데이터 웹 검색을 위한 분산형 소셜네트워크 설계

심인수*, 이홍철**
*고려대학교 산업경영공학과, **고려대학교 정보경영공학전문대학원
e-mail : siminsoo@gmail.com

A Design of Decentralized Social Network for Data Web Search

In-Soo Sim*, Hong-Chul Lee**
*Department of Industry Management Engineering, Korea University
**Graduate School of Information Managemet & Security, Korea University

요 약

오늘날 개방형 플랫폼과 정보공유공간인 소셜네트워크 서비스(Social Network)의 사용이 증가하면서 관련 웹사이트와 어플리케이션이 많이 생겨났다. 기존의 시스템은 소셜네트워크 사이트가 개인 데이터의 소유권을 가지고 있어서 정보 보호에 문제가 있고, 웹 사이트 간 정보공유가 제한되었다. 그래서 소셜네트워크의 데이터를 개인이 원하는 다른 관리 영역에서 소유할 수 있고 서로 다른 소셜네트워크에서도 자유로이 정보를 공유할 수 있는 분산형 소셜네트워크에 대한 연구가 활발히 이루어지고 있다.

본 논문에서는 시맨틱 웹 기술을 이용하여 이종 소셜네트워크 서비스, 다른 관리 영역의 데이터, 사용자들과의 관계, 데이터의 연관성들을 구조화 시키고, 데이터 웹 검색의 사용성을 높이는 데이터 웹 검색기반 분산형 소셜네트워크를 설계 하였다.

1. 서론

오늘날 소셜네트워크(Social Network)는 웹 사용자의 중요한 활동 부분을 차지하고 있으며, 대표적인 서비스 예로는 페이스북(Facebook)[1], 마이스페이스(MySpace)[2], 트위터(Twitter)[3] 등이 있다.

현재 각각의 소셜네트워크 서비스에 속해 있는 개인 정보들은 다른 소셜네트워크 서비스 안에서 자유롭게 공유할 수 없다. 즉, 사용자는 트위터에 있는 정보와 페이스북에 있는 정보를 서로 공유할 수 없다. 여기서 정보공유의 의미는 단지 데이터만 서로 보여 지는 것이 아니라, 검색을 통해서 얻을 수 있는 새로운 지식을 말한다. 또한, 웹 사용자는 소셜네트워크에 있는 개인 데이터를 마음대로 컨트롤 할 수 없다. 실제 개인 데이터의 저장장소는 서비스 제공 업체가 소유하고 있기 때문이다.[4] 소셜네트워크 서비스를 제공하는 사이트는 개인의 정보를 모으기 위해서 자신의 사이트 공간에서의 네트워크 효과를 강화시킨다. 사용자에게 자사의 소셜네트워크 공간에서만 데이터를 사용하게 제한하고 정보를 소유한다. 그러나 사용자는 이익을 추구하는 회사에게 개인정보를 맡기는 것에 불안하고 개인정보 노출의 위험성을 느낀다. 이런 문제점을 해결하기 위해서 분산형 소셜네트워크(Decentralized Social Network)에 대한 연구가 활발히 이루어지고 있다.

분산형 소셜네트워크는 데이터를 개인이 원하는 관리 영역에 소유하고 다른 사이트공간에서도 자유롭게 공유할 수 있게 해준다. 반면에, 분산형 소셜네트워크는 다른 관리 영역의 데이터, 이종 소셜네트워크 서비스, 사용자들의 관계, 데이터의 연관성이 복잡하므로 자료 공유 및 검색이 어렵다.

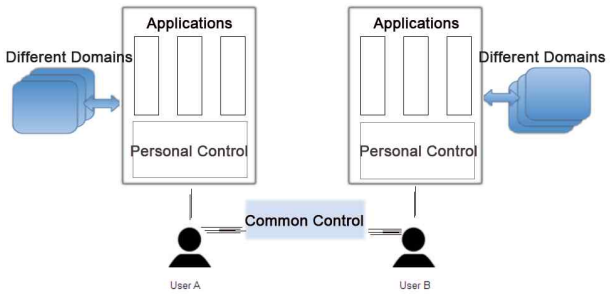
따라서 본 논문은 기존 분산형 소셜네트워크의 문서화 되어 있는 웹을 데이터 웹으로 변환시키고 구조화하는 데이터 웹 검색기반 분산형 소셜네트워크의 프레임워크를 설계하였다.

본 논문의 구성은 다음과 같다. 2장에서는 분산형 소셜네트워크와 데이터 웹 검색에 관련된 연구에 대해서 설명하고, 3장에서는 본 연구를 통해 설계한 데이터 웹 검색기반 분산형 소셜네트워크에 대해 설명하고자 한다. 4장에는 결론을 맺는다.

2. 관련 연구

2.1 분산형 소셜네트워크

분산형 소셜네트워크의 기본개념은 자신의 데이터를 직접소유하고 컨트롤 하는 것이다. 기본적인 프레임워크의 개념과 특징을 보면 다음 [그림1]과 같다.[5]



[그림 1] 분산형 소셜네트워크 기본 프레임워크

1) 분산 : 다른 영역간의 연결

웹 사용자는 데이터를 자신이 원하는 다른 영역에서 관리 하고, 개인 정보를 직접 소유할 수 있다. 그리고 사용자는 다른 소셜네트워크 공간에서도 사용자간의 정보 공유가 가능하다.

2) Open API

분산형 소셜네트워크에 접속하고 데이터를 처리하는 각각의 어플리케이션을 다른 소셜네트워크 공간에서도 사용할 수 있다.

3) 세부적 인맥관리

사용자는 각각의 인맥(개인)마다 공유영역을 설정할 수 있고, 개인별 공유레벨을 선택할 수 있다.

분산형 소셜네트워크 프레임워크를 세부적으로 구성하고 설계하는 측면은 다양하게 연구되고 있다. 사람의 개인 정보와 친구 관계를 FOAF (Friend of a Friend)를 이용해서 모델링한 프레임워크 연구가 있다. 이 방식은 정보의 자원간의 관계를 설정하기에는 부족한 부분이 있다. 이 연구에서는 자원에 대한 식별자로 URI (UniForm Resource Identifier)를 사용했다.[4] 웹에서 자신의 계정을 통합적으로 관리하는 방식인 OpenID를 통해서 공통의 ID로 여러 사이트의 소셜네트워크를 이용할 수 있다. 그리고 소셜네트워크의 정보를 RDF (Resource Description Framework)[6]로 구조화 시키고, 표준 온톨로지(ontology)를 사용해서 데이터간의 정확한 자료 공유를 할 수 있게 했다. SociaLite Language를 통해서 데이터로그, 서술형 데이터베이스 언어를 구현하여 소셜네트워크 어플리케이션 GUI에 쿼리를 쉽게 쓸 수 있게 한 연구도 진행되고 있다.[5]

2.2 데이터 웹 검색

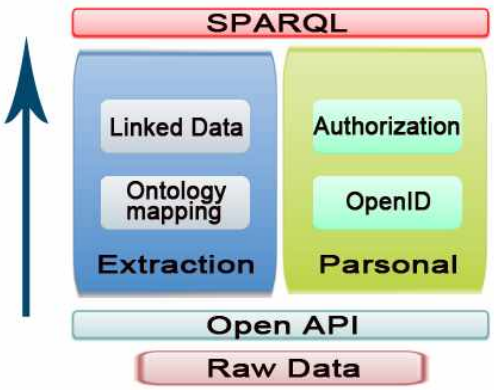
소셜네트워크 공간에 있는 많은 웹들의 자료를 공유하고 검색하기 위해서는 컴퓨터가 이해할 수 있는 데이터 웹으로 구성되어야 한다.[8] 시맨틱 기술을 이용하여 소셜네트워크 정보를 컴퓨터가 이해할 수 있는 데이터 웹으로 변환시켜야 한다.

시맨틱 웹을 구성하고 검색하는 연구는 활발히 진행되고 있다. 시맨틱 웹은 웹 데이터와의 관계를 온톨로지로 구축하여서 의미있는 검색을 할 수 있도록 한다. 온톨로지 구축방법은 RDF, OWL (Ontology Web Langage) 등이 있다. RDF는 웹 상에서 자원들의 관계를 트리플 형태(주어, 서술부, 목적부)로 표현한다.[9]

Wikipedia[10]로부터 구조적 정보를 뽑아내고, 이 정보를 웹으로 연결시켜주는 DBpedia[11] 프로젝트가 진행되었다. 이 연구는 정보 추출 프레임워크 개발을 통해서 콘텐츠를 Linked Data[12] 하였다. Wikipedia의 Infobox(기본정보)로부터 온톨로지를 매핑하여 데이터의 질을 높였다. 그리고 각각의 데이터에 대한 엔티티를 정의함으로서 데이터 웹 개발에 어려움을 해결하였다.[13]

3. 데이터 웹 검색기반 분산형 소셜네트워크의 설계

3.1 프레임워크 설계



[그림 2] 데이터 웹 검색기반 분산형 소셜네트워크의 기본 구조

본 논문에서 제안한 데이터 웹 검색기반 분산형 소셜네트워크의 기본구조는 [그림2]와 같이 크게 2부분으로 나눌 수 있다. 첫째, Extraction Management는 온톨로지 매핑을 통해서 데이터를 구조화 시키고

Linked Data화 하는 부분이다. 둘째, Personal Management는 이종 소셜네트워크 간의 연결과 권한을 설정 해주는 부분이다.

3.2 세부 설계

3.2.1 Open API

개방형 소셜네트워크의 기능을 사용할 수 있도록 Open API를 지원해준다. 플랫폼과 어플리케이션에 제약을 받지 않고 개발에 사용할 수 있다.

3.2.2 Extraction Management

Extraction Management는 문서화 되어 있는 웹을 데이터 웹으로 변환시켜 주는 부분이다. Personal Management를 통해서 인증을 받게 되면 분산형 소셜네트워크를 이용할 수 있게 된다. 이 시스템에는 기본적으로 소셜네트워크에 관련된 온톨로지가 RDF로 구축되어 있다. Raw Data의 데이터 분석을 한 후에 현재 온톨로지와 비교해서 매핑을 한다. 온톨로지 매핑을 통해서 데이터를 구조화 시켜준다. 그 다음에 Linked Data화를 시켜준다. Linked Data는 먼저 URI를 명명하고 리소스에 접근할 수 있도록 한다. URI는 웹상에 존재하는 자원을 지칭하는 스트링의 표준 형식이다. 그리고 URI로 접근 했을 때, RDF 기반의 URI가 포함되어 있는 상세 정보를 보여주고 다른 관련된 개념에도 접근 할 수 있게 해준다. 이런 과정을 거치면 데이터 웹이 생성되고 데이터가 구조화 된다.

3.2.3 Personal Management

Personal Management는 사용자에게 분산형 소셜네트워크의 접근 인증을 해준다. OpenID를 통해서 공통의 아이디로 다른 소셜네트워크를 이용할 수 있다. 각각의 다른 소셜네트워크에 접속하기 위해 공통 프록시 서버를 이용한다. 공통 프록시 서버는 클라이언트가 자신을 통해서 다른 네트워크 서비스를 간접적으로 접근 할 수 있게 해준다. Personal Management는 다른 사용자간의 데이터 공유레벨과 개인별 데이터의 공유 범위도 설정할 수 있다.

3.2.4 SPARQL

SPARQL(SPARQL Protocol and RDF Query Language)은 애플리케이션이 분산 RDF 데이터베이스에서 정교하고 복잡한 질의를 가능하게 해주며, 현재 여러 프레임워크를 지원한다. 그리고 이 언어는 RDF와 같은 트리플 구조이므로 쉽게 질의어를 만들 수 있다. Extraction Management에서 Linked Data된 데이터 웹은 구조화 되어 있어서 데이터를 검색하는 질의어를 쉽게 만들 수 있다.

4. 결론

본 연구는 데이터 웹 검색기반 분산형 소셜네트워크를 제안하였다. 이 시스템은 이종 소셜네트워크, 다른 영역의 데이터 관리 등으로 인한 데이터 검색의 어려움을 해결하기 위해서 구조화 된 데이터 웹을 생성하였다.

Extraction Management를 통해서 문서화되어 있는 웹을 온톨로지 매핑과 Linked Data 하였다. Personal Management는 OpenId를 통해서 다른 소셜네트워크 간의 정보공유 권한을 설정해 주었다. 분산형 소셜네트워크에서 각자 개인이 원하는 장소에 정보를 소유할 수 있고, 공간의 제약 없이 정보를 공유 할 수 있게 되었다. 그리고 소셜네트워크의 데이터가 구조화된 데이터 웹 형식으로 되어있기 때문에 질의를 통해서 원하는 정보를 쉽게 검색할 수 있게 되었다.

참고문헌

[1] http://www.facebook.com/
[2] http://www.myspace.com/
[3] http://www.twitter.com/
[4] Ching-man Au Yeung, "Decentralization: The Future of Online Social Networking", W3C Workshop on the Future of Social Networking, Jan. 2009.
[5] Seok-Won Seong, "PrPI: A Decentralized Social Networking Infrastructure", ACM Workshop on Mobile Cloud Computing & Service, June. 2010.
[6] http://www.w3.org/RDF/
[7] 김상균 외, “한국한의학연구원 소셜 네트워크 온톨로지 구축”, 한국콘텐츠학회논문지, vol. 9 no. 12, 2009.
[8] Tom Cruber, "Collective knowledge system: Where the Social Web meets the Semantic Web", Services and Agents on the World Wide Web, Vol. 6, pp. 4-13, July. 2008.
[9] 노상규, 박진수 공저, “인터넷 진화의 열쇠 온톨로

지", gods'Toy, 2007.
[10] http://www.wikipedia.org/
[11] http://dbpedia.org/About
[12] http://linkeddata.org/
[13] christian Bizer, Jens Lehmann, "DBpeDia - A crystallization point for the Web of Data", Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 7, pp. 154-165, July. 2009.
[14] 박지형 외, "시맨틱 웹에서의 메타데이터 검색을 위한 RDF 저장 구조 및 시스템의 개발에 관한 연구", 한국전자거래학회지, 제12권 제2호, 2007.
[15] 송창우 외, "시맨틱 웹에서 개인화 프로파일을 이용한 콘텐츠 추천 검색 시스템", 한국콘텐츠학회논문지, Vol. 8 No. 1, 2008.
[16] 유동희 외, "시맨틱 웹 기술을 이용한 온톨로지기반호텔 검색 시스템", 한국전자거래학회지, 제13권 제4호, 2008.