
소셜 네트워크 정보 수집 및 분석 시스템에 관한 연구

강민식* · 송은지**

*남서울대학교 산업경영공학과

**남서울대학교 컴퓨터학과

A Study on the System for Collecting and Analyzing of Social Network Information

Min-Sik Kang* · Eun-je Song**

*Dept. of Industry Management Engineering, Namseoul University

**Dept. of Computer Science, Namseoul University

E-mail : sej@nsu.ac.kr

요 약

최근 블로그, 카페 및 SNS(Social Network Service)와 같이 다양한 소셜 네트워크 환경을 통해 사용자 자신의 의견 등을 표현하고 있다. 이에 이러한 데이터를 수집 분석하는 시스템을 개발하는 업체에서는 소셜 네트워크 상의 정보를 수집하여 데이터의 패턴을 분석하여 마케팅 등에 이용하는 기술을 제안하고 있다. 그러나, 기존 기술에서는 소셜 네트워크 상에 존재하는 다양한 유형의 문서나 정보를 모두 수용할 수 없어서 그 분석에 제한이 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 소셜 네트워크 정보를 실시간으로 수집 및 분석함에 있어서, 소셜 네트워크상에 존재하는 다양한 유형의 문서나 정보를 모두 수용할 수 있고, 이용자의 필요에 따라 분석 결과를 제공할 수 있는 소셜 네트워크 정보 수집 및 분석 시스템을 제공하고자 한다.

키워드

Social Network Information, Collecting, Analyzing , DataBase Management System

1. 서 론

최근 인터넷으로 대표되는 네트워크 기술이 발전함에 따라 온라인 웹사이트 상의 게시판을 비롯하여 블로그, 카페 및 소셜 네트워크 서비스(SNS)와 같이 다양한 소셜 네트워크 환경을 통해 사용자 자신의 의견 등을 표현하고 있다. 이전에는 이러한 소셜 네트워크 상에 게시된 사용자 정보나 의견을 수집 및 분석하는 기술을 제공하지 않아 실시간으로 이슈(issue)를 관리하고 예측하지 못하였다. 최근에 소셜 네트워크상에서 쏟아지는 정보가 기업의 경영에 결정적인 영향을 미치고 있다는 인식이 확대됨에 따라 여러 솔루션 개발 업체에서는 이러한 정보를 수집하여 SNS 사용자나 커뮤니티 그룹에 대한 패턴을 분석하거나 혹은 마케팅 등에 이용하는 기술을 제안하고

있다. 그러나, 이와 같은 종래 기술에서는 소셜 네트워크 상에 존재하는 다양한 유형의 문서나 정보를 모두 수용할 수 없어서 그 분석에 제한이 있고, 나아가 소셜 네트워크 분석을 의뢰한 이용자의 필요에 따라 적응적으로 분석 결과를 제공하지 못한다는 문제점이 있다.

본 연구에서는 이러한 문제점을 해결하기 위한 것으로, 소셜 네트워크 정보를 실시간으로 수집 및 분석함에 있어서, 소셜 네트워크상에 존재하는 다양한 유형의 문서나 정보를 모두 수용할 수 있는 시스템을 제공하고자 한다.

II. 시스템 구성

제안하는 소셜 네트워크 정보 수집 및 분석 시스템은 소셜 네트워크상에 게시된 사용자의 의견

이나 정보를 수집하는 데이터 수집부와 데이터 수집부에서 수집한 사용자의 의견이나 정보를 입력받아 저장 및 관리하는 데이터베이스 관리 시스템(DBMS)으로 구성된다. 또한 데이터베이스 관리 시스템으로부터 사용자의 의견이나 정보를 입력받아 의미 파악이 가능하도록 처리 및 분석하는 데이터 분석부 및 데이터 분석부에서 분석된 사용자 의견이나 정보를 조합하고 평가하여 지속적인 이슈 관리 및 예측 작업을 수행하는 소셜 분석 플랫폼으로 구성된다.

데이터 수집부는 정형문서(formal document) 및 비정형 문서(informal document)를 모두 수집하고, 정보 유형 분석을 통해 정형 문서와 비정형 문서를 분류하며, 비정형 문서는 정형 문서로 변환하는 정제 기술을 포함한다. 또한, 데이터 분석부는 형태소 분석기 및 철자 오류가 있는 단어 또는 구문 분석을 위해 초성, 중성 및 어절 패턴 사전을 이용하여 보정하는 감성 분석 모듈을 포함한다. 소셜 분석 플랫폼은 사용자별 분석, 지역별 분석 및 이슈별 분석 중 어느 하나 이상을 수행하는 기능을 갖는다. 이 시스템은 소셜 네트워크 정보를 실시간으로 수집 및 분석한다. 특히, 소셜 네트워크상에 존재하는 다양한 유형의 문서나 정보를 모두 수용할 수 있어서 분석에 제한이 없다. 또한, 이용자의 필요에 따라 적응적으로 분석 결과를 제공하기 때문에 유용하고 신뢰성 높은 정보를 제공한다.

III. 시스템 기능 및 구현

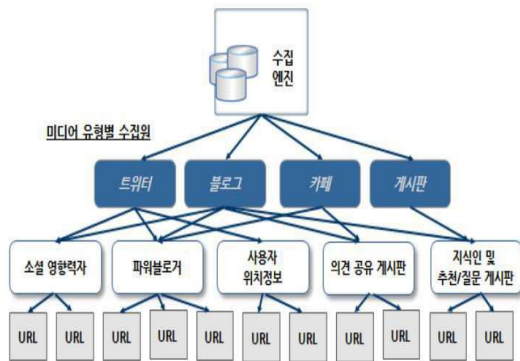


그림1. 타겟 수집원 프로파일링

데이터수집부는 실시간 수집 범위 및 정보 획득의 변화 관리 기술, 정보 유형 분석 및 수집 주기, 프로파일링 기술, 비정형 VOC(Voice of Customer) 데이터를 정형 데이터로 변환하는 정제 기술, 저비용 서버들을 이용하여 실시간 수집이 가능하도록 분산처리하는 기술을 제공한다. 실시간으로 고객 피드백 수집 기능을 제공하도록

수집 매니저 및 수집 로봇을 포함할 수 있다.

또한, 그림1과 같이 타겟 수집원 프로파일링을 수행한다. 이는, 특정 정보 수집에 특화된 타겟 사이트 추출 및 등록, 소셜 영향력자나 파워 블로거 등을 추출하여 최신 트렌드 및 여론 동향 분석, 트위터 채널 사용자 위치정보 수집을 통한 사용자 정보 분석, 이슈 정보 및 의견 공유 게시판 추출, 지식인(네이버), 추출/질문 등 지식 공유 사이트 추출, 주요 사이트 등록을 통한 유효 데이터 수집 기능을 제공한다.

데이터베이스 관리 시스템(DBMS)은 데이터 수집부에서 수집한 사용자의 의견이나 정보를 입력받아 저장 및 관리한다. 데이터 수집부의 기록 보관(archive) 기능에서는 DBMS 커넥터 모듈을 포함하고 있어서 데이터베이스 관리 시스템과 효율적인 데이터 기록을 가능하게 한다.

데이터 분석부에서는 초/중성 및 어절 패턴 사전을 참조하여 보정을 수행한다. 또한, 형태소 분석기에서 추출된 품사 정보를 기반으로 영언과 체언을 분리해서 분석 해당 품사에 종속적인 속성 데이터베이스를 구축한다. 소셜 분석 플랫폼은 데이터 분석부에서 분석된 사용자 의견이나 정보를 조합하고 평가하여 지속적인 이슈 관리 및 예측 작업을 수행한다.

IV.결 론 및 기대효과

본 연구에서는 블로그, SNS, 카페 등에 게시된 정보를 수집, 분석 및 평가하여 지속적인 이슈 관리와 예측을 가능하게 하는 소셜 네트워크 정보 수집 및 분석 시스템을 제안하였다. 인터넷과 같은 네트워크를 통해 각종 소셜 네트워크 서버(즉, 소셜 채널)에 접속하여 소셜 네트워킹을 하는 사용자의 의견이나 정보를 실시간으로 수집 및 분석하는 기술적 인프라를 제공한다. 실시간 분석된 결과를 통해 파악된 소비자 인사이트(Insight) 및 부정적 이슈 등을 사전에 감지하여 대응할 수 있는 맞춤형 서비스를 제공하며, 그 결과를 평가하여 지속적인 이슈 관리와 예측을 지원하는 기술 및 사업화 모델을 제공한다.

참고문헌

- [1] 박상원의, "온톨로지 기반 소셜 네트워크 분석을 이용한 전문가 추천 시스템", 한국정보과학회 논문지 제15권 5호 pp.390-394, 2009.
- [2] 송은지의, "고객피드백을 위한 평판분석 관리시스템에 관한 연구", 한국정보통신학회 2012년 추계 학술대회 논문집 2012.
- [3] '소셜미디어에 부는 참여형 마케팅 바람', LG경제연구원, 2010.