

# SNS를 이용한 사용자 맞춤형 정보전달 시스템

김무철 · 김용기 · 최희석

## I. 서론

1989년 Tim Berners-Lee에 의해 WWW(World Wide Web)이 제안되었고, 많은 이들이 접근성과 편리함으로 인해서 웹을 통해서 정보를 획득하게 되었다. 뿐만 아니라 누구나 쉽게 익힐 수 있는 HTML(Hyper Text Markup Language)의 장점인 쉬운 정보의 게재 및 구독으로 인해서 누구나 쉽고 빠르게 정보를 게재하고 획득할 수 있게 되었다. 그에 따라서 점점 많은 이들이 업무적으로 혹은 개인적으로 연구에 활용하기 시작했으며 업무 및 홍보에도 응용하여 활용하게 되었다.

2000년대 이후, 네트워크 성능의 향상과 더불어 스마트 장치들의 보급으로 인해서 사람들은 언제 어디서나 손쉽게 인터넷에 접속하고 자신이 원하는 정보를 생산하고 접근할 수 있게 되었다. 이를 이용할 수 있게 되었다. 그 결과 더욱 많은 사람들이 정보의 생산 과정에 참여하게 되어 정보의 양과 복잡도가 더 급격히 방대해지게 되었고, 누구나 정보를 생산하고 접근할 수 있는 환경으로 인해서 사람들은 자신이 원하는 양질의 정보를 접하는 것이 더욱 어렵게 되었다.

한편, 최근 사용자 중심의 웹(User Generated Web)인 웹2.0 패러다임의 등장 이래 온라인 소셜 네트워크를 기반으로 한 다양한 형태의 온라인 소셜 네트워크 서비스(Online Social network services)가 각광을 받고 있다. 온라인 소셜 네트워크는 새로운 형태의 협업(collaboration)과 의사소통(communication) 수단으로써 정보 시스템의 새로운 공유 방법으로 자신의 의사 표현을 언제 어디서든 자유롭게 할 수 있게 되었다[1]. 사람들은 온라인 소셜 네트워크 서비스를 통해서 이들은 SNS를 통해서 다양한 사용자와 정보를 공유하고 의사소통을 하는 등 상호작용(interaction)을 통해서 관계를 맺는다[2].

현재 전세계에는 페이스북(Facebook), 트위터(Twitter), 싸이월드(Cyworld) 등 많은 마이크로 블로그(Micro blog)형태의 소셜 네트워크 서비스가 서비스 되고 있다. 가트너의 보고서에 따르면, 트위터의 경우 가입자가 1억 7천 만명 이상이며, 매일 약 37만 명의 신규 가입자가 있으며, 일일 글 게재수가 2억 건을 돌파했다[3]. 이처럼 SNS는 전세계적으로 주목 받고 있으며, 사용자의 수, 이용시간, 유통되는 메시지의 양 등 사람들에게 큰 영향력을 가지고 있다.

특히 트위터는 다른 온라인 소셜 네트워크 서비스와 비교하여 개방적이고 단순한 기능으로 구성되어 있다. 트위터는 언제 어디서든 자신이 경험한 뉴스를 최대 140자의 짧은 텍스트를 통해 의견을 표현할 수 있다. 뿐만 아니라, 관심을 가지고 있는 다른 유저를 자신에게 등록하는 팔로우 기능, 그리고 다른 사람의 글을 자신의 의견으로 인용 혹은 추가하는 리트윗(Retweet)으로 구성되어 있다.

리트윗 기능은 자신의 의견을 다른 사람들에게 노출 시키는 역할을 해준다. 팔로우는 상호 동의가 요구되지 않기 때문에 단 방향적(One-way)인 관계를 형성하는 특징을 가지고 있다. 이는 사용자가 원하는 누구와도 쉽게 관계를 맺을 수 있다는 것을 의미하며, 사람들로 하여금 적극적으로 관계를 맺도록 유도하는 요소가 된다. 리트윗(Retweet)은 타인의 의견을 빠르고 쉽게 재생산 및 가공할 수 있게 해준다. 이는 타인의 의견이나 자신의 의견이 쉽게 확산될 수 있다는 것을 의미한다.

\* 김무철, KISTI, mckim@kisti.re.kr

\*\* 김용기, KISTI, ykkim@kisti.re.kr

\*\*\* 최희석, KISTI, choihs@kisti.re.kr

이러한 기능들로 인해 트위터는 다른 소셜 네트워크 서비스 보다 정보 전달이 빠르며, 사용자들 사이에 관계를 형성하기 쉽다는 장점이 있다. 그러나 관계를 형성하는데 어떠한 조건도 없기 때문에 각 관계의 정도를 표현할 기준이 없다. 즉, 두 사람의 관계가 친구이던, 유명인사이던 단순 follow 관계만으로 정의 될 뿐이다. 관계에 대한 이러한 문제는 자칫 소셜 네트워크의 활용성을 떨어뜨릴 수 있으며 나아가 정보의 공유 측면에서 보안성이 떨어지는 등 문제를 야기 할 수 있다.

본 연구에서는 정보 관리 시스템에서 사용자들에게 맞춤형 정보를 제공해줄 수 있는 방법으로 사용자들의 SNS정보들을 추출해서, 해당 knowledge에 관심을 가질 수 있는 사람들에게 맞춤형 정보 제공 즉 개인화 정보 전달을 제공한다.

본 논문의 구성은 다음과 같다. 2장은 관련 연구, 그리고 3장에서는 NTIS R&D 오픈지식서비스를 소개한다. 4장은 SNS에서의 정보 추출 및 맞춤형 정보 전달 과정을 설명한다. 마지막으로, 5장에서는 결론과 향후연구를 제시한다.

## II. 관련연구

소셜 네트워크는 대인간 또는 조직 간의 관계를 노드(객체)와 링크(관계)를 이용하여 도식화 한 것이다. 이것을 이용하여 심리적, 사회적 개념을 논리적인 구성으로 변환하는 것이 가능하다.

개인 또는 그룹을 의미하는 노드와 그 관계를 표현하는 링크로 이뤄진다. 링크는 개인 간의 정보의 흐름을 표현하기도 하며 정량적 표현을 통해 관계 정도(degree)를 나타내기도 한다. 연결은 방향성을 가질 수 있으며 신뢰도 등 부여하는 의미에 따라서 다양한 적용이 가능하다.

소셜 네트워크 전체 그래프 모델은 많은 노드들과 노드들을 연결하는 관계로 구성된다. 한 노드는 다른 노드와 직접 연결이 되어 있고, 다른 노드들은 또 다른 노드들과 직접 연결을 구성하고 있다. 때문에 자기 자신과 직접 연결되어 있는 노드가 아니어도 간접 연결을 이루어 하나의 네트워크를 구성 할 수 있다. 인터넷의 보급으로 오프라인에서 이루어졌던 많은 부분들이 온라인으로도 가능하게 되었다. 이에 따라 네트워크 구성원간의 관계를 다양한 범주에 적용할 수 있는 신뢰적이고 논리적인 데이터가 필요하게 되었다. 소셜 네트워크를 이용하면 대인관계 상의 사회적 개념들을 디지털 정보로 변환하는데 사용 할 수 있다. 온라인 공간이 단편적인 정보 제공만이 아닌 사회적 활동을 하는 삶의 일부가 되면서 이런 소셜 네트워크 서비스의 중요성이 높아져 가고 있다.

## III. NTIS 국가 R&D 오픈지식 서비스

이 장에서는 클라우드 소싱 기반의 백과사전인 NTIS 국가 R&D 오픈지식서비스의 백과사전을 보여준다. 제안 시스템은 웹 2.0 개념인 집단 지성을 구현하고, 소셜 태깅 기법을 적용하여 동적 카테고리를 구축한다.

우선 NTIS 국가 R&D 오픈지식서비스는 (그림 1)에서 보는 바와 같은 사용자 인터페이스를 가지고 있다. 제안 시스템은 백과사전 콘텐츠의 메타데이터(metadata), 콘텐츠(contents), 그리고 콘텐츠를 정의할 수 있는 태그 정보를 보여준다. 백과사전 콘텐츠의 메타데이터는 보이는 바와 같이 제목, 최초 집필자 및 최종 집필자 정보, 추천 정도를 포함하게 된다. 백과사전 콘텐츠는 내용을 요약한 목차와 출처정보, 그리고 변경 이력정보들을 보여준다. (그림 1)은 저자가 identity라는 주제로 백과사전 콘텐츠를 집필한 결과를 보여주고 있다. 저자의 집필 내용 뿐만 아니라 다른 사용자에게 의해서 집필 내용이 보강되었고 이것이 최종적으로 전문가에 의해서 승인된 결과를 보여주고 있다.



그림 1. NTIS 백과사전의 예

특히 변경 이력은 클라우드 소싱 기능의 핵심으로 백과사전 콘텐츠를 생성하는데 기여한 사용자들의 정보 및 활동 요약 내용을 보여주며, 변경 내용에 문제가 발견되면 롤백(roll-back)할 수 있는 가능성을 제공한다. 단, 롤백은 콘텐츠의 신뢰도를 무너뜨릴 수 있기 때문에 제한적인 사용자들에 의해서 수행 가능하다. 즉, A라는 사용자가 이 글의 일부를 수정했다면, A라는 사용자는 자신이 수정한 글의 내용에 대해서만 롤백을 수행할 수 있다.

이력			
버전	변경일시	검필자	편집요약
2	2012-12-04 오후 05:11:23	김무철	modifying contents
1	2012-11-23 오후 05:29:27	김무철	최초 등록

«처음 | <이전 | 1 | 다음> | 마지막»

그림 2. 클라우드소싱 기반 히스토리 추적 정보

#### IV. SNS연계를 통한 맞춤형 정보 전달 서비스

(그림 3)은 태그 정보 및 이를 적용한 동적 카테고리(dynamic category) 정보를 보여준다. 저자는 히그스 보손에 대한 백과사전 콘텐츠를 나노소자 및 시스템의 범위로 요약 및 정의 했다. 그 결과 히그스 보손이라는 콘텐츠는 나노소자 및 시스템이라는 동적 카테고리 안에 위치하게 된다. 즉, 나노소자 및 시스템이라는 태그를 가진 나노네트워크와 같은 성격을 띄게 되며, 이들과 군집화된다. (그림 3)에서는 히그스 보손이라는 콘텐츠가 나노소자 및 시스템이라는 태그에 의해서 나노 네트워크와 연계되어 군집화 되고 있음을 보여준다.



그림 3. 태그 기반의 동적 카테고리 정보 제공

한편, 본 시스템에서는 사용자들의 SNS정보들을 추출해서, 해당 백과사전 콘텐츠에 관심을 가질 수 있는 사람들에게 맞춤형 정보 제공 즉 개인화(personalization) 기능을 제공한다. 이를 위해서 본 시스템에서는 국가과학기술지식정보서비스의 트위터 계정으로부터 자신을 팔로우하고 있는 사용자 정보들을 추출한다.



그림 4. SNS 계정 정보 추출

한편, (그림 5)에서 보는 바와 같이 국가과학기술지식정보서비스는 국가R&D와 관련된 다양한 트위터 사용자들과 관계를 형성하고 있다. 본 시스템에서는 팔로워 정보들의 프로필 정보를 추출하여 백과사전 콘텐츠에 관심을 가질 수 있는 후보군들을 추출하는데 그 목적이 있다.



그림 5. 트위터에서의 팔로워들의 프로필 정보

리스트의 분석을 위해서 관련 정보를 리스트의 ID 값을 이용하여 수집해야 한다. 트위터는 Open API를 제공하기 때문에 간편한 인증 절차를 거치게 되면 관련 정보를 수집하는 것이 가능하다. 하지만, 하지만, 트위터에서 무분별한 트래픽 증가를 방지하기 위하여 익명의 사용자가 API로 요청 시 시간 당 150 회의 요청으로 제한하고 있으며, ID를 인증한 경우 시간당 350회의 요청이 가능하다. 물론 트위터 측에 요청하여 허가를 받을 시 제한이 없는 화이트리스트 시스템이 가능하나, 현재는 허가가 매우 제한적이다.

이와 같은 제약사항 속에서 효과적으로 데이터를 관리하기 위해서는 데이터 크롤링 시점에서 필요한 정보의 순차(Sequence)에 대한 관리가 필요하다. 한편, 트위터에는 페이지(Page)가 있는데 이는 각 데이터를 묶어서 관리하는 인덱스이다. 즉 이 페이지를 기록하고 있으면 데이터 크롤링을 한 정도를 알 수 있다. 파일 시스템을 이용하여 관련 정보들을 저장, 관리한다. 앞서 구한 리스트의 ID를 파라미터로 이용하여 관련 데이터를 수집하게 된다.

본 연구에서는 국과과학기술지식정보 서비스의 트위터 계정(그림 5)으로부터 팔로워(Follower)와 Following) 관계를 획득하고, 이들이 가지고 있는 프로필 정보들을 획득하여 매칭 여부를 확인한다.

(그림 6)은 백과사전 콘텐츠들에 대해서 이에 맞는 전달 대상을 트위터 계정의 팔로워들을 대상으로 추천해준다. 본 예시에서는 전기라는 태그를 가진 신규 콘텐츠가 업로드 되었을 때, 트위터의 프로필에 전기라는 키워드를 가진 사용자들을 추출하고, 이들을 대상으로 해당 콘텐츠를 제시할 수 있는 기능을 보여주고 있다.

**맞춤형 정보전달**

신규      검색

◀ 신규콘텐츠 | 25 건

선택	제목	태그	출처	등록일
<input type="radio"/>	아토피	아토피, 보건의료관련응용	R&D백과	2012-11-30
<input type="radio"/>	[실험] 정전기 꼼짝대 미니 검전기 만들기	FUNNY 과학, 정전기, 전기, 검전기	과학향기	2012-11-28
<input type="radio"/>	컴퓨터 바이러스는 언제 생겨났을까?	오늘의 과학, 컴퓨터 바이러스	사이언스올 열린백과	2012-11-12
<input type="radio"/>	서울대공원 멸종위기 바다악어 부화 성공 ... 7일 세 마리 공개	오늘의 과학, 바다악어, 서울대공원	사이언스올 열린백과	2012-11-09
<input type="radio"/>	사물인터넷을 가능하게 하는 기술 - 블루투스	오늘의 과학	사이언스올 열린백과	2012-11-08

1 | 2 | 3 | 4 | 5

◀ 전달대상 팔로워 | 5 건

선택	사진	ID	프로파일
<input type="checkbox"/>		60121	영상이론, 촬영, 편집, 제작에 관한 모든 것을 하고픈 영상쟁이 대전시 블로그, 이츠대전기자단, 중도일보 객원기자, 100인의 아빠다나
<input type="checkbox"/>		welovekea	대한전기협회(KEA) 공식 기업트위터
<input type="checkbox"/>		sm5910	맞팔 100% 부산/금정구/전기/대한민국의 정의 실현과 미래를 위해 나아 가겠습니다. 조국을 사랑한 남들께 사랑과 존경을 드립니다. 대한민국인 위주로 팔릭 하겠습니다.
<input type="checkbox"/>		lsh766	에너지 환경전문가,...마이크로 그리드/스마트 그리드/슈퍼 그리드/신·재생에너지/2차전지(ESS)/전기자동차,...지식의 대응합, 기술의 대응합, 즉 융복합시대를 대비하는...^^*
<input type="checkbox"/>		hht2500	저희 현대하이텍은 현대중공업(주) 과 군산대학교 풍력발전 원천기술센터와 협력 관계를 통해 신재생 에너지 분야에 대형풍력발전기, 풍력가로동, 조선분야에 철의장 생산 및 설치, 중,대형 기계가공물 설계 및 제작을 하고 있습니다.

\* 5명의 팔로워 중 5명의 팔로워에게 트윗을 작성할 수 있습니다.

그림 6. 개인 맞춤형 정보 전달

## V. 결론

웹의 발달과 더불어 무수히 많은 정보들이 생성하게 되었다. 그에 따라서, 정보의 효과적인 생산 뿐만 아니라 그 관리가 중요해지게 되었다. 따라서 제안하는 논문에서는 NTIS 오픈R&D 지식 서비스에 대해서 소개하고, SNS 기반의 개인 맞춤형 정보 전달 시스템을 제안한다. 그 결과, 정보의 바다에서 각각의 사용자들에게 맞춤형 정보를 전달할 수 있게 되었다. 본 연구에서는 태그들에 대한 블리언 매칭 정보를 이용해서 맞춤형 정보를 제공하고 있는 한계점을 가지고 있다. 따라서, 향후에는 벡터 모델을 적용하여 부분 매칭(partial matching)을 지원하는 정보 전달 서비스를 제공해야 할 것이다. 뿐만 아니라, 본 연구에서 적재된 소셜 태그들을 활용하는 개인화 검색으로의 확장 뿐만 아니라, 태그 정보의 확장 연구로 이어질 것으로 기대한다.

## 참고문헌

- Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E., Wireless sensor networks: a survey. *Computer Networks*, 38(2002), 393-422.
- Lada Adamic and Eytan Adar, "How to search a social network", *Social Networks* 27(3) 187-203, 2005.
- E. Adar and C. Re.Managing Uncertainty in Social Networks, *Data Engineering Bulletin*, 30(2):23-31, July 2007.
- Li, C., Ye, M., Chen, G., Wu, J., An energy-efficient unequal clustering mechanism for wireless sensor networks. In *Proceedings of IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, (2005), 596-604.
- Berners-Lee, T., Cailliau, R., WorldWideWeb: Proposal for a hypertexts Project, <http://w3.org/Proposal.html>, 1990.
- Berners-Lee, T., *Information Management: A Proposal*, CERN (March 1989, May 1990).
- Diverse and Exploding Digital Universe*, IDC White Paper, 2008.
- Laney, D., The Importance of 'Big Data': A Definition, Gartner. <http://www.gartner.com/resId=2057415> , 2012.
- Shepitsen. A., Gemmell, J., Mobasher, B., Burke, R. ,Personalized recommendation in social tagging systems using hierarchical clustering, In the *Proceeding of the 2008 ACM conference on recommender systems*, (2008), 259-266.
- Gupta, M., Li, R., Yin, Z., Han, J., Survey on Social Tagging Techniques, *ACM SIGKDD Explorations Newsletter*, Vol. 12, No. 1 ( 2010), 58-72.
- Ding, Y., Jacob, E.K., Fried, M., Toma, I., Yan, E., Foo, S., Milojevic, S., Upper tag ontology for integrating social tagging data, *Journal of the American society for Information Science and Technology*, Vol. 61, Issue 3(2010), 505-521.
- Cattuto, C., Baldassarri, A., Servedio, V.D., Loreto, V., Emergent community structure in social tagging systems, *Advances in complex Systems*, Vol. 11, No. 4 (2008), 597-608.