소셜 네트워크 서비스에서 콘텐츠 우선순위 선정을 위한 알고리즘 연구

김유두* · 문일영*

*한국기술교육대학교

A Study on Algorithm for Selection Priority of Contents in Social Network Service

Yu-Doo Kim* · Il-Young Moon*

*Korea University of Technology and Education

E-mail : kydman@kut.ac.kr, iymoon@kut.ac.kr

요 약

소셜 네트워크 서비스는 기존의 공급자 중심의 서비스가 아닌 서비스 이용자 간의 커뮤니케이션을 통해 서비스가 이루어지고 있다. 이러한 소셜 네트워크 서비스를 통해 생산되는 정보는 빠른 속도로 증가하고 있다. 이에 본 논문에서는 소셜 네트워크 서비스에서 표현되는 콘텐츠의 우선순위를 정하는 알고리즘을 분석한다.

ABSTRACT

SNS(Social Network Service) is not provider-oriented service unlike the existing internet communication services. So contents produced by SNS are rapidly increase. In this paper, we analyze the algorithm for selection priority of contents.

키워드

Social Network Service, SNS, Contents, EdgeRank

Ⅰ. 서 론

서비스 제공자 중심의 인터넷 포털 서비스가 사용자들 간의 콘텐츠를 생산하고 활용하는 소셜 네트워크 형태의 서비스로 급격히 변화하고 있다. 특히 페이스북[1]과 트위터[2]로 대표되는 소셜 네트워크 서비스의 가입자는 급격히 증가하고 있으며 이러한 서비스는 기존 포털보다 훨씬 사용자들의 몰입도를 증가 시키고 있다.

하지만 다양한 사람들에 의해 생성되는 콘텐츠는 그 증가 속도가 너무 빨라 일반 사용자들이 모든 내용을 접하기에 무리가 있다. 또한 기존 포털의 단순 추천 숫자나 콘텐츠 등록 순서에 따른 표시 방법은 사용자들이 원하는 정보를 제 때에 전달하여 줄 수 없다.

따라서 현재의 소셜 네트워크 서비스에서는 조금 더 지능적인 콘텐츠 우선순위 알고리즘을 사용하려고 많은 노력을 기울이고 있다.

대표적 소셜 네트워크 서비스인 페이스북의 개발자 컨퍼런스에서는 서비스를 로그인 후 첫 화면에 나오는 친구들의 최근 소식을 요약해서 전달해 주는 알고리즘인 EdgeRank[3]를 소개하였다. 또한 국내 서비스인 다음 View[4] 에서는 베스트 뉴스를 뽑는 알고리즘[5]을 활용하고 있다.

이에 본 논문에서는 소셜 네트워크 서비스에서 사용되고 있는 콘텐츠 우선순위 지정을 위한 알고리즘에 대해 알아보고 사용자가 원하는 정보에 더욱 근접할 수 있는 정보를 제공하기 위해 어떤 부분이 더 발전해야 하는지에 대한 연구를 진행하였다.

Ⅱ. EdgeRank

전 세계에서 가장 많은 사용자가 사용하고 있는 페이스북의 개발자 컨퍼런스에서 발표된 알고리즘인 EdgeRank[3]는 매우 간단하면서도 효율적인 방법을 제안 하였다.

이 알고리즘의 핵심은 페이스북에 사용자가 로그인 후 첫 화면에서 나타나는 콘텐츠의 우선순위를 지정하는 것이다. 친구관계에 있는 많은 사람의 글을 단순히 시간의 순서나 추천 수에 의한 것이 아닌 다양한 값에 의해 콘텐츠를 보여주는 순서를 정하고 있다. 아이디어의 핵심은 다음과 같이 구성되어 있다.

(1) u(Affinity score) : 새로운 소식을 업데이트 한 사람과의 친밀도
(2) w(Weight) : 각 상호작용의 가중치
(3) d(Time) : 콘텐츠 등록 시간

위 세 값의 곱의 합으로 구성하여 다음과 같은 식이 완성되며, 이 값에 의해 콘텐츠의 우선 순위를 정하게 된다.

$$\sum_{edges\ e} u_e w_e d_e \quad \text{..... (1)}$$

따라서 잘 알고 있는 최근의 여러 사람들의 댓글이나 추천을 받은 값이 주로 우선순위로 콘텐츠가 표시 된다는 것이다. 따라서 단순히 콘텐츠 자체의 투표가 아닌 기존 사람들과의 관계를 포함하여 콘텐츠를 추천하므로 친밀도 있고 신뢰성 있는 사람들의 콘텐츠가 더욱 우선순위가 높게 노출되게 되는 것이다.

Ⅲ. Daum View의 베스트 뉴스 선정

국내 업체인 다음커뮤니케이션에서 서비스하고 있는 블로거 뉴스의 경우는 열린 편집 시스템을 활용하고 있다.

여기서는 열린 편집을 통한 베스트 뉴스를 선정하는 알고리즘이 사용되고 있다[5].

이 알고리즘은 추천 버튼을 누르는 사람들의 성향을 조사하고, 추천을 받아 베스트 뉴스에 오르는 여러 과정 중의 추천 버튼을 누르는 방법을 구별하여 점수를 계산하는 것이 핵심이다.

표 1에서는 이 식을 구성하는 값들을 소개하며, 식 (2)에서는 알고리즘의 계산 식을 정의하고 있다.

$$\begin{aligned} VS_i = a(&VA_i + cVAM_i + c'VAN_i) \\ &+ b(VB_i + cVBM_i + cVBN_i) \\ &- (VD_i - VDB_i) \end{aligned}$$

..... (2)

표 1. 알고리즘 식의 구성 값

Variables	Description
VS_i	User i's voting score
VA_i	Number of Type A votes user i has
VAM_i	Number of Type A votes on mediocre blogger's postings user i has
VAN_i	Number of A votes on new blogger's postings user i has
VB_i	Number of Type B votes on mediocre blogger's postings user i has
VBM_i	Number of Type B votes on mediocre blogger's postings user i has
VBN_i	Number of Type B votes on new blogger's postings user i has
VD_i	Number of Type D votes user i has
VDB_i	Number of Type D on postings that become Best News user i has
VT_i	User i's voting trust rate
TV_i	Total number of votes user i has

이 알고리즘은 베스트글을 한 번 이라도 써본 사람(M)과 일반 블로거(N)가 특정 글을 베스트에 올릴 때 유효했던 추천(A)에 가중치를 두어 나머지 추천(B)를 토대로 그 사람이 했던 추천 점수를 계산하게 되는 것이다.

그 후 다음 사람이 충분히 읽지 않고 추천한 짧은 시간에 추천(D)과 베스트 글 이후의 유사한 추천(DB)을 뺀 값이다.

따라서 추천 점수는 내용을 충분히 읽고 추천한 글 중에 베스트 글을 쓴 사람이 베스트 글이 될만한 시점에 추천이 많은 경우 올라가게 된다.

Ⅳ. 결 론

페이스북과 트위터로 시작된 소셜 네트워크 서비스는 향후 인터넷 서비스의 중심이 될 것이다. 따라서 이를 통해 생산되는 콘텐츠를 사용자가 원하는 것을 우선하여 추천하기 위해서는 사용자 간의 관계에 대한 데이터를 추출하여 이 값을 반영하는 연구가 더 많이 연구되어야 한다.

참고문헌

[1] http://www.facebook.com
[2] http://www.twitter.com
[3] http://techcrunch.com/2010/04/22/facebook-edgerank/
[4] http://view.daum.net/
[5] Joonseong Ko et. al., "Open Editing Algorithm: A Collaborative News Promotion Algorithm based on Users' Voting History", *ICCSE*, 2009.