

A Study on the Maximizing Coverage for Recommender System

Hee Choon Lee¹⁾, Seok Jun Lee^{2)*}, Ji Won Park³⁾, Chul Seung Kim⁴⁾

Abstract

The similarity weight, the pearson's correlation coefficient, which is used in the recommender system has a weak point that it cannot predict all of the prediction value. The similarity weight, the vector similarity, has a weak point of the high MAE although the prediction coverage using the vector similarity is higher than that using the pearson's correlation coefficient. The purpose of this study is to suggest how to raise the prediction coverage. Also, the MAE using the suggested method in this study was compared both with the MAE using the pearson's correlation coefficient and with the MAE using the vector similarity, so was the prediction coverage. As a result, it was found that the low of the MAE in the case of using the suggested method was higher than that using the pearson's correlation coefficient. However, it was also shown that it was lower than that using the vector similarity. In terms of the prediction coverage, when the suggested method was compared with two similarity weights as I mentioned above, it was found that its prediction coverage was higher than that pearson's correlation coefficient as well as vector similarity.

Keywords : collaborative filtering, , MAE, pearson's correlation coefficient, recommender system, vector similarity

1. 서론

IT기술의 발전으로 전자상거래는 상품(아이템)거래의 주요한 수단으로 등장하고 있다. e-business, m-commerce 등의 등장으로 추천시스템은 전자상거래의 상품추천을

-
- 1) 강원도 원주시 우산동 660번지 상지대학교 컴퓨터데이터정보학과 교수
E-mail : choolee@sangji.ac.kr
 - 2) 강원도 원주시 우산동 660번지 상지대학교 경영학 박사과정
E-mail : crco909@yahoo.co.kr
 - 3) 경상북도 김천시 삼락동 754번지 김천대학교 병원의료행정과 교수
E-mail: jwpark@gimcheon.ac.kr
 - 4) 인천광역시 남구 도화동 177번지 인천대학교 경영학 박사과정
E-mail: fewinanalyst@kapi.re.kr

위한 방법으로 매우 유용한 방법으로 인식이 되고 있다. 또한 Digital TV의 출현의 또 다른 상거래 방법이 될 것이며 어떤 형태이든지 추천시스템이 적용될 것이다..추천 시스템을 이미 적용한 사례로 추천시스템의 활용에 대한 정량적인 연구는 Konstan과 Riedl(Konstan and Reidl (2000))의 연구에서 NetPerceptions 추천시스템을 활용하여 평균 60% 향상된 교차판매(cross-sell)의 효과를 얻었고 영국의 GUSpls(www.gus.co.uk)사의 기법을 근간으로 하는 기존의 교차판매 기법보다 50% 향상된 성과를 거두었음을 보여주고 있다. 고객의 측면에서 추천 서비스는 인터넷에서 방대한 제품 중 구매하고자 하는 제품의 선택에 드는 탐색비용을 크게 줄일 수 있으며 특히, 인터넷에서의 전자상거래에서 더욱 큰 효과를 발휘한다(예; Netflix는 25,000편 이상의 영화를 제공하고, Amazon에서는 2백만권 이상의 책을 제공하고, eBay의 경우 매일 수백만건의 새로운 경매목록이 갱신된다.). 효과적인 추천은 고객의 쇼핑 경험을 향상시킬 수 있으며 고객의 수요증대와 충성도 향상을 가능케 할 것이다. 최근 연구에서는 Amazon의 음악 추천시스템에 의해 추천을 받은 20%의 제품이 구매되었다는 것을 보여주고 있다(Swearigen and Sinha (2002)).

2. 연구의 필요성

과거의 offline business에서 e-business를 가능하게 한 것은 IT기술의 발전에 의한 것이었다. 이제 머지않아 새로운 m-commerce(mobile commerce)에 의한 상거래가 이루어질 것이다. 또한 Digital TV의 출현은 또 다른 상거래의 수단으로 등장할 것으로 기대된다. 과거의 offline 거래에서 상품 구매는 구매자의 선택이나 판매원의 추천에 의해 상품 구매가 이루어졌었다. 판매원에 의한 상품 추천은 판매원이 과거에 인지한 판매정보에 의해 상품이 추천되었다. 그러나 online 거래에서는 구매자의 선택에 의한 선호도 구매는 가능하나 판매원에 의한 상품의 추천이 불가능하다. 추천시스템은 offline 거래에서 판매원에 의한 추천 방법을 과학적인 방법을 사용하여 대신할 수 있다. 상품의 추천은 구매자의 개인정보를 이용하거나 다른 구매자들의 과거 상품구매 정보를 이용하여 상품을 추천하며 구매 희망자가 온라인으로 해당 시스템에 접근(접속)하였을 때 자동적으로 추천하는 방법이다. 또한 구매자의 과거 온라인 거래 구매정보를 이용하여 기존의 상품이나 새로운 상품에 대한 e-mail 추천을 가능하게 하는 방법이다.

Schafer 등 (2001)에서는 추천시스템이 3가지 측면에서 전자상거래에서 판매를 증대시킨다고 하였다.

1)단순 방문자를 구매자로 전환(Converting Browsers into Buyers): 웹사이트에 방문하는 고객들은 때로 구매를 목적으로 방문하는 것이 아니라 단순 방문이 목적인 고객도 있다. 추천시스템은 단순 방문 고객들이 흥미를 유도하는 제품을 찾는데 도움을 줄 수 있다.

2)교차 판매의 증대(Increasing Cross-sell): 추천시스템은 고객이 구매를 원하는 제품과 더불어 추가적으로 제품을 추천하여 교차판매를 증대시킬 수 있다. 만약 추천이 성공적이라면 평균 주문량은 증가할 것이다. 예를 들어 웹사이트에서 결재 중 이미 장바구니에 있는 물품을 기준으로 추가적인 제품을 추천하여 교차 판매를 증대

시킬 수 있다.

3)충성도 확립(Building Loyalty): 추천시스템은 웹사이트와 고객과의 관계를 가치가 부여된 관계로 확대시켜 줌으로써 고객의 충성도를 향상시킬 수 있다. 웹사이트는 고객의 흥미와 선호도에 대해 학습하고 학습된 데이터를 운영하기 위해 추천시스템을 이용하며 고객의 필요에 부합하는 인터페이스를 제공한다. 고객들은 사이트에 가장 자신들의 필요에 부합하는 제품을 추천받기 위하여 투자한다. 고객들이 추천시스템을 이용하면 할수록 시스템에 고객이 원하는 것을 잘 학습시켜주므로 사이트에 더 충성스러운 고객이 된다. 경쟁사가 동일한 능력의 시스템을 설치하더라도 고객들은 경쟁사의 시스템에 기존의 시스템에 투자한 만큼 학습의 시간과 노력을 기울여야 하므로 쉽게 경쟁사로 가지는 않을 것이다.

추천시스템의 연구의 필요성은 IT기술의 발달로 인한 글로벌 마케팅 시대가 도래한 현실에서 필요성을 충분하며 본 연구를 통한 결과가 추천시스템에 적용될 수 있을 것으로 기대된다.

3. 연구 목적

협력적 필터링에서 사용되는 특정 사용자에 대한 선호도는 근접이웃 알고리즘(nearest neighbor algorithm)이 사용되며 유사도 가중치(similarity weight)로는 피어슨상관계수(pearson's correlation coefficient)와 스피어맨의 순위상관계수(spearman's rank correlation), 벡터유사도(vector similarity)등이 있으며 선행연구에 따르면 피어슨상관계수(pearson's correlation coefficient)와 벡터유사도(vector similarity)가 자주 사용 된다

추천시스템에서 유사도 가중치인 피어슨 상관계수는 응답쌍이 1개 이거나 자료가 모두 같은 경우에는 표준편차가 0이 되어 상관계수를 구하지 못하는 경우가 있어 예측 비율(coverage)이 떨어지는 단점을 가지고 있다. 또한 응답쌍이 2개인 경우에는 피어슨 상관계수를 구하였다 하더라도 상관계수를 -1 또는 1로 계산하여 상관계수를 과대평가하는 단점을 가지고 있다. 벡터유사도는 예측비율은 높으나 MAE(mean absolute error)가 크다는 단점을 가지고 있다. 응답쌍이 1개인 경우 피어슨 상관계수는 유사도 가중치를 계산하지 못하나 벡터유사도는 1로 계산하여 유사도 가중치를 과대평가하는 단점을 가지고 있다. 본 연구에서는 예측비율과 MAE를 고려하여 예측비율을 높이며 MAE를 감소시키는 방법을 제안하고 기존의 유사도 가중치인 피어슨 상관계수와 벡터유사도를 사용하였을 때 선호도 예측치의 MAE를 비교 분석 한다.

4. 선행연구

4.1 근접 이웃 알고리즘과 유사도 가중치

GroupLens에서 제시한 사용자 이웃 기반의 예측 알고리즘을 이용한 협력적 필터링은 어떤 아이템에 대한 특정 사용자의 선호도를 예측하기 위하여 식(2)의 피어슨 상관 계수를 이용하여 유사한 선호도를 가지는 이웃들을 정하고 식(1)에 의해 특정 사용자의 선호도 예측을 계산한다.

$$U_x = \bar{U} + \frac{\sum_{J \in \text{Raters}} (J_x - \bar{J})r_{uj}}{\sum_{J \in \text{Raters}} |r_{uj}|} \quad (1)$$

여기서,

$$r_{uj} = \frac{\sum (U - \bar{U})(J - \bar{J})}{\sqrt{\sum (U - \bar{U})^2 \cdot \sum (J - \bar{J})^2}}, \quad -1 \leq r_{uj} \leq 1 \quad (2)$$

U_x 는 아이템 x 에 대한 특정 사용자 u 의 선호도 예측치이고, r_{uj} 는 유사도 가중치로 특정 사용자 u 와 이웃한 사용자 j 의 상관관계를 나타내는 피어슨 상관계수이다. J_x 는 이웃사용자 j 의 아이템 x 에 대한 선호도이고 \bar{J} 는 이웃사용자 j 의 선호도 전체의 평균이다. r_{uj} 가 1에 가까울수록 두 사용자의 선호도 경향이 매우 유사함을 나타내고 -1에 가까울수록 반대의 선호 경향을 나타낸다. Raters는 테스트 상품(아이템)에 대해 선호도를 표시한 사용자들을 의미한다. 협력적 필터링의 많은 연구들이 GroupLens에서 제시한 알고리즘을 이용하여 선호도를 예측하였다. 한편 유사도 가중치인 벡터유사도는 다음 식(3)과 같다.

식 (3)은 사용자 u 와 j 의 관계를 나타내는 코사인 벡터 유사도이다.

$$r_{uj} = \frac{U \cdot J}{|U| \cdot |J|} \quad (3)$$

다음은 GroupLens에서 제시한 문서에 대한 선호도 예측의 sample 예제와 선호도 예측 알고리즘을 이용한 계산의 예제이다(Resnick 등 (1994)).

<표 1> 문서에 대한 각 사용자의 선호도

User Item	Ken	Lee	Meg	Nan
1	1	4	2	2
2	5	2	4	4
3			3	
4	2	5		5
5	4	1		1
6	?	2	5	

$$\begin{aligned}
 K_{6_{pred}} &= \bar{K} + \frac{\sum_{j \in raters} (J_6 - \bar{J}) r_{kj}}{\sum_j |r_{kj}|} \\
 &= 3 + \frac{2r_{km} - r_{kl}}{|r_{km}| + |r_{kl}|} = 3 + \frac{2 - (-0.8)}{|1| + |-0.8|} = 4.56
 \end{aligned}$$

<표 1>에서 Ken, Lee, Meg, Nan은 각각 문서를 접한 사용자가 되고 1열은 각 사용자가 접한 문서, 즉 아이템의 종류가 된다. 예측은 사용자 Ken의 6번째 문서에 대한 선호도를 구하는 것으로 먼저 각 문서에 대한 사용자들의 선호도 평가들의 관계를 각 사용자들의 상관관계를 이용하여 관련성을 구하게 된다. 이때 사용자 간의 상관관계는 피어슨 상관관계수, 코사인 벡터 등이 이용될 수 있으며, GroupLens가 제시한 알고리즘에서는 피어슨 상관관계수가 사용되었다.

여기서 \bar{K} 는 Ken의 선호도 평균이 되고, \bar{J} 는 Ken에 이웃한 Lee, Meg, Nan 각각의 선호도 평균이 된다. r_{uj} 는 Ken과 이웃한 사용자 Lee, Meg, Nan과의 피어슨 상관관계수이다. 즉, r_{km} 은 Ken과 Meg, r_{kl} 은 Ken과 Lee의 피어슨 상관관계수이다. Ken의 선호도 평균과 이웃의 선호도 평균과 상관관계수를 이용하여 Ken의 6번째 아이템에 대한 선호도를 예측하면 4.56으로 구해진다.

4.2 선호도 예측의 평가

선호도 예측 정확도의 평가는 MAE를 사용하였으며 다음 식(4)와 같다. MAE는 선호도 예측치와 실제 평가치 사이의 편차의 절대값의 평균으로 측정한다.

$$MAE = \frac{\sum_{i=1} |\epsilon_i|}{N} \quad 4)$$

여기서, ϵ_i 는 사용자의 실제 선호도 평가치와 선호도 예측치의 편차이다.

4.3 기존 문헌 연구

협력적 필터링은 아이템의 특성이나 사용자의 프로파일과 같은 속성을 의도적으로 무시하고 사용자들이 아이টে에 대해 선호도에 대한 평가, 즉 사용자-아이템 간의 관계 데이터만을 이용하여 예측하는 접근법이다(Hill 등(1995), Resnick 등(1994), Shardanand 와 Maes(1995)). Breese 등(1998)은 사용자들 간의 유사도 가중치를 피어슨 상관계수(Pearson's correlation coefficient), 벡터 유사도(Vector similarity), 사용자가 평가하지 않은 아이টে에를 예측치 계산에서 제외시키지 않고 기본 선호도를(default voting) 부여하여 유사도를 구하는 방법, 많이 검색된 아이টে에를 찾는 것 보다 검색되지 않은 아이টে에를 찾는 방법인 역사용자빈도(Inverse user frequency)와 같은 다양한 유사도 가중치에 대해 연구하였다(Breese 등(1998)). Herlocker 등(1999)은 사용자간의 상관관계를 나타내는 상관계수가 두 사용자가 공통으로 평가한 응답 쌍에 영향을 받고 있음을 연구하였다. Herlocker 등(1999)은 응답쌍의 개수에 따른 영향을 50개로 보고 응답 쌍의 개수가 50보다 적은 상관계수에는 $n/50$ 의 가중치를 부여하고 응답쌍이 50보다 큰 경우 상관계수에 1의 가중치를 부여하여 정확도가 향상된 결과를 보였다. 협력적 필터링은 사용자와의 이웃을 형성하기 위해 사용자들의 관계 데이터를 이용하는 사용자 기반(user-based)의 협력적 필터링(Sarwar 등(1998), Claypool 등(1999))과 반대로 아이টে에의 관계 데이터를 이용하는 아이টে에 기반(item-based)의 협력적 필터링으로 나누어진다(Sarwar 등(2001), Deshpande 와 Karypis(2004)). 많은 전자상거래 사이트에서는 아이টে에 기반의 알고리즘을 이용하고 있다(Schafer 등(2001)). Herlocker 등(1999)은 협력적 필터링에 대한 다양한 평가방법을 제안하였으며 박지선 등(2001)은 협력적 필터링 기법에서 적용하고 있는 피어슨 상관계수를 이용하는 방법에서 비슷한 선호도 패턴을 가지는 사용자들을 적절히 군집화하여 이 군집에 속하는 사용자들의 평가를 기반으로 하여 협력적 필터링 기법을 수행하는 방법을 제안하였다. 김택현 등(2004)은 ordered clustering 방법을 이용하여 EachMovie dataset에 대해 예측의 정확도를 실험하였다. 이희춘(2006a)은 MovieLens dataset을 이용하여 예측의 정확도를 높이기 위해 인구통계변수와 사용자의 응답 쌍의 영향력에 대해 연구하였으며 이희춘(2006b)에서는 개선 알고리즘을 제시하여 예측의 정확도를 높이는 연구를 하였다. 이희춘 등(2006)은 유사도 가중치인 피어슨 상관계수와 벡터 유사도에 응답 쌍의 영향을 세분화시킨 유의성 가중치를 적용하여 예측의 정확도가 향상됨을 연구하였다.

5. 연구방법

본 논문에서 사용된 dataset은 GroupLens의 MovieLens dataset을 이용하여 실험을 하였다. MovieLens dataset은 943명의 평가자이 1682편의 영화에 대해 최소 20편을 평가하였으며 평가점수는 1-5점으로 평가하였다. 1682편의 영화에 943명이 평가한 평가의 수는 100,000개이다. 연구방법은 100,000개의 평가치(rating)에 대해 식(1)의 근접이웃 알고리즘(nearest neighbor algorithm)에 유사도 가중치인 식(2)의 피어슨 상관계수(pearson's coefficient of correlation)를 적용하여 모든 예측치와 MAE구하고, 식(1)의 근접이웃 알고리즘(nearest neighbor algorithm)에 유사도 가

중치인 벡터유사도(vector similarity)를 적용하여 모든 예측치와 MAE를 구하며 또한 제안방법에 의한 모든 예측치와 MAE를 구하여 MAE를 비교하며 또한 평가개수 100000개에 대한 예측개수를 비교한다.

5.1 제안방법

예측비율을 높이기 위한 방법으로 제안한 식은 다음 식(5),식(6),식(7), 같다.

$$\epsilon_i = \min \left(\sum_{i \in Raters} |R_{ij} - \widehat{R}_{ij}(\theta_i)| \right) \quad (5)$$

여기서, $\widehat{R}_{ij}(\theta_i) = \overline{u}_i \cdot \widehat{\theta}_i + \overline{m}_j$, $-\infty \leq \theta_i \leq \infty$,

$$\epsilon_i = \min \left(\sum_{j \in Movies} |R_{ij} - \widehat{R}_{ij}(\theta_j)| \right) \quad (6)$$

여기서, $\widehat{R}_{ij}(\theta_j) = \overline{m}_j \cdot \widehat{\theta}_j + \overline{u}_i$, $-\infty \leq \theta_j \leq \infty$

$$\epsilon_i = m \left[\sum_{i \in Raters} \sum_{j \in Movies} |R_{ij} - \widehat{R}_{ij}(\theta_{i1}, \theta_{i2})| \right] \quad (7)$$

여기서, $\widehat{R}_{ij}(\theta_{i1}, \theta_{i2}) = \overline{u}_i \cdot \widehat{\theta}_{i1} + \overline{m}_j \cdot \widehat{\theta}_{i2}$, $-\infty \leq \theta_{i1} \leq \infty$, $-\infty \leq \theta_{i2} \leq \infty$.

\overline{u}_i 는 사용자 i 의 평가치 평균이며 \overline{m}_j 는 영화 j 의 평가치 평균이고 R_{ij} 는 사용자 i 의 영화 j 에 대한 평가치이다. $\widehat{R}_{ij}(\theta_i)$, $\widehat{R}_{ij}(\theta_j)$, $\widehat{R}_{ij}(\theta_{i1}, \theta_{i2})$ 는 예측치를 의미하며 ϵ_i 는 사용자의 MAE를 의미한다.

즉 식(5)-식(7)은 사용자 i 의 평가치 평균 \overline{u}_i 와 사용자 i 의 평가치 평균 \overline{m}_j 와 $\widehat{\theta}_i$, $\widehat{\theta}_j$, $\widehat{\theta}_{i1}$, $\widehat{\theta}_{i2}$ 들의 1차 결합식과 사용자의 영화에 대한 평가치의 편차가 최소가 되도록 하는 예측치 $\widehat{R}_{ij}(\theta_i)$, $\widehat{R}_{ij}(\theta_j)$, $\widehat{R}_{ij}(\theta_{i1}, \theta_{i2})$ 를 찾고 그 편차를 사용자 i 의 MAE로 정의하는 것이다. 위의 방법을 편의상 예측비율 증가 방법(method of coverage increasing)이라고 부르기로 한다.

6. 연구결과 및 분석

식(5), 식(6), 식(7)을 이용하여 MovieLens 자료에 적용한 결과는 다음 <표 2>와 같다.

<표 2>MAE 비교 및 예측개수의 비교

알고리즘 결과	근접 이웃 알고리즘		제안 알고리즘		
	피어슨상관계수 적용 식(1)	벡터유사도 적용 식(1)	식(5)	식(6)	식(7)
MAE	0.62276	0.745074	0.71612	0.72826	0.69707
예측개수	99,859	99,859	100,000	100,000	100,000

<표 2>에서 근접 이웃 알고리즘을 이용한 MAE결과 비교 및 예측개수의 비교에서 피어슨상관계수를 적용한 경우 MAE가 0.62276으로 가장 작았으며 벡터유사도를 적용한 경우 MAE가 0.745074로 가장 크게 나타났다. 제안방법인 식(5), 식(6), 식(7)의 비교에서 식(5), 식(6)에 비해 식(7)의 MAE가 0.69707로 가장 작게 나타났다. 이것은 피어슨상관계수를 적용한 경우보다는 MAE가 크지만 벡터유사도를 적용한 경우보다는 MAE가 작게 나타났다. 예측치의 개수에서 피어슨상관계수를 적용할 경우 상관계수를 구하지 못하는 경우에 예측을 할 수 없어 예측개수가 줄어든다는 단점을 가지고 있으며 벡터유사도를 적용한 경우는 일반적으로 예측치를 더 많이 구할 수 있어 예측개수는 증가하지만 MAE가 크다는 단점을 가지고 있다. <표 2>에서 예측개수의 비교에서는 피어슨상관계수를 적용한 경우와 벡터유사도를 적용한 경우는 100,000개의 평가치에 대한 예측개수는 99,859개로 평가개수 100,000에 대한 모든 예측을 할 수 없었으나 식(5), 식(6), 식(7)의 경우에는 평가치 100,000개에 대해 100,000개를 모두 예측하므로 예측율이 가장 높게 나타났다. 제안방법의 적용결과 MAE는 피어슨상관계수를 적용한 경우보다는 MAE가 크며 벡터유사도를 적용한 경우보다는 MAE가 작게 나타났고 예측개수에서는 100,000개의 평가치에 대해 100,000개의 예측치를 얻을 수 있었다.

제안방법중 식(7)의 경우가 식(5), 식(6)에 비해 MAE가 가장 작았으며 벡터유사도를 적용한 경우보다도 MAE는 작으며 예측의 개수는 100,000개 모두에 대해 예측을 할 수 있다는 장점을 보였다.

7. 결론 및 제언

제안방법인 예측비율 증가 방법(method of coverage increasing)을 적용하면 MAE는 피어슨상관계수를 적용한 경우보다는 크지만 벡터유사도를 적용한 경우보다 MAE는 작으며 예측개수는 더 많이 예측을 할 수 있어 상품(아이템) 추천 목록 작성에는 장점을 가지고 있다고 할 수 있어 희소성에 의해 예측을 할 수 없는 경우에도 적용할 수 있는 방법이며 제안방법인 예측비율 증가 방법에서 θ_i, θ_j 를 사전에 계산하여 사전 정보를 가질 수 있어 추천시스템의 추천목록 작성에 빠른 계산을 할 수 있다고 본다. 전자 상거래에 적용되는 추천시스템은 추천의 정확도의 연구와 더불어 추천 시간을 단축하는 문제도 고려되어 연구되어야 한다. 아이템(상품)의 증가속도에 비해 사용자(고객)들의 단시간 폭발적인 증가로 인해 추천 시간을 단축하는 문제도 함께 고려되어 연구되어야 한다. 추천시간을 단축하는데 제안 방법과 유사한 방법을 적용하면 추천시간 단축에 도움을 줄 수 있을 것으로 기대 된다.

참고문헌

1. 김택현, 양성봉 (2004). 협동적 필터링 기반 추천 시스템을 위한 향상된 이웃 선정 방법, 제 21회 한국정보처리학회 춘계학술발표대회, 11권, 1호, 453-456.
2. 박지선, 김택현, 유영석, 양성봉 (2001). 추천 시스템을 위한 고객 클러스터링 방법을 적용한 예측 알고리즘. 한국정보과학회 2001년 춘계학술대회, 제28권, 1호, 268-270.
3. 이희춘 (2006a). An Exploratory Study for Decreasing Error of Prediction Value of Recommender System on User Based, *Journal of the Korean Data & Information Science Society*, 제17권, 1호, 77-86.
4. 이희춘 (2006b). Improved Algorithm for User Based Recommender System, *Journal of the Korean Data & Information Science Society*, 제17권, 3호, 717-726.
5. 이희춘, 이석준, 정영준 (2006). The Effect of Co-rating on the Recommender System of User Base, *Journal of the Korean Data & Information Science Society*, 제17권, 3호, 775-784.
6. Breese, J. S., Heckerman, D., Kadie, C. (1998). Empirical Analysis of Predictive Algorithms for Collaborative Filtering, In Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, 43~52.
7. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes. D., Sartin, M. (1999). Combining Content-Based and Collaborative Filters in an Online Newspaper, In Proceedings of ACM SIGIR Workshop on Recommender Systems.
8. Deshpande, M., Karypis, G. (2004). Item-based top-N recommendation algorithms, *ACM Transactions on Information Systems*, 22-1, 143-177.
9. Herlocker, J. Konstan, J. Borchers, A. Riedl, J. (1999). An Algorithmic Framework for Performing Collaborative Filtering, In Proceedings of the 1999 Conference on Research and Development in Information Retrieval, 230-237.
10. Hill, W., Stead, L., Rosenstein, M., Furnas, G. (1995). Recommending and Evaluating Choices in A Virtual Community of use, Proceedings of the SIGCHI conference on Human factors in computing systems, 194-201.
11. Konstan, J. and J. Reidl. 2000. "Tutorial Notes: Recommender Systems in E-Commerce" In the Proceeding of ACM E-Commerce 2000 Conference
12. Resnick, P., Iacovou, N., Suchak, M., Bergstorm, P., Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering

- of Netnews, In Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, 175~186.
13. Sarwar, B. M., Konstan, J. A., Borchers, A., Herlocker, J. L., Miller, B. N., Riedl, J. (1998), Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System, Computer Supported Cooperative Work, 345-354.
 14. Sarwar, B. M., Karypis, G., Konstan, J. A., Reidl, J. (2001), Item-based collaborative filtering recommendation algorithms, In Proceedings of Tenth International World Wide Web Conference, 285-295.
 15. Schafer, J. B., Konstan, J. A., Riedl, J. (2001). E-Commerce Recommendation Applications, Data Mining and Knowledge Discovery, 5-1, 115-153.
 16. Shardanand, U., Maes, P. (1995). Social Information Filtering: Algorithms for Automating "Word of Mouth", Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems, volume 1, 210~217.
 17. Swearingen, K. and Sinha, R. "Interaction design for recommender systems" In the poroceding of Designing Interactive System, 2002