

협력적 여과에서 평가 행렬의 희소성 문제를 해결하기 위한 Singular Value Decomposition의 적용 방법에 관한 연구

⁰정 준 정 대 진 김 용 환 이 필 규
인하대학교 전자계산공학과
{g1991263,g1991262,g1991260}@inhavision.inha.ac.kr

A Research for Applying Singular Value Decomposition to Collaborative Filtering for Coping With the Sparsity of Rating Matrix

⁰Jun Jeong Dae-Jin Jeong Yong-Han Kim Phill-Kyu Rhee
Dept of Computer Science and Engineering, Inha University

요 약

인터넷의 발달로 사용자들은 인터넷에서 필요한 정보를 습득할 수 있을 뿐만 아니라, 생활에 필요한 여러 가지 활동들을 할 수 있게 되었다. 특히 주목받는 부분은 구매 활동이다. 따라서 수많은 기업들이 사람들의 구매 활동에 관련된 전자상거래에 투자하고 있고, 현재 Amazon.com등과 같은 세계적인 사이트이 서비스를 실시하고 있다. 또한, 전자상거래 사이트들은 사용자들의 구매 활동을 도와주기 위해 추천 시스템의 도입을 추진하고 있다. 추천 시스템은 사용자들로부터 얻어진 정보를 학습하여 이용 가능한 상품 중에서 고객이 좋아할 만한 것은 추천해주는 시스템이다. 본 논문에서는 추천 시스템에서 사용되는 주요한 방법인 협력적 여과방법에서 초기 rating 행렬의 희소성 문제를 해결하기 위하여 Singular Value Decomposition의 적용 방법을 제안하고 한다.

1.서론

인터넷의 발달은 사람들의 생활 방식을 바꾸어 놓을 만큼 영향력을 행사하고 있다. 전화번호 검색등과 같은 단순한 정보의 획득뿐만 아니라 오락 산업, 전자상거래 등 기존의 생활 방식이 인터넷으로 전환이 가속화되고 있다.

특히, 전자상거래는 많은 관심을 받고 있으며 대규모도 이루어지고 있다. 따라서 전자상거래를 목표로 하는 사이트간의 경쟁도 치열해지고 있으며 사용자들의 관심을 형성하기 위해 많은 노력을 하고 있다.

이런 상황에서 구매 촉진을 위한 추천 시스템이 도입되기 시작했으며 Amazon.com과 cdnow.com등과 같은 일부 선진 사이트들은 기본적인 서비스를 실시하고 있다. 여기서 말하는 추천시스템이란 사용자로부터 명백하게 혹은 묵시적으로 정보를 획득하여 학습을 통해서 사용자가 선호할 만한 아이템을 추천해 주는 시스템이다[12][13]. 이것은 구전이 라는 입에서 입으로 전해지는 사회현상을 모델링한 시스템이다[15][16].

추천 시스템에 대한 두가지 접근 방법은 내용기반 여과 방법과 협력적 여과 방법으로 분류되어진다.

내용기반 여과는 추천될 대상에 대해서 자동적인 내용분석을 통한 방법이고 협력적 여과는 자동적인 내용분석이 어려운 멀티미디어 정보에 활용되어지고 있다.

추천 시스템에서 사용되는 주요한 방법인 협력적 여과방법은 사용자로부터 명시적으로 추천될 대상에 대한 평가를 입력받아 그 정보를 학습한다. 그러나 사용자들의 명시적 평가에 대한 거부감으로 인하여 초기 평가 행렬의 희소성 문제를 야기한다 [10][17]. 따라서, 본 논문에서는 희소성 문제를 해결하기 위하여 SVD(Singular Value Decomposition)의 적용 방법을 제안하고 한다.

2.관련연구

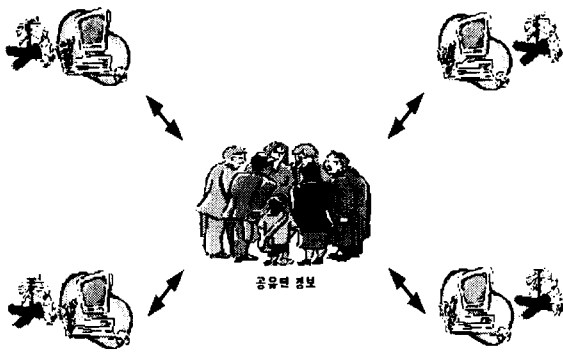
2.1.내용기반여과를이용한추천

추천을 위한 내용기반 방법은 정보검색에서 기원했고, 많은 방법들을 공유한다[1][6]. 내용 기반 방법은 사용자에게 의해서 평가된 아이템에 대한 내용과 추천될 아이템에 대한 내용을 분석해서 추천한다[4][8]. 텍스트 문서의 내용을 분석하고 추천을 하기위한 기초역할로서 내용에서 규칙성을 찾기 위한 많은 알고리즘이 제안되었다.

그러나 내용기반 여과 방법은 (a) 아이템은 기계가 파악할 수 있는 형태이어야 하며, (b) 속성들은 사람에 의해서 아이템에 할당되어야 하고, (c) 우연한 발견을 형성하기 위한 어떠한 방법도 가지고 있지 않다는 단점을 가지고 있다[4][15].

2.2. 협력적 여과를 이용한 추천

추천시스템은 추천할 아이템을 선택하기 위하여 협력적 여과(Collaborative Filtering) 방법을 사용한다. 내용 기반 여과의 문제점을 해결하기 위한 협력적 여과는 사회적 여과(Social Filtering)라고도 하며 유사한 기호를 가지는 다른 사람들의 선호도에 기반해서 아이템을 여과한다[3][9][11][14][18]. [그림1]은 협력적 여과를 이용한 추천 시스템을 간략하게 나타낸다.



[그림 1] 추천 시스템의 개념도

협력적 여과에서 사용되어지는 방법들은 기계학습, 통계학적 방법, AI 등의 분야에서 파생되었다. 예를 들면, [15]에서 방법은 통계학에서 사용되어지는 Pearson's r 방법을 협력적 여과에서 사용할 수 있도록 변형하여 사용하였다. 식(1)은 [15]에서 제시한 Constrained Pearson's r 방법이다.

$$B_{xy} = \frac{\sum (U_x - M)(U_y - M)}{\sqrt{\sum (U_x - M)^2 \sum (U_y - M)^2}} \quad (1)$$

여기서, Bxy는 Ux와 Uy 사이의 "correlation"의 새로운 측정이고 M은 평가 범위의 중간값이 된다. B의 크기는 1보다 항상 작다. 이것은 Schwarz Inequality를 사용해서 증명된다.

협력적 여과에서 사용되어지는 알고리즘들은 Constrained Pearson r 뿐만 아니라 벡터 유사

도, 베이저안 네트워크, 결정 트리, 신경망등 여러

메모리 기반 알고리즘	Correlation
	Vector Similarity
메모리 기반 알고리즘 확장	Default Voting
	Inverse User Frequency
	Case Amplification
모델 기반 방법	Cluster Models
	Bayesian Network Model
	Case Amplification

[표 1] 협력적 여과 알고리즘 분류

가지 방법들이 실험되었다.[2] [표 1]은 간략하게 알고리즘의 분류를 나타내고 있다.

2.3. 관련 시스템

추천 시스템에 대한 연구는 [12][13]에서 자세히 살펴볼 수 있다. 본 논문에서 대표적인 몇가지 시스템들에 대해서 간략하게 살펴 보겠다.

Tapestry는 아이템에 대한 rating 혹은 주석을 입력받는 정보여과 시스템이며, 협력적 여과라는 개념을 최초로 제시하였다[18]. 그러나, 자신과 유사한 사용자들을 찾아주는 기능은 지원하지 않고 단지 "구전"이라는 과정에 대한 온라인 구조만을 제공하였다[18].

GroupLens는 협력적 여과를 뉴스그룹에 대한 개인화된 선택에 적용한 시스템이다. GroupLens는 사용자들이 뉴스를 평가하고, 평가(rating)가 네트워크를 통해서 사용자 에이전트에 분배된다. 그리고 유사한 사용자들을 찾기 위하여 Pearson r 상관계수를 적용하였다[3][5][19].

Ringo는 음악가에 대한 명백한 평가를 웹 혹은 전자우편을 통해서 받아들여 사용자 프로파일을 구성하는 음악 추천시스템이다. Ringo는 GroupLens에서 사용한 Pearson r 상관계수를 수정한 Constrained Pearson r 방법을 소개하였다[15][16].

3. SVD(Singular Value Decomposition)

본 절에서는 SVD의 형식적인 정의와 예를 소개하고자 한다. 우선, A라는 행렬이 존재하면 행렬 A의 분해는 식()와 같은 형식의 m×n 대각 행렬 Σ를 포함한다.

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \leftarrow \begin{matrix} (m-r)개의 열 \\ (n-r)개의 열 \end{matrix} \quad (2)$$

여기서, D는 m과 n에서 어떤 r에 대해서 $r \times r$ 대각 행렬이다.

[정리 1]

A는 계수가 r인 $m \times n$ 행렬이라고 하면, 식(2)와 같이 $m \times n$ 행렬인 Σ 가 존재하며, D는 행렬 A의 첫번째의 r singular value이고, $\delta_1 \geq \delta_2 \geq \dots \geq \delta_r > 0$ 이 되고, 식(3)과 같은 $m \times m$ 직교 행렬인 U와 $n \times n$ 직교 행렬 V가 존재한다.

$$A = U\Sigma V^T \quad (3)$$

직교 행렬 U와 V와 Σ 를 포함하는 행렬 A의 인수분해인 $U\Sigma V^T$ 를 행렬 A의 singular value decomposition이라고 불리운다. 행렬 U와 V는 유일하지 않지만, Σ 의 대각 원소들은 행렬 A의 singular value이다.

예를 들면, 행렬 A는 다음과 같다고 하자.

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \quad (4)$$

행렬 A의 SVD를 구하려면 우선 식(5)와 같이 $A^T A$ 를 구한다. $A^T A$ 의 고유값은 18과 0이고 상응하는 고유벡터는 식(6)과 같다.

$$A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix} \quad (5)$$

$$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad (6)$$

다음으로 Av_1 을 구한다.

$$Av_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix}, \delta_1 = \|Av_1\| = \sqrt{18} = 3\sqrt{2} \quad (7)$$

따라서, u_1 은 식(8)과 같다.

$$u_1 = \frac{1}{3\sqrt{2}} Av_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} \quad (8)$$

또한, v_2 는 $A^T A$ 의 고유값인 0에 상응하기 때문에 $Av_2=0$ 이 된다.

다음으로 $\{U_1\}$ 을 R^3 의 기초로 확장한다. U_1 에 직교하는 두 개의 직교 벡터가 필요하며, 각 벡터 $u_i^T x = 0$ 는

를 만족해야 하며 그것은 방정식 $x_1 - 2x_2 + 2x_3 = 0$ 과 동일하다. 이 방정식의 해집합에 대한 기초는 식(9)와 같다.

$$w_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, w_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad (9)$$

$\{w_1, w_2\}$ 에 Gram-Schmidt 과정을 적용하면 식(10)을 얻는다.

$$u_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, u_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix} \quad (10)$$

결과적으로, $U=[u_1, u_2, u_3]$ 가 되고 $V=[v_1, v_2]$ 가 되고 Σ 는 식(11)과 같다면, A의 singular value decomposition은 식(12)와 같다.

$$\Sigma = \begin{bmatrix} 3/\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (11)$$

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad (12)$$

4. 협력적 여과에서 SVD의 적용

협력적 여과에서 SVD의 적용은 [10][17]에서 실험적으로 적용되었다. [10]에서는 사용자의 초기 rating을 SVD에 적용한 결과를 신경망을 이용하는 방법을 사용하였고 [17]에서는 웹문서를 대상으로 순수하게 SVD만을 사용하였다.

본 논문에서는 Constrained Pearson's r에서 사용자의 유사도를 계산하기 위해 상관계수를 구하는 방법 대신에 SVD를 이용하여 사용자 유사도를 계산하여 Constrained Pearson's r에서 선호도를 예측

하는 방법을 적용하여 상관계수와 SVD의 사용자 유사도 측정 성능을 비교하고자 한다.

SVD는 사용자를 비교하여 가장 유사한 이웃들을 찾는 방법을 제공한다[17]. 개인적인 아이템에 대한 rating를 비교하는 것 대신 사용자의 가중치와 각 사용자의 중요성을 비교함으로써 유사한 사용자들을 찾을 수 있다.

따라서, 식(3)의 양변에 V를 곱하여 구해진 UΣ는 사용자를 벡터 공간으로 매핑시키는 것을 의미하며, 각 행은 사용자를 나타내는 특징 벡터가 된다. 사용자간의 유사도는 특징 벡터 공간에서 cosine 값으로 구할 수 있다. 계산된 유사도에 가중치를 적용하여 사용자가 평가하지 않은 아이템에 대한 예측은 식(13)과 같이 m명의 가장 가까운 사용자의 가중치가 적용된 rating의 평균에 의해서 구할 수 있다[17].

$$P_{u,i} = \frac{w_1 n_1 + w_2 n_2 + \dots + w_m n_m}{w_1 + w_2 + \dots + w_m} \quad (13)$$

여기서, $P_{u,i}$ 는 사용자 u가 아이템 i에 대한 예측 값이다. $\{n_1, n_2, \dots, n_m\}$ 은 m명의 가장 가까운 사용자의 집합이고, $\{w_1, w_2, \dots, w_m\}$ 은 가장 가까운 사용자와의 가중치 집합이다.

예를 들어, 사용자는 {1,2,3,4,5}의 5단계로 아이템의 선호도를 평가하고 1은 선호도가 가장 낮은 값이고 5는 선호도가 가장 높은 값이라고 하고, 3명의 사용자가 5개의 아이템에 대해서 [표 2]과 같이

	아이템1	아이템2	아이템3	아이템4	아이템5
사용자1	4	1	3	2	1
사용자2	3	2	4	3	4
사용자3	1	4	4	3	3

평가를 했다고 가정하자.

[표 2] 초기 사용자 rating

[표 2]은 식(13)과 같이 행렬로 나타낼 수 있다.

$$A = \begin{bmatrix} 4 & 1 & 3 & 2 & 1 \\ 3 & 2 & 4 & 3 & 4 \\ 1 & 4 & 4 & 3 & 3 \end{bmatrix} \quad (14)$$

(14)에 대한 singular value decomposition의 결과는 (15)로 UΣ는 (16)으로 계산된다.

식(15)에서 Σ는 singular value를 나타낸다. Σ의 대각선상에 있는 값들을 보면 10.0012, 3.1247, 0.4614로 값들이 급격하게 작아지는 것을 알 수 있다. 따

라서 UΣ에서 처음에서 두 열의 값으로 이웃 사용자를 찾을 때 사용한다.

$$U\Sigma V' = \begin{bmatrix} -0.5055 & 0.6689 & 0.5448 \\ -0.6131 & 0.1656 & -0.7723 \\ -0.6069 & -0.7245 & 0.3264 \end{bmatrix} \begin{bmatrix} 10.0012 & 0 & 0 & 0 \\ 0 & 3.1247 & 0 & 0 \\ 0 & 0 & 0.4614 & 0 \end{bmatrix} \quad (15)$$

$$\begin{bmatrix} -0.4468 & 0.7835 & 0.4088 & 0.1386 \\ -0.4159 & -0.6070 & 0.6332 & 0.2223 \\ -0.6396 & -0.0709 & -0.5038 & -0.1746 \\ -0.4670 & -0.1077 & 0.5933 & 0.8530 \end{bmatrix}$$

$$U\Sigma = \begin{bmatrix} -5.056 & 2.0988 & -0.4293 & 1.9593 \\ -6.1320 & 0.5296 & -1.3021 & 2.7216 \\ -6.0702 & -2.2515 & 0.8534 & 2.8897 \end{bmatrix} \quad (16)$$

4. 실험

4.1. 실험자료

본 논문에서 제안하는 방법을 실험하기 위한 실험자료는 DEC Systems Research Center에서 제공하는 EachMovie collaborative filtering data set을 사용하였다[20]. DEC는 18개월 동안 협력적 여과 알고리즘을 실험하기 위하여 EachMovie 추천 서비스를 실행하였다. 그 결과로 수집된 자료가 EachMovie data set이다. 72916명의 사용자들이 1628개의 영화와 비디오에 대해서 2811983개의 평가값을 가지고 있고, 사용자의 중요한 정보가 제거되고 협력적 여과 알고리즘에 쉽게 적용될 수 있도록 가공하여 제공되어진다. 사용자의 rating 정보는 {0.2, 0.4, 0.6, 0.8, 1.0}의 5단계로 이루어져 있다.

EachMovie data set의 방대한 크기 때문에 본 논문에서는 전체적인 자료에서 500명의 사용자가 1628개의 영화와 비디오에 대한 3561개의 rating만을 대상으로 실험을 하였다.

4.2. 성능평가

본 논문에서는 여러 가지 평가방법중에서 mean absolute error 방법을 사용한다. mean absolute error 방법은 부호와 관계없이 개개의 오차 크기의 평균이며, mean-squared error의 값이 작을수록 더 정확한 방법이다

$$\frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n} \quad (17)$$

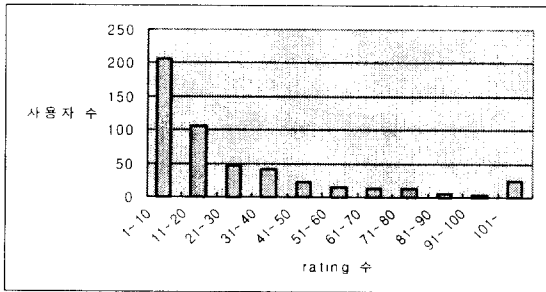
mean absolute error는 다른 오차보다 더 큰 예측 오차의 결과를 강조하는 mean-squared error의 특성에 영향을 받지 않는다. 오차의 모든 크기는 그들의 오차량에 따라서 동일하게 취급되어진다 [15][21].

4.3. 실험 방법

본 논문에서 실험 방법은 사용자의 rating이 주어졌을 때 사용자가 평가하지 않은 아이템에 대한 예상값의 정확도이다. 따라서, 각각의 사용자 rating 중에서 80%는 학습자료(training data)로 사용했고 20%는 실험자료(test data)로 사용했다. 즉, 학습자료를 가지고 실험자료에 있는 아이템에 대한 rating 값을 예측한다.

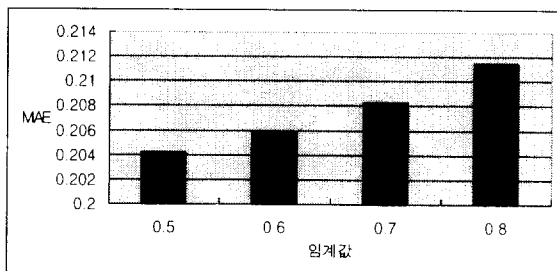
본 논문에서 제시하는 방법의 비교 평가를 위하여 동일한 자료를 대상으로 Constrained Pearson r을 이용하는 예측방법을 사용하였을 때의 오차도 계산하였다.

4.4. 실험 결과



[그림 2] rating 수 대 사용자 수

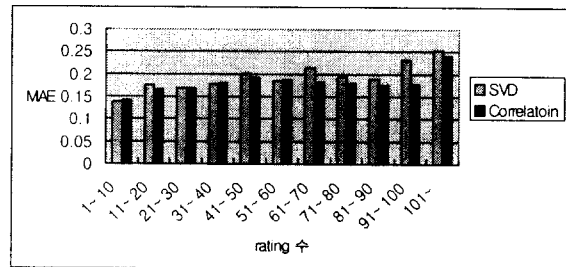
[그림 2]는 실험 자료의 rating 수에 따른 사용자의 수를 보여준다. 사용자의 62.4%는 20개 이하의 rating 수를 가지고 있고 41.2%는 10개 이하의 rating 수를 가지고 있다. 따라서, 대부분의 사용자들은 충분히 많은 rating을 하지 않는 것을 알 수 있다.



[그림 3] 임계값 대 Mean Absolute Error

전체 사용자에게 대해서 본 논문에서 제시하는 방법을 사용하여 실험한 결과는 [그림 3]과 같은 결과를 얻을 수 있었다.

[그림 3]에서 알 수 있듯이 임계값이 0.5일 때 MAE가 0.2042로서 가장 좋은 성능을 나타내고 있다. 동일한 자료에 상관계수를 이용하는 방법으로 MAE를 측정하면 0.2011으로 SVD를 이용한 방법보다는 근소하게 오차가 적은 것을 알 수 있다.



[그림 4] SVD와 Correlation의 rating 수에 따른 MAE

그러나, [그림 4]에서 보여지듯이 SVD를 이용한 방법은 rating 수가 10개 이하 일 때 MAE의 값이 가장 작은 것으로 나타났고, 약 11~80에서는 상관계수를 이용한 방법과 유사한 성능을 보였다.

위의 실험결과로서 사용자의 rating 수가 10이하 일 때 즉, rating 행렬의 희소성이 존재할 때, SVD는 상관계수를 이용한 방법보다 약 2.5% 정도의 정확도 개선 효과가 있는 것을 알 수 있었지만 사용자가 충분히 rating을 하였을 때는 상관계수를 이용하는 방법보다는 우수한 성능을 나타내지는 못했다. 따라서, SVD는 rating 행렬의 희소성 문제에 대한 보완적인 방법으로 사용될 수 있을 것이다.

5. 향후 연구 및 결론

본 논문에서는 사용자가 평가한 rating 행렬의 희소성 문제에 대한 해결책으로 Singular Value Decomposition을 이용한 방법을 제시하였다. 4절의 실험에서 알 수 있듯이 10개 이하의 rating 수를 가지는 사용자들에게 있어서 상관계수를 이용한 방법보다 우수한 성능을 얻을 수 있었다.

따라서, 향후 연구 과제로 대규모 자료에 대한 정확한 실험을 통해 SVD의 적용에 관한 연구를 진행할 것이다.

6. 참고 문헌

[1] Pattie Maes, Agents that reduce work and information overload, Communications of the

- ACM Volume 37, Issue 7 (1994) Pages 30-40.
- [2] Breese, J.S., Heckerman, D. and Kadie, C., Empirical Analysis of Predictive Algorithms for Collaborative Filtering, Proceedings of Thirteenth Conference on Uncertainty in Artificial Intelligence, Madison, WI, Morgan Kaufmann.
- [3] Miller, B., Riedl, J., and Konstan, J., Experiences with GroupLens: Making Usenet Useful Again., Proceeding of the USENIX 1997 Annual Technical Conference, pages 219-231, Anaheim, CA.
- [4] Zarko Balabanovic and Yoav Shoham., Fab: Content-based, Collaborative recommendation, Communications of the ACM, 40(3):64-72, 1997.
- [5] Konstan J.A., B.N. Miller, D. Maltz et al., GroupLens : applying collaborative filtering to Usenet news, Communications of the ACM, 40(3), 77-87, 1997.
- [6] Belkin, N.J. and Croft, B.W., Information filtering and information retrieval-two sides of the same coin, Communications of the ACM 35(2), December 1992.
- [7] Christopher Lueg, Issues in Understanding Collaborative Filtering, CHI' 99 Workshop Interacting with Recommender Systems 15/16 May 1999 Pittsburgh, Pennsylvania, USA.
- [8] Patrick Baudisch, Joining Collaborative And Content-based Filtering, CHI' 99 Workshop Interacting with Recommender Systems 15/16 May 1999 Pittsburgh, Pennsylvania, USA.
- [9] Damain Arregui, Manfred Dardanne, Xerox Research Centre Europe., Knowledge Pump : Community-centered Collaborative Filtering, Proceedings of the fifth DELOS Workshop on Filtering and Collaborative Filtering.
- [10] Daniel Billsus and Micheal J. Pazzani., Learning Collaborative Information Filters., Proceedings of the Fifteenth International Conference on Machine Learning, pages 46-54, July 1998.
- [11] Maltz and Kate Ehrlich, Pointing the way: active collaborative filtering, CHI'95 Human Factors in Computing Systems, p.202-209, 1995.
- [12] Resnick, P. and Varian, H., Recommender systems, Communications of the ACM, 40(3): 56-58.
- [13] J. Ben Schafer, Recommender Systems in E-Commerce, Proceedings of the ACM Conference on Electronic Commerce, 1999, Pages 158 - 166.
- [14] Hill, W., Stead, L., Rosenstein, M. and Furnas, G., Recommending And Evaluating Choices In A Virtual Community Of Use, Conference on Human Factors in Computing Systems-CHI '95(Denver, May, 1995).
- [15] Shardanand, U., Maes, P., Social Information Filtering Algorithms for Automating "Word of Mouth", Proceedings of the CHI-95 Conference, Denver, CO, ACM Press.
- [16] Upendra Shardanand, Social Information Filtering for Music Recommendation, MIT EECS M. Eng. Thesis, also TR-94-04, Learning and Common Sense Group, MIT Media Laboratory, 1994.
- [17] Michael H. Pryor., The Effects of Singular Value Decomposition on Collaborative Filtering, Dartmouth College Technical Report PCS-TR98-338, June 1998.
- [18] Goldberg, D., D. Nichols, B. Oki, and D. Terry, Using collaborative filtering to weave an information tapestry, Communications of the ACM, 35(12), 61-70, 1992.
- [19] Badrul M. Sarwar, Joseph A. Konstan, Al Borchers, Jon Herlocker, Brad Miller, and John Riedl, Using filtering agents to improve prediction quality in the GroupLens research collaborative filtering system, Proceedings of 1998 Conference on Computer Supported Collaborative Work.
- [20] McJones, P. (1997) EachMovie collaborative filtering data set. DEC Systems Research Center. <http://www.research.digital.com/SRC/eachmovie/>.
- [21] Ian H. Witten, Eibe Frank, Data Mining, Morgan Kaufmann Publishers, San Francisco, California 1999, p147-p150.