

빅 데이터 분석 기법을 이용한 노인의 주기별 옴니서비스 지원시스템

A Study on Omni's Senior Service Support System Using Big Data Analytics

박 경 수*, 김 광 식*, 김 창 기*, 서 정 민**
한국교통대학교*, (주)디컴즈 기술연구소**

Park Gyeongsu*, Kim Kwangsil*, Kim Chang Gi*,
Seo Jeong Min**
Nat'l Univ. of Transportation*, D-Comms Co.
Lab of Tech.**

요약

최근 빅 데이터 및 분석기술을 이용한 분야가 다양화되고 있다. 하지만 정작 사회복지분야로의 적용은 매우 미약한 것이 현실이다. 사회복지분야에서의 빅 데이터를 이용한 각종 개인화 복지 서비스의 제공은 경제적인 부분뿐만 아니라 서비스를 제공하는 기관이나 서비스를 제공받는 이들에게 모두 상호 유기적인 관계를 유지할 수 있다. 특히 사회 조직으로부터 은퇴하거나 떨어져서 개인적인 삶을 영위해야 하는 노인들에게 있어 그들의 라이프 로그와 같은 빅 데이터를 분석하여 복잡적이며 융합적인 복지 서비스를 제공하는 기법은 매우 효율적이라 할 수 있다. 이에 본 논문에서는 노인들의 빅 데이터를 이용한 시간의 흐름별 옴니 복지서비스를 제공하도록 그들의 환경을 분석하는 시스템을 제안한다.

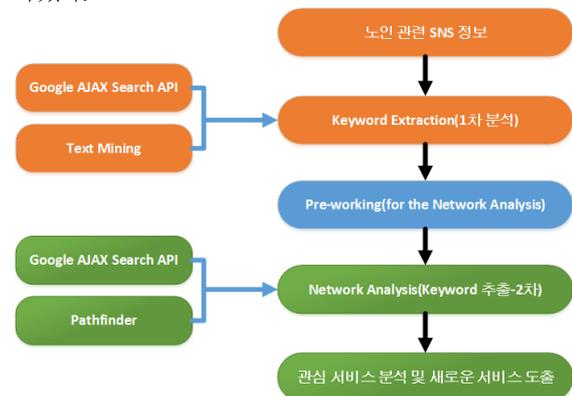
I. 서론

IT 기술의 지속적인 발전과 스마트폰, 스마트패드 등의 모바일 기기의 증가로 개인들 뿐 만 아니라 기업 및 정부도 무수히 많은 정보들을 생산하고 있다. 이에 따라 해당 시스템의 로그양이 기하급수적으로 증가하고 있다. 관련 담당자들은 관리와 의사결정을 위하여 로그들을 분석해 왔는데 이제는 분석 불가능한 정도로 많은 양이 쌓여가고 있다. 그러나 최근에는 빅 데이터 분석 기술을 이용하여 특정한 서비스를 제공하는 연구들이 많이 진행되고 있다. 그러나 사회적 약자라 할 수 있는 노인들을 위한 빅 데이터 분석기술 분야에서는 많은 연구가 진행되고 있지 않아 실제 보급하는 것은 매우 묘연한 것이 현재 상황이다. 이에 본 논문에서는 노인과 관련된 각종 빅 데이터를 수집하고 분석하여 노인들의 연령대별 주기에 맞는 옴니서비스를 추천할 수 있는 시스템을 제안한다.

II. 노인들의 SNS 빅 데이터 분석시스템

노인들과 관련된 옴니서비스를 제공하기 위한 SNS 자료를 이용하여 필요한 서비스를 예측하는 정량적인 방법으로 본 연구에서는 텍스트 마이닝과 네트워크분석 기법을 이용한 방법을 이용하였다. 그러나 이와 같은 기존의 방법들은 각각 다른 방향으로 존재하는 방식으로 필요한 서비스를 도출하는 정량적인 방법으로 활용하고 적용하기에는 어려움이 따랐다. 따라서 본 연구에서는 기존의 방식을 개선하여 중요한 키워드의 우선순위를 도출하였다. 특히, 텍스트 마이닝 방법에서는 방대한 정보에서 현

재의 트렌드를 반영한 중요한 키워드들을 도출하고, 네트워크분석 기법을 통해서 도출된 키워드들 간의 상호연관관계를 통해 우선순위를 도출하였다. 즉, 문서군내의 특정단어의 문서간 빈도수를 사용하는 것이 아니라 현재 노인들의 SNS 정보를 이용한 검색 결과값을 빈도수로 이용하고 연도별 가중치를 사용함으로써 노인들의 생애 주기별 관심 서비스를 분석이 가능하도록 하였다. 이러한 텍스트 마이닝 기법을 활용하여 방대한 정보로부터 현재 관심 분야를 반영하는 중요한 키워드를 1차적으로 추출하였다.

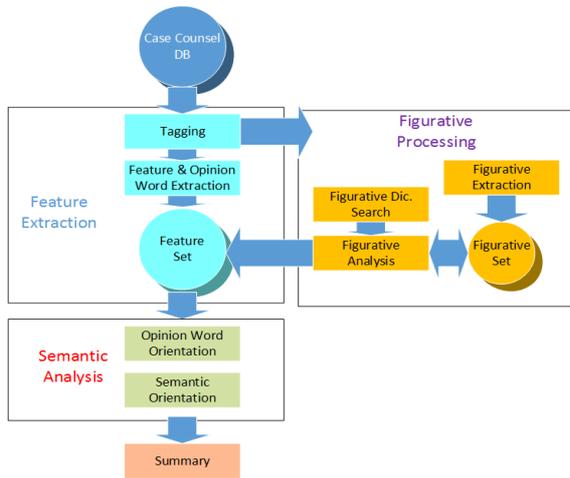


▶▶ 그림 1. 노인 SNS 자료 분석 프로세스

SNS의 노인들의 관심 서비스 정보를 반영하기 위해서 구글에서 제공하는 Open API(Application Programming Interface)를 사용하여 검색결과를 얻었으며, 검색기간,

다중 검색어 검색 등의 다양한 기능들을 활용하였다. 이렇게 도출되어진 키워드들은 네트워크 분석 방식을 활용하여 키워드간의 상호연관관계를 분석하여 우선순위를 도출하였다. 이러한 우선순위를 도출하기 위하여 네트워크 분석을 이용한 중심성 분석(Centrality Analysis)과 네트워크의 간략화 및 최적화 기법을 이용한 2가지 방식을 결합하여 분석하였다. SNS 자료를 이용한 마이닝의 실행을 위해서는 두 가지 과정이 필요하다. 첫 과정은 단어들의 패턴을 이용하여 관심 분야에 대한 문장을 구별해낸다.

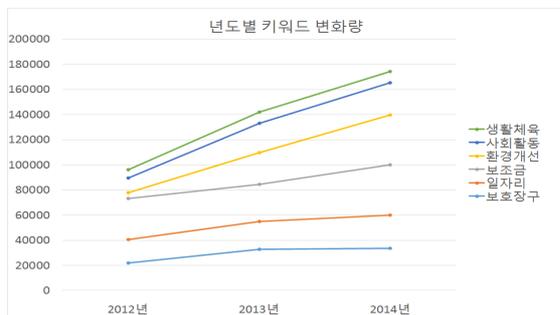
두 번째 과정은 찾아낸 관심분야에 관련된 자료들을 서비스 관련 단어를 포함한 문장과 그렇지 않은 문장으로 분류하여, 노인 서비스에 관련된 단어를 포함한 자료는 단어를 포함하지 않는 또 다른 문장에서 동일한 서비스 관련된 자료가 등장하였을 때 그 자료의 참조 대상이 된다[1, 2, 3].



▶▶ 그림 2. 관심 단어 추출 프로세스

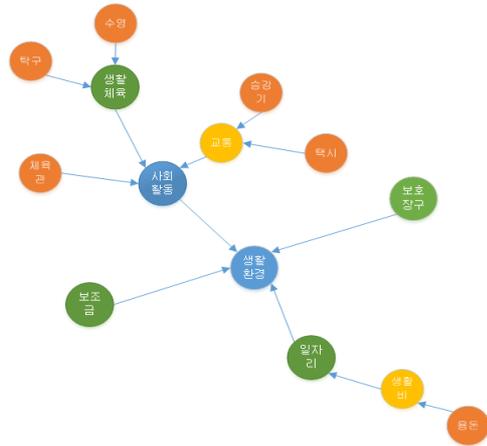
III. 실험 및 평가

본 연구에서는 텍스트 마이닝과 네트워크분석 기법을 이용하여, 노인관련 복지서비스 분야를 대상으로 트렌드 분석에 필요한 키워드를 도출하여 우선순위를 분석하였다. 특히, 노인 복지 분야의 분석대상은 노인들의 SNS를 대상으로 서비스 도출에 필요한 노인 복지에 관한 키워드들을 분석하였다. 또한, 모집단에서 분석되어진 키워드들은 노인들에게 직접 설문하여 노인 복지 분야의 관심사를 비교 분석하여 살펴보았다.



▶▶ 그림 3. 최근 노인들의 관심 분야 키워드량 변화

그림 3은 시스템을 이용한 눈 상의 키워드 변화량을 분석한 내용이다. 1단계로 현재 노인들의 관심 분야를 위한 텍스트 마이닝 분석을 위하여, SNS에서 반복적으로 노출된 키워드들을 중심으로 새로운 키워드 그룹을 형성하고, 이 키워드들의 SNS에 노출된 정보 검색 결과량을 바탕으로 키워드의 빈도수를 분석하는 방식으로 진행하였다.



▶▶ 그림 4. 주요 관심 키워드 네트워크 분석

2단계분석은 텍스트 마이닝으로 도출된 키워드들을 대상으로 서로의 우선순위를 분석하였다. 또한 구글 AJAX Search API를 이용한 프로그램, 네트워크분석의 중심성 활용, 패스파인더 방식 등 정량적으로 분석이 가능할 수 있도록 시도하였다.

IV. 결론

본 논문에서는 SNS 빅 마이닝 기술을 이용하여 sh인 관련 사회복지 서비스의 관심분야 및 트렌드를 분석하여 새로운 서비스를 도출할 수 있도록 하는 시스템을 제안하였다. 제안된 시스템은 노인 SNS의 관심 분야를 고려하여 노인 관련 SNS 사이트에 올라온 자료들을 분석하여 속성과 서비스 관련 어휘 등을 추출하고, 관심 정도를 판단하도록 하였다.

■ 참고 문헌 ■

- [1] Narayanan, R., Liu, B., and Choudhary, A., "Sentiment Analysis of Conditional Sentences", Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-09), August 6~7, 2009, Singapore.
- [2] Liu, B., "Sentiment Anlysis and Subjectivity", Invited Chapter for the Handbook of Natural Language Processing, Second Edition, To appear in Oct/Nov, 2009.
- [3] Chang Gi Kim and Jeong Min Seo, "An Opinion Mining System for A Figurative Representation of Disabilities", Proc. of KSCI Summer Conf. 2015 in Jeju, Korea, pp.95-96, 2015.