

부산 이기종 인터넷 쇼핑몰 환경에서의 벡터 모델 기반 개인화 서비스 시스템

(A Personalized Service System based on Vector Model in Distributed Heterogeneous Internet Shopping Mall Environment)

박성준[†] 김주연^{**} 김영국^{***}
(Sungjoon Park) (Juyoun Kim) (Young-Kuk Kim)

요약 본 논문에서는 서로 다른 플랫폼으로 구성된 허브 사이트 가맹점들이 지역적으로 분산되어 있는 부산 이기종 환경에서 각 가맹점들의 기존 플랫폼을 통일시키지 않고 독립적으로 고객이 관심을 가질 만한 맞춤 정보 및 광고를 선정하여 제공하기 위한 시스템 설계 및 구현 내용을 서술한다. 본 논문에서 제안하는 시스템은 각 가맹점을 방문하는 고객들의 정보를 수집하기 위한 모니터링 기능, 고객 개인의 특성에 맞는 서비스를 선정하기 위한 벡터 모델, 그리고 벡터 모델을 이용하여 선정된 서비스를 허브 사이트 또는 각 가맹점을 방문하는 고객 개인에게 제공하기 위한 기능을 지원한다. 본 시스템은 상점 입장에서 허브 사이트 가맹점이 됨으로써 기존 플랫폼을 바꾸지 않고도 통합 서비스 및 개인화 서비스 제공이 가능하며, 고객 입장에서는 한번의 고객 등록으로 맞춤 서비스를 제공받을 수 있다는 장점을 갖는다.

키워드 : 부산 이기종 환경, 벡터 모델, 개인화 서비스, 허브 사이트

Abstract In this paper, we design and implement a system that presents a method for selecting and providing personalized services independently without unifying the existing system platform with shopping malls joined in the hub site. This system provides a mechanism for gathering information left behind by many clients visiting Web sites for analysis of customers property, vector model for selecting personalized services, and mechanism for providing them to customers who visited in a shopping mall joined to the hub site. In a position of shopping mall site, this kind of personalization system can provide target advertisement, point marketing, and point share service etc. without changing existing shopping mall's environment through wrapper web server. Hub site customers can get personalized services from many shopping mall sites with only once registration for the hub site.

Key words : personalization, hub site, vector model

1. 서론

최근 인터넷 사용의 폭발적 증가와 비즈니스 트랜잭션이 웹 상에서 수행될 수 있다는 편리함으로 인해 인터넷 사용자들을 대상으로 온라인 상에서 상품과 정보, 그리고 기타 서비스를 제공하고 수입을 창출하려는 인

터넷 비즈니스 사이트 수와 거래규모가 빠른 속도로 증가하고 있다[1,2,3].

인터넷 비즈니스의 대부분은 인지도가 높은 대규모 포털 사이트나 전문 쇼핑몰을 중심으로 이루어지고 있다[4]. 이러한 대규모 포털 사이트들은 상품 및 서비스 정보들을 중앙 집중적으로 관리하기 때문에 고객들에게 효율적이고 단일화된 서비스를 제공할 수 있는 장점이 있으나, 고객이 늘어날수록 서비스 속도 및 확장성에 문제가 있으며, 기존 백화점처럼 다양한 품목의 제품을 구비하고는 있으나 특정 품목에 대하여 전문점만큼 다양한 브랜드, 다양한 모델의 제품을 구비하고 있지 못하다.

Yahoo![5]와 같은 검색 서비스로 널리 알려진 사이트들은 인지도가 높은 고객들을 계속 유지하기 위한 방법으로 기본 화면을 편집할 수 있는 기능과 고객의 취

· 본 논문은 한국과학재단이 지정한 지역협력연구센터(RRC)인 충남대학교 소프트웨어연구센터의 지원으로 수행된 과제의 결과입니다(연구 과제번호 : 01-11-02-03-A-1).

† 정 회 원 : 공주영상정보대학 컴퓨터정보계열 교수
sjpark@kcac.ac.kr

** 비 회 원 : (주)배리텍
syph@cs.cnu.ac.kr

*** 정 회 원 : 충남대학교 컴퓨터과학과 교수
ykim@cs.cnu.ac.kr

논문접수 : 2001년 3월 28일
심사완료 : 2001년 12월 27일

향에 맞는 정보를 선별하여 볼 수 있도록 하는 개인화 페이지를 제공하고 있다. 그리고 아마존[6], CDNOW [7] 등과 같은 인터넷 비즈니스로 성공한 대규모 전문 쇼핑물들은 고객의 정보를 적극적으로 활용하여 고객의 편리성과 만족도를 극대화함으로써 최고의 인터넷 쇼핑물로 성장하고 있다. 웹에서의 개인화는 특정 고객이 웹 페이지를 방문하거나, 주식거래, 검색, 물건 구입 등과 같은 고객의 행위를 조건에 맞추도록 만들어 고객이 필요로 하는 것이 무엇인지 예견하여 더 많은 즐거움을 제공하는 것에서부터 고객에 맞는 정보를 제공하는 것에 이르기까지 다양하게 시도되고 있다[1].

다른 한편으로는 중앙 집중형 포털 사이트의 단점을 극복하면서 인지도가 낮은 네티즌들을 끌어 모으기 위한 방법으로 소규모 전문 인터넷 상점들과 서비스 제공 사이트들을 연합하여 각 사이트에서 확보한 회원들을 공유하여, 공동 마케팅, 공동 포인트 적립 등과 같은 통합된 서비스를 제공함으로써 고객의 충성도(loyalty)를 높이려는 허브(Hub) 사이트[8]가 등장하고 있다. 이러한 허브 사이트의 특징은 이에 참여하는 각 가맹점들이 지역적으로 분산되어 있으며, 서로 다른 플랫폼 위에서 각 가맹점 고유의 시스템을 독립적으로 구축 운영함으로써 자율성을 제공한다. 그러나 단일 플랫폼에서 개인화에 성공함으로써 많은 고객들을 끌어 모아 성장 가도에 있는 전자상거래 업체들과 달리, 아직 이러한 환경에서 개인화 맞춤 서비스를 제공하기 위한 해법을 제공하지 못함으로써 허브 사이트를 운영하고 있는 많은 업체들이 고전을 면치 못하고 있다. 허브 사이트를 이용한 인터넷 공동 마케팅이 성공하려면 구매력이 있는 네티즌들의 인지도를 높이면서 단순 방문에서 충성도가 높은 고객으로 유도하기 위한 개인화 맞춤 서비스 제공이 절실히 요구되고 있다.

본 논문에서는 이와 같이 각 가맹점들이 지역적으로 분산되어 있으며, 서로 다른 플랫폼으로 지역적으로 분산되어 있는 분산 이기종 인터넷 쇼핑물 환경에서 고객이 인터넷 쇼핑물 방문시 실시간으로 개인화된 맞춤 정보 및 광고를 선정하여 제공하기 위한 시스템의 설계 및 구현 내용을 서술한다. 본 논문에서 제안하는 시스템은 각 가맹점을 방문하는 고객들의 행위 정보를 수집하여 허브 사이트에 저장하기 위한 모니터링 기능, 고객 개개인의 특성과 서비스 특성의 유사도(similarity)가 높은 서비스를 선정하기 위한 벡터 모델(Vector Model), 그리고 벡터 모델을 이용하여 선정된 서비스를 각 가맹점의 기존 플랫폼을 변경하지 않고 각 가맹점을 방문하는 고객 개개인에게 제공하기 위한 기능을 지원한다. 또한 제안하는 시스템을 구현하고 이에 대한 시스템 성능

평가 결과 고객의 특성에 관계없이 무작위로 광고를 선정하는 임의(random) 방식에 비해 사용자수와 광고수가 늘어나면서 벡터 모델에 의한 개인화 서비스 선정 시간이 조금씩 증가하고 있으나, 이는 고객에게 개인화 서비스를 제공함으로써 가맹점에 대한 고객의 충성도를 높이기 위한 성능 비용이라 할 수 있다. 따라서, 고객의 충성도를 높이기 위해 개인화 서비스를 제공하는데 어느 정도의 성능 비용을 감수한다면 고객의 특성에 관계없이 임의로 서비스를 제공하는 것보다 벡터 모델을 이용하여 실시간으로 고객이 관심을 가질만한 개인화 서비스를 제공하는 것이 고객이 방문한 가맹점에 대한 충성도면에서 효과적임을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구에 관하여 서술하고, 제3장에서는 분산 이기종 환경에서 개인화 맞춤 서비스를 제공하기 위한 기술적 요구 사항, 고객 행위 모니터링 기법, 그리고 벡터 모델을 기반으로 하는 개인화 서비스 선정 및 제공 기법에 대해 제안한다. 제4장에서는 구현 및 시험 결과를 기술하고, 마지막으로 제5장에서 결론 및 향후 연구 방향에 대하여 서술한다.

2. 관련 연구

개인화란 상품, 서비스 또는 이와 관련된 정보를 고객 개인의 특성에 맞춰 고객 개개인에게 제공하는 것으로 정의할 수 있다. 개인화 시스템의 기본 목표는 고객이 원하거나 선택하지 않고도 고객이 원하거나 필요로 할 만한 것들을 제공하는 것이다[9]. 웹에서의 개인화는 특정 고객이 웹 페이지를 방문하거나, 주식 거래, 검색, 물건 구입 등과 같은 고객의 행위를 조건에 맞추도록 만들어 고객이 필요로 하는 것이 무엇인지 예견하여 더 많은 즐거움을 제공하는 것에서부터 고객에 맞는 정보를 제공하는 것에 이르기까지 다양하게 시도되고 있다[1].

개인화는 서비스가 제공되는 형태 또는 서비스 선정 방법에 따라 분류할 수 있다. 제공되는 형태에 따라 추천시스템(recommender systems)[1,10,11], 고객화(customization)[12], 적응형 웹 사이트(adaptive web sites)[9]로 구분할 수 있으며, 서비스 선정 방법에 따르면 규칙 기반 필터링(rules-based filtering)[1,13], 협업 필터링(collaborative filtering)[1,14], 학습 에이전트(learning agent)[1,15] 등으로 분류할 수 있다.

추천시스템은 준비 단계에서 수집된 고객 프로파일과 고객 행위 정보를 기반으로 데이터마이닝 기법에 의해 분석된 결과를 토대로 고객에게 맞춤 서비스를 제공하는 방법으로 아마존[6], CDNOW[7], 온빛 시스템[16] 등이 있다. 고객화는 고객화 서비스를 제공하는 사이트

를 방문했을 때 자신이 이미 설정해 놓은 내용과 모양에 맞춰 서비스 및 정보를 제공할 수 있도록 하기 위한 방법으로 기본 화면을 편집할 수 있는 기능과 고객의 취향에 맞는 정보를 선별하여 볼 수 있도록 하는 개인화 페이지를 제공한다. 자신이 원하는 지역의 날씨, 좋아하는 프로팀들의 주요 경기 일정 및 경기 결과, 관심 있는 뉴스 등 자신이 원하는 유형을 미리 정의함으로써 자신에 맞는 정보를 얻을 수 있는 방법이다. 이와 같은 방법은 고객이 원하는 항목을 미리 선택해 놓아야 하는 불편한 점이 있다. 대부분의 고객은 자신에 맞는 페이지를 디자인하기 보다는 기본으로 제공하는 페이지를 더 선호하기 때문에 이와 같은 유형의 서비스를 제공하는 많은 사이트들은 기본 페이지를 잘 설계하는데 많은 노력을 기울이고 있다. 대표적인 사이트로는 마이야후! [17]가 있다. 적응형 웹 사이트는 고객이 방문하기를 원하는 페이지로 이동하는데 최적화 경로를 통해 이동할 수 있도록 웹사이트를 구성하는 방법이다. 예를 들어 레저용 RV(Recreational Vehicle)를 구입하길 원하는 고객이 어떤 종류의 RV가 있는지 알아보기 위한 일반적인 방법은 먼저 자동차 회사를 방문해서 그 회사에서 생산되는 차를 구경하고 난 후, 또 다른 회사에서 제공하는 RV를 보기 위해서는 해당 회사의 URL를 이용하여 관련 사이트를 다시 접속하여야 한다. 고객이 원하는 제품을 선택하면 고객에게 이러한 번거로움을 덜어주기 위해 여러 회사에서 제공되는 다양한 RV를 볼 수 있도록 하면 고객은 빠르게 접근할 수 있으므로 매우 만족하게 될 것이다. 이와 같은 방법은 항해(navigation) 패턴 분석에 의해 효율적으로 제공할 수 있다. 대표적인 사이트로는 Music Machines [18]가 있다.

규칙 기반 필터링은 고객에게 등록시 개인 신상, 관심 분야, 선호도 등에 대한 몇 가지 질문을 통해 고객의 프로파일을 수집한다. 이렇게 수집된 고객의 인구통계학적/심리적 정보와 고객의 선호도 정보에 알맞은 정보 및 상품을 추천 혹은 제공하는 것이다. 이와 같은 방식은 Broadvision [19]을 예로 들 수 있다. Broadvision은 웹 사이트 관리자가 등록 과정을 통해 수집된 인구 통계학적 정보 또는 정적인 프로파일 정보를 기반으로 규칙을 생성하도록 하며, 이렇게 생성된 규칙은 특정 고객에게 적합한 정보를 선정하는데 이용된다. 협업 필터링은 비슷한 취향을 가진 고객들에게 서로 아직 구매하지 않은 상품들을 교차 추천하거나 분류된 고객의 취향이나 생활 형태에 따라 관련 상품을 추천하는 형태로 서비스를 제공한다. 이와 같은 방식은 Firefly [20], 그리고 Net Perceptions [21]을 예로 들 수 있다. 학습 에이전트는

고객이 어떤 페이지를 방문하는지, 어떤 광고를 선택하는지, 어떤 제품을 검색하는지 등과 같은 고객의 행동을 기준으로 고객의 선호도와 관심을 알아내고 이를 바탕으로 고객에게 적절한 맞춤 서비스를 제공하는 방법으로 Learn Sesame [22,23], WebWatcher [24] 등이 있다.

개인화 맞춤 서비스를 제공하기 위한 추천 시스템으로는 지금까지 협업 필터링 알고리즘이 널리 이용되어 왔다 [20,25]. 협업 필터링 기반 추천 시스템은 과거에 비슷한 행위를 보였던 고객들 간의 인접도(neighborhood)를 계산한 후, 인접도가 높은 고객들을 그룹으로 묶어 고객이 구입하지 않은 항목 중 같은 그룹에 속하는 다른 고객들이 주로 구입하는 상위 N개의 항목을 추천하는 방식이다. 그러나, 이와 같은 방식은 희소성(sparsity)과 확장성(scalability)에 문제가 있다 [20]. 한 사이트가 보유하고 있는 전체 항목 중 고객이 구입하는 항목의 수는 아주 적다. 각각의 고객과 항목에 대한 구매 정보의 양이 한정되어 있기 때문에 고객들간의 인접도를 정확하게 계산할 수 없고, 잘못 추천할 가능성이 높아진다. 또한, 고객들간의 인접도를 계산하기 위한 알고리즘은 고객과 서비스 항목 수의 증가에 따라 매번 인접도를 계산해야 하는 어려움이 있다.

이와 같은 문제를 해결하기 위한 한가지 방법이 고객의 군집을 구성하고 군집에 속하는 고객들 간에 인접도를 계산하여 광고나 서비스 항목을 추천하는 방식이다 [26,27]. 이러한 접근 방법은 추천 엔진의 속도는 증가시킬 수 있지만, 추천의 질이 떨어지는 경향이 있다. 또 다른 접근 방법은 고객 개인별 선호도 점수에 의한 방식이다 [28]. 고객 개인별 선호도 점수 방식은 고객 개인별로 제품군 트리의 말단 노드에 존재하는 제품군에 대한 선호도 점수를 테이블 형태로 관리하는 선호도 테이블 방식과 제품군 간의 인접도를 광고 선정에 반영하기 위한 선호도 트리를 이용하는 방법을 적용하고 있다. 이와 같은 방법은 제품이 한 두개로 제한적일 경우에는 매우 효과적이지만 각 제품마다 분리된 제품군의 수가 많아지게 되면, 제품간의 인접도를 반영하지 못하는 단점을 가지고 있다. 또한 각각의 고객은 모든 제품에 대한 제품군의 선호도 점수를 가지고 있어야 하며, 선호도 점수가 가장 높은 제품군에 속하는 광고가 선정되도록 되어 있다. 광고 측면에서 보면 제품군에 따라서만 광고가 분류되어 있고, 그 외의 특성에 대해서는 광고에 반영되어 있지 않으므로 고객의 선호도와 광고 특성 간의 유사성을 충분히 반영할 수 없다. 따라서, 본 논문에서는 이와 같은 문제점을 보완하기 위한 방법으로 고객의 프로파일과 제공될 서비스를 분석한 후, 벡터 모델을 적

용하기 위한 표현 방법에 따라 고객 성향 및 서비스 성향을 모두 반영할 수 있는 개인화 서비스 선정 기법을 제안한다.

3. 개인화 서비스 제공 시스템

3.1 시스템 구성

본 논문에서 제안하고 있는 분산 이기종 환경에서 개인화 서비스를 제공하기 위한 시스템 구성[29]은 [그림 1]과 같다.

허브 사이트는 소규모 인터넷 전문 상점들과 서비스 제공 사이트들이 연합하여 각 사이트에서 확보한 회원들을 공유하여 공동 마케팅, 공동 포인트 적립 등의 통합된 서비스를 제공함으로써 고객의 충성도를 높이기 위한 가상 공동체(virtual community)이다. 허브 사이트에 참여하고 있는 각각의 가맹점들은 물리적으로 서로 다른 지역에 분산되어 있으며, 시스템, 운영체제, 데이터베이스 등과 같은 서로 다른 이기종 환경에서 독립적으로 운용된다. 각 가맹점에 등록된 회원들이 허브 사이트에서 제공하는 개인화 맞춤 서비스를 제공 받기 위해서는 허브 사이트에 회원으로 가입해야 한다. 허브 사이트는 제3의 광고주로부터 광고를 수수받아 각 가맹점 페이지의 일정 영역에 고객의 특성에 따라 개인화 맞춤 광고를 제공할 수 있으며, 광고를 제공한 가맹점들은 고객에게 제공한 횟수에 따라 제3의 광고주로부터 광고 수입을 얻을 수 있다.

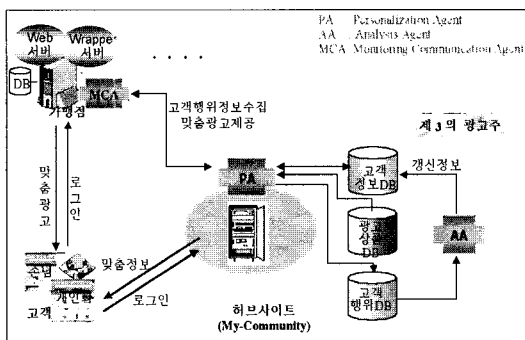


그림 1 전체 시스템 구성도

본 시스템은 허브 사이트에 가입된 각 가맹점 사이트에 존재하는 래퍼(wrapper) 서버와 모니터링 및 통신 에이전트(MCA), 그리고 허브 사이트에 존재하는 개인화 에이전트(PA), 분석 에이전트(AA)로 구성된다. 래퍼 서버는 각 가맹점의 기존 플랫폼을 변경하지 않고 개인

화 서비스를 제공하기 위해 추가적으로 설치한 웹 서버이다. 모니터링 및 통신 에이전트(MCA)는 각 가맹점 래퍼 서버에 존재하면서 허브 사이트 회원으로 가입된 고객이 가맹점을 직접 방문하거나 허브 사이트를 통해 가맹점을 방문하여 머무르는 동안 고객이 행한 일련의 정보들을 수집하여 고객이 가맹점을 떠남과 동시에 실시간으로 허브 사이트에 있는 고객 행위 데이터베이스에 기록하는 역할을 수행한다. 개인화 에이전트(PA)는 허브 사이트 회원이 각 가맹점 또는 허브 사이트 방문 시 고객의 특성에 가장 적합한 광고를 선정하는 일과 허브 사이트와 각 가맹점에 선정된 정보를 제공하는 기능을 수행한다. 그리고 분석 에이전트(AA)는 모니터링 및 통신 에이전트에 의해 수집된 고객 행위 정보와 초기에 얻어진 개인 프로파일 정보를 이용하여 고객 특성의 각 항목에 대한 가중치를 분석하는 기능을 가진다. 분석 에이전트는 본 논문의 범위에서 제외하기로 한다.

3.2 개인화 서비스 제공 요구 조건

허브 사이트에 가입하는 각 가맹점들은 지역적으로 분산되어 있으며, 이질적인 서로 다른 시스템들이 독립적으로 운용되는 특징을 가지고 있다. 이와 같은 분산 이기종 환경에서 개인화 맞춤 서비스를 제공하기 위해서는 다음과 같은 요구 조건을 만족시켜야 한다.

첫째, 새로운 쇼핑물이 허브 사이트에 추가될 때, 쇼핑물이 가지고 있는 기존 머천트 시스템을 변경하거나 수정하지 않고 개인화 서비스를 쉽게 제공할 수 있어야 한다.

둘째, 이미 허브 사이트에 가입된 각 가맹점 사이트와 독립적으로 새로운 쇼핑물이 허브 사이트에 추가될 수 있어야 한다.

셋째, 각 가맹점에서 제공되는 페이지의 일정 영역에 개인화 맞춤 서비스가 제공될 수 있어야 한다.

넷째, 가맹점을 방문하는 고객은 이전에 방문한 페이지와 다르다는 느낌을 받지 않으면서, 고객의 관심도가 높은 서비스를 제공받을 수 있어야 한다.

다섯째, 가맹점 사이트에 대한 상품 정보나 광고 등을 허브 사이트에 등록하고 이들 방문 고객들에게 선별적으로 제공되어야 한다.

3.3 고객 행위 모니터링 기법

고객 행위 모니터링 기법은 허브 사이트 회원인 고객이 각 가맹점에 직접 방문하는 경우와 허브 사이트를 통해 각 가맹점을 방문하는 경우에 고객 행위를 수집하기 위한 방법이다. 수집하게 되는 고객 행위 로그 정보는 고객의 성향을 분석하는데 필요한 정보들로 고객 ID, 방문페이지, 구매 정보, 선택된 광고 정보, 검색어, 그리

고 웹 페이지를 방문한 시간과 떠난 시간 등이다.

허브 사이트에서 수집된 고객 행위 정보를 기반으로 고객 개개의 성향을 분석하기 위해서는 각 가맹점으로 부터 모니터링된 로그 정보를 한곳으로 통합하여야 한다. 따라서 각 가맹점에서는 로그 정보를 모니터링하여 이들 정보를 허브 사이트로 제공하기 위해서는 고객의 상태 정보를 유지하면서 고객 행위를 수집하는 방법과 수집된 정보를 허브 사이트로 보내기 위한 방법이 필요하다. 고객의 상태 정보를 유지하면서 고객 행위를 수집하는 방법으로는 숨은 서식 항목(hidden form field), 쿠키(cookie) 또는 세션(session)을 이용할 수 있다. 그리고 수집된 정보를 허브 사이트로 보내기 위한 방법으로는 각 가맹점에 로그 파일로 일단 저장 한 후, 오프라인으로 주기적으로 로그 파일을 가져다 허브 사이트의 고객 행위 데이터베이스에 저장하는 방법과 고객이 가맹점을 방문한 후, 가맹점을 떠나자마자 고객이 가맹점에 머무는 동안 수집된 로그 정보를 실시간으로 허브 사이트에 보내는 방법이 있다. 다음은 고객 행위를 수집하기 위한 각각의 방법들에 대한 장단점과 본 논문에서 이용된 기법을 기술하였다.

HTTP 프로토콜은 상태 정보를 가지고 있지 않은 프로토콜이기 때문에 고객이 웹 페이지 방문시 웹 서버에서 클라이언트로의 전송이 끝나면 연결도 곧바로 끊어지기 때문에 고객이 다른 페이지로 이동하게 되면 동일 고객인지 알 수 없다. 이와 같은 문제를 해결하기 위한 전형적인 방법으로는 숨은 서식 항목, 쿠키, URL 재작성, 그리고 세션 등이 있다[30].

숨은 서식 항목은 클라이언트에게 제공되는 웹 페이지의 모습에 영향을 끼치지 않으면서, 클라이언트의 입력에 관계없이 서버에 보내지는 고정된 이름과 값을 저장하는 방식이다. 숨은 서식 항목은 세션에 관한 정보를 저장하는데 이용될 수 있지만 각 페이지마다 동적으로 생성될 때만 작동되고 브라우저 종료와 함께 사라지므로 숨은 항목에 특정 값을 입력해서 서버에 전송하면, 서버는 이 숨은 항목의 내용을 바탕으로 동작하는데 이용될 수 있으며, HTML 태그 내에서 숨은 서식 항목을 이용하여 고객의 링크 클릭이나 이미지 클릭, 또는 검색 행위 등을 구별할 수 있다. 그러나 이와 같은 기능은 JSP(Java Server Pages)[31]의 등장으로 JSP 빈즈를 이용하여 해결할 수 있다.

쿠키란 클라이언트가 웹 페이지 방문시 웹 서버가 적은 양의 텍스트 정보를 브라우저에 보내고 브라우저는 이 정보를 보관하고 있다가 추후 동일 웹 페이지 방문시 웹 서버에 URL과 함께 제공하는 방법이다. 그러나, 쿠키

의 문제점은 사이트들이 민감한 데이터에 대해 쿠키에 의존할 때 발생한다. 어떤 사람의 컴퓨터 또는 쿠키 정보를 얻은 공격자는 가치있는 개인 정보에 온라인 접근을 할 수 있다. 최악의 경우 무자격 사이트들이 서버상에 있는 실제 사용들에게만 연결된 식별자들을 이용하는 것이 아니라 쿠키 자체에 직접 신용 카드나 다른 민감한 정보를 끼워넣을 수 있다. 따라서 쿠키의 이용은 매우 제한적이며 이를 보완하기 위한 방법으로 보안이 강화된 전자 지갑과 같은 것들이 등장하고 있다[32].

URL 재작성은 URL 끝에 세션을 식별하는 부가적인 정보를 추가하는 방법으로 브라우저가 쿠키를 지원하지 않거나, 클라이언트가 쿠키를 사용할 수 없도록 한 경우 좋은 방법이다. 그러나 이와 같은 방법은 웹 사이트를 참조하면서 클라이언트에 반환된 모든 URL마다 부가적인 정보가 추가되어야 하기 때문에 주의해야 한다. 그리고 클라이언트가 세션을 떠났다가 즐겨찾기 또는 링크를 통해 돌아온 경우 세션 정보를 잃어 버릴 수 있다.

세션이란 클라이언트의 지속적인 접속을 의미한다. 세션을 설정하여 이어지는 접속이 이전 클라이언트로부터 온 것인지 확인하기 위해 웹 서버는 브라우저에 세션 고유 번호를 쿠키 값으로 건네주고, 이후 접속시 마다 이 값을 이용하여 어떤 클라이언트로부터 접속인지 확인하는 방법이다. 만약 세션 ID가 없다면 세션을 설정하지 않았으므로 고유 번호를 생성해서 브라우저에 보내줄 것이고, 세션ID가 이미 있으면 그 리스트를 살펴서 출처를 확인한다. 세션ID가 현재 리스트에 없으면 그 세션은 유효하지 않은 것으로 간주한다. 이와 같은 방법을 쿠키만으로 구현하게 되면 모든 문서마다 쿠키 값을 주고 받고 유효한지 확인하고 새로운 고유번호를 생성하고 이 번호들을 관리해야 할 것이다. 자바에서는 이와 같은 세션 기능을 매우 간단하게 처리한다. 특히, JSP 세션 빈즈를 이용하면 효율적으로 세션을 관리할 수 있다.

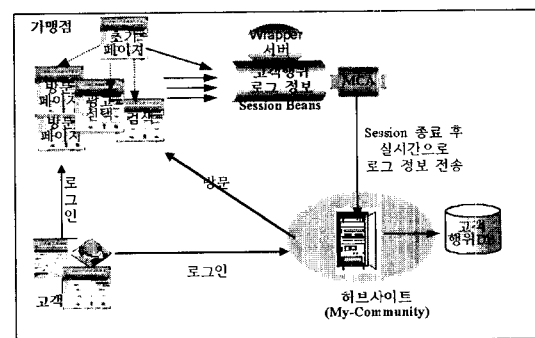


그림 2 고객 행위 모니터링 구성도

본 논문에서는 고객 행위를 수집하기 위한 방법으로 JSP에서 제공되는 세션 빈즈를 이용하여 세션 동안 고객 행위가 발생할 때마다 관련 로그 정보를 일시적으로 벡터에 저장한다. 세션 트래커(Session Tracker)[33]는 일정 시간 동안 동일 고객으로부터 요청에 대한 상태 정보를 유지하기 위한 자바 웹 서버에 의해 제공되는 기능이다. 세션 트래커에 의해 일정 시간 페이지 요청이 없는 세션 객체들은 자동으로 세션이 무효화 된다. 세션이 무효화 되면, 세션이 가지고 있던 모든 정보도 시스템에서 제거되므로, 세션이 무효화 될 때, 이벤트를 받아 처리하기 위한 인터페이스로 서블릿에서 제공하는 Http SessionBindingListener 인터페이스의 valueUnbound 메소드를 통해 세션이 무효화 될 때 객체가 신호를 받아 실시간으로 허브 사이트에 연결된 데이터베이스에 일시적으로 벡터에 저장된 로그 정보를 보내는 기법을 이용한다. 또한, 세션 동안 고객 행위 정보를 얻기 위해 각 가맹점의 기존 페이지마다 세션 관련 코드를 추가하는 것이 아니라 래퍼 서버를 두어 이곳에서 새로 추가되거나 변경되는 내용을 반영하도록 한다.

[그림 2]는 고객이 가맹점을 직접 방문하거나 허브 사이트를 통해 가맹점을 방문하는 경우 고객 행위를 모니터링하여 수집된 정보를 허브 사이트로 전송하기 위한 구성도이다. 고객이 가맹점을 직접 방문하거나 허브 사이트를 통해 가맹점을 방문하게 되면 JSP에서 제공되는 세션을 이용하여 고객과의 세션을 설정하고, 가맹점 래퍼 서버는 고객의 쿠키로부터 E-mail 주소를 가져와 고객ID와 세션ID를 얻어오고, 세션 빈즈를 이용하여 고객 방문 페이지 정보를 일시적으로 벡터에 저장한다. 벡터에 저장된 정보는 고객이 가맹점에 머무르게 되는 세션기간 동안만 존재하며, 세션이 종료되면 모니터링 에이전트는 허브 사이트에 존재하는 고객 행위 데이터베이스와 JDBC를 통해 직접 연결하여 벡터에 저장된 로그 정보를 전달한다.

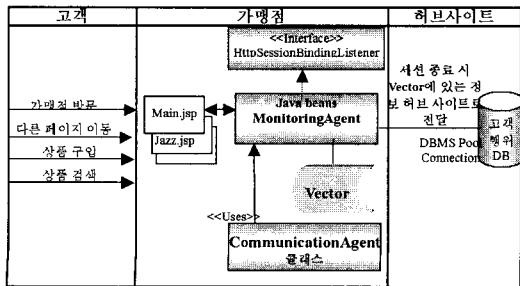


그림 3 고객 행위 모니터링 과정

[그림 3]은 각 가맹점에서의 고객 행위를 모니터링하기 위한 일련의 과정을 묘사한다. [그림 4]는 고객이 가맹점 방문을 종료한 경우 모니터링 정보를 허브 사이트 고객 행위 데이터베이스에 저장하기 위한 모니터링 및 통신 에이전트 클래스 구성도이다. MonitoringAgent 클래스는 고객 고유 값과 로그인 페이지를 설정하고, 각각의 고객 행위를 벡터에 추가하기 위한 메소드들로 구성된다. CommunicationAgent는 모니터링 정보를 허브 사이트에 보내고, 허브 사이트로부터 개인화 서비스를 제공받는 역할을 수행하는 메소드들로 구성된다.

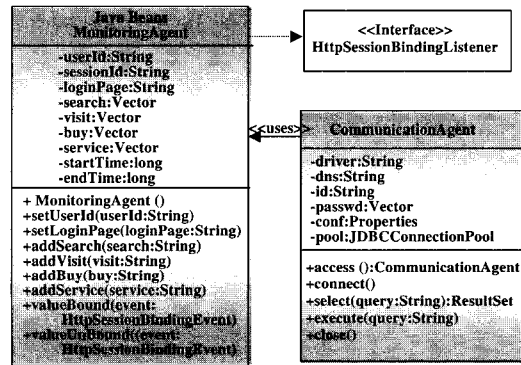


그림 4 모니터링 및 통신 에이전트 클래스 구성도

3.4 개인화 서비스 선정 기법

고객의 성향에 맞는 개인화 서비스를 선정하기 위한 과정은 [그림 5]와 같이 먼저 고객 개인의 프로필과 수집된 고객 행위 정보를 이용하여 고객 성향을 표현하고, 의뢰 받은 서비스의 내용을 분석하여 서비스 내용에 따라 서비스 성향을 표현한다. 그 다음 이렇게 표현된 결과를 이용하여 고객 성향과 서비스 성향간의 상관 관계를 계산하여 상관 관계가 높을수록 고객의 성향에 적합한 서비스라 생각할 수 있다.

본 논문에서는 고객 행위가 수집되어 분석되기 전 단계로 고객 프로필만을 이용하여 고객 성향을 표현하였으며, 서비스는 내용 분석을 위해 의뢰자로부터 서비스되어지기를 원하는 고객의 특성과 광고 내용을 수집하여 분석한 후 개인화 에이전트가 처리할 수 있는 형태로 표현할 수 있다고 가정한다. 이와 같은 가정에서 주어진 고객 성향과 서비스 특성을 비교하여 고객에 가장 적합한 서비스를 선정하기 위한 벡터 모델을 제안한다.

벡터 모델

광고와 고객이 가지는 특성에 가중치(non-binary weight)를 할당함으로써 0과 1로 할당된 binary

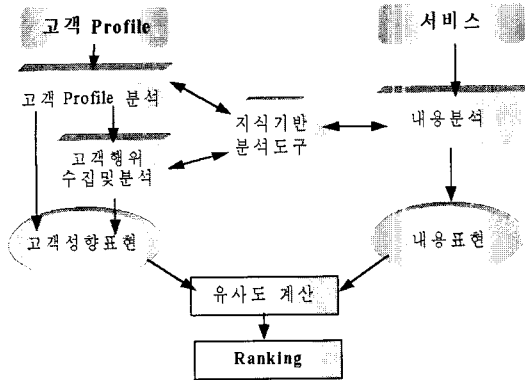


그림 5 개인화 맞춤 서비스 선정 과정

weight의 한계를 보완하는 방법으로 시스템 전체에 있는 각각의 광고와 고객 특성 벡터간의 상관 관계로써 유사성을 계산하는데 벡터 모델을 이용할 수 있다. 벡터 모델은 다음과 같이 정의할 수 있다.

정의 1

벡터 모델에 대해서 가중치 $W_{i,j}$ 는 $(k_i, ad_j) \geq 0$ 쌍으로 구성된 비이진 값(non binary)이다. 여기서,

k_i : 광고가 가지는 각각의 특성(인구통계적 특징 및 관심사항) 중 i 번째 특성

ad_j : j 번째 광고

이며, 고객 성향에 대한 가중치 $W_{i,cust}$ 는 $(k_i, cust) \geq 0$ 쌍으로 구성된다고 하자. 그러면, 고객 벡터는 다음과 같이 표현할 수 있다. $cust = (W_{1,cust}, W_{2,cust}, W_{3,cust}, \dots, W_{t,cust})$ 여기서, t 는 허브 사이트에 존재하는 모든 고객 성향의 수, 광고 벡터도 고객과 마찬가지로 $ad_j = (W_{1,j}, W_{2,j}, W_{3,j}, \dots, W_{t,j})$ 로 표현할 수 있다. 따라서, 광고 ad_j 와 고객 $cust$ 는 [그림 6]과 같이 t 차원 벡터로 생각할 수 있다.

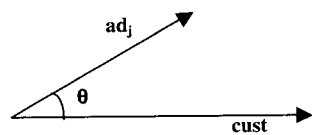


그림 6 Cosine $\theta = sim(ad_j, cust)$

광고 벡터 ad_j 와 고객 벡터 $cust$ 사이의 코사인 상관 계수[34]를 바탕으로 하는 벡터 부합 연산을 사용하여 다음 식과 같이 광고와 고객간의 유사성을 계산할 수 있고, 이 유사성을 이용하여 광고에 대한 순위를 부여할

수 있다. 즉, 광고와 고객간의 유사성을 계산하는 식은 다음과 같다.

$$sim(ad_j, cust) = \frac{ad_j \cdot cust}{|ad_j| \times |cust|} = \frac{\sum_{i=1}^t (W_{i,j} \times W_{i,cust})}{\sqrt{\sum_{i=1}^t W_{i,j}^2} \times \sqrt{\sum_{i=1}^t W_{i,c}^2}}$$

여기서, $sim(ad_j, cust) = 1$ 이면 고객과 해당 광고는 매우 밀접한 관계를 가지고 있다고 할 수 있으며, $sim(ad_j, cust) = 0$ 이면 연관성이 전혀 없다고 할 수 있다.

표 1 벡터 모델에 의한 순위 부여 접근법

| 특성 벡터 | 결혼 | | 성별 | | 나이 | | | | 직업 | | 관심사항 | | | | | | | | |
|-------|----|----|----|----|----|----|----|----|----|----|------|----|----|----|----|---|---|---|---|
| | 기혼 | 미혼 | 남성 | 여성 | 1대 | 2대 | 3대 | 4대 | 안정 | 취업 | 취업 | 취업 | 취업 | 취업 | 취업 | | | | |
| 고객 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 2 | 0 |
| 광고1 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 2 | 1 | 1 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 |
| 광고2 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 |
| 광고3 | 1 | 0 | 1 | 0 | 0 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |

예를 들어, 고객 벡터와 광고 벡터가 [표 1]과 같이 가중치를 가지는 경우 광고1에 대해서 다음과 같이 계산할 수 있다.

$$\sum_{i=1}^t (W_{i,j} \times W_{i,cust}) = 7 \text{ 여기서, } t = 19$$

$$\sqrt{\sum_{i=1}^t W_{i,j}^2} = \sqrt{27} = 5.19, \quad \sqrt{\sum_{i=1}^t W_{i,c}^2} = \sqrt{12} = 3.46$$

$$sim(ad_j, cust) = 7 / 17.95 = 0.39$$

같은 방법으로 나머지 광고들에 대해서도 유사도를 계산하면 고객과 광고2의 유사도 값은 0.31이 되며, 고객과 광고3의 유사도 값은 0.72가 된다. 이상에서 보는 것처럼 가중치 부여 방법에 따라 고객과 광고간의 유사도가 매우 달라짐을 알 수 있다.

가중치 부여 방법은 첫째, 광고 대상 품목에 대한 구매 고객들에 대한 고객 특성 정보(결혼유무, 성별, 직업 등)를 분석하여 각 특성마다 구매 고객 수를 합한 후, 이를 전체 구매 고객으로 나눈 결과를 가중치로 이용하는 방법이 있다. 둘째는 광고주가 허브 사이트에 광고 의뢰시 프로파일에 광고 특성을 명시하도록 하는 사람에 의해 주관적으로 모델링하는 방법이 있다[35]. 첫 번째 가중치 부여 방법에 대한 예를 들면, 광고1에 대한 품목을 구매한 고객이 전체 100명이라 하고, 각각의 구매 고객에 대한 인구통계학적 특성을 분석한 결과 [표 2]와 같다고 가정하고 가중치를 계산할 수 있다. 또한 피

드백(feedback)을 통해 구매 고객에 대한 축적된 정보를 주기적으로 변경할 수 있다.

표 2 광고1 구매 고객 특성 분석(전체 100명)

| 특성 | 결혼 | | 성별 | |
|-----|-----|-----|-----|-----|
| | 기혼 | 미혼 | 남 | 여 |
| 광고1 | 80 | 20 | 70 | 30 |
| 가중치 | 0.8 | 0.2 | 0.7 | 0.3 |

3.5 개인화 서비스 제공 기법

분산 이기종 환경에서 개인화 맞춤 서비스를 제공하기 위한 요구 조건을 충족시키면서 개인화 맞춤 서비스를 제공하기 위해 각 가맹점이 가지고 있는 기존 웹 서버와 독립적으로 래퍼 서버를 두었다. 래퍼 서버는 JSP와 JSP 빈즈를 이용하여 기존의 가맹점 웹 페이지의 일정 영역에 개인화된 광고를 제공하기 위한 기능과 JSP 세션을 이용하여 방문 고객의 행위를 모니터링하고 모니터링된 정보를 세션이 끝날 때 실시간으로 허브 사이트에 있는 고객 행위 데이터베이스에 전송하기 위한 기능을 제공한다.

개인화 서비스를 제공하기 위한 시나리오는 다음과 같이 3가지로 구분할 수 있다. 첫째, 허브 사이트를 방문하는 고객에게 개인화 서비스를 제공하는 방법, 둘째, 가맹점 사이트를 직접 방문하는 고객에게 개인화 서비스를 제공하는 방법, 그리고 셋째는 허브 사이트를 먼저 방문한 후 가맹점 사이트를 방문하는 고객에게 개인화 서비스를 제공하는 경우로 구분할 수 있다.

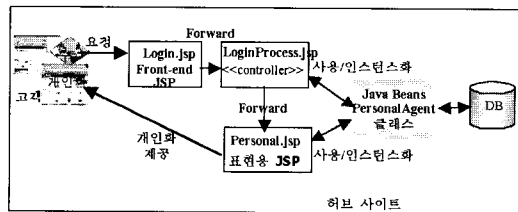


그림 7 허브 사이트 개인화 제공 구조

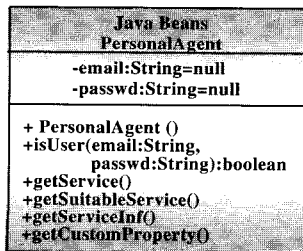


그림 8 PersonalAgent 클래스 구성도

[그림 7]은 허브 사이트를 방문하는 고객에게 개인화 서비스를 제공하는 구조이다. Login.jsp 페이지를 통해 고객이 로그인하면 Login.jsp는 컨트롤을 LoginProcess.jsp로 넘겨 PersonalAgent 클래스에서 고객 인증 후, 고객에 맞는 개인화 서비스 선정 과정을 거쳐, Personal.jsp를 통해 고객에게 개인화 서비스를 제공한다. [그림 8]은 PersonalAgent 클래스 구성도이다.

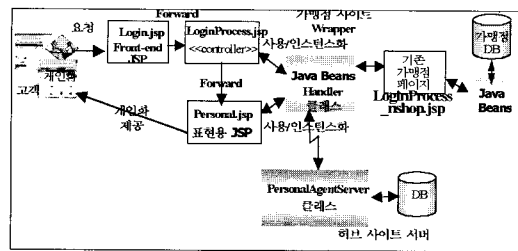


그림 9 가맹점 사이트 개인화 제공 구조

[그림 9]와 [그림 10]은 허브 사이트 회원이 가맹점을 직접 방문하는 경우 개인화 서비스를 제공하기 위한 구조와 처리기(handler), PersonalAgentServer 클래스 구성도이다. PersonalAgentServer는 분산 환경에 있는 각 가맹점으로부터 개인화 서비스 요청을 받으면, 벡터 모델을 이용한 개인화 선정 알고리즘을 통해 고객 개인의 특성에 맞는 개인화 서비스를 선정하여 각 가맹점에 제공하기 위한 데몬 프로세스로 허브 사이트에 존재한다.

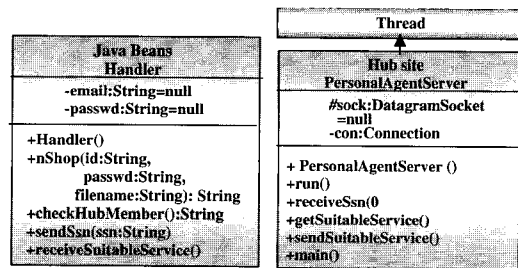


그림 10 처리기, PersonalAgentServer 클래스 구성도

[그림 11]은 허브 사이트 회원이 허브 사이트를 먼저 방문한 후, 허브 사이트에 가입된 가맹점을 방문하는 경우 개인화 서비스를 제공하기 위한 구조이다. 먼저, 고객이 방문한 가맹점의 방문 페이지의 빈즈 클래스 처리기를 통해 허브 사이트의 PersonalAgentServer에 개인화 서비스를 요청한다. 그러면 PersonalAgentServer는 개인화 서비스 선정 과정을 거쳐 선정된 서비스를

가맹점에 보내면 Personal.jsp 페이지를 통해 고객에게 개인화 페이지를 제공한다.

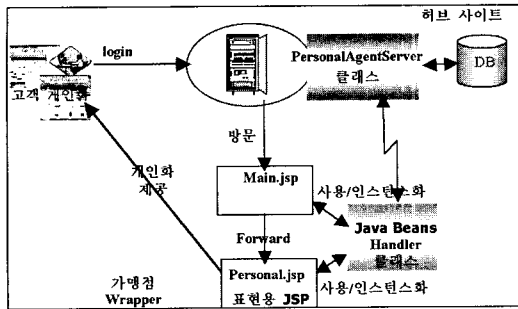


그림 11 허브 사이트를 통한 가맹점 방문시 개인화 서비스 제공 구조

4. 구현 및 시험

4.1 시스템 구현 환경

본 논문에서 제안한 시스템을 구현하기 위한 구현 환경은 [표 3]과 같다. 구현 언어로는 JSP를 이용하였으며, 허브 사이트에 가입된 가맹점으로는 LAN으로 구성된 인터넷 환경에서 이질적이고 독립적으로 수행되는 CD Shop과 Video Shop으로 구성된다.

허브 사이트와 CD Shop은 Linux 운영체제에서 데이터베이스는 Oracle8을 이용하였으며, 웹 서버로는 최근 많이 이용되고 있는 Apache1.3을 이용하였다. 그리고 Video Shop은 Windows NT 운영체제에서 데이터베이스는 MySql 3.21, 웹 서버로는 Microsoft 사의 IIS를 이용하였다. 이와 같은 서로 다른 환경에서 동작하는 서로 다른 가맹점들을 허브 사이트와 연합하여 각 가맹점들이 제공하는 기존 시스템을 변경하지 않고 개인화 맞춤 서비스를 제공하기 위한 방법으로 각 가맹점에 래퍼 서버를 두었다. 래퍼 서버로는 Apache1.3 웹 서버를 이용하였으며, 웹 서버에서 JSP를 사용하기 위해 Tomcat

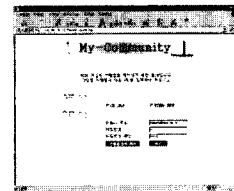
표 3 시스템 구현 환경

| 환경 \ 가맹점 | 허브 사이트 | CD Shop | Video Shop |
|----------|----------------------|----------------------|----------------------|
| 운영체제 | Linux 2.2.9 | Linux 2.2.9 | Windows NT |
| DBMS | Oracle 8.0.5 | Oracle 8.0.5 | MySql 3.21 |
| 웹 서버 | Apache1.3 | Apache 1.3 | IIS 5.0 |
| AP 서버 | Tomcat 3.1 JDK1.3 | Tomcat 3.1 JDK1.3 | Tomcat 3.1 JDK1.3 |
| 기타 | Oracle JDBC | Oracle JDBC | MySql JDBC |

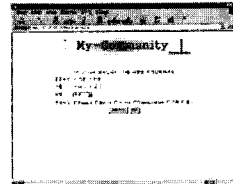
3.1과 JDK1.3을 이용하였다. 그리고 각 가맹점에서 허브 사이트에 있는 Oracle 데이터베이스에 접근하기 위해 JDBC 드라이버를 이용하였다.

4.2 구현 결과 및 실험 시나리오

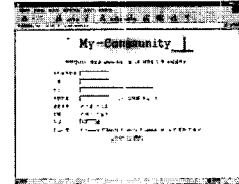
[그림 12]는 본 논문에서 개발한 인터넷 쇼핑몰 분산 환경에서 개인화 서비스를 제공하기 위한 시스템의 실행 화면 예를 나타낸다. [그림 12]의 (a)는 회원 가입 초기 화면으로 자신의 E-Mail주소와 패스워드만을 입력하게 되면 기존 가맹점 회원 여부를 파악한 후, 기존 가맹점 회원이면 가맹점으로부터 고객의 기본 정보를 가져오게 되며, 고객 성향을 분석하는데 필요한 기본 정보만을 [그림 12]의 (b-1)과 같이 요청한다. 기존 가맹점



(a) 회원 가입 초기 화면

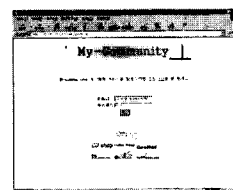


(b-1)

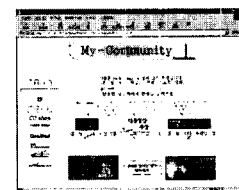


(b-2)

(b) 상세 가입 화면



(c-1)



(c-2)

(c) 로그인 및 허브 사이트 개인화 페이지



(d) 가맹점 개인화 페이지

그림 12 개인화 서비스 실행 예

회원이 아니면 (b-2)와 같이 상세 정보를 요청한다. (c-1)은 허브 사이트에 로그인 하기 위한 인터페이스를 나타낸다. (c-2)는 인증된 고객에게 개인화 서비스를 제공하는 실행 예를 나타내며, (d)는 고객이 가맹점을 방문한 경우 제공되는 개인화 서비스 실행 예이다.

다음은 고객이 허브 사이트를 방문하여 회원 가입 후 가맹점을 방문하기까지 개인화 맞춤 서비스를 고객에게 제공하는 일련의 실행 시나리오이다.

- 1) 고객이 웹 브라우저를 통해 My-Community 허브 사이트에 접속한다(c).
- 2) 비 회원인 경우 회원 가입 버튼을 선택한다(c).
- 3) 초기 가입에서 자신의 E-Mail 주소와 패스워드를 입력한다(a).
- 4) 가맹점 회원인 경우 가맹점 회원용 상세 가입 화면에 결혼 유무, 성별, 직업, 관심 사항 등을 입력하고 가입 버튼을 선택한다(b).
- 5) 가맹점 회원이 아닌 경우 비회원용 상세 가입 화면에 주민등록번호, 이름, 주소, 우편번호, 결혼 유무, 성별, 직업, 그리고 관심 사항을 입력하고 가입하기 버튼을 선택한다(b).
- 6) 등록 절차가 완료되었으면 로그인 화면으로 간다(c).
- 7) 로그인 화면에서 E-Mail 주소와 패스워드를 입력하고 로그인 버튼을 선택한다(c).
- 8) 그 결과 허브 사이트 개인화 페이지가 나타난다. 화면에 나타난 자신의 성향과 화면에 나타난 광고가 타당한지 확인해 본다(c).
- 9) 허브 사이트 개인화 페이지에서 CD Shop 광고를

선택한다(c).

10) 그 결과 CD Shop 가맹점 개인화 페이지가 나타난다. 개인화 페이지에 제공된 자신의 특성과 제공된 제 3의 광고 특성이 타당한지 확인한다(d).

11) 가맹점에서 제공하는 상품을 선택한다(d).

12) 가맹점 방문을 종료하고 허브 사이트에 있는 고객 행위 데이터베이스에 저장된 정보를 데이터베이스 스키마를 이용하여 고객 행위에 대한 로그 정보가 저장되었는지 확인해 본다.

4.3 개인화 서비스 선정 기법 성능 분석

고객 성향에 따른 개인화 서비스를 선정하기 위한 기법으로 제안한 벡터 모델에 대한 성능을 비교하기 위한 방법으로 고객 성향에 관계없이 임의로 광고를 선정하는데 걸리는 시간과 벡터 모델을 이용하여 개인화된 광고를 선정하는데 걸리는 시간에 대한 성능 비용(cost)을 이용하여 비교 실험을 수행하였다. 성능 비용이란 고객 성향에 관계없이 임의로 광고를 선정하는데 걸리는 시간을 100이라 가정할 때, 벡터 모델을 이용하여 개인화된 광고를 선정하는데 걸리는 시간이 어느정도 인지 알기 위한 상대적인 비율이다. 시험 환경은 CPU 펜티엄III 550 MHz, 운영체제는 Linux 2.2.9, 데이터베이스 Oracle8 환경에서 자바를 이용하였으며, 벡터 모델과 임의 방식에 대해 각각 사용자수, 광고수의 변화에 따른 평균 응답 시간을 측정하였다.

[그림 13]은 사용자수, 광고수의 증가에 따라 벡터 모델과 임의 방식을 이용하여 광고를 선정하는데 걸리는 시간을 각각 10회씩 수행한 결과에 대한 평균 응답 시간을 측정한 결과이다. [그림 14]는 벡터 모델을 이용하여 개

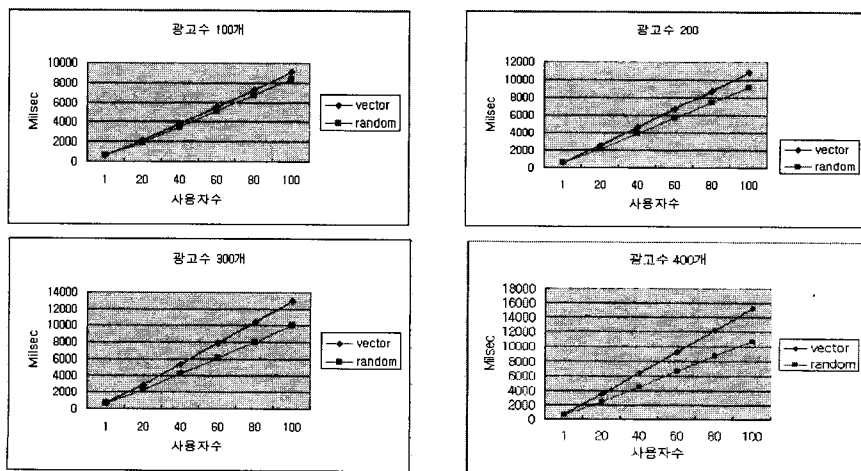


그림 13 사용자수, 광고수에 따른 벡터모델과 임의방식에 대한 응답시간

인화 서비스를 제공하는데 필요한 성능 비용이 임의 방식을 이용할 때 보다 어느정도 더 소요되는지 알아보기 위한 방법으로 사용자수, 광고수에 따른 성능 비용을 측정 한 결과이다. 성능 비용을 계산하는 식은 다음과 같다.

$$C = ((T_v / T_r) - 1) * 100$$

여기서,

C : 성능 비용

T_v : 벡터 모델을 이용하여 개인화 서비스를 선정하는데 소요되는 응답시간

T_r : 임의 방식을 이용하여 서비스를 선정하는데 소요되는 응답시간

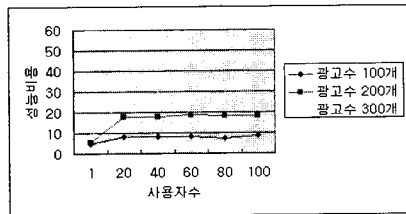


그림 14 개인화 정보 제공에 필요한 성능비용

시험 결과 고객의 특성에 관계없이 무작위로 광고를 선정하는 임의 방식에 비해 사용자수와 광고수가 늘어나면서 벡터 모델에 의한 개인화 서비스 선정 시간이 조금씩 증가하고 있음을 볼 수 있다. 이는 고객에게 개인화 서비스를 제공함으로써 가맹점에 대한 고객의 충성도를 높이기 위한 성능 비용으로 trade-off가 있다. [그림 14]에서 동시 사용자수 약 100명에 200개 정도의 광고를 제공하는 경우 개인화 서비스 제공에 필요한 성능 비용이 임의 방식에 비해 약 18% 정도의 성능 비용이 추가됨을 알 수 있다. 예를 들어 고객의 특성에 관계없이 임의로 광고를 제공하는데 걸리는 시간이 1초이면, 고객의 특성에 따라 개인화 광고를 제공하는데 걸리는 시간은 1.18초 걸린다는 의미이다. 따라서, 고객의 충성도를 높이기 위해 개인화 서비스를 제공하는데 20% 정도의 성능 비용을 감수할 수 있다고 가정하면 동시 사용자수 약 100명에 200개 정도의 광고를 제공하는 규모의 시스템에서는 고객의 특성에 관계없이 임의로 서비스를 제공하는 것보다 벡터 모델을 이용하여 실시간으로 고객이 관심을 가질만한 개인화 서비스를 제공하는 것이 고객이 방문한 가맹점에 대한 충성도면에서 효과적이라 할 수 있다. 성능 비용은 광고수의 증가에 따라 큰 차이를 보이고 있으므로, 광고수가 많은 큰 규모의 시스템에서는 모든 광고를 개인화 서비스 제공 대상으

로 하지 않고, 클러스터링(clustering)을 통해 적은 단위로 분류함으로써 큰 규모의 시스템에도 벡터 모델을 효율적으로 적용할 수 있을 것이다.

5. 결론 및 향후 계획

본 논문에서는 분산 이기종 환경에서 고객 모니터링을 통한 고객 관심 정보를 얻기 위한 기법과 벡터 모델에 의한 개인화 서비스 선정 기법을 제안하였다. 그리고 서로 다른 이기종 환경에서 독립적으로 운용되고 있는 가맹점 사이트의 기존 플랫폼을 변경하지 않고 개인화 맞춤 광고를 제공하기 위한 시스템 구조와 제공 방법을 구현하여 그 가능성을 확인하였다. 또한, 개인화 광고 선정 기법으로 제안한 벡터 모델에 대한 성능을 임의로 제공하는 기존 방식과 비교 평가하였다.

본 논문에서 제시된 시스템은 첫째, 온라인 상에서 고객의 행위를 모니터링하고 분석하여 고객 프로파일을 각 가맹점 사이트에 생성 관리하는 기능을 제공하며, 둘째, 가맹점 사이트에 대한 상품정보나 광고 등을 허브 사이트에 자동으로 등록하고 이를 허브 사이트를 방문하는 고객들에게 이들의 고객 프로파일을 활용하여 선별적으로 고객의 성향에 맞춰 노출시키는 기능을 제공한다. 셋째는 제3의 광고주로부터 허브 사이트가 수주한 상품 정보나 광고를 타 가맹점 사이트 고객들에게 이들의 고객 프로파일을 활용하여 선별적으로 각 가맹점 사이트의 특정 광고란에 노출시키는 기능을 가진다. 본 논문에서 제시한 맞춤 서비스 제공 시스템은 상점 입장에서 허브 사이트 가맹점이 됨으로써 기존 플랫폼을 바꾸지 않고 통합 서비스가 가능하며, 분산 이기종 환경의 다수 인터넷 쇼핑물 사이트를 대상으로 개인화 서비스 제공이 가능하다. 그리고 고객 입장에서는 한 번 등록으로 다수의 가맹점으로부터 맞춤 정보를 제공받을 수 있다.

향후 연구로는 각 가맹점으로부터 고객 행위를 모니터링하여 수집한 정보와 고객 프로파일 정보를 이용한 통계학적 성향 분석 결과를 고객 특성 항목 각각에 가중치를 부여하기 위한 가중치 계산 방법과 실시간으로 규칙 매칭(rule matching)을 위한 연구와 고객 정보 및 고객 행위 모니터링 정보 등과 같은 데이터를 저장하거나 교환 시 표준화를 위해 W3C에서 표준으로 제안하고 있는 XML(eXtensible Markup Language)을 적용하기 위한 연구가 필요하다.

참고 문헌

- [1] Mobasher, B., et al., Automatic Personalization

- Based on Web Usage Mining, *Comm. of the ACM*, Vol 43(8), Aug. 2000.
- [2] G.-H. Yan, et al., Toolkits for a Distributed, Agent-Based Web Commerce System, *In International IFIP Working Conference on Trends in Distributed Systems for Electronic Commerce(TrEC '98)*, Hamburg, Germany, June 1998.
 - [3] Allen Cliff, et al., *Internet World Guide to One-to-One Web Marketing*, John Wiley & Sons, Inc., 1998.
 - [4] 통계청 서비스업통계과, 전자상거래 통계조사 결과 2000년 7, 8월 사이버쇼핑몰조사, 통계청, 2000. 10, <http://www.nso.go.kr/report/data/suec0010.htm>
 - [5] Udi Manber et al., Experience with Personalization on Yahoo!, *Comm. of the ACM*, Vol. 43(8), Aug. 2000.
 - [6] 아마존, <http://www.amazon.com>
 - [7] CDNOW, <http://www.cdnnow.com>
 - [8] Wholesale Hub, <http://www.wholesalehub.com>
 - [9] Maurice D. Mulvenna et al., Personalization on the Net using Web Mining, *Comm. of the ACM*, Vol. 43(8), Aug. 2000.
 - [10] J. Ben Schafer, Joseph Konstan, John Riedl, Recommender Systems in E-Commerce, *In Proceedings of ACM E-Commerce*, 1999.
 - [11] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, Analysis of Recommendation Algorithms for E-Commerce, *In proceedings of ACM E-Commerce*, 2000.
 - [12] Myra Spiliopoulou, Web Usage Mining for Web Site Evaluation, *Comm. of the ACM*, Vol. 43(8), Aug. 2000.
 - [13] Gediminas Adomavicius, Alexander Tuzhilin, User Profiling in Personalization Applications through Rule Discovery and Validation, *Conference on Knowledge Discovery in Data*, Aug. 1999.
 - [14] Good, N., Schafer, B., Konstan, J., Borchers, A., Sarwar, B., Herlocker, J., and Riedl, J., Combining Collaborative Filtering with Personal Agents for Better Recommendations, *In Proceedings of the AAAI-99 conference*, 1999.
 - [15] Ee-Peng Lim, Wee-Keong Ng, An Overview of the Agent-Based Electronic Commerce System (ABECOS) Project, *Bulletin of the Technical Committee on Data Engineering*, Vol. 23(1), Mar. 2000.
 - [16] 온넷시스템, <http://www.onbit.com>
 - [17] 마이 야후!, <http://www.myyahoo.com>
 - [18] Music Machines, <http://machines.hyperreal.org>
 - [19] Broadvision, <http://www.broadvision.com>
 - [20] Shardanand, U. and Maes, P. Social information filtering: algorithms for automating word of mouth., *In proceedings of the ACM CHI Conference*, 1995.
 - [21] Net Perceptions, <http://www.netperceptions.com>
 - [22] Alper Caglayan et al., Learn Sesame a Learning Agent Engine *Applied Artificial Intelligence*, Vol. 11, pp. 393-412. 1997.
 - [23] Michelle A. Hoyle and Christopher Lueg, Open Sesame!: A Look at Personal Assistants, *Proceedings of the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM 97)*, April. 1997.
 - [24] Robert Armstrong et al., WebWatcher: A Learning Apprentice for the World Wide Web, *AAAI Spring Symposium on Information Gathering*, March. 1997.
 - [25] J.Konstan et al., GroupLens:Applying collaborative filtering to Usenet news, *Communications of the ACM*, 40(3), pp. 77-87, 1997.
 - [26] Bamshad Mobasher et al., Discovery of aggregate usage profile for web personalization, *In Proceedings of the WebKDD Workshop*, 2000.
 - [27] Lyle H et al., Clustering methods for collaborative filtering, *In Workshop on Recommendation Systems at the 15th National Conference on Artificial Intelligence*, 1998.
 - [28] 김종우, 이경미, 김영국, 유관중, 인터넷 상점에서의 실시간 개인화된 광고 제공 기법, *경영정보학연구* 제9권 제4호, 1999년 12월.
 - [29] 박성준, 김주연, 김영국, 인터넷 쇼핑물 허브 사이트를 위한 개인화된 맞춤 서비스 제공 시스템, *한국정보과학회 추계학술발표논문집* 제27권 2호, 2000년 10월.
 - [30] Karl Avedal et al., *Professional JSP*, Wrox Press, 2000.
 - [31] Marty Hall, *Core Servlets and JavaServer Pages*, Prentice-Hall, Inc., 2000.
 - [32] 윤혜숙, 김영국, 최훈, 홍승재, 복수의 인터넷 쇼핑물에서 통용되는 안전한 전자상거래 지불수단으로서 로열티 시스템, *한국정보처리학회 논문지* 제7권 제5호, 2000년 5월.
 - [33] Cynthia Bloch and Stephanie Bodoff, The Java Tutorial Continued: Servlets, <http://java.sun.com/docs/books/tutorial/servlets/index.html>
 - [34] Ricardo Baeza-Yates and Berthier Ribeiro-Neto, *Modern Information Retrieval*, Addison Wesley Ltd., 1999.
 - [35] Sander M. Bohte et al., On Current Technology for Information Filtering and User Profiling in Agent-Based Systems, Part I: A Perspective, short reports on Current Tech. for information Filtering and User Profiling in Agent-Based Systems, Jan., 2000.



박성준

1985년 동국대학교 통계학과 학사. 1987년 동국대학교 통계학과 석사. 2001년 충남대학교 컴퓨터과학과 석사. 2001년 ~ 현재 충남대학교 컴퓨터과학과 박사과정. 1989년 ~ 1998년 한국전자통신연구원 선임연구원. 2002년 ~ 현재 공주영상정보대학 전임강사. 관심분야는 개인화, 웹 에이전트, 데이터마이닝, 분산정보시스템



김주연

2000년 충남대학교 컴퓨터과학과 학사. 2002년 충남대학교 컴퓨터과학과 석사. 2002년 ~ 현재 (주)베리텍. 관심분야는 지능형 네트워크, 전자상거래



김영국

1985년 서울대학교 계산통계학과 졸업. 1987년 서울대학교 계산통계학과 석사. 1995년 버지니아대학교 컴퓨터과학과 박사. 1995년 VTT (Technical Research Centre of Finland) 방문연구원. 1995년 SINTEF Telecom & Informatics, Norway 방문연구원. 1996년 ~ 현재 충남대학교 정보통신공학부 부교수. 관심분야는 실시간데이터베이스시스템, 전자상거래시스템, 분산정보시스템 등