

Extraction Method of Multi-User's Common Interests Using Facebook's 'like' List

Yeonju Lim[†] · Sangwon Park^{**}

ABSTRACT

The today's rapid spread of smartphones makes it easier to use SNS. However, it reveals only their daily life or interest. Therefore, it is hard to really get to know the detailed part of multi-user's common interests. This paper proposes a content recommendation system which recommends people wanted by identifying common interests through SNS. Recommendation system includes proposal formula considering people wanted and deviation in group. After simulation, the proposed system provide high-quality adapted contents to many users by recommendation item according to the common interest. Number of cases about formula are four. It recommend contents that they have many number of 'like' and few number of deviation in users. The proposed system proves by simulations of four cases and read user's 'likes' data. It provide high-quality adapted contents to many users by recommendation item according to the common interest.

Keywords : Common Interests Recommendation, Facebook 'like' List

페이스북의 ‘좋아요’ 리스트를 이용해 다중 공통 관심사항을 추출하는 기법

임 연 주[†] · 박 상 원^{**}

요 약

최근 스마트폰 발달로 인터넷 접근이 쉬워짐에 따라 소셜 네트워크 서비스(SNS)의 이용이 손쉬워졌다. 하지만 현재 SNS는 개인의 일상 또는 관심사 공유에 그치며 여러 사용자 간의 공통관심사 파악은 어렵다. 본 논문에서는 SNS를 통해 개인이 아닌 여러 사용자 간의 공통관심사를 파악하여 스마트폰을 통해 원하는 것을 추천해주는 콘텐츠 추천 시스템을 제안한다. 추천 시스템은 그룹 내 사용자들의 선호도와 편차를 고려하여 제안한 공식을 포함한다. 시뮬레이션 후 공식에 대해 나올 수 있는 경우는 4가지로 간추려졌다. 그 결과 개인의 선호도를 나타내는 ‘좋아요’ 수가 많으면서 페이스북 사용자들 간 선호도 편차가 적은 콘텐츠를 추천한다. 제안한 방법은 공식에 대한 4가지 경우의 시뮬레이션과 실제 페이스북 사용자들의 ‘좋아요’ 데이터로 증명한다. 제안 시스템은 그룹 내에서의 선호도와 편차를 고려하여 공통관심사를 추천해주기 때문에 양질의 맞춤형 콘텐츠를 제공한다.

키워드 : 공통관심사 추천, 페이스북 ‘좋아요’ 리스트

1. 서 론

인터넷과 IT기술의 급속한 발전으로 다양한 콘텐츠 생산 및 보급이 가능해지고 시공간을 초월하여 정보가 공유될 수 있는 스마트 생활환경으로 진입하게 되었다. 음성과 문자메시지와 같은 기본적인 커뮤니케이션 기능과 더불어 스마트

폰의 빠른 확산은 공유, 참여, 개방, 창조를 근간으로 하는 소셜 네트워크 서비스(Social Network Service, SNS) 이용을 손쉽게 만들었다[1].

대다수의 사람들이 웹 콘텐츠를 얻기 위해 더 이상 검색을 사용하지 않고, SNS 상에서 자신과 관심을 공유하는 지인이나 전문가들에게 의존하는 방식을 선호하게 되었다[2]. SNS 이용률은 전년과 비교하여 다소 높게 조사되었다. 이는 SNS가 공유 및 확산기능을 가지고 있어 사람들 간에 원활한 쌍방향 소통이 용이하고 정보 공유가 편리하다는 장점이 있기 때문이다[2].

하지만 현재의 SNS는 사용자들이 자신의 생각과 좋아하

※ 본 연구는 2015학년도 한국외국어대학교 교내학술연구비에서 지원되었음.

† 준 회 원 : 한국외국어대학교 정보통신공학과 학사과정

** 종 신 회 원 : 한국외국어대학교 정보통신공학과 교수

Manuscript Received: November 5, 2014

First Revision: January 27, 2015

Accepted: January 30, 2015

* Corresponding Author: Sangwon Park(swpark@hufs.ac.kr)

는 관심사 등을 표현하고 공유하는 데만 그친다. 오늘날 우리 사회는 개인생활뿐 아니라 그룹생활 전반에 걸쳐 이미 스마트 환경으로 진입했기 때문에, SNS 활용이 그룹 내의 소통방식에 큰 변화를 가져올 수 있을 것으로 추정된다[3]. 그리고 그룹 내 소통은 조직문화를 형성하는 전제 조건으로서도 매우 중요하다. 그룹 간 의사소통이 중요한 현대사회에서 개인이 아닌 여러 사람 간의 공통분모를 추천해주는 메커니즘이 필요하다.

본 논문에서 소개하는 콘텐츠 추천 애플리케이션은 페이스북(Facebook)을 기반으로 사용자들의 공통관심사를 추출하여 콘텐츠를 추천해주는 시스템이다. 페이스북을 선정하는데는 2가지 이유가 있다.

첫째, 국내 한 커뮤니케이션 전문 매체의 조사 결과에 따르면, 국내 기업 및 기관 소셜미디어 담당자 120명을 대상으로 소비자 커뮤니케이션을 위해 가장 비중을 두는 SNS에 대해 조사한 결과 페이스북이 전체 응답자의 91%로 압도적으로 높게 나타났다[4]. 이처럼 페이스북은 많은 대중에 사용하고 있다.

둘째, 페이스북에는 페이스북에만 존재하는 강력한 피드백의 기능인 '좋아요' 버튼이 있다. 사용자는 '좋아요' 버튼을 클릭해서 쉽게 페이지의 게시물을 뉴스피드에서 받아볼 수 있고, 사용자들이 관심 있어 하는 정보를 '좋아요' 버튼 하나로 표현할 수 있다. '좋아요' 버튼의 클릭만으로 담벼락을 통해 친구들의 뉴스피드로 자신의 관심사를 빠르게 전파시킬 수 있다. 또한, 담벼락과 댓글 등록 시 실시간 알림서비스가 제공되어 피드백의 신속한 확인이 가능하다[5].

본 논문에서는 페이스북을 이용하여 추출한 개인의 선호도 정보를 바탕으로 여러 사용자들이 공통으로 관심 있어 하는 것을 제공해주기 위해 방법을 제안하였다.

본 논문에서 제안하는 시스템은 페이스북을 이용해 추출한 여러 사용자의 관심사를 제안한 공식에 적용하여 공통관심사 콘텐츠를 추천하는 형태이다. 현재까지의 SNS는 개인의 선호도만 파악하는 데 중점을 두었다. 하지만 본 논문은 개인이 아닌 그룹 내 공통관심사를 파악함으로써 원하는 것을 추천해준다는 장점을 가지고 있다.

논문의 구성은 다음과 같다. 2절에서는 관련 연구에 대하여 기술하고 3절에서는 페이스북을 이용한 개인별 관심사 수집 방법과 Architecture에 대해 설명한다. 4절에서는 본 논문에서 제안한 공통관심사 추출과정을 자세히 설명한다. 5절에서는 본 논문에서 제안한 방법으로 특정 콘텐츠를 예시로 들어 실험한 결과와 결과에 대한 검증을 한다. 마지막으로 6절에서는 결론에 대해 소개한다.

2. 관련 연구

클러스터 기반의 다중 콘텐츠 추천은 사용자의 콘텐츠 소비성향을 파악하여 사용자별로 선호하는 콘텐츠들을 예측하는 알고리즘이다. 각 사용자들의 사용 히스토리를 분석하여 추출한 패턴들을 기반으로 해당 사용자의 선호정보를 유추하

여 추출한다[6]. 본 논문에서는 사용자들의 히스토리를 분석함으로써 그룹 내 공통관심사 콘텐츠 추천을 제공할 수 있도록 한다.

사용자기반 협업적 필터링기법은 사용자 간의 유사성을 측정하여 추천 대상 사용자와 유사한 이력을 보이는 이웃 사용자들의 아이템에 대한 선호를 바탕으로 추천 대상 사용자에게 아이템을 추천하는 방법이다. 사용자는 의사결정을 하기 위해 자신의 관심사와 비슷한 다른 사용자의 의견에 큰 영향을 받는다[7]. 하지만 본 논문에서는 사용자 기반 협업적 필터링 개념을 기반으로 개인이 아닌 그룹에 공통관심사를 제시하고자 한다.

James[8]는 세계에서 가장 인기 있는 온라인 커뮤니티인 Youtube를 사용하여 비디오를 추천해주는 시스템이다. Youtube 사이트에서 활동했던 데이터(비디오 시청 기록, '좋아요' 누른 비디오)를 기반으로 개인에게 맞춤형 비디오를 추천해준다[8].

동적경쟁 추천 알고리즘(DCR 알고리즘)은 환경 변화에 취약한 단일접근 방식의 한계를 극복하기 위한 알고리즘이다[9]. 기존 추천 알고리즘은 사용자의 특성을 고려하지 않아 사용자 특성이 변할 때 일관성 있고 안정된 추천을 제공하기 어렵다. 하지만 DCR 알고리즘은 여러 추천 알고리즘 중 경쟁력 있는 방법을 선택하여 최상의 추천 결과를 제공하는 것이 목표이다[9]. 예를 들어, 사용자의 선호도 특성이 변할 때 기존의 추천 방법인 클러스터 기반의 다중 콘텐츠 추천 기법과 사용자 기반 협업적 필터링 기법 중 더 정확한 추천 방식을 선택하여 최상의 콘텐츠를 추천해준다.

Map-Reduce video 추천 시스템은 사용자에게 의미 있고 흥미로운 비디오를 추천해주기 위해 아이템 기반 협업적 필터링을 포함한다[10]. 사용자 기반의 협업적 필터링은 추천 대상 사용자와 비슷한 선호를 가진 사용자의 의견이 의사결정에 큰 영향을 미친다는 것에 기반했다면, 아이템 기반 협업적 필터링은 추천 대상 사용자가 선호했던 아이템과 비슷한 아이템을 선호할 것이라는 것에 기반한 아이템 선별 방법이다. MySpace의 모든 비디오 데이터를 분석하여 개인에게 최적화된 맞춤형 비디오를 추천해준다[10].

Table 1은 기존 관련 연구를 분석한 표이다. 기존 관련 연구는 모두 추천 대상이 개인이다. Shin[6]은 추천 콘텐츠 예시로 음악, 방송, 광고, 기사를 사용하였고 Jean[7]은 추천 콘텐츠 예시를 들지 않았다. James[8]와 Yu[9]의 논문은 Video 예시로 사용하였다. Jean[7]의 목적은 기존 추천 시스템의 성능의 향상이지만 Jean[7]을 제외한 나머지 논문은 개인에게 최적화된 콘텐츠를 추천해주는 것이 목적이다. Shin[6]과 Jean[7]을 제외한 나머지 논문은 SNS를 이용하여 논문에 대한 검증이 이루어졌다.

Table 2는 기존 관련 연구와 본 논문의 차이점에 대해 정리한 표이다. 기존 관련 연구는 특정한 사용자에게 콘텐츠를 추천하였다면 본 논문은 여러 사용자 간 공통관심사를 파악하여 그룹 내 콘텐츠를 추천해준다. 그 수단으로는 페이스북을 이용한다. 그룹 내 사용자 간의 관심사 유사도와 편차를 고려하여 그룹에게 공통관심사를 추천해준다.

Table 1. Related Research Comparison

	The multiplex contents recommendation based on clusters	User-based Collaborative Filtering	James[8]	DCR Algorithm	Map-Reduce Video recommendation system
Recommendation type	Personal	Personal	Personal	Personal	Personal
Examples of recommended contents	Music, Broadcast, Advertising, Article	X	Video	X	Video
Purpose	Development of Personalized content Intelligent Recommendation Agent	Improvement of existing recommendation system	Personalized video recommendation	Personalized contents recommendation	Personalized video recommendation
data for verification	X	A-Data	Youtube Data	Twitter Data	MySpace Data

Table 2. Difference between Existing Related Researches and this Paper

Existing Relative Researches	This Paper
Contents recommendation for a certain user	Contents recommendation for group
Similarity measurement after analyzing a user’s past purchase information	Identify user preferences by analyzing the list of Facebook ‘Like’
Contents recommendation for recommended target users based on the similar history	Contents recommendation with consideration for degree of similarity and difference between the group user preference

3. 페이스북을 이용한 개인 관심사 수집

본 논문에서는 개인이 아닌 그룹 내 공통관심사를 추천하고자 한다. 그러기 위해서는 개인의 선호도를 먼저 파악해야 한다. 여러 SNS 중에서도 페이스북을 선택한 이유는 국내 SNS 사용자 중 91%가 페이스북을 이용하고 있으며 페이스북은 개인의 선호도 파악이 가능하다는 장점을 가지고 있기 때문이다[4].

본 논문에서는 페이스북 사용자들의 ‘좋아요’ 리스트 분석 결과를 개인의 선호도 파악에 활용하고, 이를 통해 그룹 내 공통관심사를 추출할 것이다. 다음은 페이스북 사용자들의 ‘좋아요’ 리스트를 추출하는 방법에 대한 설명이다.

3.1 페이스북 데이터 수집을 위한 API

온라인 혹은 오프라인 상에서 관계를 가지고 있는 사용자들은 많은 생각과 콘텐츠들을 서로 공유하게 되고, 어느 정도 상관관계를 보이게 된다. 그러므로 본 연구는 페이스북에서 관계를 맺는 사용자들을 그룹으로 만들어 이들의 사용 히스토리를 분석해 추천에 반영하고자 한다.

페이스북 데이터를 수집하기 위한 API로는 Graph API가 있다. 페이스북 Graph API를 사용하려면 App ID(앱 ID)가 필요하다. App ID는 마치 회원가입을 하듯이 이루어진다.

페이스북의 핵심은 Social Graph이다. Fig. 1은 페이스북의 모든 객체(object)들에 대한 연결(connection) 관계를 나타낸 것이다. 페이스북의 Social Graph는 꼭지점에 해당하는 객체(object)와 변에 해당하는 연결(connection)로 구성되어 있다.

객체에 해당하는 것은 사용자, 페이지, 이벤트, 그룹, 어플, 상태 메시지, 사진, 동영상, 노트 등이다. 이들은 독립적인 객체(object)로 저마다 고유한 ID를 가지고 있다.

객체(object)는 각각 따로 존재하는 것이 아니라 서로 간에 연결(connection)되어 있다. 각각의 객체마다 연결될 수 있는 객체의 종류가 정해져 있다.

Fig. 2는 페이스북 사용자의 ‘좋아요’ 리스트를 추출한 그림이다. category는 Food 또는 Cuisine인 것으로 추출하였다.

하나의 객체가 가진 연결(connect)들의 목록을 얻기 위해서는 다음 주소로 접속한다.

[http://graph.facebook.com/\(객체의 ID\)/\(객체의 연결 이름\)](http://graph.facebook.com/(객체의 ID)/(객체의 연결 이름))

Fig. 2의 경우에는 [http://graph.facebook.com/\(100002297444231\)/\(like\)](http://graph.facebook.com/(100002297444231)/(like)) 주소로 접속한다. 페이스북 서버에서는 해당 ID의 ‘좋아요’ 리스트를 담은 json파일을 출력한다.

3.2 페이스북 데이터 수집 방법

여러 사용자의 공통관심사를 제시하기 위해서는 각 사용자의 ‘좋아요’ 리스트 추출이 필요하다. 본 논문에서는 페이스북 PHP SDK를 이용하여 사용자의 ‘좋아요’ 리스트를 추출할 것이다.

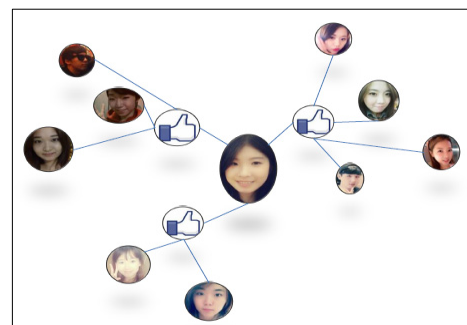


Fig. 1. Social Graph of Facebook

```
{
  "category": "Food/beverages",
  "category_list": [
    {
      "id": "187462324610381",
      "name": "Food Consultant"
    },
    {
      "id": "193831710644458",
      "name": "Italian Restaurant"
    },
    {
      "id": "108366235907857",
      "name": "Newspaper"
    }
  ],
  "name": "Italian Food",
  "created_time": "2014-08-16T02:52:32+0000",
  "id": "94422420656"
},
{
  "category": "Cuisine",
  "name": "Pizza",
  "created_time": "2014-08-16T02:49:50+0000",
  "id": "169671689754628"
},
}
```

Fig. 2. 'Like' List Using Graph API

페이스북 사용자들은 고유한 App ID와 Secret code값을 갖고 있다. App ID와 Secret code값으로 로그인한 상태의 사용자 정보를 요청한다. Facebook Database에 저장된 '좋아요' 리스트를 추출해내어 웹 사이트로 보이게 된다[11].

추출된 '좋아요' 리스트는 원하는 대로 추출 가능하다. 본 논문에서는 그룹 내 공통관심사 추출 방법을 제안하는 데 음식을 예로 든다. 즉 사용자의 '좋아요' 리스트 중에서 음식만 추출하여 text파일로 저장하였다.

3.3 Architecture

Fig. 3은 다중 공통관심사항 추천 설계를 위한 전체 구상도이다. 스마트폰을 이용하여 개인의 선호도 추출을 가정한다. 페이스북 사용자들이 주위에 가까이 있다면 블루투스 등을 이용하여 서로를 감지하여 그룹을 이룰 수 있다. 하지만 본 논문에서는 페이스북 ID를 서로 공유함으로써 그룹을 이룬다.

페이스북 사용자들은 자신만의 페이스북 ID를 가지고 있다. 각각의 페이스북 ID를 한 명의 페이스북 사용자에게 모두 보낸다. 그룹 내 페이스북 사용자들의 ID를 모두 가지고 있는 사람은 차례로 페이스북 ID를 페이스북 서버로 보낸다. 페이스북 서버에서는 해당 ID에 맞는 '좋아요' 리스트를 보내준다. 본 논문에서는 수집된 사용자들의 '좋아요' 리스트를 분석하여 그룹 내 공통관심사를 추천하는 프레임 워크를 제안하고자 한다.

4. 공통관심사 추출

그룹 내에서 선호하는 관심 사항을 추천해주기 위해 4단

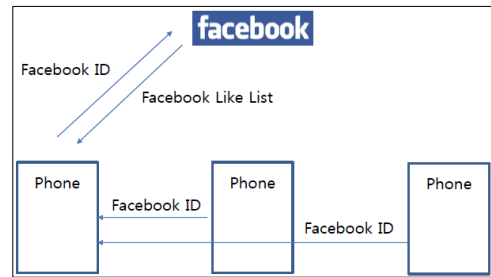


Fig. 3. Whole Diagram

계를 제안한다. 본 논문에서 쉽게 설명하기 위해 음식을 한 예로 설정하였다.

페이스북 사용자들은 각자 자신이 선호하는 음식을 가지고 있고 선호하는 음식에 대해 '좋아요' 버튼을 누른다.

첫 번째 단계에서는 페이스북 사용자들의 '좋아요' 리스트를 카테고리화한다. 음식이라는 큰 범주를 추천해주시기보다는 좀 더 세부적으로 한식, 중식, 양식, 일식, 패스트푸드, 분식으로 카테고리화한다.

두 번째 단계에서는 페이스북 사용자들마다 '좋아요' 개수가 모두 다를 수 있으므로 총 개수를 1로 노멀라이징한다.

Table 3. Example of Normalizing

	Korean Food		Chinese Food		Total	
	n	rate(%)	n	rate(%)		
user A	2	0.2	8	0.8	10	1
user B	20	0.2	80	0.8	100	1

Table 3은 노멀라이징이 왜 필요한지를 설명해주는 표이다. 카테고리를 한식과 중식으로 선정하고, 사용자 A의 총 '좋아요' 개수는 10개이며 한식은 2개, 중식은 8개라 가정하였다. 사용자 B의 '좋아요' 총 개수는 100개이며 한식은 20개, 중식은 80개라 가정하였다. 한식의 개수만 본다면 사용자 A는 2개, 사용자 B는 20개로 사용자 B의 한식에 대한 선호도가 상당히 높은 것처럼 보인다. 하지만 '좋아요' 총 개수를 고려하여 비율로 따진다면 한식에 대한 선호도 비율은 사용자 A와 사용자 B 모두 0.2%로 동일하다. 따라서 총 '좋아요' 개수를 1로 노멀라이징해주는 과정이 꼭 필요하다.

세 번째 단계에서는 공통관심사를 추출한다. 본 논문에서는 그룹 내 각 페이스북 사용자들의 카테고리별 편차와 '좋아요' 개수를 고려하여 연산하였다.

$$c_k = \frac{likes_k}{\sigma_k} \tag{1}$$

- * c_k : k 번째 카테고리 결과
- $likes_k$: k 번째 카테고리의 '좋아요' 수
- σ_k : 모든 사용자에 대한 k 번째 카테고리 표준편차

c_k 는 페이스북 사용자의 '좋아요' 수와 그룹 내 선호도 편차를 고려하여 계산한 결과값이다. k 번째 카테고리에 대한

Table 4. Validation Data of Experiment 1

Food / %	user 1	user 2	user 3	user 4	$likes_k$	σ_k	c_k	u_i
Korean	0.188	0.188	0.063	0.063	0.502	0.063	7.968	0.078
Chinese	0.188	0.250	0.312	0.188	0.938	0.052	18.04	0.042
Western	0.063	0.063	0.125	0.250	0.502	0.240	2.092	0.164
Japanese	0.063	0.063	0.063	0.188	0.377	0.055	6.855	0.046
Fast food	0.250	0.312	0.375	0.250	1.187	0.052	22.83	0.000
Flour based	0.250	0.125	0.063	0.063	0.501	0.076	6.592	0.052
Total	1	1	1	1				

Table 5. Validation Data of Experiment 2

Food / %	user 1	user 2	user 3	user 4	$likes_k$	σ_k	c_k	u_i
Korean	0.188	0.250	0.250	0.188	0.876	0.032	23.37	0.001
Chinese	0.125	0.063	0.125	0.313	0.626	0.094	6.660	0.787
Western	0.188	0.312	0.188	0.188	0.876	0.055	15.93	0.007
Japanese	0.125	0.125	0.250	0.063	0.563	0.212	2.656	0.066
Fast food	0.313	0.063	0.125	0.063	0.564	0.102	5.530	0.057
Flour based	0.063	0.188	0.063	0.188	0.502	0.063	7.968	0.024
Total	1	1	1	1				

‘좋아요’ 수는 많고, 모든 사용자에게 대한 k 번째 카테고리 표 준편차는 작은 카테고리를 추천하는 방식이다. 결과적으로 가장 큰 c_k 를 선택한다.

$$m_k = \frac{\sum_{i=1}^t p_k(i)}{t} \tag{4}$$

$$likes_k = \sum_{i=1}^t p_k(i) = \sum_{i=1}^t \frac{n_k(i)}{\sum_{j=1}^s n_j(i)} \tag{2}$$

- * t : 사용자 수
- $p_k(i)$: 사용자 i 의 모든 카테고리에 대한 k 번째 카테고리 비율
- $n_k(i)$: 사용자 i 의 k 번째 카테고리 ‘좋아요’ 수
- s : 카테고리 종류 수

$likes_k$ 는 k 번째 카테고리 ‘좋아요’ 수를 나타낸다. 페이스북 사용자 각각의 모든 카테고리에 대한 k 번째 카테고리 비율을 모두 구한 후 더한 결과이다.

$p_k(i)$ 는 사용자 i 의 k 번째 카테고리 ‘좋아요’ 수를 사용자 i 의 모든 카테고리 ‘좋아요’ 수로 나눈 결과이다. 즉 해당 사용자 i 의 k 번째 카테고리 선호도 비율과 같다.

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^t (p_k(i) - m_k)^2}{t}} \tag{3}$$

- * m_k : 모든 사용자의 k 번째 카테고리 ‘좋아요’ 평균

σ_k 는 k 번째 카테고리에 대해 그룹 내 페이스북 사용자 간 편차를 나타낸다. σ_k 가 크다면 그룹 내 페이스북 사용자들의 k 번째 카테고리에 대한 선호도 차이가 크다는 것을 의미한다.

본 논문에서는 그룹 내 공통관심사를 파악해주는 메커니즘을 제안하는 것이 목표이기 때문에 그룹 내 페이스북 사용자들의 k 번째 카테고리에 대한 선호도 차이가 크다면 공통 선호도라고 간주하지 않는다.

m_k 는 모든 페이스북 사용자의 k 번째 카테고리 ‘좋아요’ 평균을 의미한다. k 번째 카테고리에 대한 그룹 내 페이스북 사용자들의 선호도 평균을 기준으로 선호도 차이가 큰지 작은지를 파악할 수 있다.

네 번째 단계는 c_k 가 가장 큰 카테고리를 선택하여 그룹 내 공통관심사로 추천해준다. c_k 가 클수록 $likes_k$ 는 c_k 에 비례하면서 σ_k 는 반비례한다.

본 논문은 4단계에 걸쳐 그룹 내 공통관심사를 파악하는 메커니즘에 대해 제안하였다. 결과적으로 그룹 내 페이스북 사용자들의 ‘좋아요’ 수는 크면서 그룹 내 각 페이스북 사용자들의 ‘좋아요’ 수 차이는 크지 않은 카테고리를 추천해주는 방법이다.

$$U_k = (\sum_{i=1}^t |p_k(i) - p_s(i)|) \times \sigma_k \tag{5}$$

- * $p_s(i)$: 사용자 i 의 추천된 카테고리 비율

u_k 는 k 번째 카테고리에 대한 사용자들의 불만족도 비율이다. Equation (1)에 대한 정확도 검증을 위한 공식이다. 가장 낮은 u_k 값을 가진 카테고리(가장 높은 c_k 값을 가진 카테고리)가 일치함을 보여줌으로써 Equation (1)의 정확도를 보여준다.

5. 실험 및 평가

페이스북 월간 사용자 수는 10억 명이 넘는다. 사용자들의 50% 이상이 글을 게재한 후 20분 내에서 1시간 이내에

Table 6. Validation Data of Experiment 3

Food / %	user 1	user 2	user 3	user 4	$likes_k$	σ_k	c_k	u_i
Korean	0.125	0.438	0.188	0.188	0.939	0.120	7.825	0.022
Chinese	0.125	0.063	0.188	0.125	0.501	0.045	11.13	0.017
Western	0.188	0.125	0.250	0.188	0.751	0.045	16.69	0.009
Japanese	0.188	0.063	0.063	0.188	0.502	0.063	7.968	0.024
Fast	0.250	0.188	0.250	0.188	0.876	0.032	27.38	0.006
Flour based	0.125	0.125	0.063	0.125	0.438	0.017	25.76	0.007
Total	1	1	1	1				

Table 7. Validation Data of Experiment 4

Food / %	user 1	user 2	user 3	user 4	$likes_k$	σ_k	c_k	u_i
Korean	0.333	0.100	0.260	0.250	0.943	0.241	3.914	0.070
Chinese	0.100	0.100	0.080	0.133	0.310	0.048	6.458	0.023
Western	0.100	0.100	0.110	0.090	0.400	0.007	57.14	0.003
Japanese	0.100	0.200	0.170	0.130	0.600	0.038	1.579	0.011
Fast	0.100	0.300	0.133	0.130	0.660	0.079	8.354	0.030
Flour based	0.222	0.222	0.230	0.210	0.884	0.009	98.27	0.002
Total	1	1	1	1				

반응을 보인다[12]. 페이스북 사용자 수가 많고 ‘좋아요’ 버튼을 누르는 반응도 모두 제각각이기 때문에 사용자들의 ‘좋아요’ 총 개수를 1로 노멀라이징하여 추출하였다.

페이스북 사용자가 누른 ‘좋아요’ 리스트에서 음식이라는 범주를 모두 추출하여 6가지 카테고리(한식, 중식, 양식, 일식, 패스트푸드, 분식)로 분류하였다. 카테고리는 ‘전국 음식점 주소록 2014’를 기반으로 분류하였다.

4명을 하나의 그룹으로 가정하고 시뮬레이션하였다. 시뮬레이션한 데이터를 분석한 결과 공식에 대해 나올 수 있는 경우는 4가지로 간추려졌다. 본 논문에서 제안하는 메커니즘은 시뮬레이션 데이터와 실제 페이스북 사용자의 데이터로 검증할 것이다.

실험 1: 그룹 내 페이스북 사용자 간 카테고리 선호도 편차가 동일할 경우 ‘좋아요’ 수를 고려한다.

Table 4는 실험 1에 대한 검증 데이터이다. 중식과 패스트푸드는 σ_k 가 0.052로 동일하다. 즉, 4명의 페이스북 사용자들이 중식을 선호하는 편차와 패스트푸드를 선호하는 편차가 동일하다는 것을 의미한다. 편차가 가장 작으면서 동일하기 때문에 어떤 카테고리를 그룹 내 공통 선호도로 추천해주어야 할지 선택하기 어렵다. 이런 경우 ‘좋아요’ 수를 고려해 제안한 Equation (1)에 적용하면 공통으로 선호하는 카테고리를 그룹 내 추천해줄 수 있다.

중식 카테고리를 Equation (1)에 적용했을 때 c_k 는 18.4로 두 번째로 높다. 패스트푸드 카테고리를 Equation (1)에 적용했을 때 22.83으로 가장 높다. 따라서 그룹 내 공통 선호도는 패스트푸드이다. 또한 패스트푸드는 u_i 가 0으로 가장 낮은 불만족도를 보인다.

실험 2: 카테고리에 대한 ‘좋아요’ 수가 동일할 경우 그룹 내 페이스북 사용자 간 카테고리 선호도 편차를 고려한다.

Table 5는 실험 2에 대한 검증 데이터이다. 한식과 양식은 $likes_k$ 가 0.876으로 동일하다. 즉 4명의 페이스북 사용자들이 한식을 좋아하는 ‘좋아요’ 수와 양식을 좋아하는 ‘좋아요’ 수가 동일하다는 것을 의미한다.

‘좋아요’ 수가 가장 많은 카테고리를 공통관심사로 선택하면 한식과 양식이 될 수 있다. 하지만 페이스북 사용자들의 해당 카테고리에 대한 선호도 편차가 크다면 공통관심사라고 결정할 수 없다. 따라서 선호도 편차를 모두 고려해야 한다.

한식 카테고리를 제안한 Equation (1)에 적용했을 때 c_k 는 23.37로 가장 높기 때문에 그룹 내 공통 선호도로 추천해준다. 또한 u_i 가 0.001로 불만족도가 가장 낮다. 양식 카테고리를 제안한 Equation (1)에 적용했을 때 c_k 는 15.93으로 한식보다 낮은 것을 볼 수 있다.

실험 3: 그룹 내 페이스북 사용자 간 카테고리 선호도 편차가 가장 작은 카테고리가 무조건 공통 선호도는 아니다.

Table 6은 실험 3에 대한 검증 데이터이다. 분식 카테고리는 σ_k 가 0.017로 가장 낮다. 그룹 내 4명의 페이스북 사용자들이 분식 카테고리에 대해 고루 좋아한다는 것을 의미한다. 하지만 편차가 가장 낮은 분식을 그룹 내 공통관심사로 제시한다면 이는 정확하지 않다. 왜냐하면 분식의 $likes_k$ 는 0.438로 다른 카테고리에 비해 선호도가 낮기 때문이다.

분식 카테고리를 제안한 Equation (1)에 적용했을 경우

Table 8. Validation Data of Experiment 1 to 5 Using Real Data

Food / %	user1	user2	user3	user4	user5	user6	user7	user8	user9	user10
Korean	0.333	0.077	0.091	0.727	0.625	0.429	0.286	0.500	0.400	0.333
Chinese	0.083	0.231	0.000	0.091	0.000	0.000	0.000	0.000	0.200	0.000
Western	0.250	0.308	0.181	0.000	0.250	0.286	0.143	0.167	0.000	0.000
Japanese	0.167	0.000	0.181	0.091	0.000	0.000	0.000	0.000	0.200	0.500
Fast	0.083	0.154	0.364	0.000	0.000	0.000	0.571	0.167	0.000	0.000
Flour based	0.083	0.231	0.181	0.091	0.125	0.286	0.000	0.167	0.200	0.167
Total	1	1	1	1	1	1	1	1	1	1
Food / %	user11	user12	user13	user14	user15	user16	$likes_k$	σ_k	c_k	u_i
Korean	0.444	0.571	0.667	0.600	0.500	0.375	6.958	0.181	38.44	0.173
Chinese	0.111	0.143	0.000	0.000	0.167	0.125	1.151	0.062	18.57	0.380
Western	0.111	0.143	0.167	0.000	0.000	0.000	2.006	0.078	25.72	0.436
Japanese	0.111	0.000	0.000	0.200	0.333	0.375	2.158	0.143	15.10	0.760
Fast	0.000	0.000	0.000	0.200	0.000	0.125	1.664	0.137	12.15	0.890
Flour based	0.222	0.143	0.167	0.000	0.000	0.000	2.063	0.061	33.82	0.328
Total	1	1	1	1	1	1				

c_k 는 25.76으로 두 번째로 높다. 패스트푸드를 제안한 Equation (1)에 적용했을 때는 27.38로 가장 높다. 따라서 패스트푸드를 그룹 내 공통관심사로 추천해준다. 패스트푸드는 u_i 가 0.006으로 불만족도가 가장 낮다.

한식 카테고리는 $likes_k$ 가 0.939로 가장 높지만, σ_k 는 다른 카테고리에 비해 상당히 높기 때문에 c_k 가 작게 나타났다. 즉, 한식 카테고리에 대한 선호도는 높지만 4명의 페이스북 사용자가 고루 좋아하지는 않는다.

양식 카테고리는 $likes_k$ 가 0.751, σ_k 가 0.045, c_k 는 16.69로 나타났다. c_k 가 세 번째로 큰 수이기 때문에 4명의 페이스북 사용자가 3번째로 공통으로 선호하는 카테고리인 것을 알 수 있다.

실험 4: 카테고리 선호도를 나타내는 ‘좋아요’ 수가 가장 많다고 해서 무조건 공통 선호도는 아니다.

Table 7은 실험 4에 대한 검증 데이터이다. 한식 카테고리의 $likes_k$ 수는 0.943으로 가장 높다. 그룹 내 4명의 페이스북 사용자가 가장 많이 선호한다는 것을 의미한다. 하지만 가장 많이 선호하다고 해서 한식을 그룹 내 공통관심사라고 제시한다면 이는 정확하지 않다. 왜냐하면 그룹 내 페이스북 사용자 4명 중 한 명의 선호도가 0.943일 수도 있기 때문이다. 공통관심사라는 것은 모두가 고루 좋아하는 것을 의미하기도 한다.

한식 카테고리를 제안한 Equation (1)에 적용했을 경우 c_k 는 39.14로 다소 낮은 편이다. 가장 많이 선호하는 카테고리이긴 하지만 페이스북 사용자 4명이 고루 선호하지는 않기 때문이다.

분식 카테고리를 제안한 Equation (1)에 적용했을 경우 c_k 는 98.27로 가장 높게 나타났다. 분식 카테고리에 대한 선호도가 가장 높지는 않지만 높은 편이면서 고루 좋아하기 때문에 그룹 내 공통관심사로 추천한다.

양식의 $likes_k$ 수는 0.400, σ_k 는 0.007, c_k 는 57.14로 c_k 는 2

번째로 높다. 그룹 내 4명의 페이스북 사용자가 2번째로 선호하는 카테고리인 것을 알 수 있다.

4가지 실험의 검증데이터를 통해 개인의 선호도를 나타내는 ‘좋아요’ 수가 많으면서 페이스북 사용자들 간 선호도 편차는 적은 콘텐츠를 추천하는 것을 볼 수 있다.

Table 8은 시뮬레이션 결과가 아닌 20대에서 50대까지 연령대별로 4명씩 16명을 그룹으로 이루어 실험한 실제 데이터이다.

실험 3은 그룹 내 페이스북 사용자 간 카테고리 선호도 편차가 가장 작은 카테고리가 무조건 공통 선호도는 아니라는 것을 보여준다. Table 8이 실험 3과 일치하는 것을 볼 수 있다.

분식 카테고리는 σ_k 가 0.061로 가장 낮다. 그룹 내 4명의 페이스북 사용자들이 분식 카테고리에 대해 고루 좋아한다는 것을 의미한다. 하지만 편차가 가장 낮은 분식을 그룹 내 공통관심사로 제시한다면 이는 정확하지 않다. 왜냐하면 분식의 $likes_k$ 는 2.063으로 다른 카테고리에 비해 선호도가 낮기 때문이다.

분식 카테고리를 제안한 Equation (1)에 적용했을 경우 c_k 는 33.82로 두 번째로 높다. 한식을 제안한 Equation (1)에 적용했을 때는 38.44로 가장 높다. 따라서 한식을 그룹 내 공통관심사로 추천해준다. 한식의 불만족도 u_i 는 0.173으로 가장 낮다. 그룹 내 공통관심사로 추천한 카테고리인 것과 불만족도가 가장 낮은 카테고리가 일치함으로써 제안한 방법의 정확도를 보여준다.

6. 결론

최근 스마트폰의 빠른 확산은 참여와 공유를 바탕으로 하는 SNS의 이용을 쉽게 한다. 이에 따라 SNS에 대한 연구들의 활발히 진행되고 있으며 이 논문에서도 SNS를 활용한다.

SNS 중에서도 페이스북은 많은 대중들이 사용하고 있고, 강력한 피드백 기능인 ‘좋아요’ 버튼은 페이스북에만 있다. 사용자들이 관심 있어 하는 정보를 ‘좋아요’ 버튼 하나로 표현할 수 있는 장점 때문에 페이스북을 선정하였다.

현재 페이스북에서는 여러 사람들의 관심이나 활동을 지향하는 사람들이 자신의 신상정보를 드러낸다. 또한 페이스북에는 자신이 올린 글뿐 아니라 친구들의 생각도 실시간으로 게시된다. 이것은 개인의 일상 또는 관심사에 대한 콘텐츠를 획득할 수 있게 해준다[13]. 하지만 여러 사용자 간 공통관심사 파악은 어렵다. 본 논문에서는 페이스북을 통해 개인이 아닌 여러 사용자 간 공통관심사를 파악하여 아이템을 추천하는 콘텐츠 추천 시스템을 제안하였다.

본 논문에서는 페이스북 사용자들의 페이지 ‘좋아요’를 추출하여 개인의 선호도를 파악하였다. 페이스북 사용자로 그룹을 만들어 각 사용자들의 선호도와 표준 편차를 고려하여 공통관심사를 제공해주는 공식을 제안하였다. 4명을 한 그룹으로 가정하고 시뮬레이션한 결과 제안한 공식에 대해 나올 수 있는 경우의 수는 4가지로 간추려졌다. 그 결과 개인의 선호도를 나타내는 ‘좋아요’ 수가 많으면서 페이스북 사용자들 간 선호도 편차가 적은 콘텐츠를 추천한다.

시뮬레이션 데이터와 실제 페이스북 사용자의 데이터를 이용하여 실험한다. 두 가지 데이터를 통해 그룹 내 추천된 콘텐츠는 가장 낮은 불만족도를 가진 콘텐츠와 일치함을 보여줌으로써 제안한 공식의 정확성을 증명한다.

References

[1] Park, H. J., Rho, S. K., “Finding Influential Users in the SNS Using Interaction Concept: Focusing on the Blogosphere with Continuous Referencing Relationships,” *The Journal of Society for e-Business Studies*, Vol.17, No.4, pp.69-93, 2012.

[2] Jong-Gook Bae, Jae-Dong Yang, and Ho-Sang Jo, “Expert Recommending System with Extended Object-Based Thesauri(XOT) for Social Network Service,” *Journal of Korean Institute of Information Scientists and Engineers*, Vol.39, No.6, pp.473-487, 2012.

[3] Lee, H. G., “The exploratory study of the meaning of Facebook : Users’ cognition of communication context and the mode of communication,” *Journal of Cyber Communication Academic Society*, Vol.28, No.4, pp.129-172, 2011.

[4] Myong-Ok Kim, Mi-Sun Lee, “A Study on the Effect of Communication Using Facebook on Organization Culture and Emotional Labor: Focusing on K QuasiNon-Governmental Organization,” *Journal of the Society for e-Business Studies*, Vol.18, No.2, pp.131-152, 2013.

[5] Gyu-Dong Bak, Seung-Jae Oh, Bo-Ra Gang, and Hyo-Jeong So, “Synchronous Feedback Mechanism in Online Social Rating Systems: Examining Social Interaction Effect,” *Journal of the HCI Society of Korea*, Vol.2014, No.2, pp.785-788, 2014.

[6] Saim Shin, Sung-Ju Park, Se-Jin Jang, and Seok-Pil Lee, “Developing the contents recommendation agent using the social user clusters in the social network services on the multiple media convergence environment,” *Journal of korea information processing society*, Vol.38, No.2, 2011.

[7] Seo-Hun Jean, Gi-Hwan Gim, and Jeon-Su Jeon, “Study on Personalized Recommendation Optimization Plan,” *korea university*, 2013.

[8] James Davidson, Benjamin Liebald, and Junning Liu, “The YouTube Video Recommendation System,” *RecSys, “10Proceedings of the fourth ACM conference on Recommender system,”* pp.293-296, 2010.

[9] SeokJong Yu, “The dynamic competitive recommendation algorithm in social network services,” *Information Sciences*, pp. 1-14, 2012.

[10] Yohan Jin, Mingqing Hu, Harbir Singh, Daniel Rule, and Mikhail Berlyant, “Myspace Video Recommendation with Map-Reduce on Qizmt,” *Semantic Computing(ICSC), IEEE Fourth International Conference on Semantic Computing*, pp.126-133, 2010.

[11] FacebookDeveloper [Internet], <https://developers.facebook.com/docs/plugins/>, 2014.

[12] Simplify360, “1 million Facebook fans brings in an average of 826 likes and 309 comments per post,” [Internet] <http://thenextweb.com/facebook/2011/05/17/1-million-facebook-fans-brings-in-an-average-of-826-likes-and-309-comments-per-post/>

[13] Je Hyok Rew, Young Hwan Choi, and Een Jun Hwang, “Human Computer Interaction: A Facebook Page Ranking and Highlight Contents Selection Scheme based on User Interests,” *korea information processing society*, Vol.3, No.2, pp.101-108, 2014.



임연주

e-mail : limyeonju92@naver.com
 2011년~현재 한국외국어대학교
 정보통신공학과 학사과정



박상원

e-mail : swpark@hufs.ac.kr
 1994년 서울대학교 컴퓨터공학과(학사)
 1997년 서울대학교 컴퓨터공학과(석사)
 2002년 서울대학교 컴퓨터공학과(박사)
 2002년~2003년 세종사이버대학교 디지털
 콘텐츠학과 전임강사
 2003년~현재 한국외국어대학교 정보통신
 공학과 교수

관심분야 : Flash Memory, Embedded Database, Mobile Computing