

소셜 네트워크에서의 사용자관계 시각화 방안

양승연*, 조현성*, 이천희*, 박석천**
*가천대학교 모바일소프트웨어학과
**가천대학교 컴퓨터공학과 (교신저자)
e-mail:201240184@gc.gachon.ac.kr

Visualization Method of User Relationship in Social Networks

Seungyeon Yang*, Hyunsung Jo*, Chunhee Lee*, Seok-Cheon Park**

*Dept. of Mobile Software, Gachon University

**Dept. of Computer Engineering, Gachon University, Corresponding Author

요 약

데이터 시각화 기술을 이용하여 데이터에 대한 접근성과 이해도를 높여 사용자가 빅데이터 가치를 효율적으로 활용할 수 있도록 소셜 네트워크에서의 사용자 관계 시각화 기술에 대해 연구하였다. 본 논문은 소셜 네트워크의 사용자관계를 시각화하기 위해 소셜 네트워크 데이터 스트리밍 기술, 동적 애니메이션 적용을 위한 시각화 기술 그리고 사용자 관계를 분석하기 위한 데이터마이닝 기술을 적용하여 소셜 네트워크 사용자들 간의 관계를 본인 중심으로 시각화하는 방안을 제시하였다.

I. 서론

정보의 홍수, 빅데이터 시대이지만 데이터에 접근하기 어렵고 이해할 수 없다면 수많은 데이터는 가치 있게 쓰이지 못할 것이다. 데이터를 쉽게 이해하기 위한 기술적인 노력이 필요하다. 시각화(Visualization)란 보이지 않는 것이 일정한 형태로 나타나 보이게 하는 것이다. (그림 1.1)과 같이 데이터 시각화란 사용자에게 더 효과적으로 정보를 전달하기 위하여 그래픽 요소를 활용하여 데이터가 정보로서 의미가 생성되도록 형상화 하는 것을 말한다. 데이터 시각화의 주요 목적은 그래픽적 의미를 이용하여 명확하고 효과적으로 정보를 전달하기 위한 것이다. 이를 위해 데이터 시각화 기술은 보여 지는 미적 형태와 정보 전달하는 기능성 두 가지를 모두 가져야 한다[1].

네트워크에서 형성된 사용자들 간의 관계를 시각화하여 보여주고자 한다. 소셜 네트워크를 통해 사용자는 타 사용자와의 일정한 관계를 지속적으로 형성하고 하고 있다. 데이터마이닝과 관계 가중치 부여 알고리즘을 통해 사용자에게 의미 있는 사용자들 간의 관계를 시각화하여 보여줌으로써 현재 소셜 네트워크에서 본인을 중심으로 이루어진 타 사용자와의 관계 상태를 확인할 수 있다.

현재 데이터 시각화 관련 분야는 인포그래픽스, 정보 시각화, 과학적 시각화, 통계 시각화로 나뉜다. 관계 시각화 분야에 대해서는 상관관계, 영향도 등의 인과관계 중심의 데이터간의 관계를 중점으로 연구되었다. 본 논문에서는 사용자 관계를 중심으로 데이터 시각화 기술을 특화시킨다. 1차 사용자는 소셜네트워크 사용자로 정의하며, 2차 사용자는 국가, 기업 등으로 보다 정확도와 신뢰성이 높은 데이터를 활용할 수 있을 것이다.

관계 시각화 기술을 구현하기 위해 2D Quartz 라이브러리를 이용한다. 사용자의 동작에 반응하는 시각화 기능과 부드러운 애니메이션 동작 기능을 적용하여 동적 데이터를 표현할 것이다. 관계의 특성을 시각화 개념으로 정의하여 애니메이션 동작으로 구현하여, 미적 형태에 대한 사용자의 만족도를 높이고, 관계에 대한 이해도를 향상 시킨다.

II. 관련 연구

본 장에서는 본 연구에 필요한 빅데이터를 이용한 시각화 기술동향 및 산업동향 그리고 발전방향에 대해 알아보고자 한다. 소셜 네트워크와 같이 빅데이터로부터 정보를 받아서 시각화 기술을 통해 새롭게 가치를 창조하는 다양한 시각화 기술 적용 시스템이다.



(그림 1.1) 데이터 시각화

소셜 네트워크에서의 사용자 관계 시각화 기술은 소셜

2.1 기술동향

정보 시각화 개념이 체계적으로 연구가 되기 시작한 것은 비교적 최근의 일이며, 스마트폰의 보급으로 소셜 네트워크가 활성화 되면서 폭증한 데이터양에 따라 간단한 방법만으로 더 이상 복잡한 정보를 쉽게 이해하거나 전달하기가 어려워졌다. 정보 시각화는 미디어 아티스트, 프로그래머, 작가, 디자이너 등 다양한 분야를 아우르며 전문적인 영역에서 활용 발전되고 있다.

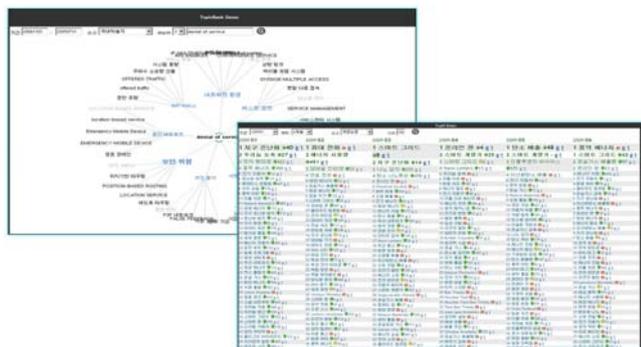
국내 빅데이터를 이용한 시각화 사례는 학술, 기술, 국방, 기업, 통신 등으로 데이터 분야의 범위에 제한 받지 않고 응용 구축되고 있다[2].

(그림 2.1)은 부산지식네트워크 서비스이다. 부산 내 거주하거나 부산 출신 지식인 전문가로 구성되는 네트워크, 지식인들의 논문과 기고문 등을 분석해 전문 지식 네트워크를 도출, 이와 관련된 전문가들을 연결, 추천하는 기능과 전문가들 간의 친밀도나 소셜 네트워크를 분석, 보고하는 기능을 제공한다.



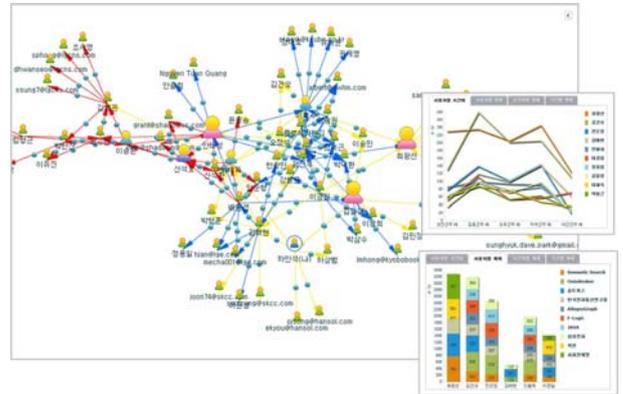
(그림 2.1) 부산 지식 네트워크 서비스

(그림 2.2)는 삼성전자 신기술 센싱 및 트렌드 분석에 시각화 기술을 사용하였다. 삼성전자 KMS에 축적되어 있는 내부 지식 콘텐츠들, 해외학술자료, 국내 학술회의 정보, 각종 연구보고서, 해외과학기술 동향분석 정보 등 대용량 기술문서들의 트렌드 분석, 특정 기술의 연관기술관계망 분석, 다양한 기술 검색 서비스 등 제공한다.



(그림 2.2) 삼성전자 신기술 센싱 및 트렌드 분석

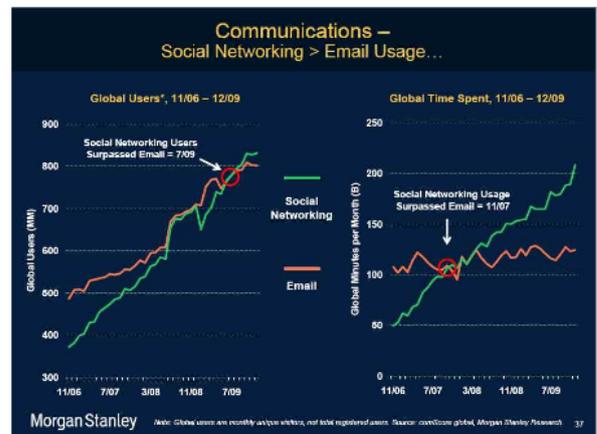
(그림 2.3)은 기업의 SNS, 이메일과 첨부문서, 논문/특허와 같은 비정형 정보로부터 시맨틱 소셜 네트워크를 분석, 지식 유통 구조를 분석하는 시스템으로 조직 내의 문서들로부터 의미적 관계를 발견하여 통합된 지식을 생성, 다양한 관점에서의 분석을 통해 사용자에게 의미 있는 결과를 제공한다.



(그림 2.3) 기업 이메일 네트워크 분석 화면

2.2 산업동향

소셜 네트워크 서비스를 기반으로 하는 기업의 마케팅 시장의 스마트 폰 단말의 기술의 발달과 통신 서비스의 확장으로 급성장하고 있다. (그림 2.4)는 소셜 네트워크 서비스와 이메일 사용 추이 비교 그래프이다. 2007년 11월 소셜 네트워크 서비스 사용 시간이 이메일 사용 시간을 추월하였고, 2009년 7월부터 소셜 네트워크 서비스 사용자의 수가 이메일 사용자의 수를 넘어서면서 소셜 네트워크 서비스를 활용하는 마케팅 전략은 지속적으로 성장하였다. 미국 등 영어권 지역에서 시작되어 확대된 소셜 네트워크 서비스는 비영어권국가인 아시아뿐만 아니라 아프리카, 중동 지역까지도 확대되었다. 소셜 네트워크 서비스 사용자와 이용시간이 확대됨에 따라 이를 이용한 마케팅 전략은 지역의 한계성을 극복하고 파급력을 가진다.



(그림 2.4) 소셜 네트워크 서비스와 이메일 사용 추이

<표 2.1> 국내·외 산업동향

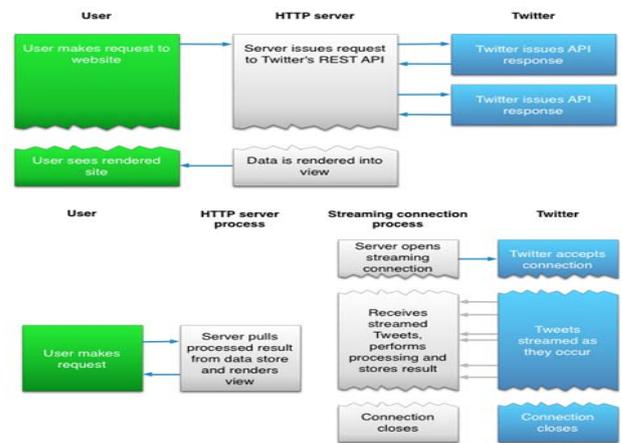
구분	산, 학, 연	특이사항	비고
국내동향	미래창조과학부	빅데이터 분석·시각화 전문가 양성 과정 개설 확대	인력양성사업
	SAS 코리아	데이터 시각화 지원 'SAS 비주얼 애널리틱스 6.1' 출시	기업용SW
	한국테라데이터	분석엔진 통합·데이터 시각화 '애스터 디스커버리' 발표	기업용SW
	오픈SNS 연구소	웹에서 생성되는 데이터 수집 및 시각화	KTEN-KNA
	투비소프트	시각화용 리치 인터넷 애플리케이션 개발	기업용SW
	공공데이터 연구센터	투표기록 데이터 시각화 인터랙티브 웹사이트 개발	폴리티즈
국외동향	YouTube	인포그래픽을 통한 영상 순위 데이터 시각화	일반사용자용SW
	스플링크	컴퓨터 로그와 기계에서 나오는 데이터 시각화	나스닥 상장
	비비시모	데이터를 검색 후 그룹으로 나누어 보여주는 시각화	IBM인수
	SAP	주요 정보를 한눈에 보여주는 시각화 솔루션 '비주얼 인텔리전스' 발표	기업용SW
	HP	데이터 시각화 기술 보유기업 인수	버티카 (Vertica)
	Intel	데이터간 관계를 시각하기 위한 라이브러리 개발	타 회사 투자
	Google	퓨전 테이블 애플리케이션 업데이트	자사 서비스
	Yahoo	시각화 웹사이트 (visualize.yahoo.com)	트래픽 시각화

<표 2.1>은 국내·외 산업동향이다. 먼저 국내에서는 국가 주도의 기술 개발이 이루어지고 있으나 대중화되지 않은 미흡한 시각화 솔루션이며, 빅데이터를 수집하고 관리하는 단계에 진입중인 기업이 대부분이다. 시각화 솔루션을 가지고 있는 기업이 대부분이 외국계 기업이다. 대표적인 소셜 네트워크인 페이스북은 전 세계적으로 6억 명 이상 사용하고 있고 사용자의 수가 현재 증가 추세를 이루 있다. 글로벌 기업인 SAS, Oracle, IBM, HP 등 대기업들이 데이터 시각화 기술력을 가진 중소기업체를 인수하거나 공동 기술연구를 통해 투자 개발하고 있다. 현재 소셜 네트워크 기반한 시각화 기술은 일반 사용자가 사용할 수 있는 시각화 솔루션은 부족한 상황이다.

III. 소셜 네트워크에서의 사용자관계 시각화 방안

3.1 소셜 네트워크 데이터 스트리밍 API 기술

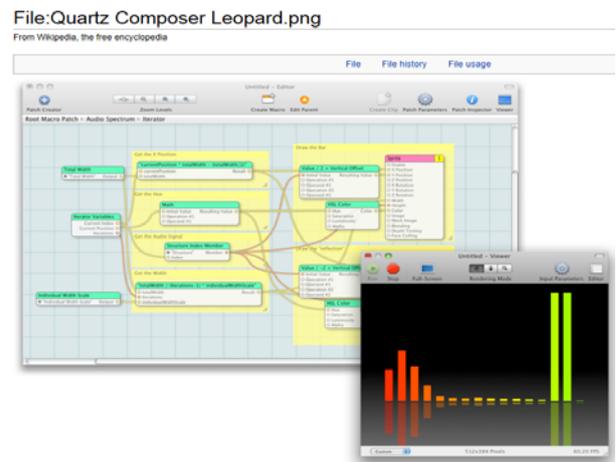
소셜 네트워크에서는 2종류의 API를 제공한다. REST API와 스트리밍 API이다. 스트리밍 API에는 3가지 종류의 수집 방식이 있다. 공개 스트림(Public stream)은 트위터에 공개된 데이터 스트림으로, 특정 사용자나 특정 주제와 관련된 데이터를 수집하여 데이터 마이닝하기에 적합하다. 사용자 스트림(User stream)은 한명의 사용자와 관련된 거의 모든 데이터 스트림 사이트 스트림(Site stream)은 다수의 사용자를 대신해서 트위터에 접속한 서버를 위한 데이터 스트리밍 기술이다. (그림 3.1)은 트위터의 데이터 스트리밍 API의 구조이다. 본 논문에서는 아래 스트리밍 API방식을 사용하여 소셜 네트워크로부터 관련 정보를 전송받아 적용한다.



(그림 3.1) 트위터 데이터 스트리밍 API

3.2 동적 애니메이션 적용 사용자관계 시각화 기술

2D Quartz 라이브러리는 사용자 동작에 반응하는 시각화 기능과 부드러운 애니메이션 동작 기능을 사용하여 사용자가 관계를 이해하기 쉽도록 한다.



(그림 3.2) 2D Quartz 라이브러리

(그림 3.2)는 애플리케이션을 만들 때 드로잉을 구현할 수 있는 독립적인 라이브러리로 코어 그래픽스 프레임워크에 속한다. 2D Quartz 라이브러리는 메모리상의 뷰나 이미지에 직접 그릴 수 있도록 설계된 함수, 자료형, 객체로 구성되어 있다. 이것은 뷰나 이미지를 가상 캔버스 위에 그리며, 페인터 모델이라는 드로잉 방식을 사용한다. 페인터 모델은 물감을 가지고 종이에 그림을 그리는 것과 같은 과정으로 드로잉 명령을 가상캔버스에 적용시키는 방식이다.

3.3 사용자 관계도 구현을 위한 데이터 마이닝

단어 기반의 키워드 분석을 통한 가중치 부여 알고리즘을 설계하여 사용자간의 관계에 가중치 적용한다. 가중치 부여를 한 사용자 성향 패턴을 분석할 수 있다. 소셜 네트워크상의 글을 분석하기 위해 텍스트 마이닝 기법을 적용한다.

일반적인 방법은 특성벡터이며, 특성추출 과정을 통해 텍스트에 대한 특성 벡터를 생성하게 된다. 즉, 텍스트 분석의 기반이 되는 것이 바로 특성추출에 의한 특성벡터이며 이의 통계수치는 각 분석기법들의 근거가 된다. 특성추출은 텍스트에서 중요한 용어를 인식하여 추출해 내는 것으로 추출된 용어들은 일반적으로 단어의 원형으로 변형되어 특성 벡터를 구성하게 된다. 이러한 특성벡터는 문서를 분류하거나 요약하는데 기초정보로 사용되며, 특성의 중요성을 나타내는 가중치는 함수와 지지도 함수의 계산은 단어가 발생한 위치와 발생한 횟수에 기반 한다.

예를 들어 한 문서 내에서 여러 번 나타나는 단어의 중요도는 높다고 가정하지만, 여러 문서에서 걸쳐서 발생도가 높다면 이 단어의 중요도는 낮다고 간주한다. 따라서 가중치 부여 함수의 계산에는 단어 발생한 문서 개수의 역함수 값이 사용된다. 각 방법에 따라 크게 문서의 클러스터링화, 분류화, 요약의 세 가지로 분류된다. 분류의 수행에 앞서 클러스터링화를 먼저 수행시켜 전체 문서집합의 개요를 획득하고 분류를 위한 판단 기준을 얻어낸다. 이에 기반한 기법은 텍스트에서 정보나 지식을 발견하고 추출하는데 사용된다.

상기와 같이 소셜 네트워크에서의 사용자관계 시각화를 통해 사용자가 빅데이터 가치를 효율적으로 활용할 수 있도록 소셜 네트워크에서의 사용자관계 시각화 방안을 연구하였다. 본 논문에서 제시하는 방안으로는 먼저, 사용자의 이용 환경이 모바일 기반이므로 모바일 애플리케이션 형태로 구현하는 것이 사용자의 접근성과 기술의 활용성을 높일 수 있다. 다음으로 사용자 스트림 API를 이용하여 사용자 관련 데이터를 전송 받는다. 그리고 소셜 네트워크상의 글을 분석하기 위해 텍스트 마이닝 기법을 적용하여 단어 기반의 키워드 분석을 한다. 분석결과에 대해 가중치 부여 알고리즘을 설계하여 사용자관계에 대해 가치를 부여한다. 그 결과를 바탕으로 2D Quartz 라이브러리를 이용하여 시각적으로 표현한다. 현재 소셜 네트워크

에서 본인을 중심으로 이루어진 타 사용자와의 관계 상태를 한눈에 확인할 수 있으며, 사용자의 동작에 반응하는 동적 애니메이션 동작으로 구현하여, 미적 형태에 대한 사용자의 만족도를 높이고, 관계에 대한 이해도를 향상시킨다. 또한 분석 결과를 통해 사용자가 인지하지 못하는 자신의 인간관계에 대해 새로운 정보가 제공된다. 본 논문에서 제안하는 소셜 네트워크에서의 사용자관계 시각화 방안을 통해 빅데이터에 대한 사용자의 접근성과, 데이터에 대한 사용자의 이해도를 높임으로써 정보의 전달력을 향상시킨다.

IV. 결론

정보의 홍수, 빅데이터 시대이지만 데이터에 접근하기 어렵고 이해할 수 없다면 수많은 데이터는 가치 있게 쓰이지 못할 것이다. 데이터를 쉽게 이해하기 위한 기술적인 노력이 필요하다. 시각화란 보이지 않는 것이 일정한 형태로 나타나 보이게 하는 것이다. 데이터 시각화란 사용자에게 더 효과적으로 정보를 전달하기 위하여 그래픽 요소를 활용하여 데이터가 정보로서 의미가 생성되도록 형상화하는 것을 말한다.

현재 데이터 시각화 관련 분야는 인포그래픽스, 정보 시각화, 과학적 시각화, 통계 시각화로 나뉜다. 관계 시각화 분야에 대해서는 상관관계, 영향도 등의 인과관계 중심의 데이터간의 관계를 중점으로 연구되었다.

본 논문에서는 소셜 네트워크안의 사용자 관계를 분석하여 사용자에게 다루기 쉬운 형태의 정보를 제공하기 위해 소셜네트워크 데이터 스트리밍 기술, 2D Quartz를 이용한 사용자 관계 시각화 기술, 사용자 관계도 구현을 위한 데이터 마이닝 기술을 제시하였다.

향 후 연구로는 본 논문에서 제안한 소셜 네트워크에서의 사용자 관계 시각화 방안을 적용하여 모바일 기반 애플리케이션을 구현하고 사용자의 피드백에 관한 연구가 필요하다.

사사의 글

본 연구는 2013년도 지식경제부의 SW전문인력양성사업의 재원으로 정보통신산업진흥원의 고용계약형 SW석사과정 지원사업(HB301-13-1003)으로부터 지원받아 수행되었습니다.

참고문헌

- [1] 박지범 "온라인 소셜 네트워크에서 사용자 간의 관계 강도 변화 추정 기법" 2013 한국정보과학논문지, 2013 40(1) 44-53
- [2] 설광수 "소셜 네트워크 서비스 사용자 간의 친밀도 측정 기법 및 실험" 2012 한국정보과학논문지, 2012 39(4) 335-341