 특성을 사용자가 선호하는 특성으로 간주하게 되어 사용자가 선호하는 제품군의 전체적인 특성에 반영하게 된다.



[그림 6] 사용자의 아이টে에 대한 특성 분포도

[그림 6]에서는 사용자가 각 제품들의 카테고리에서 선호하는 특징들을 대분류별로 고려하였다. 이렇게 고려되어진 특징들은 사용자가 해당하는 품목을 선택했을 때 그 품목에서 사용자가 선호하는 특성들을 기준으로 물품을 추천해주게 된다. 또한 [그림 6]은 사용자가 선호하는 특징들에 대해 한 눈으로 확인할 수 있어, 해당 사용자의 심리를 파악할 수 있다.

6. 성능 평가

[표 1] 시스템 성능 평가

(◎ : 지원 불됨, △ : 지원 가능, X : 지원 안 됨)

구분	제안하는 시스템	오이유직	Visual Web Engine	전자도서관 시스템
초기추천기능	△	△	X	△
구매진추천	◎	△	△	△
구매성향파악	◎	◎	△	◎
개인성향파악	◎	△	△	△
성향성도입여부	◎	△	△	X
실시간추천성	△	X	△	X

7. 결론 및 향후 연구방향

본 논문은 사용자에게 자신만의 특정한 관심정보에 따른 각 제품군들에 대한 선호도 비율을 고려한 선호하는 아이টে을 추천받기 위한 에이전트 시스템을 제안하였다. 따라서 각 고객은 본인의 개인 관심 정보가 자신의 관심 정보 데이터베이스에 축적됨으로써 웹 쇼핑을 할 때마다 관심 컴포넌트들이 재생성되어지도록 지속적인 고객 관심도에 대해 시스템은 지속적인 재계산을 통하여 사용자의 관심 정보를 실시간적으로 갱신하게 된다. 향후 연구 과제로는 각각의 아이টে들에 따른 개별적인 특성을 고려한 아이টে을 추천 하게 된다면 사용자의 관심 정보 추출에 대해 보다 정확한 추천성을 높일 수 있을 것으로 사료된다.

참고문헌

[1] Trinity College Dublin, *Broadcom Eireann research Ltd, "Software Agent : A Review", 27, May, 1997.*
 [2] Maes,P, and Wexelblat, A., *"Interface Agents: A Tutorial."* In: *Conference Companion, Conference on Human Factors in Computing Systems, CHI 96, ACM Press, 1996.*
 [3] Dunja Mladenic, *Personal WebWatcher : Implementation and Design, Technical report IJS-DP-7472, October, 1996.*
 [4] Thorsten Joachims, *"A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization", March 1996.*
 [5] Bruce Krulwich, *"Learning document category description through the extraction of semantically significant phrases", Center for Strategic Technology Research Andersen Consulting LLP 100 South Wacker Drive, Chicago, IL 60606, 1995.*
 [6] Yezdi Lashkari, Max Metral, Pattie Maes, *"Collaborative Interface Agents", Conference of the American Association for Artificial Intelligence, Seattle, August 1994.*