
소셜 북마킹 서비스의 태그를 이용한 개인화 콘텐츠

Personalized Contents using the Tags of the Social Bookmarking Service

한주현, Juhyeun Han*, 정문열, Moonryul Jung**

요약 웹 2.0 이라 불리는 현 웹의 패러다임은 개방, 공유, 참여로 압축하여 말할 수 있다. 이 속에서는 사용자의 참여와 공유로 콘텐츠가 생산 또는 재생산된다. 이러한 콘텐츠는 사용자의 관심을 반영하기 때문에 사용자가 어떠한 콘텐츠를 만들어 냈는지, 수집했는지 등을 분석하면 사용자의 관심 범주를 추출할 수 있다. 본 논문에서는 사용자가 소셜 북마킹 서비스를 이용하며 생성한 태그를 바탕으로 사용자의 관심 범주를 추출하여 이를 통해 개인화 콘텐츠 제공 서비스를 제안한다. 우선, 웹 서비스에서 제공하는 피드를 이용하여 사용자가 생성한 태그 중 가장 많이 쓰인 10개의 태그와 그것들과 관련 있는 태그들만 모아서 관심 범주를 추출하기 위한 태그 집합을 구성한다. 구성된 태그 집합을 바탕으로 피어슨 상관 계수를 통해 태그 간 동시 사용률을 조사한다. 이후 사용자 흥미에 부합하는 콘텐츠를 검색하기 위해 조사된 동시 사용률을 바탕으로 검색 키워드 그룹을 추출한다. 이렇게 만들어진 키워드 그룹들은 사용자의 평소 관심사와 관련된 콘텐츠를 검색하는데 사용되며, 이를 통해 사용자의 관심 있는 내용의 콘텐츠를 사용자의 특별한 검색 절차 없이 제공받는다. 이러한 방식을 통해 사용자가 원하는 정보를 입력하는 절차 없이도 웹에 축적된 사용자의 정보를 사용하여 자동으로 개인화된 콘텐츠를 제공할 수 있을 것으로 기대 된다.

핵심어: *Personalize Contents, Collaborative Filtering, Social Bookmarking, Tag, Mash-up, Web 2.0*

*주저자 : 서강대학교 영상대학원 미디어공학과 석사과정 e-mail: hjh1019@sogang.ac.kr

**교신저자 : 서강대학교 영상대학원 미디어공학과 교수 e-mail: moon@sogang.ac.kr

1. 서론

최근의 웹 서비스는 셀 수 없을 정도로 다량의 정보를 보유하여 사용자는 그 속에서 사용자가 원하는 정보를 얻기 위해 많은 시간을 소모하게 된다. 이에 많은 웹 서비스들은 사용자에게 그들이 만족할 수 있는 정보를 제공한다. 이를 위해 추천 시스템을 이용한다.

추천 시스템은 사용자의 선호도를 분석하여 정보들에 대한 사용자의 선호도를 예측하는 시스템이다. 대표적인 온라인 상점인 '아마존닷컴(Amazon.com)'은 추천 시스템을 도입하여 'Your Recommendation'이라는 서비스로 성공적인 상용화 사례를 보였다. 추천 시스템은 사용자의 평가값을 예측하는 방법에 따라 크게 내용기반 추천(Content-Based Collaborative Recommendation)[1] 방법과 협업 필터링(Collaborative Filtering)[2] 방법이 있다.

내용기반 추천 방법은 검색 또는 추천하고자 하는 대상 정보나 자료의 내용을 설명하는 정보들과 사용자가 선호 또는 요구하고자 하는 내용을 비교하여 이에 부합되는 대상을 추천 결과물로 제시하는 방법이다. 이 방법은 문자로 구성된 정보를 적합한 주제에 부합시키는 추천 시스템을 통하여 그 우수성이 입증된바 있다.

협업 필터링 방법은 과거에 사용자들이 명시적으로 정보들에 대하여 선호하는 정도에 따라 평가한 값들에 근거하여, 사용자 혹은 정보 사이의 유사한 정도를 계산한다. 유사한 사용자들이 평가한 선호들의 가중 평균을 계산하여 대상 사용자의 평가값을 예측할 수 있으며, 이러한 협동적 여과 방식은 앞에서 언급한 아마존닷컴의 추천 시스템 등에서 사용되고 있다.[3]

협업 필터링 방법을 사용하여 추천 시스템을 제공하는 방식은 기존 웹 서비스들이 웹 2.0 패러다임으로 전환하며 보편적인 방법으로 사용되고 있다. 웹 2.0 패러다임은 참여, 공유, 개방의 특징을 갖고 있다. 이 속에서는 사용자의 참여로 콘텐츠가 생산되며, 사용자들 간의 공유를 통해 콘텐츠의 재생산이 이루어진다. 대부분의 이러한 콘텐츠는 사용자의 관심 범주에서 비롯된다. 즉, 사용자의 참여로 생성된 콘텐츠는 사용자의 관심사가 반영된다. 이에 협업 필터링 방법은 사용자에게 의해 생성된 콘텐츠와 정보를 이용하여 사용자의 흥미에 부합하는 정보를 추천하기 위한 방법으로 알맞다.

따라서 사용자가 만들어낸 콘텐츠 혹은 데이터를 협업 필터링 방법으로 분석하면 사용자의 관심 범주에 대한 데이터를 추출할 수 있다. 최근 이러한 방법에 대한 연구는 널리 이루어지고 있다. 특히 소셜 북마킹과 태그를 기반으로 협업 필터링을 사용하여 사용자의 정보를 분석하거나 콘텐츠의 정보를 분석하는 여러 연구결과가 발표되고 있다.[4][5] 또한 이를 사용하여 개인화된 정보를 제공하는 연구 결과도

이러지고 있다.[6]

본 연구는 사용자의 관심 범주를 알아내기 위해 소셜 북마킹 서비스를 선택하였으며, 사용자가 서비스를 이용하며 만들어낸 태그를 바탕으로 사용자의 관심 범주 데이터 추출 방법을 소개한다. 또한 추출된 데이터를 통해 사용자의 흥미에 부합하는 콘텐츠를 검색하기 위한 검색 키워드 조합 방법을 소개한다.

본 논문의 구성은 다음과 같다. 2장에서는 태그와 소셜 북마킹을 설명한다. 3장에서는 소셜 북마킹 서비스를 통해 태그 수집 방법을 설명한다. 4장에서는 태그 간 동시 사용률을 조사하기 위한 방법을 설명한다. 5장에서는 동시 사용률을 바탕으로 한 검색 키워드 조합 방법을 설명한다. 6장에서는 제안한 방법에 대한 구현 및 결과를 보여준다. 마지막으로 7장에서는 결론을 제시한다.

2. 소셜 북마킹과 태그

웹 서비스를 통해 사용자의 관심 범주를 추출하기 위해서는 사용자의 관심사에 대한 데이터를 저장하는 웹 서비스를 이용한다. 사용자의 관심 범주는 일반적으로 태그라고 불리는 키워드로 표시한다. 즉, 우리의 목적을 위해서 태깅(Tagging) 서비스를 지원하는 웹 서비스를 사용한다.

2.1 태그(Tag)

태그는 인터넷 북마크, 이미지, 동영상, 파일 등의 여러 콘텐츠의 정보를 나타내는 키워드이다. 태그는 어떤 콘텐츠가 담고 있는 정보를 표현하는 메타데이터로서 음악, 이미지, 동영상 등과 같이 검색이 어려운 멀티미디어 콘텐츠를 열람 및 검색하는데 쓸 수 있다. 시스템(서비스)에 따라 콘텐츠를 생산한 생산자 혹은 콘텐츠 사용자에게 의해 태그는 자유로운 형태로 주관적인 관점에서 기록될 수 있다. [7]

많은 사용자들이 많은 콘텐츠를 태깅(Tagging)한 웹사이트의 태그 집합은 폭소노미(Folksonomy)가 된다. 폭소노미는 토마스 반더 월(Thomas Vander Wal)에 의해 제창된 용어로 '사람들'을 지칭하는 'folk'와 '분류'를 지칭하는 'taxonomy'의 합성어로 '사람들에 의한 분류'로 정의될 수 있다.[8] 폭소노미의 중요한 점은 용어들 사이의 계층이나 부모-자식 관계, 형제 관계가 없다는 것이다. 공통의 콘텐츠를 바탕으로 한 태그 집합, 즉 '관계된(related)' 태그들이 자동적으로 생성된다.[9] 하나의 콘텐츠에 여러 태그가 붙어 그 콘텐츠를 다양한 면에서 연관성을 보여 줄 수 있으며, 태그를 통해 손쉽게 콘텐츠를 검색하고, 분류하거나, 다른 콘텐츠와 엮어 네트워크로 만드는 일을 쉽게 해 준다. 네트워크를 가진 태그들이 온라인 통해 공개된다면 사용자의 개인

정보(개인의 흥미, 관심사에 대한 정보)를 관리할 수 있는 좋은 방법이 될 수 있다.[10]

2.2 소셜 북마킹(Social Bookmarking)

인터넷 북마크(Internet Bookmark)는 사용자가 되찾을 수 있는 웹 페이지의 위치를 저장한다. 사용자가 저장한 북마크는 사용자의 흥미, 관심에 대한 기록이다.[11] 즉, 북마킹(즐거찾기 추가)은 개인 한 사람의 기준으로 가치가 있다고 판단을 하여 이루어지는 개인적 행위이기 때문에 저장된 북마크들은 개인의 관심에 대한 집합으로 볼 수 있다.

소셜 북마킹 서비스라 불리는 웹 기반 북마킹 서비스는 사용자가 웹 서핑 중 북마크 할 가치가 있다고 생각되는 사이트, 웹 페이지 등을 발견하면 웹 브라우저가 아닌 원격 웹 서버(북마킹 서비스 서버)에 북마크를 추가하는 방법이다. 이에 사용자는 언제 어디서나 자신의 북마크를 관리할 수 있다. [12]

소셜 북마킹 서비스는 사용자에게 북마킹 할 수 있는 공간을 제공하며, 사용자가 저장한 북마크들을 정리, 공유한다. 사용자가 저장한 북마크는 일반적으로 공개되어 다른 사용자들과 공유된다. 대부분의 소셜 북마킹 서비스는 사용자에게 전통적인 브라우저 기반의 폴더 시스템 대신 태그로 사용자의 북마크들을 정리하도록 하고 있다. 이를 통해 사용자는 특정 태그를 선택하면 태그와 관련된 다른 사용자의 북마크를 볼 수 있으며, 다른 사용자와 네트워크를 맺을 수 있다. 또한 사용자들의 북마크들을 웹 피드(Web Feeds)¹⁾로 제공한다. 이는 새로운 북마크가 저장, 공유되며, 다른 사용자에 의해 태그가 달린 사실을 알 수 있도록 한다.

이런 서비스들의 발전과 사용자 수가 늘어남에 따라 북마크 고유의 기능을 넘어 북마크 평가, 주석달기 등의 추가 기능이 부여되며, 브라우저로부터 북마크들을 이출, 이입하는 기능, 웹 인용, 소셜 네트워크를 잇는 기능 등을 갖는다.[13]

2.2절에서 소개한 소셜 북마킹 서비스에 저장한 북마크의 태그들을 이용하면 태그가 가진 특성에 따라 사용자의 흥미에 부합하는 콘텐츠를 검색하여 제공할 수 있으며, 문자 정보뿐만 아니라 이미지, 동영상 등의 멀티미디어 정보도 제공할 수 있다.

3. 태그 수집

소셜 북마킹 서비스의 데이터를 수집하기 위해서는 각 서비스에서 제공하는 웹 피드를 이용할 수 있다. 많은 소셜 북

¹⁾ 웹 피드(Web Feed): 업데이트된 콘텐츠를 주기적으로 사용자에게 제공하는데 사용되는 데이터 형식

마킹 서비스는 웹 피드를 제공하여 서비스가 가진 데이터를 이용할 수 있도록 한다. 따라서 제공되는 웹 피드를 이용하여 사용자가 사용한 태그들을 수집한다.

사용자의 관심 범주를 추출하기 위해 가장 많이 사용된 순으로 상위 10개의 태그를 선정한다. 선정된 태그와 관련된 있는 태그들을 추출하기 위하여 10개의 태그들과 관련된 북마크들의 모든 태그들을 수집한다. 수집된 태그들은 각 북마크에 쓰인 빈도수에 따라 표 1과 같이 정리한다.

표 1. 북마크들의 태그 집합에 대한 예

태그 \ 북마크	A	B	C	D	...
cook	1	1	0	0	
recipe	1	1	1	0	
baking	1	0	1	0	
cake	0	0	1	0	
blog	1	1	0	1	
...					

4. 태그간의 동시 사용률 조사

3절에서 수집된 태그 집합을 바탕으로 태그들끼리의 동시 사용률을 조사한다. 이는 콘텐츠를 검색하는데 있어서 검색 키워드 조합을 위한 데이터로 사용된다. 동시 사용률 조사 방법으로는 통계학에서 사용되는 상관분석을 이용한다.

4.1 상관분석(Correlation Analysis)

상관분석이란 연구하고자 하는 변수들 간의 관련성을 분석하기 위해 사용되는 방법으로서, 하나의 변수가 다른 변수와 상관성이 있는지, 있다면 어느 정도 상관성이 있는지 알아보기 위한 통계기법이다. 상관분석에서 의미하는 상관관계는 선형적인 상관관계 즉, 어떤 변수가 A만큼 증가하거나 감소할 때 다른 변수도 A만큼 증가하거나 감소하는 것을 의미한다. 따라서 상관관계에는 양의(+) 상관과 음의(-) 상관 두 종류가 있다.[14]

본 논문에서는 두 태그간의 상관성 즉, 태그 간 동시 사용률을 조사하기 위해 피어슨 상관계수를 이용한다.

4.2 피어슨 상관계수(Pearson's Correlation Coefficients)

두 확률 변수 X 와 Y 간의 선형관계의 강도를 나타내는 측도인 피어슨 상관계수는 식(1)과 같이 정의된다.

$$\rho_{X,Y} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (1)$$

상관계수 $\rho_{X,Y}$ 의 범위는 -1에서 1사이의 값을 가지며, $\rho_{X,Y}$ 의 값에 1에 가까울수록 X 와 Y 사이에 양의 방향으로 직선관계가 강하게 존재한다는 것을 의미하며, $\rho_{X,Y}$ 의 값이 -1에 가까울수록 음의 방향으로 직선관계의 정도가 심하다는 것을 의미한다.[15] 일반적으로 $\rho_{X,Y}$ 가 0.5 이상이면 높은 상관관계를 가진다.

4.3 피어슨 상관계수를 통한 동시 사용률 조사

3절에서 수집된 태그 집합을 바탕으로 각 태그들끼리 피어슨 상관계수를 계산한다. 계산된 상관계수는 태그 간 동시 사용률로 볼 수 있으며 상관계수가 1인 경우 두 태그가 사용된 북마크들은 같다.

높은 상관관계를 의미하는 0.5 이상의 상관계수를 갖는 태그들은 하나의 북마크에 동시에 사용된 빈도수가 높음을 의미한다. 실제로 본 연구에서도 하나의 북마크에 동시에 사용된 태그들끼리 0.5 이상의 상관계수를 갖고 있으며, 복수의 북마크에서 동시에 사용된 태그들일수록 1에 가까운 결과가 나타났다.

0 이하의 상관계수(음의 상관계수)를 갖는 태그들은 같이 쓰이지 않았음을 의미한다. 따라서 이러한 태그들은 검색 키워드 조합 시 고려하지 않는다.

5. 동시 사용률을 이용한 검색 키워드 추출

사용자의 흥미에 부합하는 콘텐츠를 찾기 위한 키워드를 4절에서 조사된 동시 사용률을 바탕으로 태그들을 그룹화하여 만든다. 검색 키워드 그룹에는 검색 시 '모두 포함' 되어야 하는 키워드들과 '적어도 하나 포함' 되어야 하는 키워드들로 나눌 수 있다.

표 2. 태그 간 동시 사용률의 예

태그 \ 태그	A	B	C	D	E
A	1	0.9	0.3	0.6	0.1
B	0.9	1	0.5	0.3	0.7
C	0.3	0.5	1	-0.1	0.5
D	0.6	0.3	-0.1	1	0.2
E	0.1	0.7	0.5	0.2	1

5.1 검색 시 모두 포함해야 하는 키워드

표 2의 예에서 태그 A를 기준으로 B와 D가 0.5 이상의 동시 사용률을 갖는다. 이는 A와 B, D 키워드가 비교적 많은 빈도수로 같이 사용됐음을 의미한다. 따라서 검색 시 [A,B], [A,D], [A,B,D] 등과 같이 검색 키워드 그룹을 구성할 수 있다. 하지만, [A,B,D]의 경우 B와 D간의 동시 사용률을 비교했을 때 0.5 미만이기 때문에 [A,B,D] 세 키워드를 같이 사용하지 않는다. 따라서 A는 [A,B], [A,D]의 그룹만 구성이 가능하다.

태그 E의 경우 B와 C가 동시 사용률 0.5 이상이며, B와 C간의 동시 사용률 또한 0.5 이상이기 때문에 [E,B], [E,C], [E,B,C] 그룹 모두를 포괄하는 [E,B,C] 키워드 그룹을 갖는다.

5.2 검색 시 적어도 하나 포함해야 하는 키워드

태그 A를 기준으로 C와 E는 0.5미만 0이상의 동시 사용률을 보인다. 이 경우 5.1절에서 구성된 태그 그룹과 조합한다. [A,B] 그룹과 조합 시 C와 A, C와 B의 동시 사용률을 비교했을 때 0 이상으로 [A,B|C]의 그룹을 갖는다. E 역시 A,B 각각 0 이상의 동시 사용률을 갖기 때문에 [A,B|C,E] 그룹으로 조합된다. [A,B|C,E]의 의미는 A,B 키워드를 모두 포함하고 C와 E 중 적어도 하나를 포함하는 콘텐츠를 검색하는 키워드 그룹이다.

태그 E의 경우 A와 D가 0이상의 동시 사용률을 갖는다. 따라서 [E,B,C]그룹과 조합될 수 있는 후보가 된다. A는 B,C,E 모두 0이상의 동시 사용률로 [E,B,C|A]그룹을 갖지만 D는 C와 -0.1의 동시 사용률을 갖기 때문에 [E,B,C] 그룹에 속할 수 없다.

5.1절과 5.2절의 키워드 조합 규칙을 통해 표 2의 예로부터 나올 수 있는 키워드 그룹은 [A,B|C,E], [A,D|E], [B,C,E|A] 이다.

6. 구현 및 실험

본 연구에서 방법을 평가하기 위하여 소셜 북마킹의 태그를 이용한 개인화 콘텐츠를 제공하는 웹 어플리케이션을 개발하였다. 개발 환경으로는 Adobe Flex 기술을 이용하였으며, 소셜 북마킹 서비스 데이터로는 달리셔스 서비스(Delicious, <http://delicious.com/>)를 이용하여 한 사용자의 북마크 및 태그 데이터를 수집하였다. 또한 유튜브(YouTube, <http://youtube.com/>) 서비스의 API²⁾를 이용하

여 추출된 키워드로 동영상을 검색하여 개인화된 콘텐츠를 제공하였다.

cook	
recipe	0.8528028654224418
baking	0.47809144373375745
beef	0.33567254331867563
chicken	0.33567254331867563
shopping	0.33567254331867563
food	0.33567254331867563
blog	0.1712373798553933

baking	
shopping	0.7021094974994094
cook	0.47809144373375745
recipe	0.2548235957188128
blog	0.038525706426470976

recipe	
cook	0.8528028654224418
chicken	0.39361094683048226
beef	0.39361094683048226
baking	0.2548235957188128
blog	0.17180203318601237

food	
cook	0.33567254331867563
blog	0.1487709540894462

chicken	
recipe	0.39361094683048226
cook	0.33567254331867563
blog	0.1487709540894462

beef	
recipe	0.39361094683048226
cook	0.33567254331867563

shopping	
baking	0.7021094974994094
cook	0.33567254331867563

그림 2. 태그 간 동시 사용률 조사 결과

그림 1은 딜리셔스 서비스에서 수집된 한 사용자의 북마크와 태그를 바탕으로 동시 사용률을 조사한 결과이다.

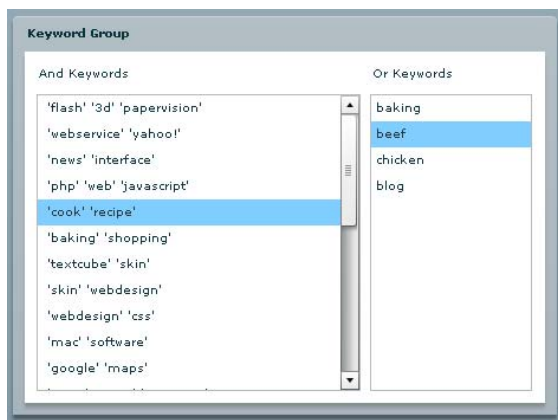


그림 2. 동시 사용률을 바탕으로 조합된 키워드 그룹

그림 2는 동시 사용률을 바탕으로 검색 키워드를 추출한 결과이다. 왼쪽 리스트는 '검색 시 모두 포함해야 하는 키워드' 그룹이며, 왼쪽 리스트 중 하나의 그룹을 선택하면 선택된 그룹의 '검색 시 적어도 하나 포함해야 하는 키워드' 들이 오른쪽 리스트에 나타나도록 하였다.

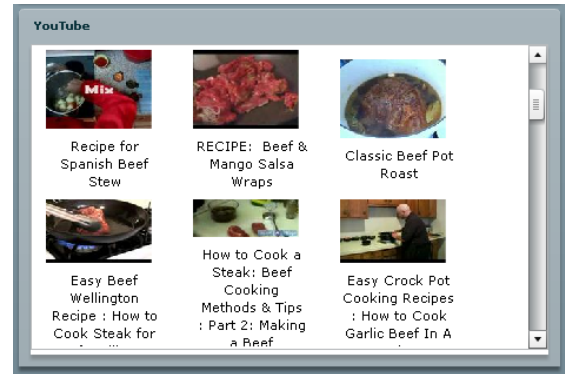


그림 3. 키워드 그룹을 통해 유튜브 동영상 검색 결과

그림 3은 유튜브 서비스에서 'cook'과 'recipe'를 모두 포함하고 'beef'를 적어도 하나 포함하는 동영상을 검색한 결과이다. 그림에서 나타난 바와 같이 소고기(beef) 요리(cook)에 관련된 조리방법(recipe)을 담은 동영상이 검색되었다.

이처럼 딜리셔스에 저장한 북마크와 태그 데이터를 바탕으로 자동으로 사용자의 흥미에 부합하는 동영상을 검색하여 제공할 수 있었다.

7. 결론

본 논문에서는 개인화 콘텐츠 제공을 위해 사용자가 인터넷에 축적한 북마크의 태그를 이용하여 사용자의 관심 범주를 추출하고 이를 바탕으로 사용자의 흥미에 부합하는 콘텐츠를 검색하기 위한 검색 키워드 조합 방법을 제시하였다. 이를 통해 사용자의 흥미와 관련 동영상 콘텐츠를 제공하는 어플리케이션을 개발하였다. 이는 사용자가 소셜 북마킹 서비스 이용만으로 사용자가 관심 있는 다른 콘텐츠를 제공받도록 하는데 의의가 있다.

이러한 방법으로 기존 사용되었던 개인화 콘텐츠를 제공하기 위해 사용자에게 미리 입력 받는 절차를 없애고 사용자가 웹을 사용하면서 축적한 데이터를 바탕으로 자동으로 개인화된 콘텐츠 서비스를 제공할 수 있을 것으로 기대된다.

2) API: Application Programming Interface.

참고문헌

- [1] Marko Balabanovic, Yoav Shoham, "Content-Based, Collaborative Recommendation", Communications of the ACM, pp.66~72, 1997.
- [2] Daniel Billsus, Michel J. Pazzani, "Learning Collaborative Information Fliters", the Fifteenth International Conference on Machine Learning, 1998.
- [3] Greg Linden, Brent Smith, Jeremy York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering", IEEE Internet Computing, 2003.
- [4] Yanfei Xu, Liang Zhang, "Personalized Information Service Based on Social Bookmarking", ICADL 2005, LNCS 3815, pp. 475~477, 2005.
- [5] Douglas Eck, Paul Lamere, Thierry Bertin-Mahieux, Stephen Green, "Automatic Generation of Social Tags for Music Recommendation", Advances in Neural Information Processing Systems, 2008.
- [6] Michael G. Noll, Christoph Meinel, "Web Search Personalization Via Social Bookmarking and Tagging", ISWC/ASWC 2007, LNCS 4825, pp. 367~380, 2007.
- [7] "Tag (metadata)",
[http://en.wikipedia.org/wiki/Tag_\(metadata\)](http://en.wikipedia.org/wiki/Tag_(metadata))
- [8] Thomas Vander Wal, "Folksonomy",
<http://www.vanderwal.net/folksonomy.html>, 2007.
- [9] Adam Mathes, "Folksonomies-Cooperative Classification and Communication Through Shared Metadata", Computer Mediated Communication, LIS590CMC, pp.4. 2004.
- [10] Adam Mathes, "Folksonomies-Cooperative Classification and Communication Through Shared Metadata", Computer Mediated Communication, LIS590CMC, pp.9. 2004.
- [11] Ben Lund, Tony Hammond, Martin Flack and Timo Hannay, "Social Bookmarking Tools (I): A General Review", D-Lib Magazine Vol. 11, No. 4, 2005
- [12] "Internet bookmark",
http://en.wikipedia.org/wiki/Internet_bookmark
- [13] Ben Lund, Tony Hammond, Martin Flack and Timo Hannay, "Social Bookmarking Tools (II): A Case Study - Connotea In", D-Lib Magazine Vol. 11, No. 4, 2005.
- [14] 김렬, 성도경, 이환범, 송건섭, 조태경, 이수창, "통계분석의 이해 및 활용", 대명, 2005. pp.163~165.
- [15] 조진관, "피어슨 상관계수 측정을 이용한 추천 알고리즘의 성능 비교분석", 동국대학교 석사학위논문, 2006.
- [16] Toby Segaran, Programming Collective Intelligence, O'Reilly Media, 2007, pp.54~82