

소셜미디어 재난이슈 탐지 및 모니터링

최선화, 최우정, 이종국
국립재난안전연구원 재난정보분석센터
e-mail:shchoi333@gmail.com

Monitoring and Sensing Disaster Issues from Social Media

Seon-Hwa Choi, Woo-Jeong Choi, Jong-Kook Lee
Disaster Information Analysis Center,
National Disaster Management Institute

요 약

최근 IT기술발전으로 데이터 저장 및 처리비용은 하락하고 소셜네트워크 서비스 등이 확대되어 데이터의 양적 팽창이 가속화되면서 빅데이터에서 가치를 창출하려는 시도가 늘어나고 있다. 특히, 국내 스마트폰 이용자가 3천만을 넘어서면서 모바일 기기는 대량 데이터 생산의 원천이 되고 있다. 모바일 기기는 스마트시대 대표 통신채널인 소셜미디어 활용을 가속화 시키며 다양한 분야에서 영향력을 발휘하고 있다. 특히 일본과 아이티 지진피해, 강남역 침수 등 재난발생 시 피해상황 파악 및 경보, 정보교환 등에 소셜미디어 채널이 활발히 활용되면서 새로운 매체로 떠오르고 있다. 본 논문에서는 언어분석을 통해 소셜미디어 속에 재난이슈를 모니터링하고 소셜미디어가 재난예방 및 대응에 적극 활용되기 위해 필요한 기술개발 방안을 제시하고자 한다.

1. 서론

정보사회 패러다임은 과거 PC시대에서 현재 모바일 시대를 지나 미래에는 사물 정보화, 지능화, 개인화로 상황인식, 미래전망, 개인 맞춤형 서비스 등이 실현되는 스마트 시대로 변화할 것[1]으로 보아 정보사회의 핵심 연료인 데이터의 가치는 날로 증가하고 있다. 세계적 컨설팅기관인 IDC가 디지털 정보량이 기하급수적으로 증가하여 2020년에는 관리해야 할 정보량이 50배 급증하고, 10배 많은 서버가 필요할 것이라고 전망할 만큼 대규모 데이터가 중대 이슈로 부각되면서 ‘빅 데이터(Big Data)’라는 용어가 등장하였다. 최근 기술발전으로 데이터 저장 및 처리비용은 하락하고 소셜네트워크 서비스 등은 확대되어 막대한 데이터가 생산되고 있다. 더불어 상황 모니터링을 위한 각종 센서(강우계, 수위계, CCTV 등) 뿐만 아니라 도로, 건축물에 내장된 센서에서 데이터가 만들어지고 있어서 데이터 양적 증가세는 가속될 것이다.

국내 스마트폰 이용자가 3천만을 넘어[2]서면서 모바일 기기는 대량 데이터 생산의 원천이 되고 있다. 세계적으로 스마트폰 이용자의 60~70%가 소셜미디어를 적극 활용[3]하고 있다는 보고는 앞으로 소셜미디어의 중요성이 더욱 증가하고, 다양한 분야에서 영향력을 발휘할 것으로 예측된다. 최근에는 소셜미디어 기반의 인터랙션 데이터를 분석

하여 각 분야에 활용하고자 하는 요구가 급증하고 있다 [2].

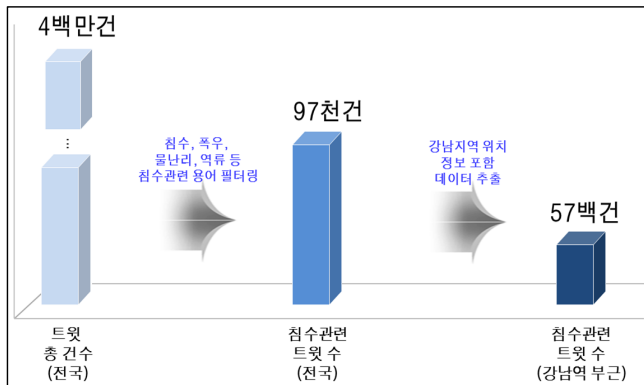
일본과 아이티 지진피해 당시 피해상황 파악 및 경보, 정보교환 등에 소셜미디어 채널이 활발히 활용되면서 신속·정확한 재난관리 활용 가능성을 열어 주었다. 우리나라 경우, 2010년 9월 추석연휴 첫날, 서울 및 경기지역에 집중호우가 발생했고, 당시 곧바로 물에 잠긴 홍대 전철역 사진이 트윗에 올라오면서 피해상황이 최초로 알렸으며, 부산 해운대 아파트 화재사건 당시 불꽃이 살아있고 구경하는 사람도 담겨있는 이른바 화재 사진으로는 특종감이었던 화재사진이 트윗에 올라왔다. 부산발 KTX가 급정역에 멈춰선 사건 또한 코레일 본사에 사고보도가 안된 시점에 사고를 알리는 트윗이 먼저 올라왔다.



(그림 1) 동일본 지진당시 트윗발생 현황(2011)

1) IDC(2011), 'Digital Universe study'
2) 방송통신위원회, 2012년 2월 기준
3) 한국, 중국, 미국, 일본 국가별 모바일 주요서비스 이용률 조사, ourmobileplanet.com 2011.12

트윗은 정보전달의 신속성뿐만 아니라 국민 스스로 긴급정보 공유의지의 수단으로 활용되었다. 강남역 침수가 일어난 2011년 7월 26일 12시부터 7월 27일 12시까지 하루동안 전국적으로 발생된 트윗은 약 4백만 건, 이 중 침수, 폭우, 물난리, 역류 등과 같은 침수관련 용어를 포함한 트윗은 약 9만7천 건이었다. 트윗에 강남지역 위치정보를 포함한 것은 약 5천7백건으로 하루동안 침수를 알리는 엄청난 정보가 발생하였다. 그 트윗안에는 침수의 전조, 침수피해 상황, 대피방법 등 물리적 센서로는 획득할 수 없는 정성적 정보가 전달되었다. 이것은 피해현장 혹은 인근에 사람이 재난 상황정보 전달의 센서가 되는 휴먼센서(Human Sensor)라 할 수 있다.



(그림 2) 강남역 침수관련 트윗발생 현황(2011)

국가정보화전략위원회가 발표한 빅데이터를 활용한 스마트 정부구현 방안에는 다양한 공공부문 활용사례 및 가능성을 제시[3]하고 있다. 특히, 피해 및 위험을 사전에 예측하고 대응할 수 있는 선제적 재난관리 활용사례는 적용가능성 검토를 위해 행정안전부가 주관하여 파일럿 시스템을 구축하였다. 도시 내수침수 전조예측으로 위험지역을 분석하고 신속한 경보를 통해 국민 피해 최소화를 목표로 강우정보와 침수예상도 등 정형데이터와 소셜네트워크 서비스 및 뉴스 등 비정형데이터를 연계하여 도시 내수침수 전조를 예측할 수 있는 모델을 수립하였다. 국립재난안전연구원은 모델 설계에 참여하였으며 모델은 2011년도 침수피해 발생지역인 강남구(역삼분구), 종로구(효자분구)를 대상으로 검증을 실시하였다. 파일럿을 통해 비정형데이터를 활용한 모델이 국지적 폭우, 지역적 특수상황 등에 기인하는 예측되지 못한 침수 전조를 감지하고 예측된 침수 위험에 대한 신뢰도를 더욱 향상시킬 수 있는 새로운 예-정보체계로 활용 가능함을 확인하였다[4].

이처럼 스마트시대에 소셜미디어는 관계기관보다 한발 빠른 피해상황 보고매체이며 국민과 소통하는 채널로서 활용되고 있기 때문에 재난관리에 초점을 맞추어 소셜미디어의 내용 및 위치정보를 추출하고 관련 정형데이터와 연계분석을 통해 재난 전조의 위험 가능성을 평가하는 기술개발이 필요하다. 본 논문에서는 대표 소셜미디어인 트윗 데이터를 분석하여 겨울철 재난인 대설, 한파, 강풍 발

생과의 상관관계 분석 결과를 제시한다. 또한 소셜미디어 등의 텍스트 데이터가 재난예방 및 대응에 적극 활용되기 위해 선행되어야 할 기술개발 방안을 제시하고자 한다.

2. 트윗으로 보는 재난이슈

2.1 언어분석

트윗으로부터 재난이슈를 모니터링하기 위해서 트윗 데이터 수집하여 언어처리과정을 거쳐 관계 데이터 추출 과정을 수행하였다[4].

데이터는 트위터의 API를 이용해 한국인 이용자로 판별된 계정의 트윗을 수집하였으며, 수집된 데이터의 URL 분석과 자연어 처리를 동시에 병렬적으로 처리하였다. 이용자가 특정 URL을 인용하면서 트윗에 포함된 shortened URL을 분석해 실제 URL과 해당 URL의 유형을 판단한다. 즉, 사진, 뉴스, 동영상, 블로그, 커뮤니티 등으로 분류하고 뉴스나 블로그, 커뮤니티 등 어떤 제목을 분석하여 내용을 파악하였다.

언어처리는 형태소 분석을 통해 문장을 형태소 단위로 구분하고 품사(Part Of Speech)를 결정한 뒤, 명사(Noun)만을 대상으로 복합 색인어를 추출한다. 복합명사는 2개 이상의 명사로 구성되어 의미를 구성할 수 있는 단위로서 예를 들면, “지금 출근길입니다”라는 트윗에서 “출근”과 “길”은 각각 명사지만 “출근길”로 합쳐졌을 때도 하나의 의미를 구성할 수는 복합명사이므로 “출근길”도 인식하는 것이다. 마지막으로 트윗에서 리트윗(RT), 리플(Reply) 등의 관계 데이터를 추출하여 트윗의 확산력, 노출도 분석을 수행하였다.

재난유형은 겨울철 재난인 대설, 한파, 강풍으로 한정하고 2012년 11월부터 2013년 2월까지 트윗속의 재난이슈를 모니터링 하였다. 각 재난유형의 키워드는 전문가조사를 통해 <표 1>과 같이 정의하였다.

<표 1> 재난유형별 키워드 집합

재난유형	키워드
대설	폭설, 눈폭탄, 대설, 눈보라, 눈&교통, 눈&통제, 도로&눈, 눈&비닐하우스, 눈&축사, 눈&붕괴, 피해&눈, 빙판길, 정전&눈, 제설, 눈&고립, 낙상&눈길, 잔설
한파	강추위, 최저기온, 한파, 맹추위, 추위&정전, 추위&피해, 추위&사고, 온도&떨어, 체감온도, 영하권, 동파, 추위&전력, 결빙, 전력난, 전력수급
강풍	강풍, 강한바람, 돌풍&추위, 간판&바람, 바람&양식장, 정전&바람, 바람&가로수, 바람&전봇대

2.2 재난이슈 모니터링 시범서비스

국립재난안전연구원은 2012년 11월부터 5개월간 트윗에서 겨울철 재난인 대설, 한파, 강풍을 모니터링하는 서비스[5]를 시범적으로 운영하였다. 서비스에서는 트윗 원문

4) 소셜미디어 분석 서비스 『소피언』 활용, (주)사이람
5) 국립재난안전연구원 『트윗으로 보는 재난이슈』(www.ndmi.go.kr),

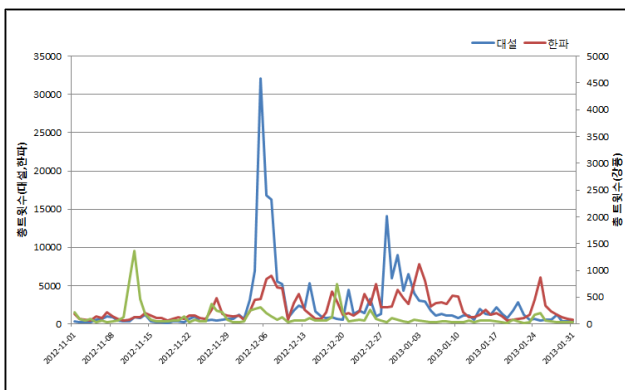
분석의 결과를 일별·시간별 트윗량 추이, 월간·일간 핫 이슈, 그리고 관련 사진 및 뉴스 등을 제공한다(그림 3). 심층언어분석 과정을 수행하지 않고 단순 키워드기반 발생빈도 중심의 분석결과이므로 정확성은 떨어지지만 소셜미디어 분석이 재난관리에 활용되기 위한 시사점을 얻을 수 있었다.



(그림 3) 『트윗으로 보는 재난이슈』 시범서비스 화면

재난이슈 트렌드를 보면, 대설에 대한 언급이 가장 많았으며, 한파, 강풍 순으로 트윗에서 많이 이야기하고 있었다(그림 4).

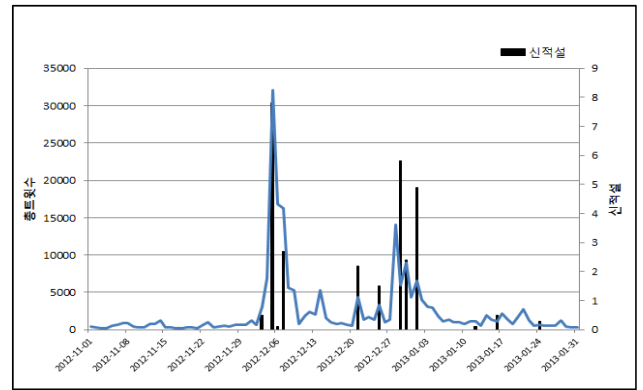
재난관련 트윗량의 추이는 기상 변화와 관련이 있었다. 즉, 대설에 대한 트윗과 서울지역 적설량을 비교해 보면 적설량이 많을 때 대설관련 트윗량도 증가하는 패턴을 보였다. 즉, 적설량이 가장 많았던 2012년 12월 5일에 트윗량도 가장 많은 것으로 분석되었다(그림 5).



(그림 4) 재난관련 트윗량 추이(2012.11 ~ 2013.1)

반면에 한파는 기온이 낮을수록 트윗량이 증가하였다. 최저기온 영하 16.4도를 기록했던 2013년 1월 3일 트윗량이 약 8천여건으로 가장 많이 발생한 것을 확인할 수 있었다. 강풍관련 트윗 또한 평균풍속이 강할 때 트윗량도 증가하는 걸 확인할 수 있었다. 다만, 전국단위 트윗량을 서울지역의 기상데이터와 비교해서 트윗량과 기상데이터

의 패턴이 정확히 일치하는 것은 아니지만 유의미한 상관성을 얻을 있어서 재난관리 활용 가능성을 높여주었다. 향후 지역단위 기상데이터와 트윗발생 데이터를 비교하여 정밀한 상관성 비교가 필요할 것이다.



(그림 5) 적설량과 대설 트윗량 추이(2012.11 ~ 2013.1)

3. 결론 및 향후계획

다양한 분야에서 영향력을 발휘하고 있는 소셜미디어는 각계분야에서 사람들 간의 관계를 토대로 소위 세상을 바꾸는 힘으로 등장하고 있다. 특히, 재난 및 안전관리분야에 긍정적 사례들이 보도되면서 그 힘이 막강해지고 있다.

국립재난안전연구원은 재난 및 안전관리 기술개발을 총괄하는 국립연구기관으로서 사회 패러다임의 변화를 선도적으로 반영하기 위해 지난해 11월부터 5개월간 『트윗으로 보는 재난이슈』 서비스를 시범운영하면서 재난에 대한 국민의 목소리를 분석하였다. 단순 텍스트 분석 및 발생빈도 분석의 한계는 있었지만 5개월간의 이슈 모니터링 결과, 재난관리에 활용성이 높다는 것을 알 수 있었다.

향후, 심층 언어분석과 위치정보 매핑 기술 등을 적용하여 실시간으로 소셜미디어의 이슈를 탐지하고 모니터링한다면 재난의 전조감지가 가능할 것이다. 또한 소셜미디어 모니터링 정보와 기존의 정형 데이터 즉, 기상, 센서, 시뮬레이션 등의 데이터를 융합·분석하여 좀 더 정확하고 정밀도 높은 의사결정 지원기술 개발이 필요하다.

이에 국립재난안전연구원에서는 2013년부터 빅데이터 기반 재난 전조감지 및 대응기술 개발 연구를 진행 중이다. 비정형 데이터와 정형 데이터를 묶은 빅데이터 융합·분석으로 재난 예방 및 대응 기술을 단계적으로 개발한다. 1단계는 비정형 빅데이터 수집·분석을 통한 재난 전조감지 기술을 개발하고 2단계, 재난 빅데이터 융합·분석을 통한 재난 대응·전과기술 개발을 목표로 하고 있다. 이는 민간·공공의 빅데이터 융합 분석을 통해 재난전조를 사전에 감지하여 재난위험 사각지대를 해소하고 소셜네트워크를 통한 실시간 제보 분석으로 맞춤형 신속대응 및 상황전파를 가능하게 할 것이다.

참고문헌

- [1] 한국정보화진흥원, “선진국의 데이터기반 국가미래전략 추진현황과 시사점”, 2012
- [2] Informatica, “Harnessing Big Data”, <http://www.informatica.com/us/vision/harnessing-big-data/>
- [3] 국가정보화전략위원회, 빅데이터를 활용한 스마트 정부 구현(안), 2011
- [4] 국립재난안전연구원, “재난안전과 빅데이터”, 재난안전지, 2012년 12월