

개인화 추천시스템을 위한 효율적 연관 규칙 방법

고병진*, 유영훈*, 조근식*

*인하대학교 전자계산학과

e-mail:{kbj111,yhyu}@eslab.inha.ac.kr , gsjo@inha.ac.kr

Effective Association Rule Method for Personalized Recommender System

Byoung-Jin Ko*, Young-Hoon Yu*, Geun-Sik Jo*

*Dept. of Computer Science & Engineering, Inha University

요 약

인터넷 특성상 방대한 양의 정보와 상품 등으로 사용자들이 원하는 정보를 찾기 위해서 많은 시간을 낭비하고 있는 실정이다. 이러한 사용자의 시간 소모를 줄이기 위해서 추천 시스템이 개발되었다. 현재 인터넷 상의 추천 기술 중에서 가장 많이 사용하는 기법으로는 협력적 여과(Collaborative filtering) 방법이다. 그러나, 협력적 추천 방법으로 추천 받기 위해서는 특정수 이상의 아이টে에 대한 평가가 필요하며, 또한 비슷한 성향을 가지는 일부 사용자 정보에 근거하여 추천함으로써 나머지 사용자 정보를 무시하는 경향이 있다. 이러한 문제점이 발생되므로 최근에는 데이터 마이닝(Data Mining) 기법 중 연관 규칙(Association Rule)을 이용한 추천 시스템이 개발되고 있다[1,10].

그러나, 연관 규칙 기법은 개인별 사용자의 성향을 반영하지 못하는 단점이 있다[4]. 연관 규칙은 단지 대용량 데이터 베이스에서 아이টে간의 지지도(Support)와 신뢰도(Confidence)에 근거하여 규칙을 발견하는 특징을 가지고 있기 때문이다. 즉 개인성향을 무시하고 아이টে간의 연관성만을 근거로 하여 아이টে를 추천하기 때문이다. 본 논문에서는 효율적인 연관 규칙을 이용한 개인화 추천 시스템을 구현하기 위해서 연관 규칙과 협력적 여과 방법을 통합한 시스템을 제안한다. 본 시스템과 기존 추천 시스템에 대하여 성능 비교 실험을 수행함으로써 제안한 방법의 타당성을 제시한다.

1. 서 론

컴퓨터 네트워크 발달과 더불어 전자상거래의 급성장에 따라 온라인 구매의 비중이 급속도로 증가하고 있으며 이를 효율적으로 추천하기 위한 시스템에 대한 연구가 활발히 이루어지고 있다.

추천을 위한 방법으로 많은 알고리즘들이 연구되어 왔으며 그 대표적인 예가 협력적 여과(collaborative filtering), 혹은 사회적인 정보 여과(social information filtering) 라는 방법이 제안되었다[3].

그러나 이러한 협력적 여과 방법의 문제점은 사용자가 평가한 아이টে의 집합이 전체 아이টে 집합에 비해서 극히 작으면 유사한 사용자를 찾기 위해서는 강제적으로 임계치 이상의 아이টে에 대해서 평가해야 한다는 것이다. 또 다른 문제점으로는 협력적 여과 방식은 추천 대상자(Target User)와 유사한 히스토리(History)를 가지는 특정 n명의 이웃이 가진 정보를 바탕으로 추천을 해주므로 국소적 추천(local recommendation)에 머물고 나머지 이웃들에게서 이끌어 낼 수 있는 전역적 추천(global recommendation)을 놓칠 수 있다. 또한 협력적 여과 방법은 대용량의 데이터 정보를 가지고 추천하기에는 아주 많은 시간과 정확도의 문제점이 나타나고 있다[4].

이러한 협력적 여과 방식의 단점들을 보완하기 위해 데이터 마이닝의 연관 규칙 방법을 이용한 추천 시스템이 연구되고 있다[1,8].

그러나, 연관 규칙은 단순히 아이টে간의 표면적인 연관관계에 의해 추천되기 때문에 개인 성향이 반영이 되지 않는다.

따라서 본 논문에서 제시하는 시스템은 효율적인 연관 규칙 방법을 이용하여 개인 성향을 반영한 추천 시스템을

구축함으로써 기존의 추천 시스템들보다 성능을 향상시키고 있다.

본 논문의 구성은 다음과 같다. 2장에서는 연관 규칙과 협력적 여과 방법에 대한 관련 연구를 분석하고, 3장에서는 본 논문에서 제시한 ACR 시스템을 설명한다. 4장에서는 성능 평가를 위한 실험 환경 및 실험 결과 분석을 기술한다. 끝으로 5장에서는 결론을 맺는다.

2. 관련 연구

2.1 연관 규칙 방법(Association Rule Method)

연관 규칙(Association Rule)은 한 항목 그룹과 다른 항목 그룹 사이에 존재하는 연관성을 규칙의 형태로 표현한 것이다[1,2]. 연관 규칙 탐사는 사용자에게 적절하게 입력된 지지도, 신뢰도라는 척도를 이용하여 데이터 상호간의 연관성을 파악할 수 있다. 연관 규칙은 일반적으로 다음과 같이 기술되어질 수 있다. $I = \{ i_1, i_2, i_3, i_4, \dots \}$ 를 아이টে(Items)들의 집합이라고 하고 D 를 트랜잭션(Transaction)들의 집합이라 하자. 그러면 각 레코드는 I 에 속하는 아이টে들의 부분집합으로 구성되어진다. 이때 연관 규칙은 다음과 같은 형태로 기술되어 질 수 있다.

$$X \rightarrow Y (X \subset I, Y \subset I, \text{ AND } X \cap Y = \emptyset)$$

여기서 X, Y 는 각각 연관 규칙의 바디(Body), 헤드(Head) 라고 불린다. 연관 규칙이 내포하고 있는 의미는 트랜잭션 내에 아이টে X 가 존재하면 어느 정도에 확률을 가지고 Y 역시 트랜잭션 내에 존재함을 의미한다. 이들 연관 규칙은 주어진 트랜잭션들의 집합과 관련해 지지도

와 신뢰도라는 두 가지 측정값이 존재한다. 신뢰도는 아이터밋집합 X를 포함하고 있는 트랜잭션 중 아이터밋집합 Y역시 포함하고 있는 트랜잭션의 비율이며, 지지도는 모든 트랜잭션에 대해 아이터밋집합 X, Y를 둘 다 포함하고 있는 트랜잭션의 비율을 나타낸다. 다른 말로 말해서 연관 규칙의 신뢰도는 아이터밋 사이에 연관 정도(Correlation)를 나타내며 지지도는 아이터밋 사이에 연관 정도의 중요도를 나타낸다. 이러한 연관 규칙은 주로 장바구니 분석(Market Basket Analysis)에 주로 사용되고, 또한 구매도 분석, 상품의 교차 매매(Cross-Marketing), 카탈로그 디자인, 열가 매출품(Loss Leader)분석, 상품진열, 구매 성향에 따른 고객 분류 등등 다양하게 사용된다[2].

2.1.1 Apriori 알고리즘

연관 규칙 발견 알고리즘 중 하나인 Apriori는 지지도를 이용하여 동시에 자주 나타나는 항목(빈발 항목 집합)들을 정제하고 빈발 항목 집합에서 생성된 규칙들은 신뢰도를 이용하여 정제하는 방식이다. Apriori는 후보 항목 집합에서 각각의 지지도를 계산한 후 사용자가 정의한 지지도보다 크거나 같은 조건을 만족하는 데이터로 빈발 항목 집합을 구성한다. 그리고 후보 항목 집합은 전 단계의 빈발 항목 집합의 조인연산을 통해 구성된다. Apriori 알고리즘은 효율을 높이기 위해 AprioriTid, AprioriHybrid 등으로 확장되어 연구되고 있다[2].

2.2 협력적 여과

협력적 여과 방식(Collaborative Filtering)은 오늘날 웹 상에 대부분의 성공적인 추천 시스템에 쓰이는 대표적인 추천 기술이다. 협력적 여과 방식은 다른 사용자들의 선호도 정보를 바탕으로 해서 유사한 성향을 가지는 이웃 사용자를 찾고, 그 이웃 사용자에 의해 높은 선호도를 보인 구매 아이터밋을 사용자에게 추천하는 방식이다. 이들 시스템은 추천 대상자(Target User)와 유사한 히스토리 정보를 가지는 이웃들의 집합을 찾기 위해 통계적 방법을 사용한다. 우선 이런 유사한 이웃들이 찾아지면 특정 아이터밋을 추천 대상자(Target User)에게 추천해주기 위하여 여러 알고리즘들을 사용하게 된다[5,7].

2.2.1 GruopLens

초기에 유즈넷 뉴스를 대상으로 개발되었으며 (식1)의 Pearson 상관관계 계수를 통하여 사용자 a와 사용자 k사이의 유사성의 정도를 결정하고 이렇게 계산된 유사도를 이용하여 (식 2)와 같이 사용자 a의 정보 y에 대한 선호도의 예측을 하게 된다[5,7]. 여기서 j는 사용자 a와 사용자 k가 모두 선호를 매긴 정보를 의미하고 $r_{x,j}$ 는 사용자 x의 정보 y에 대한 선호도를 나타내며 \bar{r}_x 는 사용자 x의 전체 정보에 대한 평균 선호도이다.

$$w_{a,k} = \frac{\sum_j (r_{a,j} - \bar{r}_a)(r_{k,j} - \bar{r}_k)}{\sqrt{\sum_j (r_{a,j} - \bar{r}_a)^2 + \sum_j (r_{k,j} - \bar{r}_k)^2}} \quad (1)$$

$$\hat{p}_{a,j} = \bar{r}_a + \frac{\sum_i w_{a,i}(v_{i,j} - \bar{r}_i)}{\sum_i w_{a,i}} \quad (2)$$

2.2.2 Ringo

Ringo[6]는 음반과 서적 등을 추천하기 위하여 개발되었으며 Pearson 알고리즘을 사용하여 현 사용자와 기존 사용자들간의 유사도를 결정하고 후 유사도가 일정 임계치(Threshold)를 넘는 기존 사용자들만 현 사용자의 이웃(neighbors)으로 인정하고 이러한 이웃들만 (식 2)에 적용하여 선호도 예측을 한다.

2.2.3 best n-neighbors

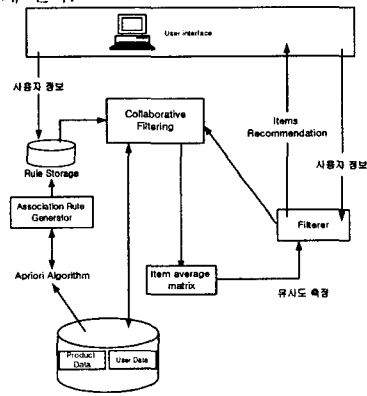
이 방법은 Pearson 알고리즘으로 현 사용자와 기존 사용자들 사이에 유사도를 결정하고 후 유사도가 가장 큰 n명의 사용자들만 현 사용자의 이웃으로 인정하고 (식 2)를

통하여 선호도를 예측을 한다. 특히 기존 사용자들을 모두 이웃으로 인정하는 경우는 Pearson All but 1으로 지정한다.

3. ACR(Association-Collaborative Recommender) 시스템

본 연구에서 제시한 추천 시스템은 사용자들의 히스토리 데이터를 Apriori 알고리즘을 사용하여 아이터밋간의 연관성을 발견하고 연관성 중에서 사용자에게 추천될 규칙을 추출한다. 추출된 규칙에 부합되는 기존 사용자들의 등록 정보와 사용자의 등록 정보를 협력적 여과 방법을 사용하여 유사성을 분석한다. 분석된 값들 중에서 임계치보다 큰 기존 사용자들의 등록정보를 평균값으로 변화한다. 등록정보의 평균과 사용자의 등록정보의 유사도를 계산해서 유사도가 큰 순서대로 아이터밋을 추천한다.

(그림 1)는 시스템 전체 구조를 나타내고 있다. 연관 규칙과 협력적 여과 방법을 이용한 추천 시스템은 데이터 베이스에서 아이터밋간의 연관성 규칙을 발견하고 발견된 규칙은 룰 저장소에 저장된다. 만약 룰 저장소에 A → B (support: 55%, confidence: 71%)이 발견 되었고 추천을 원하는 사용자가 A를 봤다면 일반적인 연관 규칙은 B를 추천하게 된다. 그러나, 본 논문의 시스템은 B를 추천하는 것이 아니고 A → B를 본 기존 사용자들의 등록정보를 가지고 추천을 원하는 사용자의 유사성을 분석한다. 분석된 각각의 유사성 중에서 임의의 임계치 이상인 것을 통합한다. 통합한 정보의 평균값과 사용자의 등록정보의 유사성을 분석하고 유사성이 가장 높은 아이터밋 순서로 추천을 하게 된다.



(그림1) ACR 시스템 구조

3.1 연관 규칙 생성

본 연구의 ACR 시스템은 사용자들의 히스토리 정보를 가지고 연관 규칙을 생성하기 위해 연관 규칙 생성기(Association Rule Generator) 모듈을 만들었다. 연관 규칙 생성은 Apriori Algorithm[2]을 사용하였으며 외부로부터 적당한 지지도와 신뢰도 값을 입력으로 받는다. 연관 규칙 생성기는 입력받은 지지도와 신뢰도를 만족시키는 규칙들을 룰 저장소(Rule Storage)에 저장하며, 저장된 규칙들을 이용하여 사용자에게 아이터밋을 추천하는데 이용한다.

본 ACR 시스템에서는 연관 규칙을 생성하는데 있어서 사용자 입력에 대해 실시간으로 룰을 생성하지 않으며 주기적으로 데이터 베이스로부터 연관 규칙을 생성하여 룰 저장소에 업데이트를 시킨다.

3.2 협력적 여과

협력적 여과 방법은 룰 저장소에서 사용자에게 추천될 아이터밋을 추출한다. 아이터밋을 이용한 기존 사용자들의 등록 정보와 사용자의 정보를 협력적 여과 방법을 통해 유사성을 계산한다. 유사성이 임계치 초과하는 사용자들을 통합하여 평균을 구한다. 기존 사용자들의 평균 정보와

사용자의 유사도를 계산한 후 가장 큰 유사도 값을 나타내는 아이템을 사용자에게 추천한다.

3.3 유사도 측정 방법

사용자의 유사도를 측정하기 위해서는 유사 매트릭스 또는 여러 차원으로 유사도를 판단할 수 있는 방법이 있어야 한다[9]. 본 논문에서는 사용자로부터 추출된 등록 정보와 협력적 여과 방법을 통해 추출된 기존 사용자들의 평균 정보를 가지고 다음과 같은 식에 의해 계산된다.

$$RI_i = CY_i + CS_i - \epsilon I_i + \dots + C5_i + CB_i \quad (3)$$

- RI_i: 추천 아이템 유사도의 합
- CY_i: 나이의 유사도
- CS_i: 학력의 유사도
- CL_k_i: 장르별 관심도의 유사도
- CB_i: 사용자의 선정 값

(식3)에서 구한 RI의 값이 가장 큰 값부터 사용자에게 추천을 하게 된다. 나이, 학력, 장르별 관심도, 사례 선정 값들의 유사도는 다음과 같은 식에 의해 계산을 하게 된다.

$$IF |Y - AY_i| \leq 20 \quad Then \quad CY_i = W_y - \frac{|Y - AY_i|}{2}$$

$$ELSE |Y - AY_i| > 20 \quad CY_i = 0$$

$$CS_i = W_s \times (1 - \frac{|S_i - AS_i|}{3})$$

$$CI_{1_i} = W_1 \times (1 - \frac{|I_{1_i} - AI_{1_i}|}{5})$$

$$CI_{2_i} = W_2 \times (1 - \frac{|I_{2_i} - AI_{2_i}|}{5})$$

:

$$CI_{5_i} = W_5 \times (1 - \frac{|I_{5_i} - AI_{5_i}|}{5})$$

$$CB_i = \frac{TC_i}{SMax} \times 10$$

I_k: 사용자 장르별 관심도

AI_k: 평균 장르별 관심도

W_y: 나이 가중치

W_s: 학력 가중치

W_i: 관심도 가중치

SMax: 성공횟수 최대값

각각의 가중치값은 다음과 같이 설정하였다.

$$W_y = 10 \quad W_s = 3 \quad W_i = 5$$

$$SMax = 100$$

4. 실험 및 평가

4.1 실험 환경

본 논문에서 제시한 ACR 시스템을 구현하기 위해 연관 규칙 및 협력적 여과 방법은 C++로 구현했다. 시스템의 환경은 Windows 2000 server에서 MS-SQL 2000 서버와 Microsoft Visual Studio 6.0 환경하에서 실험했다. 실험 데이터는 현재 온라인 만화 서비스를 하고 있는 한 회사의 실 데이터를 사용했다. 이 데이터는 230명의 사용자 고객 정보와 각각의 장르별 관심도, 그리고 사용자들의 히스토리 데이터와 1392권의 만화 데이터 정보로 구성되어 있다. 사용자의 관심도는 사용자가 동적으로 수정이 가능하게 만들었다.

4.2 실험 방법

4.2.1 연관 규칙

연관 규칙은 지지도와 선호도를 바탕으로 아이템들의 표면적 연관성을 발견되기 때문에 너무 작은 값을 지정하게 되면 룰 베이스에 저장되는 규칙들이 많아진다. 이러한 문제로 본 실험은 지지도는 10%부터 25%까지 5%씩 증가하고 선호도 값은 40%의 값으로 정의했다. 이렇게

정의된 지지도와 선호도를 바탕으로 연관 규칙을 발견하기 위해서 먼저 데이터 베이스 안에 있는 사용자의 히스토리 데이터를 전처리 과정을 거쳐서 flat파일로 작성하였다.

(그림3) flat 파일

작성된 flat 파일을 가지고 연관 규칙방문을 적용하게 되면 다음 (그림4) 같은 규칙을 생성하게 되며 생성된 규칙은 룰 베이스에 저장된다.

(그림4) 연관 규칙 추출

4.2.2 협력적 여과

연관 규칙을 통해 추천을 원하는 사용자의 추천 아이템을 결정하고, 결정된 아이템들의 사용자들을 2.2.1절의 (식1) 사용하여 다른 사용자들간의 유사성을 측정한다. 측정된 유사성의 값이 임계치 값 보다 큰 사용자를 이웃으로 인정하고 이웃으로 인정된 사용자들에 대한 각각의 나이, 학력, 장르별 관심도의 평균을 구한다. 결과는 4.2.3절의 <표2>와 같이 표현된다.

4.2.3 유사도 측정

유사도 측정은 사용자의 나이와 학력 그리고 장르별 관심도를 바탕으로 유사도를 측정하게 된다. 여기서 학력에 대한 지식을 표현하기 위해서 대학생:1, 대학생:2, 고등학생:3로 정의 하였다.

나이 (Y)	학력 (s)	학원만화 관심도(I11)	스포츠만화 관심도(I12)	액션만화 관심도(I13)	공포만화 관심도(I14)	무협만화 관심도(I15)
20	1	2	5	3	1	5

<표1> 사용자 매트릭스

<표1>과 같은 사용자 매트릭스가 구성된다. 사용자 매트릭스가 완성되면 협력적 여과 방법을 이용하여 나타난 사용자간의 유사성의 평균들을 <표2>와 같은 평균 매트릭스로 나타낼 수 있다.

학력 (AS)	나이 (AY)	학원 (AI1)	스포츠 (AI2)	액션 (AI3)	공포 (AI4)	무협 (AI5)	만화 코드	사용 자수 (IC)
1	15	1	3	2	1	5	102	55
2	20	2	5	2	4	5	204	60
1	25	5	1	2	3	4	150	45
3	20	2	4	3	4	5	520	50

<표2> 아이템별 평균 매트릭스

유사도는 3.3의 (식3)을 이용해서 구한 결과 값은 다음 <표3>과 같다.

만화 코드	CS	CY	CL1	CL2	CL3	CL4	CL5	CB	RI
102	3	7.5	4	3	4	5	5	5.5	37
204	2	10	5	5	4	2	5	6	39
150	3	7.5	2	1	4	1	4	4.5	27
520	2	10	5	4	4	1	4	5	35

<표3> 유사도 계산의 예

위의 유사도의 값이 결정되면 가장 큰 값들로 정렬하고 가장 큰 값의 순서로 추천하게 된다. 추천의 개수는 사용자의 정의에 따라 개수를 정의 할 수 있다. 만약 <표3>에

서는 사용자가 추천 아이템을 2개로 정의하게 되면 204와 102의 만화책을 추천하게 된다.

4.3 평가 방법

4.3.1 성능 평가 방법

실험을 위해 우선 전체 실험 자료를 70%, 30%로 임의로 분류하여 70%에 속하는 자료를 학습자료(Training Data)로 사용하고 30%에 속하는 자료를 검사 데이터(Test Data)로 사용하였다[8]. 우리는 기존의 협력적 여과 방식의 추천 시스템과 연관규칙 추천시스템 그리고 ACR 시스템을 각각 비교 분석하여 성능의 차이를 비교하였다. 협력적 여과 방법의 이웃 선택 임계치 값은 0.2에서부터 0.5 까지 변화를 주면서 성능을 분석하고, 연관 규칙 시스템은 지지도를 10%부터 25%까지 변화를 주고 지지도는 40%로 정해져서 성능을 평가했으며, ACR 시스템은 이 두 개의 방법을 통합하여 성능을 분석하였다. 각각의 시스템에서 추천되어질 아이템 수(top-N)를 20개의 아이템으로 제한하였다.

본 논문에서는 추천되어진 top 아이템 N개(top-N)가 검사 자료 내에서 얼마나 정확하게 예측되어졌나를 알아보기 위해 정보 검색(Information Retrieval)에서 사용되는 재현율(Recall), 정확도(Precision)을 사용하였다[4].

Recall : 검사데이터(testset)에 대한 매칭 데이터 (hit set) 의 비율로 다음과 같다.

$$Recall = \frac{test \cap top - M}{test} \dots \text{식}(4)$$

Precision : 추천 대상자에게 추천되어진 베스트 아이템 N개에 대한 매칭 데이터의 비율로 다음과 같다.

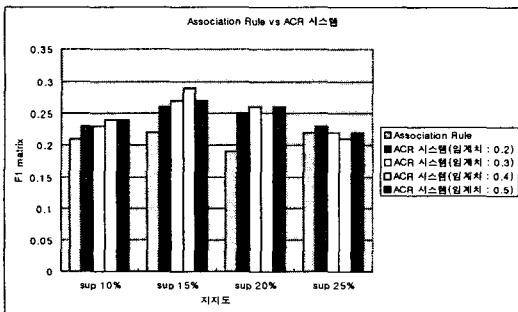
$$Precision = \frac{test \cap top - M}{N} \dots \text{식}(5)$$

그러나 위 측정 방법은 서로 반비례 관계가 존재한다. 예를 들면 N이 증가하면 Recall은 증가하고 Precision은 감소하는 경향이 있다. 그러므로 Recall과 Precision에 동일한 가중치를 주어 하나로 통합한 standard F1 metric 방식을 사용한다.[4]

$$F1 = \frac{2 * Recall * Precision}{Recall + Precision} \dots \text{식}(6)$$

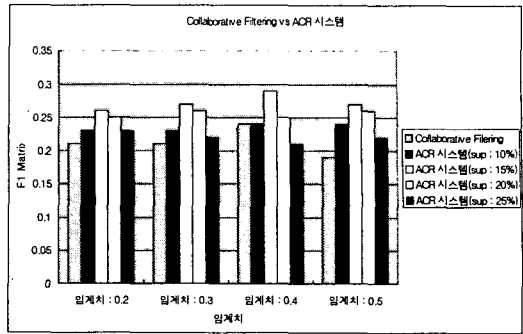
4.3.2 성능 평가

(그림5)는 연관 규칙과 ACR 시스템의 성능을 각각 분석한 결과이며, (그림6)은 협력적 여과 방법과 ACR 시스템의 성능을 분석한 결과 있다. (그림5,6)에서 보면 본 ACR 시스템의 성능이 현재의 추천시스템들 보다 향상된 결과를 나타내고 있다.



(그림5) Association Rule vs ACR 시스템

(그림5)는 기존의 연관 규칙 추천시스템의 선호도를 5%씩 증가시키면서 실험한 결과이다. ACR 시스템은 연관 규칙의 선호도와 임계치 변화에 대해서 실험을 했다.



(그림6) Collaborative Filtering vs ACR 시스템

(그림6)는 기존의 협력적 여과 방법 추천시스템의 임계치를 0.2부터 0.5까지 0.1씩 증가시키면서 실험한 결과이다. ACR 시스템은 연관 규칙의 선호도와 임계치 변화에 대해서 실험을 했다. (그림5,6)의 실험결과를 보면 ACR 시스템의 임계치 값이 0.4와 지지도가 15%일 때 성능이 극대화가 된다.

5. 결론

본 논문은 연관규칙의 단점을 보완하기 위해 연관 규칙 방법과 협력적 여과 방법을 사용하였다. 연관 규칙은 사용자의 선호도를 반영하지 못하므로 협력적 여과 방법을 사용 사용자의 선호도가 반영될 수 있도록 시스템을 설계하였다. 제안한 방법에 대한 성능을 기존의 추천 시스템과 비교 실험한 결과 정확도가 향상된 것을 알 수 있다.

참고 문헌

- [1] Weiyang Lin, Sergio A. Alvarez, and Carolina Ruiz, " Collaborative Recommendation via Adaptive Association Rule Mining" WEBKDD 2000, Boston
- [2] Jiawei Han, Micheline Kamber " Data mining : Concepts and Techniques", Morgan Kaufmann, 2001
- [3] X. Fu, J. Budzik, and K. J. Hammond "Mining navigation history for Recommendation" In Proc. 2000 interational conf. intelligent user interfaces, pages 106-112, 2000
- [4] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl " Analysis of e-commendation Algorithms for E-Commerce" Minnesota, 2000
- [5] Resnik, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. "GroupLens: An Open Architecture for Collaborative Filtering of Netnews." In Proceedings of CSCW, pgs 175-186, 1994
- [6] John S. Breese, David Heckerman, and Carl Kadie "Empirical analysis of predictive algorithms for collaborative filtering." Proceeding of the 14th Conference on Uncertainty in Artificial Intelligence, pages.43-52, 1998
- [7] Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., and Riedl, J., "GroupLens: Applying Collaborative Filtering to Usenet News." Communications of the ACM, 40(3), pages. 77-87, 1997
- [8] Bamshad Mobasher, Honghua Dai, " Effective Personalization Based on Association Rule Discovery from Web Usage Data", 3rd ACM, 2001
- [9] 성백균, 김상희, 박덕원, "전자상거래를 위한 사례기반의 판매지원 에이전트", 정보처리학회 제 7권 pp.1649-1656, 2000.5
- [10] 이기현, 고병진, 조근식, "연관규칙과 협력적 여과 방식을 이용한 추천시스템", 지능형 정보시스템 학회 춘계, 2002.5