

소셜 네트워크 사용자 행위의 속성 분석을 통한 맞춤형 커뮤니티 서비스 시스템 설계

신은세* · 김명준* · 한소라* · 오은지* · 이강환

*한국기술교육대학교 컴퓨터공학부

A design of Customized Community Service System

based on user-behavior analysis on social network

Eun-se Shin · Myung-june- Kim · So-ra Han · Eun-ji Oh · Kang-whan Lee

Korea University of Technology and Education

E-mail : deer@kut.ac.kr

요 약

최근 소셜 네트워크 서비스는 언제 어디서나 정보를 누구라도 손쉽게 전달하고 볼 수 있는 수단으로 각광받고 있다. 소셜 네트워크 서비스의 주요한 특징은 사람과 사람 사람과 정보, 정보와 정보 간의 관계 네트워크로서 사용자가 능동적으로 참여한다는 것이다 하지만 범람하는 수많은 정보들 속에서 사용자가 직접 정보를 검색 및 분류해야 하는 과정은 사람과 정보간의 관계 네트워크 측면에서 소셜의 의미를 충족하지 못한다 이러한 기존의 정보 활용법은 사용자의 선호도에 따른 맞춤형 정보의 수용과 공유를 제시하지 못하고 있다

본 연구에서 설계된 사용자 맞춤형 서비스 시스템은 사용자의 상황인식 속성정보와 이에 따른 선호도를 평가하는 알고리즘을 기반으로 하여 보다 효율적인 커뮤니티 공간이 제공될 수 있는 맞춤형 커뮤니티 서비스 시스템을 설계 제안한다 제안된 시스템에서는 소셜 네트워크 서비스에서 사용자가 텍스트를 읽거나 작성하는 행위를 바탕으로 사용자의 관심사를 제공된 알고리즘으로 분석하여 사용자의 선호도에 따른 정보를 분류하고 사용자의 인적정보로부터 선별한 유사 사용자들을 통해 신뢰성이 높은 정보를 우선적으로 선출한다 따라서 사용자의 속성과 선호도를 고려한 상황인식 정보를 제공함으로써 사용자가 직접 정보를 검색 및 분류하는 과정을 단축하고 정보의 신뢰성을 향상할 수 있는 방법을 제시한다 이러한 상황인식 기반의 맞춤형 커뮤니티 서비스 시스템은 실시간으로 많은 정보가 공유되는 서비스에서 다양하게 적용되어 인터넷 신문 타겟 마케팅 광고 등의 응용분야에서 다양한 정보제공 서비스 시스템으로 적용될 수 있을 것으로 본다

키워드

소셜 네트워크 서비스, 타겟 마케팅 광고, 맞춤형 정보

1. 서 론

소셜 네트워크는 웹상에서 개인 또는 집단이 하나의 노드가 되어 각 노드들 간의 상호의존적인 관계에 의해 만들어지는 사회적 관계 구조를 말한다[1]. 91.6%의 소셜 네트워크 서비스 사용자는 소셜 네트워크 서비스 사용 목적을 자료 및 정보 획득으로 꼽았다[2]. 하지만 현재의 소셜 네트워크 서비스는 이용하는 사용자 수만큼 불필요한 데이터들의 발생 빈도도 높아졌다 이러한 기존의 정보 활용법은 사용자의 선호도에 따른 맞춤형 정보의 수용과 공유를 제시하지 못하

는 문제점이 있다.

본 연구에서는 소셜 네트워크 서비스에서 이루어지는 사용자들의 행위를 바탕으로 상황인식 속성정보를 이용한 맞춤형 커뮤니티 서비스 시스템을 제안하였다.

II. 제안 시스템

2.1 사용자 관심사 도출

사용자의 선호도는 시간이 흐름에 따라 변화한다. 과거에는 스포츠에 관한 관심사를 선호하

였으나, 근래에는 시사에 관한 관심사를 좋아하게 된 사용자가 있다고 할 때, 이전에 선호하였던 관심사에 대한 가중치는 낮춰줄 필요가 있다 [3]. 시간에 따른 사용자의 최근 관심사를 도출하기 위해 사용자가 최근 선택한 관심사에 대한 기록을 저장한다. 하지만 기존의 시간 가중치를 통한 사용자의 선호도 변화 분석은 단순히 시간에 의존하므로, 사용자가 실제로 선호하는 관심사라 하더라도 최근 선택한 관심사에 대하여 더 높은 선호도를 출력한다. 따라서 본 연구에서는 사용자가 실제로 선호하는 관심사를 고려한 사용자 관심사 도출 알고리즘을 제안한다

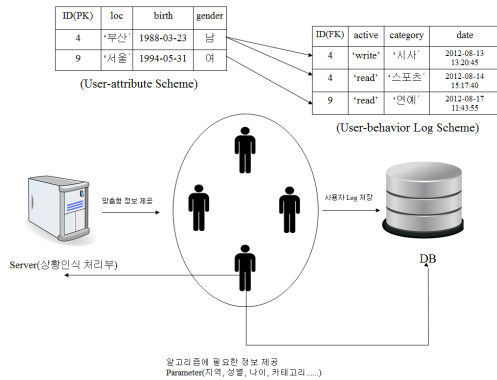


그림 1. 사용자 관심사 분석 구조도

그림 1은 사용자의 행위를 기반으로 사용자의 관심사를 도출하는 구조도이다 사용자가 소셜 네트워크 서비스에서 작성하거나 구독한 정보들로부터 관심사, 작성 시간 등을 추출하여 Log를 저장한다. DB에 저장된 Log를 기반으로 상황인식 처리부에서는 각 사용자의 최근 관심사를 분석하기 위해 시간에 따라 감소하는 비율인 σ 를 정의한다. 그리고 시간 감소율과 사용자에게 의하여 정의된 관심사에 대한 중요도를 이용하여 사용자의 관심사별 선호도(Preference)는 다음의 수식으로 표현 될 수 있다.

$$Preference = \sum_{i=1}^n \sigma^{\alpha} * w_c \quad (1)$$

제안된 사용자 선호도 분석을 위한 알고리즘의 식(1)은 관심사별 선호도를 구하는 수식으로써 특정 관심사에 해당하는 로그의 개수가 n개 일 경우, 해당로그별로 감소율에 시간차를 적용시키고, 그 수치에 해당 관심사의 가중치를 곱함으로써 선호수치를 도출한다 따라서 시간차를 적용시키는 감소율의 알고리즘은 다음과 같다

$$\alpha = (date_{current} - date_{target}) / T \quad (2)$$

현재의 날짜에서 분석 대상이 되는 행위가 발생한 날짜의 차인 α 값을 구하는 식이다. 주기 T

는 기간계수에 해당하는 값으로, 예를 들어 주기 계수(T)의 값을 7로 설정한다면 시간 감소율에 따른 단위를 1주일로 설정 할 수 있다. 또한 사용자의 관심 아이템($w_c(i)$)는 사용자의 관심사별 가중치로 정의 하고 이의 전체 계수의 합은 1로 정의 한다.

$$\sum_{i=1}^m w_c(i) = 1 \quad (3)$$

이렇게 도출된 n개의 로그의 모든 선호수치를 합함으로써 특정 관심사의 선호수치를 계산함으로써, 사용자의 최대 관심사를 도출할 수 있다

2.2 사용자 맞춤형 정보 제공 알고리즘

사용자 맞춤형 정보 제공 알고리즘은 사용자의 속성에 기반을 두고 정보와 사용자간에 유사도를 통해 정보의 적합도를 선출한다 기존의 맞춤형 정보 제공 서비스에서는 Jaccard similarity 알고리즘[4]을 바탕으로 단순히 두 집합 간의 유사도를 구하는 방식이다 하지만 본 연구에서는 도출된 사용자의 관심사와 인적정보를 패키징하고, 패키징된 각 속성에 따른 가중치를 고려하여 보다 적합한 정보를 사용자에게 연계될 수 있도록 제안한다. 서버에서는 전달 받은 패키지를 이용하여, 사용자가 원하는 정보를 선출한다 사용자 맞춤형 정보를 제공하기 위하여, 정보의 적합도를 도출하는 알고리즘은 다음과 같다

$$Priority = \sum_{i=1}^n Param(i) * w(i) \quad (4)$$

제안된 알고리즘의 식(4)는 정보에 대한 적합도를 구하는 수식으로 해당 정보의 속성별 값을 반환하는 함수 $param(i)$ 에 각 속성에 해당하는 가중치 $w(i)$ 를 곱하여, 대상 정보의 적합도를 도출한다. 함수 $param(i)$ 는 사용자의 속성과 대상 정보의 속성을 비교하여 속성이 일치한다면 1, 일치하지 않는다면 0을 반환한다. 예를 들어 A 라는 사용자의 '나이' 속성의 인덱스(i) 값이 2라고 가정한다면 A의 나이가 20대이고, 해당 정보가 20대가 많이 열람한 정보라 할 때 사용자의 속성과 해당정보간의 '나이' 속성에 대한 교집합이 {20대}가 되어 $param(2) = 1$ 이 도출된다. 하지만 두 속성이 일치하지 않는 경우, 교집합은 \emptyset 이 되어 $param(2) = 0$ 이 되게 된다. 따라서 $param(i)$ 는 해당 속성에 대한 값이 사용자와 정보가 동일한 경우는 1, 다른 경우는 0의 값을 가지게 된다.

$$\sum_{i=1}^n w(i) = 1 \quad (5)$$

제안된 알고리즘의 식(5)는 사용자의 속성의

수를 n 개라 했을 때, 모든 사용자 관심사의 가중치의 합은 1이라는 것을 표현하고 있다. 이에 따라 사용자는 각 사용자의 속성에 가중치를 부여할 수 있다. 본 알고리즘을 통해 도출한 정보의 적합도를 내림차순으로 정렬시키면 사용자 맞춤형 정보를 우선적으로 제공할 수 있으므로, 맞춤형 정보 제공 서비스 구현이 가능해진다.

2.3 제안된 알고리즘의 구현 및 모의실험

본 연구에서 제안한 사용자 맞춤형 정보 제공 알고리즘의 정확도에 대한 증명을 위해 신문 서비스 시스템을 이용하여 실험 하였다. 도출된 사용자 관심사를 인자로 하여 맞춤형 정보 제공 알고리즘을 통해 사용자에게 적합한 기사를 선출하도록 시스템에 적용하였다.

표 1. 모의실험 환경

항목	내용	가중치
성별	남자	0.1
나이	20대	0.1
지역	서울	0.1
관심사	스포츠	0.7
기사의 수	20~100개	
사용자 선택	10개	

실제 사용자의 기사 선호도를 확인하기 위해 기사의 목록을 보여주고 [표 1]과 같은 속성을 가진 사용자에게 직접 선호하는 기사 10개를 선출하도록 하였다. 제안된 알고리즘에 의해 선출된 기사목록과 사용자가 선출한 기사목록을 비교하여 어느 정도의 적중률을 보이는지 실험하였다.

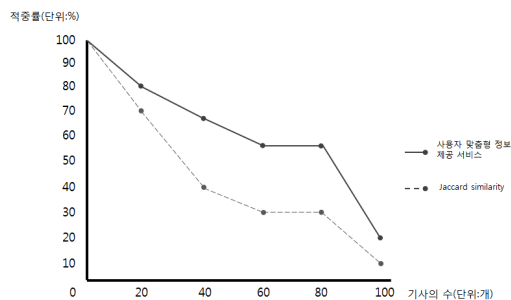


그림 2. 기사의 수에 따른 적중률 변화

그림 2에서 보여주는 바와 같이 사용자 맞춤형 정보 제공 알고리즘을 적용한 경우가 Jaccard similarity 알고리즘을 적용하였을 경우 보다 정보의 양에 따른 적중률이 평균적으로 약 22%가량 더 높다는 것을 알 수 있다. 이를 통해 사용자 맞춤형 정보 제공 알고리즘이 Jaccard similarity 알고리즘보다 상대적으로 정보연계의 정확도가 좀 더 우수함을 확인할 수 있다.

V. 결 론

본 연구에서는 사용자상황인식을 통해 도출된 사용자의 관심사와 사용자의 인적정보를 활용하여, 맞춤형 정보를 제공하는 알고리즘을 기반으로 한 맞춤형 커뮤니티 서비스 시스템을 제시하였다. 제안된 시스템에서는 소셜 네트워크 서비스 사용자가 선호하는 정보를 파악하고 신뢰성 있는 정보 선출이 가능한 알고리즘을 제시하였다. 이는 사용자가 직접 정보를 검색 및 분류하는 노력과 시간을 줄일 수 있도록 하며 제공된 정보에 대해 높은 신뢰를 가질 수 있도록 한다.

향후 제안된 시스템은 인터넷 신문과 소셜 네트워크 서비스와 같이 실시간으로 많은 정보가 공유되는 서비스에 적용되어 타겟 마케팅 광고와 같은 응용분야로 활용될 수 있다.

참고문헌

[1] S. Wasserman, and K. Faust, Social Network Analysis: Methods and Applications, Cambridge University Press, 1994.
 [2] "2010인터넷실태조사", 한국인터넷진흥원.
 [3] "시간 가중치와 가변형 K-means 기법을 이용한 개인화된 음악 추천 시스템, 김재광·윤대복·김동문·이지형, 성균관대학교 전자전기컴퓨터공학과
 [4] Jaccard, Paul (1901), "Étude comparative de la distribution florale dans une portion des Alpes et des Jura", Bulletin de la Société Vaudoise des Sciences Naturelles.
 [5] "스마트폰 사용자를 위한 사용자 맞춤형광고 서비스 모델", 박민영*·권혁철, 부산대학교 컴퓨터공학과, 2011.
 [6] "사용자의 인적정보를 이용한 소셜 네트워크 활성화 방안" 김민수*·유홍진·채진석, 인천대학교 컴퓨터 공학과
 [7] "소셜 네트워크 서비스 클라이언트의 사용행태 연구" 이재홍·전주경
 [8] "소셜 북마킹 시스템 내 유사 사용자 정보를 이용한 웹 콘텐츠 랭킹 알고리즘" 박수진·이시화·이만형·황대훈, 경원대학교 인터넷미디어학과
 [9] "SNS(소셜네트워크 서비스)를 활용한 유통업체 온라인 마케팅 활성화 방안"에 관한 연구 최재용