

소셜 빅데이터의 감성과 재난전조의 연관성에 관한 연구

배병길, 이보람, 최선화

국립재난안전연구원 재난정보연구실

e-mail: baze11984@korea.kr, boram1024@korea.kr, shchoi333@gmail.com

The Study on the Relationship between Disaster Signs and Sentimental of the Social Bigdata

ByungGul Bae, BoRam Lee, SeonHwa Choi
Disaster information Research Division
National Disaster Management Institute

요 약

여러 가지 예측하기 힘든 요소에 의해서 발생하는 재난을 미리 감지하는 것은 매우 어려운 일이다. 특히, 일부라도 예측할 수가 있는 자연재난이 아닌 복합재난의 경우, 측정될 수가 있는 정형적인 데이터가 존재하지 않기 때문에 재난을 예측하기 위한 데이터가 없는 것이 현실이다. 본 논문에서는 재난에 대한 전조를 감지하기 위해 소셜미디어에서 사람들이 직접 생성하는 소셜 빅데이터를 활용하여 재난과 관련된 메시지의 감성이 재난전조와 연관성이 있다는 것을 알아보고자 한다. 그래서 실제 사람들이 작성한 재난과 관련된 트윗을 수집하고 감성분석하여 재난발생 전후의 감성변화를 분석하였다.

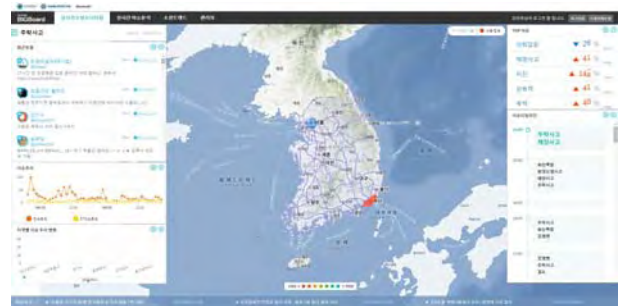
1. 서론

재난은 예측하기 힘든 요소에 의해서 발생하기 때문에 기존의 정형적인 데이터의 분석만으로 재난의 전조를 감지하거나 예측하기에는 많은 어려움이 존재한다. 특히, 정형적인 데이터가 존재하지 않는 인적재난과 복합재난의 경우 재난 예측과 전조를 연구하기 위한 올바른 데이터가 존재하지 않는 것이 사실이다. 이러한 데이터의 부재로 인하여 재난전조에 대한 연구보다는 재난의 대처와 빠른 감지에 대한 연구가 많았던 것이 사실이다. 하지만 사람들이 직접 작성한 비정형 데이터를 기반으로 만들어진 소셜 빅데이터는 여러 가지 정보와 의미를 가지고 있고, 동시에 재난에 대한 정보를 가지고 있기 때문에 기존에 존재하지 않았던 재난전조를 감지할 수 있는 가능성을 가진 데이터의 영역이다.

기존에 연구되었던 소셜 빅데이터의 재난전조에 대해서 살펴보면, 많은 사람들에 의해서 특정한 재난이 많이 언급될 경우에 그 재난이 발생되거나 발생할 가능성이 높은 재난전조로 활용하였다. 하지만 유명인을 비롯한 빅마우스에 의해서 메시지가 많이 언급되어지는 소셜미디어의 특성에 따라서 재난과 관련된 키워드를 가진 메시지가 많이 발생하더라도 재난의 전조와 연관성이 없는 경우가 많았던 것이 사실이다. 그래서 소셜 빅데이터의 감성과 재난전조의 관련성에 대해서 알아보고자 한다[1].

소셜미디어에서 사람들이 작성한 메시지를 수집해서 만들어진 소셜 빅데이터를 활용하기 위해서 국립재난안전연구원에서는 총 71개의 재난유형에 대해서 재난을 모니

터링하고 분석하는 시스템, 소셜빅보드를 개발하였다. 소셜빅보드는 다양한 종류의 재난에 대해서 모니터링 하고 실시간으로 발생하는 재난과 관련된 트윗을 감지하여 수집하고 수집된 트윗에 대해서 연관어분석과 감성분석을 과정을 거쳐서 나온 결과물을 이용하여 재난에 대해서 트윗을 활용한 심층적인 분석을 도와주는 분석 시스템이다. 본 논문에서는 소셜빅보드의 감성분석기능을 활용하여 재난전조와 연관성에 대해서 분석해 보았다[2].



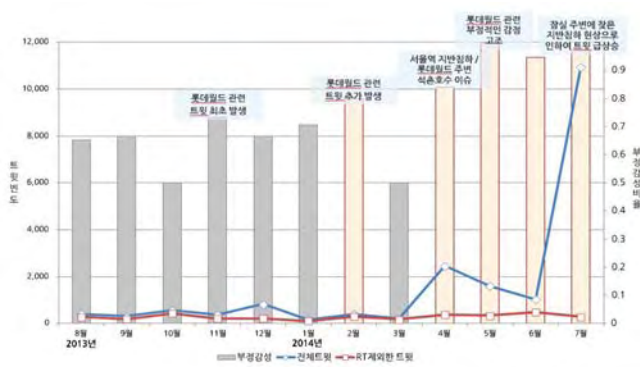
(그림 1) 소셜빅보드 실시간모니터링

2. 본론

사람들이 직접 작성한 재난과 관련된 메시지의 감성이 앞으로 발생할 재난에 대한 전조가 될 수 있다는 가능성을 확인하고자, 사람들이 작성한 재난과 관련된 트윗에 대해서 감성분석을 하여 재난전조와 연관성에 대한 분석을 진행해 보았다. 그래서 본 논문에서는 사전연구의 사례로

현재 이슈가 되고 있는 잠실의 싱크홀, 지반침하에 대한 자료를 수집하고 분석하였다. 자료는 작년 8월부터 올해 7월까지 소셜빅보드에서 수집한 지반침하와 관련된 트윗을 분석한 자료이며, 이 과정에서 지반침하와 관련되어 있다고 감지된 트윗은 총 19280건이다.

수집된 지반침하 관련 트윗의 빈도를 분석해보면 2013년에는 지반침하 트윗의 수가 많이 발생하지 않았다. 5개월동안 총 2492개의 트윗이 발생하였으며 2014년에는 16788건이라는 많은 수의 지반침하 트윗이 발생하였다. 트윗의 감성과 재난전조의 연관성에 대한 연구를 진행하기 위해서 소셜빅보드를 활용하여 수집된 트윗에 대한 감성분석을 실행하였다. 소셜빅보드에 의해 2013년에 발생한 트윗의 감성중에서 평균 부정감성의 비율은 0.6364의 비율이 나왔으며 2014년에는 평균 부정감성비율이 0.9415였다.



(그림 2) 분석 기간(2013.8-2014.7) 그래프

2014년 7월에 이슈가 되었던 잠실지역의 지반침하 전조와 부정감성의 연관성을 찾기 위해 잠실 싱크홀 사건에 초점을 맞추어 분석하였다. 2013년에는 전반적으로 지반침하와 관련된 트윗이 많이 발생하지 않았고 잠실의 싱크홀이 부각되는 시기도 아니었기 때문에 관련 트윗도 많지는 않다. 하지만 그럼에도 2013년 11월에는 잠실에 있는 석촌호수와 관련하여 수위가 낮아지고 있다는 트윗이 발생하기 시작하였다. 그리고 이러한 트윗이 발생한 11월에 2013년에 가장 높은 부정비율인 0.7333이 나왔다.

2014년에는 초반에는 비교적 적은 수의 지반침하 관련 트윗이 발생하였으며 부정적인 감성의 수치도 그리 높지가 않았다. 2월에 잠실 제2롯데월드의 안전사고가 발생함에 따라서 석촌호수가 같이 언급되어 지반이 불안하다는 언론보도가 발생하여 관련 트윗이 발생하면서 2월의 부정감성 비율은 0.8까지 올라갔지만, 지반침하에 대한 관심보다는 잠실 제2롯데월드 자체에 대한 안전에 대한 관심이 초점되어져 있었던 시기이다.

본격적으로 잠실지역의 싱크홀에 대해서 사람들이 관심을 가지기 시작한 4월부터 관련트윗이 급증하기 시작하였고 부정적인 감성도 급증하였다. 4월에는 트윗의 수가 전월에 비해서 무려 약 12배가 증가한 2455건이 발생하였고 이 트윗들의 부정감성 역시 0.84로 상당히 높은 수치를

보여주었다. 그리고 5월과 6월에도 천건이 넘는 수의 트윗이 발생하며 0.9를 넘어서는 높은 부정감성비율을 보였다. 마지막으로 석촌호수 주변에 실질적인 싱크홀이 발생한 6월 마지막주 이후인 7월에는 관련트윗의 수가 무려 10949건이 발생하였으며 부정감성비율 또한 0.9724가 나왔다.

이번 감성분석과 재난트윗을 분석해보면, 재난과 관련된 트윗이기 때문에 기본적으로 부정적인 감성의 비율이 높지만, 그 중에서도 재난의 전조에 해당되는 트윗이 나타나면 부정감성비율이 더욱 증가하는 것을 관찰할 수가 있었다. 그리고 트윗의 내용을 살펴보면 싱크홀이 발생하기 전에는 사람들이 우려하는 트윗이 부정적인 감성을 차지하고 있었으며, 석촌호수 주변의 이슈가 집중되는 4월과 5월에는 우려하는 트윗이 점점 고조되면서 불안의 감정으로 변화했다. 그리고 사건이 발생 직전인 6월과 직후인 7월에는 사건에 대한 심각성을 언급하며 안전에 대한 불만과 관련된 트윗이 급증하면서 사람들의 부정적인 트윗이 급격히 증가하였다.

3. 결론

본 논문에서는 소셜 빅데이터에 대한 감성분석 데이터와 재난전조의 연관성을 알아보기 위한 연구를 진행하였다. 그 과정에서 소셜빅보드 시스템을 사용하여 지반침하에 대한 실제 발생한 1년간의 트윗 데이터를 사용하였으며 분석한 결과는 다음과 같다.

정형적인 데이터와 마찬가지로 사람들이 작성한 메시지가 데이터가 되어서 재난 전조를 감지할 수가 있다. 특히, 잠실과 같이 유동인구가 많거나 인구가 많은 지역의 경우에는 문제가 발생시에 그 재난과 관련된 트윗을 많이 작성하였다. 특히, 재난 전조에 해당되는 메시지는 부정적인 감성의 비율이 압도적으로 높았다. 그래서 지반침하와 관련된 트윗에서도 사건이 발생하기 전에는 사건 발생을 우려하는 부정적인 감성의 비율이 조금씩 증가하였으며 실제 싱크홀이 발생한 이후에는 트윗수가 급증하면서 부정적인 감성의 높은 비율이 유지되는 현상이 보여졌다.

본 논문에서는 잠실 싱크홀과 관련된 트윗과 이 트윗의 감성비율에 대해서 분석하였다. 하지만 데이터를 분석하는 과정에 있어서 감성에 있어서 구체적인 지수나 세부적인 분류가 추가적으로 필요하다는 점을 느꼈으며, 이 필요를 해결하기 위해서 추가적으로 세부적인 감성지수를 연구가 필요함을 인지하였다. 이후에 더욱 세분화된 감성을 연구하여 트윗에 대해서 분석을 한다면, 구체적이고 정확한 감성을 이용한 재난전조를 감지할 수 있을 것이다.

참고문헌

[1] 최선화 외, “빅데이터 기반 재난 전조감지 모델” 한국방재학회, Vol. 14, pp149-157
 [2] 최원석 외, 비정형 빅데이터 수집 및 분석모델 개발, 국립재난안전연구원, 2013