빅데이터 기반 재난 전조감지 모델

The Model for Sensing of Disaster Signs Based on Big Data

최선화*

Choi, Seon Hwa*

Abstract

In recent years, social media has become ubiquitous and important for social networking and content sharing. Especially, in the disaster area, social media supports backchannel communications, allowing for wide-scale interaction that can be collectively resourceful, self-policing, and generative of information that is otherwise hard to obtain. At the time of great Japan and Haiti earthquake, social media channels actively were utilizing to grasp the damage, to warn, and to exchange information. This paper is to introduce the model for sensing the signs of inundation and detecting inundation risk areas by analyzing big data related to disasters. This model is comprised of three steps: a sign sensing step through monitoring and analyzing unstructured data such as social media, a risk detecting step through comparing and analyzing structured data such as precipitation, inundation hazard maps, and so forth, and a disaster status dissemination step to disaster related organizations, local governments, and the public. By applying our model to Gangnam inundation damage, 2011, in Korea, we substantiated that there is the potential for utilization on our model.

Key words : Big data, Disaster signs, Social media

요 지

최근 소셜미디어는 사회적 소통과 자료공유를 위해 언제 어디서나 사용되는 중요한 통신매체가 되고 있다. 특히, 재난분야에서 소셜미디어는 다른 방법으로 얻기 힘든 집단지성과 정제된 종합정보를 광역적으로 소통할 수 있는 비공식 통신채널이다. 아이티와 동일본 대지진 당시, 소셜미디어는 피해를 파악하고 경고 및 정보를 교환하는 채널로 적극 활용되었다. 본 논문에서는 정부의 빅데이터 마스터플랜의 활용가능성을 검증하기 위해 개발된 파일럿 모델을 소개하고 재난관리에 실제 적용하기 위한 방안 및 국립재난안전연구원의 관련기술 개발 현황을 설명하고자 한다. 이 모델은 최근 피해가 급증하고 있는 도시 내수침수 전조를 초기에 감지하여 위험지역을 분석하고 신속한 경보를 위해 강우정보와 침수예상도 등 정형데이터와 소셜네트워크서비스(SNS) 및 뉴스 등 비정형데이터를 연계하여 도시 내수침수 전조를 감지할 수 있는 모델이다. 2011년 발생한 두 개 분구의 침수사례를 통해 모델의 활용 가능성을 검토하였으며, 향후 국지적 폭우, 지역적 특수상황 등에 기인하는 예측되지 못한 침수 전조를 감지하고 예측된 침수 위험에 대한 신뢰도를 더욱 향상시킬 수 있는 새로운 예·경보체계로 활용 가능할 것이 기대한다.

핵심용어 : 빅데이터, 재난 전조, 소셜미디어

1. 서 론

최근 고성능·고집적, 모바일 기술 발전, 하드웨어 비용저하, 소셜네트워크서비스(SNS) 발달, 각종 사물·그림자 정보증가 등으로 정보기술(IT)의 패러다임은 데이터의 양적·질적 팽창 속 혼돈을 해결하고 새로운 가치를 창출하려는 방향으로 변화하고 있다(한국정보화진흥원, 2011). 즉, 과거 쓰레기처럼

버려지고 관심 갖지 않았던 막대한 데이터가 새로운 가치가 되어 문제해결의 핵심자원이 된다는 빅데이터 기술이 화두가 되면서 경제적 가치창출, 사회적 문제해결, 새로운 IT 패러다임의 견인이 될 것으로 주목받고 있다.

특히, 모바일 기술 발전으로 스마트폰 사용자는 국내의 경우 2012년 말 기준 약 3,500만대¹⁾에 육박했고 세계적으로 12억 시대²⁾를 맞이하고 있다. 스마트폰 사용자의 스마트폰 활용

*교신저자. 정회원. 안전행정부 국립재난안전연구원 공업연구사(Tel: +82-2-2078-7741, Fax: +82-2-2078-7744, E-mail: shchoi333@gmail.com)
Corresponding Author. Member. Principal Researcher, National Disaster Management Institute of MOSPA

1) 방송통신위원회, 2012년 2월 기준

2) 한국, 중국, 미국, 일본 국가별 모바일 주요서비스 이용률 조사, ourmobileplanet.com 2011.12

Fig. 1. Hurricane Sandy and the Geographies of Flooding on Twitter.

현황을 조사한 결과 약 70%가 소셜미디어를 적극 활용한 것으로 나타났으며, 단순히 지인들과의 친분을 나누는 장에서 벗어나 집단지성과 정제된 종합정보를 광역적으로 소통할 수 있는 비공식 통신채널임이 입증되고 있다. 특히, 재난분야에서 소셜미디어는 국민 스스로 피해상황을 파악하고 경고 및 정보를 교환하는 새로운 채널로 활용되면서 그 가치가 매우 중요하게 되었다. 아이티와 동일본 대지진, 우리나라 강남지역 침수피해, 미국 허리케인 샌디, 최근 아시아나 항공기 착륙 사고 등 소셜미디어를 통해 재난발생 현장을 생생히 보도하고 피해 최소화화를 위해 정보를 공유하였다. 이 같은 소셜미디어의 재난관리 잠재력을 이용해 신속·정확하고 지능적인 재난관리 체계 마련 연구가 필요하다.

미국 동부를 강타한 허리케인 샌디(2012.11) 발생 당시, 뉴욕은 샌디와 관련된 트윗의 본부좌도 같았다. Fig. 1(a)³⁾는 허리케인 샌디를 언급했던 트윗의 발생 위치와 빈도를 지도에 맵핑한 것으로 파란 원이 클수록 해당 트윗의 빈도가 높은 것을 의미한다. 실제 가장 큰 피해가 있었던 맨하탄의 송전탑 부근에 가장 많은 트윗이 발생했음을 알 수 있다. 또한, 샌디 피해가 발생할 당시 “flooding”이라는 키워드가 언급된 트윗의 위치와 건수를 지도에 맵핑한 결과 동부 전역에 해당 트윗이 다수 발생했으며 특히 허리케인 샌디의 경로를 따라 다수 트윗이 발생했음을 알 수 있다(Fig. 1(b)).⁴⁾

동일본 대지진(2011.03) 당시, 국가 통신망, 기관망 등이 마비된 상태에서 국민들 스스로 다양한 소셜미디어를 활용하여 지진 피해상황 및 대피소 정보를 공유하여 효과적으로 대응한 사실은 소셜미디어의 위력을 잘 설명해주는 사례이다(Fig. 2).⁵⁾

지금까지 사례들이 재난이 발생한 이후에 소셜미디어 활용

Fig. 2. Social Media During Earthquake, 2011 in Japan.

현황을 분석하고 그 가치를 제시하는 사례였다면 특정사건을 해결하기 위한 목적을 두고 커뮤니티 단위로 활용한 사례가 있다. 허리케인 샌디 당시 송전탑 피해로 인해 전력공급이 어려워 집안에 발전기를 가동하기 위해 기름 확보를 위한 주유대란이 벌어졌고, 이 사건은 주유소 매플러 서비스라는 새로운 접근방식으로 해결되었다. 즉, 주유소를 방문한 일반 시민들이 페이스 북, 트위터 등을 통해 주유소의 기름 보유상태, 연락처, 대기시간 등을 매플러 학생들에게 전달하고 이 정보를 지도에 매핑하여 공유하는 서비스를 제공함으로써 주유대란이 진정되면서 해결된 것이다. 이처럼 커뮤니티 기반 특정 목적을 갖는 소셜미디어 활용은 신뢰성 높은 집단지성을 활용할 수 있는 장점이 있다.

소셜미디어 속, 재난관리 사례는 국내도 예외는 아니다. 2010년 9월 추석연휴 첫날, 경기지역에 기록적인 폭우가 발생하였을 때 곧바로 홍대 전철역이 물에 잠겼고 물에 잠긴 사진과 함께 침수사실을 보도한 것은 트위터의 한 트윗이었다. 이는 어느 언론매체에서도 보도되지 않았던 상황이었으며, 이미 홍대 전철역이 범람한 후에 TV에서 재난방송이 시작되었다.

3) <http://www.floatingssheep.org/2012/10/hurricane-sandy-and-geographies-of-tweet/>

4) <http://www.floatingssheep.org/2012/10/the-urban-geographies-of-hurricane.html>

5) <http://www.aljazeera.com/indepth/features/2011/03/2011131474027904222.html>

또한, 부산 해운대 아파트 화재사건 당시 불꽃이 살아있고 구경하는 사람도 담겨있는 이른바 화재 사진으로는 특종감이었던 화재사진이 트윗에 올라왔다. 부산발 KTX가 금정역에 멈춰선 사건 또한 코레일 본사에 사고보도가 안된 시점에 사고를 알리는 트윗이 먼저 올라왔다. 최근 미국 샌프란시스코공항에 아시아나 항공기가 추락 사고는 우리나라 시간으로 7월 7일 오전 3시27분에 발생했다. 대형 사고였지만 미국에서 일어난 데다 시간대도 새벽이어서 실시간으로 파악하기 어려웠지만 1시간 남짓 후 사건발생을 최초 보도한 것은 트위터에서 이루어졌다. 당시 항공기사고 현장과 대피하는 승객들의 모습이 담긴 사진도 제공하면 사건현장을 생생히 보도한 것이다. 이처럼 트윗은 재난상황시 정보전달의 신속성뿐만 아니라 국민 스스로 긴급정보 공유의지의 수단으로 활용되었다(최선헌 등, 2013). 2011년 강남지역 침수피해 역시 광화문 침수와 유사한 사례로 소셜미디어를 통한 피해상황 공유로 인명피해를 최소화 할 수 있었던 사례였다.

현 사회는 사회문제, 환경문제, 안전문제 등 지금까지의 일하는 방식으로 해결하기 어려운 국가사회 현안문제에 직면하고 있어 미래를 예측하고 대응할 수 있는 선제적 국정운영 요구가 증대하고 있다. 이미 다양한 산업분야에서 새로운 가치창출에 기여하고 있는 빅데이터 기술을 도입하여 정부를 포함한 공공분야에서도 데이터 활용을 통해 사회현안 및 국민의 니즈를 파악하고 미래전략 수립, 선제적 공공서비스 제공 등 정부 혁신이 요구된다. 재난관리도 정부중심 대응의 기존 방식에서 벗어나 소셜미디어 등의 빅데이터를 활용하여 국민이 참여하는 재난대응과 정치·사회·경제 등 제반 이슈와 연계된 분석·예측 방식으로 변화하여야 하며 이를 위한 기술개발이 필요하다.

본 논문에서는 안전행정부의 빅데이터 마스터 플랜의 활용가능성을 검증하기 위해 개발된 파일럿 모델을 소개하고 재난관리에 실제 적용하기 위한 방안 및 국립재난안전연구원의 기술개발 현황을 제시하고자 한다. 이 모델은 최근 피해가 급증하고 있는 도시 내수침수 전조를 조기에 감지하여 위험지역을 분석하고 신속한 경보를 위해 강우정보와 침수예상도 등 정형데이터와 소셜네트워크서비스(SNS) 및 뉴스 등 비정형데이터를 연계하여 도시 내수침수 전조를 감지할 수 있는 모델이며, 3단계 프로세스로 구성되어 있다. 1단계는 소셜미디어와 같은 비정형 데이터를 모니터링하고 분석하는 전조감지 단계이며, 전조가 감지되면 침수와 관련된 정형데이터, 즉 강우, 침수예상도, 침수 흔적도 등과 비교분석하여 위험을 판단하는 단계를 거쳐 피해위험이 있다고 판단되는 경우 관련 기관, 지자체, 대국민 등에 전파하는 단계로 구성되어 있다. 이 모델은 2011년 발생한 침수사례를 통해 활용 가능성을 검토하였으며, 향후 국지적 폭우, 지역적 특수상황 등에 기인하는 예측되지 못한 침수 전조를 감지하고 예측된 침수 위험에 대한 신뢰도를 더욱 향상시킬 수 있는 새로운 예·경보체제로

활용 가능할 것이 기대한다.

2. 재난 전조감지 모델

2.1 재난관리와 소셜미디어

소셜미디어는 신속한 전파능력과 함께 잘 정리된 정확한 정보, 즉 집단지성의 결정체라고 할 수 있다. 이런 잠재능력은 재난분야 뿐만 아니라 다양한 산업분야에서 가치창출을 위해 각광받고 있다. 소셜미디어의 대표주자인 트윗은 특히, 재난 분야에 있어서 정보전달의 신속성뿐만 아니라 국민 스스로 긴급정보를 공유하려는 의지를 실현하는 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 9만7천건 발생했다. 이 트윗들 중에서 강남지역 위치정보를 포함한 것만 추출하였더니 약 5천7백건 트윗으로 하루동안 특정지역의 침수를 알리는 정보로는 엄청난 양이 발생한 것이다(Fig. 3). 그 트윗들 안에는 침수의 전조, 침수피해 상황, 대피방법 등 물리적 센서로는 획득할 수 없는 정성적 정보가 있었다. 또한, 피해현장 혹은 인근에 사람들이 재난 상황정보를 전달하는 센서, 즉 휴먼센서(Human

Fig. 3. The Number of Tweets For a Day (2011).

Fig. 4. Tweet Locations Occurred in Gangnam.

Sensor)가 되어 소셜미디어라는 소통채널에 정보를 보낸 것이다(최선화 등, 2013). Fig. 4의 붉은 도트는 강남역 침수피해 관련 트윗이 발생한 위치를 표시한 것이다. 본 장에서는 정부의 빅데이터 마스터 플랜의 활용가능성을 검증하기 위해 개발된 파일럿 모델을 소개하고자 한다. 이 모델은 소셜미디어와 같은 비정형 데이터를 모니터링하고 분석하는 전조감지 단계, 전조가 감지되면 침수와 관련된 정형데이터, 즉 강우, 침수예상도, 침수 흔적도 등과 비교분석하여 위험을 판단하는 단계, 피해위험이 판단되는 경우 관련기관, 지자체, 대국민 등에 전파하는 단계로 구성되어 있다.

2.2 비정형 데이터 분석

소셜미디어⁶⁾에서 침수피해의 전조를 감지하기 위해서는 Fig. 5에서 설명하는 것과 같은 몇 가지 절차를 수행한다. 우선, 트위터에서 침수관련 실시간 데이터를 수집하고, 언어처

리(불용어와 스팸데이터를 제거하기 위한 형태소 분석) 과정을 거쳐 의미있는 데이터만을 추출하기 위해 불필요한 데이터를 필터링한다. 다음 단계로 트윗 원문에서 위치정보를 추출하여 위치기반 데이터 분석을 수행한 후, 트윗 발생빈도와 결합하여 침수 전조가능성을 판단한다.

2.2 정형 데이터 분석

비정형 데이터분석 과정을 통해 침수 전조가 감지되면 좀 더 상세한 침수 위험지역을 판별하기 위해 강우량, 침수예상도, 침수실적도 등과 같은 침수관련 정형데이터와 융합·분석을 수행한다.

침수예상도는 모의와 해석 작업을 기반으로 생성되는 데이터로 전조가 발생된 지역의 침수위험을 실시간으로 획득하기에는 시간·성능의 한계가 있다. 따라서, 사전에 시범지역의 강우량, 도시관망도 등의 자료를 기반으로 침수해석 모델을

Fig. 5. The Procedures Sof Inundation Sign Sensing.

Fig. 6. An Inundation Hazard and History Map in Gangnam.

6) 본 연구에서 다루는 소셜미디어는 주로 트위터이고 이후 소셜미디어는 트위터로 언급하겠음.

Fig. 7. The Procedure of Building Scenario Database of Inundation Hazard Maps.

구축하고 빈도별 · 지속시간별 강우 시나리오를 설정한 후 침수예상도를 작성한다. 이렇게 작성된 침수예상도들을 데이터베이스로 구축하고 실제 전조가 발생되면 실 강우상황과 가장 적합한 침수예상도를 데이터베이스에서 검색하여 침수예상도를 의사결정에 활용한다. 과거 해당지역의 침수피해 이력은 지자체에서 제작한 침수실적도를 활용하였다.

2.3 도시 내수침수 전조감지 모델의 개념

도시 내수침수 전조감지 모델은 3단계로 구성되어 있다. 1단계는 소셜미디어와 같은 비정형 데이터를 모니터링하고 분석하는 전조감지 단계이다. 실시간 강우량과 침수관련된 소셜미디어 글들을 분석하여 이상패턴을 감지하고 이 사실을 알린다. 중앙정부와 지자체는 해당 시스템에 접속한 상태에서 일상업무를 수행하다가 전조발생 알림 팝업이 발생하면, 사실확인을 위해 시스템에서 소셜미디어와 강우 등의 전체 추이

를 파악한다. 위험상황이 의심되면 좀 더 상세한 정형데이터, 즉 강우, 침수예상도, 침수 흔적도, CCTV 등과 비교분석하여 위험을 판단하는 2단계를 거친다. 마지막으로 피해위험이 판단되는 경우 관련기관, 지자체, 대국민 등에 전파하는 단계를 수행한다.

이 모델은 2011년도 침수피해 발생지역인 강남지역 A분구와 종로지역의 B분구를 대상으로 검증을 실시하였다. 강남지역의 경우, 침수가 발생한 7월27일 오전 7시전부터 침수피해 우려의 트윗 글들이 발생했으며, 약 2시간 전, 강우량과 트윗 발생 패턴을 분석한 결과 침수전조를 감지할 수 있었다. 피해 전에 위험 가능성을 미리 알려주는 전조감지 모델을 실용화하기 위한 세부기술들을 지속적으로 개발하고 재난관리 시스템으로 활용된다면, 향후 국지적 폭우, 지역적 특수상황 등에 기인하는 예측되지 못한 침수 전조를 감지하고 예측된 침수 위험에 대한 신뢰도를 더욱 향상시킬 수 있는 새로운 예 · 경

Fig. 8. The Conceptual Diagram of Our Model.

Fig. 9. A Screenshot of the Pilot System of Our Model.

보체계로 활용가능 할 것 이다.

이 모델은 기존방식인 예보 강우량 기반의 단순예측및 광역적 위험분석에서 강우량, 소셜미디어를 활용한 실시간 전조감지 및 상세지역의 위험분석까지 가능하게 되어 정확하고 신속한 상황파악 및 대응이 가능하게 될 것이다. 또한, 평시 중앙정부는 광역단위 전조 모니터링, 지자체는 지역단위 전조 모니터링을 수행하다가 전조가 발생되면 동시에 위험판단 및 의사결정이 가능하여 최종 전파단계를 단축시킬 수 있을 것이다. 무엇보다 재난 전조 그리고 실 상황에서 사후 상황까지 소셜미디어를 통해 국민과 소통하며 대응하고 문제를 해결함으로써 지금까지의 해결방식이 아닌 정부 3.0이 지향하는 공개, 공유, 소통, 협력을 통한 재난관리 3.0이 가능해 질 것이다.

도시 내수침수 전조감지 모델 서비스 화면은 세 개 구성요소로 설계되었다. 오른쪽 상단은 해당 지역의 시간당 강우량을 확인할 수 있는 도표가 표시되고, 오른쪽 하단에는 시간대별 트윗발생 빈도를 보여줌으로써 동시간대의 강우량과 트윗빈도를 함께 비교할 수 있으며 시간대별 추이를 확인할 수 있다. 화면 왼쪽 영역은 트윗발생 위치, 침수예상도, 침수흔적도, CCTV 위치 등을 지도기반으로 표출하고 해당정보를 클릭하거나 지도를 확대·축소하여 세부내용을 확인할 수 있다.

3. 재난관리 실용화를 위한 향후연구 및 결론

3.1 재난관리 실용화를 위한 시사점

도시 내수침수 전조감지 모델은 시범지역의 침수사례로 실효성을 검증하였지만 실용화를 위해 갈 길이 아직도 멀다. 국민이 참여하는 재난관리의 새 모델인 만큼 세부기술 개발뿐만 아니라 소셜미디어 등의 새로운 소통채널을 재난관리에 선순환적으로 활용하기 위한 운영체계 수립 또한 확고히 해

야 할 것이다.

본 연구에서 제시한 모델뿐만 아니라 최근 빅데이터를 활용한 재난관리는 소셜미디어 데이터를 분석하여 재난관리에 활용하려고 시도하는 수준이다. 빅 데이터를 활용함으로써 한발 앞서 국민의 안전을 책임진다는 관점에서 재난관리 활용은 파급 효과가 매우 크지만 진정한 빅 데이터 기반 재난관리를 위한 체계적 연구개발이 부족한 상태이다. 빅데이터 기술의 재난관리 실용화를 위해서는 정부와 각 기관들은 다음 내용에 우선순위를 두고 중점추진 해야 한다.

정부가 빅데이터 활용을 통해 비용절감과 스마트 정부 구현의 가능성을 강조하고 있지만 빅데이터 시장의 초입에 위치한 우리나라는 빅데이터 인프라 구축 및 분석기술 도입·개발을 위해 갈 길이 멀다. 따라서 정부는 주도적으로 공공정보를 개방하고 이를 활용하여 분석할 수 있는 빅데이터 분석 플랫폼 기술 개발에 적극 투자하여야 할 것이다. 우선, 정부는 빅데이터 활용을 위한 인프라 구축과 처리 및 분석기술을 도입하여 범국가적 빅데이터 플랫폼을 구축하여야 한다. 이미 IT 기업들의 전쟁터가 되고 있는 빅데이터 시장에서 최적 인프라 및 솔루션을 도입하여 표준 플랫폼을 구축하고 각 기관이 자유롭게 수평적 데이터 분석이 가능하도록 이끌어야 할 것이다. 국가정보화전략위원회(2012)는 ‘빅 데이터 마스터 플랜’을 수립하여 빅 데이터를 범정부적으로 사용하기 위한 인프라 구축과 이를 기반으로 한 빅 데이터 분석·활용 과제를 도출하였다. 안전행정부는(2013) 빅데이터 인프라를 통해 공공데이터와 민간데이터를 융합하여 분석 서비스를 제공하는 빅데이터 공동기반 플랫폼 구축과 이 플랫폼을 기반으로 시범과제를 수행하였다. 타 부처에서도 ‘빅데이터 서비스 활성화 방안’을 마련하여 민간의 빅 데이터 활용을 촉진하기 위한 계획을 발표하는 등 기업과 국가의 경쟁력을 강화하고, 사회 현

안 해결과 스마트 라이프 구현을 기대하고 있다(국립재난안전연구원, 2012).

위와 같은 정부의 투자와 함께 각 기관은 빅데이터 표준 플랫폼을 활용하여 해당기관에서 도출할 수 있는 가치를 정의하고 이를 운영하기 위한 체계, 데이터, 그리고 기술개발이 필요하다. 재난관리의 경우, 유관기관들은 지난 수년간 각 기관에서 생성되는 데이터를 정형화하여 의사결정에 활용하여 왔다. 우리나라 국가재난관리 총괄기관인 소방방재청은 신속한 재난상황 파악 및 초동대처를 위해 각급 기관에서 보유·관리하고 있는 재난관리정보를 온·라인으로 연계·공유하는 재난정보공동활용시스템⁷⁾을 개발하여 활용하고 있다. 하지만 날로 복잡·다양해지는 재난상황파악 및 대응, 선제적 정책 수립 지원을 위해서는 관련 기관내 공공정보의 범위를 벗어나 사회, 환경, 문화, 경제 등까지 분야를 확대하고 민간정보까지 수평적 융합·분석기반이 마련되어야 한다(국립재난안전연구원, 2012). 대표적 민간정보인 소셜미디어는 신속·정확한 피해상황 및 대피정보의 공유매체임과 동시에 재난관리 전반에 국민적 소통과 참여를 가능하게 하는 새로운 매개체다. 따라서 재난관리를 위한 정형·비정형 데이터를 정의, 재난 빅데이터 운영체계 및 분석기술 개발 중·장기계획 수립 등 지속적인 연구개발에 투자한다면 가시적 효과를 앞당길 수 있을 것이다. 또한, 현재 각 부처에서 마련하고 있는 빅데이터 정책들과 재난관리 정책과의 유기적 협력으로 인력 및 기술 기반을 공고히 해야 할 것이다.

3.2 연구현황 및 향후계획

국립재난안전연구원은 앞장에서 설명한 전조감지 모델을 실용화하기 위해 『빅데이터 기반 재난 전조감지 및 대응기술 개발』이라는 주요 연구사업을 수행하고 있다. 재난 빅데이터 융합·분석을 통한 재난 전조감지 및 대응·전파기술 개발을 목표로 2013년부터 단계적으로 수행되는 이 사업은 우선, 재난분야 비정형 빅데이터⁸⁾ 기반 재난 전조감지 기술 개발을 목표로 비정형 데이터 분석기술, 모니터링 기술을 개발하고 재난 전조감지 시스템 개발을 1단계 목표로 설정하였다. 1단계

에서 개발된 비정형 데이터 분석기술과 그동안 재난관리에 활용되었던 정형데이터를 융합·분석하는 기술을 2단계에서 개발하고 재난 상황대응시스템 개발을 최종 산출물로 계획하고 있다. 단, 단계별 사업추진 방안은 국립재난안전연구원내 관련사업과 논의를 통해 병행추진을 검토하고 있어 계획했던 것 보다 조속한 연구성과가 기대된다.

2013년 연구내용은 크게 두 영역으로 나누어 수행된다. 비정형 빅데이터 즉, 텍스트 데이터 수집·분석 모델 개발과 소셜 빅데이터 재난관리 운영방안 및 이슈탐지 기법 연구로 나누어 진다.

텍스트를 분석해서 그 안에 내재된 가치를 추출하기 위해서는 언어처리, 마이닝 기술 등이 도입되어야 한다. 해당 기술은 인공지능분야의 한 계류로서 언어학과 전산학이 결합되어 만들어진 언어공학 분야로 기술개발이 꾸준히 지속되어 왔다. 세계적인 검색전문 기업인 구글은 문서수집·검색, 의미처리 등을 위해 언어처리기술을 활용하고 있으며, 우리나라 대표 포털인 네이버 역시 자체 언어처리 솔루션을 보유하여 엄청난 텍스트 데이터를 수집·분석하고 있다. 텍스트 분석의 성패는 언어처리 기초기술 개발보다는 활용하고자 하는 분야의 기초 언어지식이 얼마나 정교하고 커버리지가 넓은지에 달려 있다. 사람이 문서를 읽고 그 안에 의미와 가치를 순식간에 판단하는 것은 이미 머릿속에 학습된 언어지식을 활용하기 때문이다. 따라서, 재난분야의 텍스트 분석을 컴퓨터를 통해 자동화하기 위해서는 재난분야 언어지식 즉, 분류체계, 용어, 용어와 용어간 관계를 정의하고 디지털화해야 한다. 국립재난안전연구원에서는 재난분야 언어지식 모델을 개발하여 이를 기반으로 비정형 빅데이터를 자동 모니터링하는 모델 개발을 진행 중이다. 언어지식 모델구축은 다음과 같은 세부내용으로 구성된다.

우선, 재난 비정형 빅데이터 모니터링을 위한 분류체계를 정의한다. 재난분야 분류체계는 범국가적 표준이 없고 법적근거, 기관, 사용목적에 따라 다소 차이를 두고 정의하여 사용하고 있다. ‘국가과학기술위원회’ 재난유형 분류, ‘재난 및 안전관리 기본법’ 상 재난유형 분류 등을 검토하되, 실시간 비정

Fig. 10. Road-map of Research and Development.

7) 중앙정부 및 재난관리 책임기관 등 43개 기관의 223종(기상, 수위, 우량, 재난위험시설, 의료, 교통사고·통제, 구호 등) 정보 제공.

8) 본 사업의 비정형 데이터 범위는 텍스트 데이터로서 소셜미디어, 언론보도 자료로 한정하여 수행된다.

Fig. 11. The Plan of Disaster Management Using Social Bigdata.

형 빅데이터 모니터링을 목적으로 하는 분류체계를 정의하고자 한다. 즉, 트위터와 같은 소셜미디어 속 재난을 모니터링하기 위해서는 전통적이고 학문적으로 체계화된 분류체계보다는 소셜미디어 현상을 잘 반영하는 분류체계로 정의할 필요가 있다. 따라서 트위터의 실 데이터를 기반으로 분류체계를 정의하되 재난 및 안전관리 전 범위를 수용할 수 있는 유연성을 고려하여 설계할 계획이다.

다음으로 재난관련 용어 모델을 정의한다. 재난관련 용어는 물론 용어간의 관계 모델정의를 목표로 재난분류체계를 기준으로 관련용어를 수집하고 이들 간의 관계모델을 정의한다. 용어간의 관계는 향후 재난유형별 연관재난, 피해, 원인 등을 동시에 모니터링하는데 활용될 것이다. 재난분야 언어지식모델은 시범적용, 검토, 확장과정을 거쳐 향후 타 시스템 혹은 연구에서 공동 활용할 수 있도록 제공될 계획이다.

재난 전조감지를 위해 활용되는 소셜미디어는 사람이 인지할 수 없는 속도로 전 세계에서 실시간으로 발생하고 있어 소셜 빅데이터라 해도 과언이 아니다. 이런 소셜 빅데이터를 재난 관리에 효율적으로 활용하기 위해서는 데이터 생성자인 소셜 미디어 사용자들의 적극적 정보공유 및 참여유도, 신뢰있는 정보게시 등의 의식과 문화를 형성해야 한다.

현재 트윗을 재난관리에 활용하기 위해서 넘어야 할 한계도 많다. 우선 재난관련 트윗량과 신뢰도 문제이다. 평상시와 재난시 모두 재난관련 트윗량은 약 1%로 일반 트윗에 비해 현저히 낮다. 물론 하루 동안 발생하는 트윗량을 생각하면 적은 양은 아니지만 필터링 과정을 거친 후 의미있는 정보를 주는 트윗들이 필요하다. 무엇보다 재난관리에 중요한 정보로 활용될 위치정보를 공개하는 트윗은 평상시와 재난시 모두 1%미만으로 나타나 실효성에 의문이 제기되고 있다. 다만, 트윗 원문에 위치를 언급하는 키워드를 포함한 트윗이 재난시 약 35% 발생한다는 사실은 고무적이다. 이러한 한계를 해결하기 위해서는 ①재난관리를 목적으로 하는 트윗 운영방안 마련과 ②트윗 원문의 신뢰성 검증 및 내용기반 위치추출 기법 고도

화가 필요하다.

소셜미디어를 활용한 재난관리 사례를 보면 크게 두 가지 유형으로 분류할 수 있다. 첫째, 불특정 다수의 자발적 트윗을 활용하는 경우이다. 동일본 대지진, 허리케인 샌디 당시 트윗을 매핑한 사례가 이에 해당하는데, 이 경우는 정보생성의 즉시성과 소스가 다양해 트윗량은 많으나, 내용의 신뢰도는 낮고 재난관리 활용 가치를 분석하기가 매우 어렵다. 둘째, 커뮤니티 중심의 목적성 트윗을 활용하는 경우는 내용의 신뢰도는 높아 분석이 용이하지만 사전에 커뮤니티를 구성하고 별도의 공유플랫폼을 운영·관리해야 한다.

국민 스스로 신뢰도 높은 다양한 정보(위치, 상황, 위험도, 피드백 등)를 자발적으로 생성하도록 유도하고 소셜미디어가 재난관리에 유용한 수단으로 활용되기 위한 범정부적 운영체제는 두 유형의 장점을 취하여 마련되어야 할 것이다. 국립재난안전연구원에서는 이를 위해 소셜 빅데이터 속의 재난통계를 살펴보고 재난관리 소셜미디어 활용현황 및 문제점을 분석한 후 선순환적 재난관리 활용을 위한 소셜미디어 운영방안을 정책과 기술적인 부문으로 구분하여 제시할 계획이다.

또한, 소셜미디어에서 발생하는 재난이슈를 탐지하는 재난이슈탐지 기법 연구를 통해 재난 위험전조를 예측하여 피해를 최소화하고 발생한 사건을 최초 감지하여 신속한 대응이 가능하도록 활용할 계획이다. 재난 이슈는 크게 두 가지로 구분하여 탐지기법을 개발할 계획이다. 즉, 재난이 일어나지는 않았으나 일어날 가능성 및 전조를 알리는 전조이슈와 실제 재난이 발생했음을 알리는 사건발생 이슈로 구분하여 접근한다. 전조이슈는 평상시와 다른 패턴을 감시하고 정확성보다는 트렌드 파악을 중시하는 반면에 사건발생 이슈는 최초 사건을 알리는 시점을 감지하고 내용의 정확성과 진위를 파악하는 것이 중요하다. 따라서 성격이 다른 이슈를 분리하여 감지할 수 있도록 탐지 기법을 연구하고 있다. 이런 성질은 재난 유형과도 관련이 있다. 즉 호우, 폭염 등과 같은 재난은 강우량과 기온이라는 전조를 알리는 변수들이 존재하므로 전조파

악이 가능하지만 화재, 폭발 등은 불특정 기간에 발생하는 재난이므로 전조파악이 매우 어렵다. 이 연구에서는 재난유형을 전조파악이 가능한 정형과 전조파악이 불가한 비정형 유형으로 구분하여 차별화된 모니터링 체계와 탐지기법을 적용하여 제시할 것이다.

4. 결 론

소셜미디어는 신속한 전파능력과 함께 잘 정리된 정확한 정보, 즉 집단지성의 결정체라고 할 수 있으며 소셜미디어의 잠재능력은 재난분야 뿐만 아니라 다양한 산업분야에서 가치창출을 위해 각광받고 있다. 본 논문에서는 정부의 빅데이터 마스터 플랜의 활용가능성을 검증하기 위해 개발된 파일럿 모델인 도시 내수침수 전조감지 모델을 소개하였다. 모델은 소셜미디어와 같은 비정형 데이터를 모니터링하고 분석하는 전조감지 단계, 전조가 감지되면 침수와 관련된 정형데이터, 즉 강우, 침수예상도, 침수 흔적도 등과 비교분석하여 위험을 판단하는 단계를 거쳐 피해위험이 있다고 판단되는 경우 관련 기관, 지자체, 대국민 등에 전파하는 단계로 구성되어 있다. 모델은 2011년 발생한 침수사례를 통해 활용 가능성을 검토하였으며, 향후 국지적 폭우, 지역적 특수상황 등에 기인하는 예측되지 못한 침수 전조를 감지하고 예측된 침수 위험에 대한 신뢰도를 더욱 향상시킬 수 있는 새로운 예·경보체계로 활용 가능할 것이 기대된다.

시범지역 침수사례로 모델의 실효성을 검증하였지만 실용화를 위해서는 국민이 참여하는 재난관리의 새 모델인 만큼 세부기술 개발뿐만 아니라 소셜미디어 등의 새로운 소통채널을 재난관리에 선순환적으로 활용하기 위한 운영체계 수립 또한 확고히 해야 할 것이다. 이를 위해 국립재난안전연구원은 『빅데이터 기반 재난 전조감지 및 대응기술 개발』이라는 주요 연구사업을 수행하고 있다. 재난 빅데이터 융합·분석을 통한 재난 전조감지 및 대응·전파기술 개발을 목표로 2013년부터 단계적으로 수행되는 이 사업은 재난 전조 그리고 실 상

황에서 사후 상황까지 소셜미디어를 통해 국민과 소통하며 대응하고 문제를 해결함으로써 지금까지의 해결방식이 아닌 정부 3.0이 지향하는 개방, 공유, 소통, 협력을 통한 재난관리 3.0 실현에 이바지 할 것으로 기대한다.

References

- Choi, S.-H., Choi, W.-J., and Lee, J.-K. (2013) Monitoring and Sensing Disaster Issues from Social Media, *Proc. of the 40th Conference of the KIPS*, Vol. 20 no. 1, pp. 332-335.
- Choi, S.H. (2013) The Model for Sensing of Disaster Signs based on Big data, *Proceedings of the International Conference on Computer Applications and Information Processing Technology*, pp.169-172.
- Hurricane Sandy and the Geographies of Flooding on Twitter, floating sheep, <http://www.floatingsheep.org/2012/10/hurricane-sandy-and-geographies-of.html>.
- Jeannette, S., Leysia, P., and Irina S. (2008) Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires, *Proceedings of the 5th International ISCRAM Conference*, Washington DC., USA.
- Lee Y.J., Park Y.S., and Choi S.H. (2011) Developing an Intelligent System for the Analysis of Signs of Disaster, *Proceedings of TIEMS Annual Conference*.
- National Committee of Information Strategy (2011) The realization plan for Smart Government utilizing Big Data.
- National Disaster Management Institute (2012) Disaster & Safety, and Big Data, Disaster & Safety.
- The Urban Geographies of Hurricane Sandy in New York City, floating sheep, <http://www.floatingsheep.org/2012/10/the-urban-geographies-of-hurricane.html>.
- Track Japan's disaster map, <http://www.aljazeera.com/indepth/features/2011/03/201131474027904222.html>

Received July 18, 2013

Revised July 30, 2013

Accepted December 23, 2013