

실시간 기상 빅데이터를 활용한 홍수 재난안전 시스템 설계 및 구현

Design and Implementation of a Flood Disaster Safety System Using Realtime Weather Big Data

김연우*, 김병훈**, 고건식**, 최민웅**, 송희섭**, 김기훈*, 유승훈*, 임종태*, 복경수*, 유재수*
충북대학교 정보통신공학과*, 충북대학교 빅데이터협동과정**

Yeonwoo Kim(ywkim106@chungbuk.ac.kr)*, Byoungsoon Kim(bhkim@chungbuk.ac.kr)**,
Geonsik Ko(kgs@chungbuk.ac.kr)**, Minwoong Choi(mwchoi@chungbuk.ac.kr)**,
Heesub Song(heesub@chungbuk.ac.kr)**, Gihoon Kim(gihoonkim@chungbuk.ac.kr)*,
Seunghun Yoo(yshun2487@chungbuk.ac.kr)*, Jongtae Lim(jtlim@chungbuk.ac.kr)*,
Kyungsoo Bok(ksbok@chungbuk.ac.kr)*, Jaesoo Yoo(yjs@chungbuk.ac.kr)*

요약

최근 빅데이터 분석 기술을 통해 새로운 정보를 도출하기 위한 분석 기법들과 이를 활용한 다양한 서비스가 개발되고 있다. 그 중에서도 재난안전은 생활에 밀접한 서비스로 가장 중요하게 연구되고 있다. 본 논문에서는 실시간 기상 빅데이터 분석을 이용한 홍수 재난안전 시스템을 설계하고 구현한다. 제안하는 시스템은 실시간으로 수집되는 방대한 양의 정보를 검색하고 처리한다. 더불어 실시간 정보와 과거에 수집된 정보들을 결합하여 위험요인을 분석하고, 예측 정보를 사용자에게 제공한다. 또한, 제안하는 시스템은 사용자 메시지 및 뉴스와 같은 실시간 정보와 태풍·홍수 등으로 인한 하천 범람 등과 같은 재난 위험요인을 분석한 위험 예측 정보를 제공한다. 따라서 사용자는 제안하는 시스템을 통해 향후 발생 가능성이 있는 재난안전 사고 위험에 대비할 수 있다.

■ 중심어 : | 빅데이터 분석 | 재난안전 | 실시간 서비스 | 재난안전지도 서비스 | 기상 |

Abstract

Recently, analysis techniques to extract new meanings using big data analysis and various services using them have been developed. A disaster safety service among such services has been paid attention as the most important service. In this paper, we design and implement a flood disaster safety system using real time weather big data. The proposed system retrieves and processes vast amounts of information being collected in real time. In addition, it analyzes risk factors by aggregating the collected real time and past data and then provides users with prediction information. The proposed system also provides users with the risk prediction information by processing real time data such as user messages and news, and by analyzing disaster risk factors such a typhoon and a flood. As a result, users can prepare for potential disaster safety risks through the proposed system.

■ keyword : | Big Data Analysis | Disaster Safety | Realtime Service | Disaster Safety Map Service | Weather |

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학CT연구센터육성 지원사업(IITP-2016-H8501-16-1013), 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원(No. 2016R1A2B3007527), 정보통신·방송 연구개발 사업의 일환으로 수행하였음[B0101-15-0266, (담당-1세부) 실시간 대규모 영상 데이터 이해·예측을 위한 고성능 비주얼 디스커버리 플랫폼 개발]

* 본 논문은 한국콘텐츠학회 2016 춘계 종합학술대회 우수논문입니다.

접수일자 : 2016년 11월 07일

심사완료일 : 2016년 11월 21일

수정일자 : 2016년 11월 21일

교신저자 : 유재수, e-mail : yjs@chungbuk.ac.kr

I. 서론

국내에서 발생하는 자연재난은 주로 여름철에 집중되어 있다. 재난이라 함은 태풍·홍수·호우·폭풍·폭설·가뭄·지진·황사 등 자연현상으로 인하여 국민의 생명·신체 및 재산과 국가에 피해를 주는 것을 말한다[1]. 실제로 얼마 전 국내의 남부지역에서도 제 18호 태풍 ‘차바’의 영향으로 강풍과 폭우로 인한 농경지 유실, 주택가 침수, 해안가 선박의 전복 등 큰 피해가 발생했다. 또한 국내에서 유래가 드문 강도 높은 지진의 발생으로 사회적 불안감과 경제적 손실을 가져왔다. 재난안전에 대한 중요성은 항상 강조되어 왔지만, 최근의 재난 소식을 접하면서 재난안전에 대한 중요성은 더욱 부각되고 있다. 이에 국내에서는 정부 및 지자체의 공공 기관 주도하에 각종 재난에 대처하고 예방하기 위한 활동을 꾸준히 진행하고 있다. 최근에는 컴퓨팅 기술이 급속히 발전함에 따라 빅데이터를 활용하려는 시도가 활발히 진행되고 있다.

빅데이터는 인터넷의 발달과 더불어 스마트 기기의 급속한 보급, 소셜 네트워크 서비스(SNS) 사용자의 폭발적인 증가로 인하여 데이터가 기하급수적으로 증가하면서 만들어진 결과물이다[2]. 이러한 빅데이터는 데이터의 수집과 관리, 활용의 문제가 중요한 이슈로 대두 되고 있다. 또한 그 데이터를 가공하고 분석하는 방법에 따라 상황인식, 의사결정, 그리고 가까운 미래 예측을 가능하게 하는 영역까지 확대되고 있다.

재난관리를 위해서는 기존의 정부중심 재난 대응 및 전파에서 벗어나 정형 데이터와 비정형 데이터를 총칭하는 빅데이터의 통합분석을 수행하고, 실시간 모니터링을 통한 선제적 대응 및 피해 최소화를 위한 노력이 필요하다[3]. 더불어 소셜 미디어의 재난 관리에 대한 잠재력을 이용하여 지능적인 맞춤형 재난 관리를 실현해야 할 뿐만 아니라, 공공 및 민간영역에서 발생하는 재난관련 빅데이터가 재난안전 관리의 예방과 대비, 대응 및 복구에 큰 도움이 될 것이라는 것에 이견이 없다[4]. 이러한 이유로, 최근에는 빅데이터 분석 기술을 통해 새로운 정보를 도출하기 위한 분석 기법들과 이를 활용한 다양한 서비스들이 개발되고 있다.

재난안전에 있어서 엄청난 양의 정형 및 비정형 데이터들을 의미 있는 정보로 가공하는 빅데이터의 활용은 매우 중요하다. 이 정보를 이용하여 자연재해는 물론, 인적재난과 같은 대용량 정보를 빠른 시간 내에 분석하여 예측 가능한 재난으로부터 피해를 경감시키고 예방을 위한 방재 활동 등의 선제적 대응이 가능하다. 또한, 재난 발생 후에는 48시간에서 72시간에 해당하는 대규모 재난의 초기 단계에서 시민들의 대피와 공공의 안전이 가장 크게 중요하다[5]. 그러므로 빅데이터의 분석은 재난 발생 후의 재난 대응 방식에도 중요한 실시간 빅데이터 분석 정보를 제공함으로써 2차 피해로 이어질 수 있는 각종 상황에 대처하고 예방할 수 있도록 도움을 준다. 더불어 재난 복구 활동에 있어서도 유선 통신이 불가능한 상황에서 GPS를 이용한 위치정보 및 소셜 네트워크 서비스와 같은 정보전달체계는 실시간 분석을 통해 생존자의 안전 확보, 원조물자 공급, 자원봉사자 조직 등과 같은 재난복구 활동을 전방위적으로 지원할 수 있다[6]. 기존의 재난안전 관련 서비스는 일반적으로 공공 데이터를 활용한 정형 데이터 분석 서비스를 제공하고 있다. 실시간 데이터 분석을 통한 정보 제공은 아직 미흡한 상황이고, 사용자의 요구 사항을 충분히 반영하지 못하고 있다.

본 논문에서는 다양한 자연재난 중 기상과 관련된 실시간 빅데이터를 활용한 홍수 재난안전 시스템을 설계하고 구현한다. 제안하는 시스템은 정형화된 공공데이터를 활용하여 통계 및 분석 정보를 제공하며, 사용자 메시지 및 뉴스와 같은 실시간으로 수집되는 정보를 검색하고 분석하여 실시간 정보를 제공한다. 또한 실시간 정보와 과거의 이력 정보를 수집 결합하여 태풍·홍수 등으로 인한 하천 범람등과 같은 위험요인을 분석하고, 모바일 네트워크의 GPS 위치정보를 이용하여 위험 예측 정보를 사용자에게 제공한다. 제안하는 시스템 서비스를 통해 사용자는 자신이 위치한 지역에서 자주 발생하는 기상 관련 재난에 대해 사전 방재활동을 할 수 있고, 향후 발생 가능성이 있는 재난안전 사고 위험에 대비할 수 있도록 실시간 기상 빅데이터 분석을 활용한 다양한 재난안전 시스템 서비스를 제공한다.

본 논문의 구성은 다음과 같다. II장에서는 본 논문과

관련된 기존의 시스템을 분석하고, III장에서는 제안하는 실시간 기상 빅데이터를 활용한 홍수 재난안전 시스템의 주요 기능을 설명한다. IV장에서는 제안하는 시스템의 구현 환경과 다양한 서비스 구현 결과를 기술한다. 마지막 V장에서는 논문의 결론에 대해 기술한다.

II. 관련 연구

국민안전처는 기존의 기관별로 운영되었던 ‘재난안전알리미’ 앱을 기반으로 11개 기관 15개의 재난안전정보를 통합·연계하여 ‘안전디딤돌’이라는 모바일 앱을 통하여 다양한 재난안전정보 서비스를 제공하고 있다 [7]. 이 서비스는 재난문자 및 기상특보에 대해 지역별·기지국 기반의 문자 알림 서비스와 국민행동요령 조회, 시설물 정보 조회, 민방위 교육일정 안내 등의 국민 맞춤형 재난안전서비스를 제공한다. 다만, 다양한 재난안전정보를 제공하고 문자와 같은 알림 서비스를 제공하는 반면에 비재난 상황에서 사용자들에게 관심 있는 정보를 제공함으로써 일상에서의 접근성을 높이는 데는 다소 부족함이 있다.

행정자치부는 2014년 9월부터 전국 15개 시범 지역을 대상으로 인터넷과 모바일을 통해 ‘생활안전지도 서비스’를 제공하기 시작했다[8]. 이 서비스는 치안, 교통, 재난, 맞춤형 등과 같은 국민생활안전과 밀접한 빅데이터 정보를 각각 등급화하여 구분하고 해당지역의 위험 정도를 보여주는 서비스이다. 생활주변의 안전정보는 각 기관에서 분산·관리되고 있는 통계데이터를 활용하여 사고 발생 빈도 및 사고 발생 가능성에 따라 발생 빈도가 높은 지역은 붉은색, 낮은 지역은 흰색에 가까운 색으로 표시하여 직관성을 높였다. 치안·범죄사고 정보도 역시 통계데이터의 빅데이터 분석을 통해 위험도 구분을 색으로 표시 구분하는 지도정보를 제공하고 관련 관공서 및 주민의 자구노력을 유도하는데 활용한다. 이 서비스는 주로 빅데이터 통계분석을 통해 우리 지역의 재난발생 및 피해 관련 정보를 제공하고 있다. 이 정보를 활용하여 사용자는 재난 대응을 위한 사전 예방활동 등을 수행 할 수 있으나 실시간으로 변화

하고 예측이 어려운 현재의 재난상황에서 필요한 정보들을 구하기는 쉽지 않다.

국립재난안전연구원은 2013년 5월 기상정보, CCTV, 현장 센서 뿐만 아니라 트위터 등의 SNS 정보를 분석한 Big data, 스마트폰, 무인헬기 등의 하드웨어와 ICT 기술을 이용하여 현장영상, 위성영상, 시뮬레이션 등을 통합한 ‘스마트 재난상황실(SBB)’을 개발 운영하고 있다[9]. 스마트 재난상황실은 다양한 기관 및 정보를 취합·분석한 재난정보를 제공해 주고 있다. 그러나 이 서비스는 정부 및 지자체 재난상황실에서 모든 정보를 통합해 현장의 상황을 모니터링 하고 분석·대응하는 재난관리 시스템으로 인프라가 구축된 환경에서 각종 재난상황을 입체적으로 파악하는 시스템이다. 따라서 사용자들의 접근성과 이용의 보편성에 어려움이 있고 이와 유사한 모바일 서비스는 제공되지 않고 있다.

해외의 경우를 살펴보면, 미국은 2012년 3월 오바마 대통령이 ‘빅데이터 연구개발 이니셔티브(Big Data R&D Initiative)’를 발표함에 따라 주요 연방기관들이 빅데이터 관련 연구개발을 추진하였다. 이에 지질조사소는 지구시스템 과학 분야에 빅데이터 활용방안 연구를 추진하고 과거 100여 년 동안 발생한 지진들을 유형별, 크기별로 조사하여 그 피해정도를 분석하였다[10]. 이를 통해 지진으로 인한 재난상황을 미리 시뮬레이션해 봄으로써 마련된 메뉴얼을 이용하여 큰 사고에 미리 대처할 수 있도록 하고, 다양한 지진 형태 및 환경 분석을 통한 방재 대책을 마련하였다.

영국의 환경청은 영국의 지역별 홍수 발생 가능성을 예측하여 지도에 표시한 ‘실시간 홍수 경고 지도(Live Flood Warning map)’ 서비스를 제공하고 있다[11]. 이 서비스는 지역별 강우량, 지표면의 고도, 토양, 지질, 배수, 홍수 이력 등의 다양한 정보를 통해 위험지역을 도출하고, 홍수 발생 가능성을 경보에서 해제까지 4단계로 구분하여 웹서비스로 제공한다. 홍수 예측에 필요한 다양한 요인 간의 빅데이터 관계 분석을 통한 시스템을 구축하였다.

일본의 재난관리체계는 사회적 여파가 큰 대규모 자연재난을 계기로 변화해 왔다. 일본 정부의 빅데이터 활용 전략의 기원은 2011년 동일본 대지진으로 인한 지

진, 쓰나미, 원전사고의 3중 복합 대형재난을 겪은 후, 총무성의 'Active Japan ICT 전략'에서 비롯되었다고 볼 수 있다[12][13]. 이동 통신 기술, 인공 위성망 등을 통해 재난 현장에서 발생하는 다양한 센서 데이터를 분석·대응하는 인프라를 구축하여 재난대응 능력 강화를 통한 재난 피해 최소화를 목표로 한다. 특히 일본 기상청이 운영하는 국가지진네트워크 및 국가방재과학연구소(NIED)는 지진네트워크를 구축하고 수백, 수천 개의 관측소를 통해 지진에 대한 진도, 해수면 변동, 지각 변형 정보 등을 수집한다. 수집된 정보는 기상청의 지진 현상 관측시스템(Earthquake Phenomena Observation System, EPOS)에 의해 분석되고 유효한 정보들을 중앙 및 지방, 주요 매체 등에 통보한다. 통보는 2분 이내로 지진 정보가 전파 되고 관계 기관들은 피해 현황 파악 및 대응체계를 구축한다. 재난관리에 있어서 사전 예측과 예방을 위한 빅데이터 활용도 중요하지만 지진과 같은 예측이 어려운 자연재해에 대해서 재난 발생 후 피해 최소화를 위한 빅데이터 활용의 한 예를 볼 수 있다.

국내의 서비스는 주로 빅데이터 통계분석을 활용한 재난안전 정보를 제공하며, 국내뿐만 아니라 국외 서비스의 대부분이 국가 공공기관에서 갖춰진 시스템 인프라를 기반으로 정보를 수집하고 분석하여 서비스하고 있다. 재난재해 상황에 직접적으로 노출되어 있는 사용자는 정보에 대한 접근이 용이하고, 재난에 대비하여 사전 예방 및 재난상황 시 피해 최소화를 위한 대응이 필요하다. 본 논문에서 제안하는 시스템은 사용자들이 평소에 관심을 가질 수 있는 기상 정보를 제공하여 서비스 접근성을 높이고 모바일 기기를 통한 실시간 기상 빅데이터 분석 정보를 제공한다.

III. 재난안전 시스템 설계

1. 시스템 구조

재난안전과 관련된 많은 양의 정보들은 인명보호와 경제적 자산을 지키는 등 생활에 밀접한 정보로 가장 중요하게 연구되고 있다. 제안하는 시스템에서는 기상

과 관련된 정형데이터 및 비정형데이터들을 수집 저장하고 사용자에게 의미 있는 기상 정보와 통계정보를 제공하며, 빅데이터 분석을 통해 실시간 사용자메시지 정보를 더하여 다양한 기상 정보를 제공한다. 또한 과거의 기상 이력을 통계 분석하여 사용자가 위치한 지역의 재난이력정보와 현재의 기상상황에 따라 발생할 수 있는 위험 예측정보를 제공함으로써 사용자에게 재난피해로부터 대비할 수 있도록 한다.

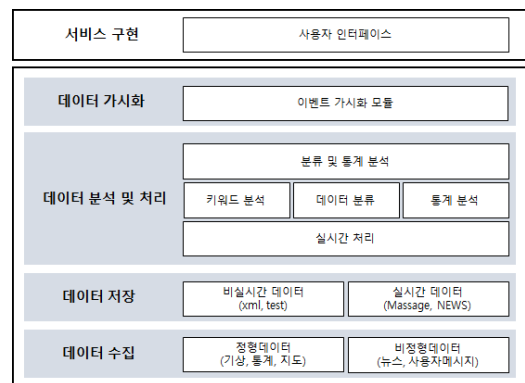


그림 1. 시스템 구조

[그림 1]은 제안하는 시스템구조를 나타낸다. 데이터 수집단계에서는 기상, 통계, 지도 데이터와 같은 정형데이터와 제안하는 시스템에서 사용되는 뉴스, 사용자메시지와 같은 비정형데이터를 수집 한다. OpenAPI를 통하여 일정 주기 마다 기상 상황, 수문 상황, 뉴스, 지도 정보 등을 수집하여 처리하고, 실시간 데이터 수집은 사용자 메시지와 클라이언트에서 서버로 전송되는 위치정보 등을 수집하여 처리한다. 데이터 저장 단계에서는 데이터 수집단계에서 수집된 데이터를 2가지 형태로 분류하여 데이터베이스에 저장 한다. 실시간 서비스를 위한 데이터는 In-Memory 데이터베이스인 Redis에 저장을 하고, 정제 및 가공이 필요한 비실시간 데이터는 관계형 데이터베이스인 Hbase에 저장하는 기능을 수행한다. 데이터 분석 및 처리 단계에서는 실시간 처리 모듈을 사용하여 데이터 저장단계에서 수집된 데이터를 동작에 맞춰 처리한다. 예를 들어 사용자 메시지 분석은 키워드 분류 모듈을 통해 사용자 메시지에서 서

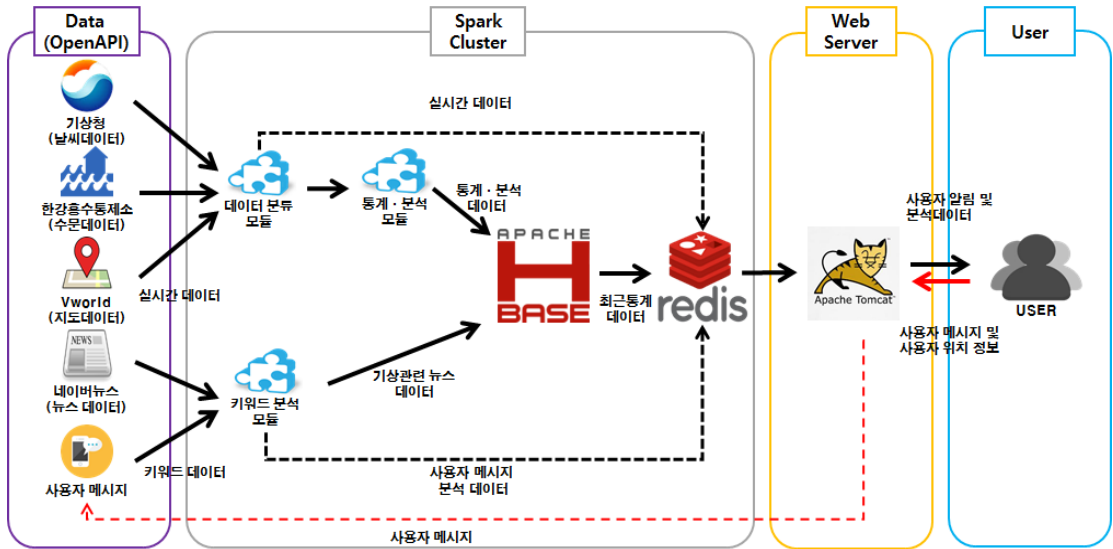


그림 2. 시스템 흐름도

비스 관련 키워드를 분류한다. 사용자 메시지 분류는 한국어 텍스트를 형태소 단위로 분석하여 품사와 함께 출력해주는 한나눔 형태소 분석기[14]를 통해 사용자 메시지를 형태소 단위로 분류한다. 분류된 형태소들은 사전에 구축해 둔 온톨로지[15]를 통해 서비스 관련 키워드를 추출한다. 객체와의 관계나 객체만의 의미를 표현하는 자료인 온톨로지는 해당 서비스와 관련된 단어들로 구축하였다. 데이터 분류 모듈과 통계 분석 모듈은 OpenAPI를 통하여 수집된 데이터를 사용한다. 기상 상황, 수문 상황 등 서비스에 필요한 데이터만을 추출하여 통계분석 모듈에서 통계 분석 기능을 수행한다. 데이터 가시화에서는 데이터 분석 및 처리, 데이터 저장에서 사용된 데이터를 클라이언트로 전송하고, Android, HTML을 활용하여 사용자에게 가시화한다. 제안하는 시스템은 서비스에 필요한 데이터를 받아오는 Data 수집 단계, 수집된 데이터를 서버에서 정제하여 다양한 데이터베이스에 저장하는 Spark 클러스터링, 사용자에게 제안하는 시스템의 서비스를 제공하기 위한 WebServer, 그리고 사용자에게 원하는 정보 등의 서비스를 위한 클라이언트 가시화로 구성되어 있다.

[그림 2]는 데이터 수집부터 클라이언트로의 가시화 서비스까지 제안하는 시스템의 구성과 전체적인 흐름

을 보여준다. 제안하는 시스템의 서비스에 대해 간략히 설명하면, 사용자는 클라이언트 디바이스를 통해 메시지를 전송하고, 전송된 사용자 메시지는 키워드 분석 모듈로 전송된다. 메시지를 전송 받은 키워드 분석 모듈은 해당 메시지를 분석하고, 분석된 정보에 따라 Hbase나 Redis에 전송이 된다. 여기에서 Redis는 전송 받은 분석 정보를 다시 WebServer를 통해 사용자가 요구한 정보를 제공한다.

2. Data 수집

Data 수집 단계에서는 다양한 OpenAPI와 제안하는 시스템 서비스에서 제공하는 사용자 메시지 등에서 필요한 정보들을 수집한다. 제안하는 시스템에서 사용된 OpenAPI 중에서 기상관련 데이터는 기상청 OpenAPI를 이용하여 데이터를 수집한다. 수집된 기상관련 데이터에서는 예보 기간과 행정구역을 시·공간적으로 세분화하여 3시간 단위로 기온, 최고기온, 최저기온, 상대습도, 풍향, 풍속, 강수량 등 기상예보 자료를 제공한다.

홍수와 관련된 서비스를 위한 데이터로는 한강 홍수통제소의 OpenAPI를 통해 전국의 수문의 현황을 매 시간 수집한다. 수집된 데이터에서는 각 수문의 위치와 현재의 수문 수위 상태, 증가량 등의 데이터를 실시간

제공한다. 재난안전 시스템 서비스를 위한 데이터는 국가공간정보와 다양한 콘텐츠를 제공하는 공간정보 오픈 플랫폼 지도서비스인 Vworld[16]의 Open API를 통해 전국의 지도데이터 및 홍수 등으로 인한 과거 범람 지역, 안전 대피소 위치 등을 수집한다. 이렇게 수집된 데이터들은 데이터 분류 모듈로 전송된다.

뉴스 데이터는 네이버 뉴스 OpenAPI를 이용하여 제안하는 시스템에서 제공하는 기상 및 재난 관련 뉴스 데이터를 실시간으로 수집한다. OpenAPI를 통해 수집되는 데이터들은 각 데이터가 저장된 외부 데이터베이스에서 제안하는 시스템에 필요한 데이터를 XML 형식으로 수집한다. 외부 데이터베이스에 필요한 데이터를 질의 요청하면 외부 데이터베이스에서는 요청받은 질의를 통해 질의에 맞는 데이터를 XML 형태로 전송한다. XML 형태로 전송받은 데이터는 XML 파서를 통해 파싱한다. 또한 사용자 인터페이스를 통해 수집되는 데이터는 요청이 있을 때 마다 데이터를 수집한다. 사용자가 사용자 인터페이스를 통해 데이터를 전송 시 실시간으로 데이터를 전송받으며 키워드 분석 모듈로 전송되어 수집된 데이터를 분석한다.

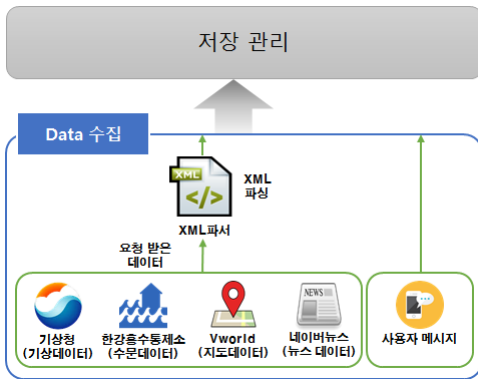


그림 3. Data 수집

3. 저장관리

저장관리 단계에서는 Data 수집 단계에서 수집된 데이터를 각각의 모듈을 통해 정제 작업을 진행한다. 데이터 분류 모듈은 수집된 데이터를 활용 방법에 따라 실시간 처리를 위한 기상 데이터 및 수문 데이터 등과

통계분석 처리를 위한 지도 데이터 및 텍스트 데이터 등으로 분류한다. 또한 실시간 처리 후 지속적으로 발생하는 데이터들을 추후 통계분석에 활용될 수 있도록 분류하여 저장한다. 즉 실시간 처리를 통해 클라이언트에 전송하여 사용자에게 보여주는 데이터는 In-Memory 데이터베이스인 Redis에 저장되고, 통계 및 분석을 위한 데이터는 RDBMS인 Hbase에 저장한다.

Redis에 저장되는 데이터는 사용자 인터페이스를 통해 사용자에게 보여주는 데이터인 일일 기상데이터, 주간 기상데이터, 통계데이터 등이 있으며, 이러한 데이터는 Redis의 특징에 맞추어 Key-Value형태로 저장된다. 이는 사용자 인터페이스에서 웹서버를 통해 사용자가 원하는 정보를 Key만으로 빠르고 용이하게 검색할 수 있는 장점이 있다.

Hbase에는 통계분석에 필요한 데이터를 저장한다. 과거 10년간의 전국 기상데이터를 대용량 빅데이터 저장에 효율적인 Hbase를 통해 저장하였다. RDBMS형태에 맞추어 Table로 데이터를 저장하였으며 각각의 테이블은 지역을 구분하여 지역별 Table을 작성하였다. 이 데이터는 통계분석 모듈을 통해 분석을 수행한 후 실시간 처리를 위한 데이터로 Redis에 저장되기도 한다. 이처럼 각각 데이터의 특성에 맞추어 데이터를 저장하고 요청 시 데이터를 검색한다.

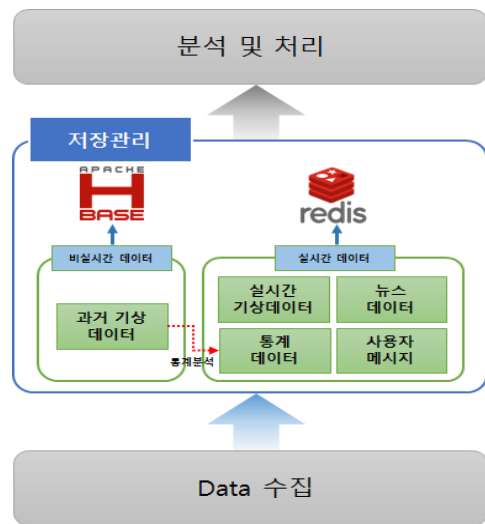


그림 4. 저장관리

4. 분석 및 처리

데이터 분석 및 처리는 키워드 분석 모듈, 통계분석 모듈로 나누어진다. 키워드 분석 모듈은 사용자 메시지 데이터와 뉴스데이터를 형태소 단위로 나누어 분석한다. 사용자 메시지가 입력되면, 해당 메시지를 제안하는 시스템에서 빅데이터 분석에 필요한 메시지인지 분류하기 위해 형태소 단위로 나누어 분석한다. 즉, 메시지가 입력되면 한나눔 형태소 분석기를 통해 형태소 단위로 메시지가 분할되고, 분할된 데이터는 Jena에서 SPARQL질의를 통해 사전에 구축된 온톨로지와 비교하여 같은 의미의 데이터 존재 여부를 확인 후 일치하는 데이터가 있으면 데이터 저장 모듈로 전송하여 데이터를 저장한다. 만일 일치하는 데이터가 없으면 메시지는 삭제된다. 뉴스 데이터 또한 사용자 메시지와 마찬가지로 형태소 분석기를 통해 형태소 단위로 분할되며 Jena와 온톨로지를 통해 적합한 뉴스 데이터인지 확인한다.

통계 분석 모듈에서는 사용자 인터페이스를 통해 보여주는 통계 데이터와 지도 서비스를 나타내기 위해 사용한다. 사전에 데이터 수집 모듈을 통해 수집된 기상 데이터와 데이터 저장 모듈에 저장된 지역별 과거 10년간의 기상 이력 데이터를 통계 분석하여 사용자들에게 보여준다. 기상 이력 데이터는 지역별, 기상별, 월별 통계분석을 통해 월 단위 평균값을 산출한다. 산출된 값은 Redis에 저장되어 사용자에게 과거 10년 동안의 기온, 강수, 풍속, 풍량 등의 속성으로 분류된 통계데이터를 그래프와 표로 가시화하여 제공한다.

재난안전지도 서비스에서는 사용자 위치에 따른 위험도를 등급으로 분류하여 표시하며, 통계 데이터를 활용한 하천 주변의 위험도, 대피소 위치정보 등을 제공한다. 사용자 인터페이스를 통해 실시간으로 사용자의 현재 위치데이터(GPS)를 수신하며 위치 데이터에 따른 주변 과거 범람 지역데이터, 실시간 수문별 수위 데이터, 기상데이터 등을 분석한다. 태풍·폭우 등으로 인한 기상 이변 시 각각의 수문별 수위에 따라 분석되는 해당 지역의 위험도를 판별하고 GPS 위치 정보와 함께 고려하여 사용자의 위험 여부를 색깔별 등급으로 구분하여 사용자에게 서비스 하는 기능을 수행한다. 위험

여부를 분석한 데이터는 데이터 저장 모듈에 의해 저장이 되지 않고 즉시 사용자 인터페이스를 통해 사용자에게 전달하여 위험상황에 신속히 대처할 수 있도록 정보를 제공한다.

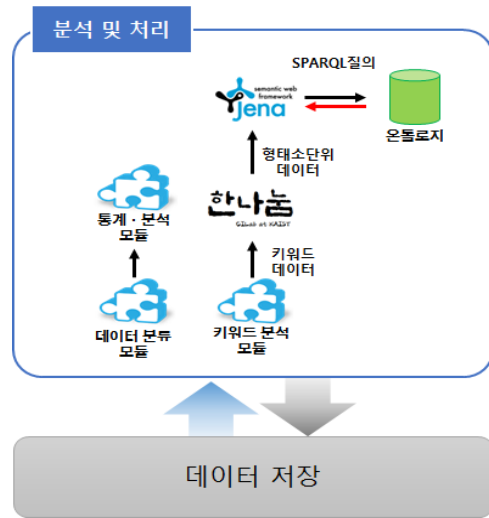


그림 5. 분석 및 처리

IV. 재난안전 시스템 구현

1. 구현 환경

본 논문에서 제안하는 시스템 구현 환경은 [표 2]와 같다. 분산 처리를 위해 3개의 노드를 구성하였고 WebServer 구축은 Apache Tomcat을 사용 하였으며, 데이터 저장을 위해 RDBMS인 Hbase와 In-Memory 데이터베이스인 Redis를 사용하였다. 또한 데이터 수집하고 정제하며 가공하기 위해 Spark 플랫폼을 통해 Spark Cluster를 만들었다. 데이터 수집은 제안하는 시스템의 클라이언트에서 실시간으로 수집되는 사용자 메시지와 기상청 API, Vworld API, 네이버 뉴스 API 등 다양한 OpenAPI를 사용하여 수집하였다.

표 1. 제안하는 시스템 구현 환경

구성	사용 요소
프로세서	Intel i3-2100, 3.10GHz, 2Core
메모리	8GB DDR3
노드 수	3
서버(웹, 온톨로지)	Tomcat, Jena
데이터베이스 (NoSQL, RDBMS)	HBase 1.1.3 Redis 2.8.2
분산처리 시스템	Spark 1.5.2
데이터 수집	기상청 API, Vworld API, 사용자 메시지, 한강통제소 API, 네이버 API

2. 서비스 기능

2.1 오늘의 날씨

제안하는 시스템의 현재 기상정보 서비스 기능은 일 상에서 주로 접하는 관심 있는 정보로서 사용자의 접근 성을 향상시킨다. 사용자가 현재 접속한 지역이나 사용 자가 설정한 관심 지역의 현재의 기상관련 실시간 정보 를 보여준다. 이 서비스에서는 현재 날씨 정보, 사용자 추천 날씨 정보, 그리고 시간별 날씨 정보를 제공한다.

현재 날씨 정보는 기상청 예보 API 에서 제공하는 기 상 상태와 강수 확률 정보를 이용하여 기상 상태를 시 각적으로 표현하고, 온도데이터를 이용하여 현재 기온 을 보여준다. 시간별 날씨 정보 서비스를 위한 기상청 중기 예보 API는 3시간 단위의 시간별 날씨 정보와 실 시간 온도를 제공한다. 우리는 이 데이터를 통해 향후 24시간 동안의 날씨 정보를 사용자에게 제공함으로써 일상생활에서 필요한 일일 기상정보를 시각적인 표현 으로 구현하였다.

사용자 추천 날씨 정보 서비스 기능은 사용자 메시 지를 통해 분석된 해당 지역의 기상 상태를 보여준다. 어 플리케이션 사용자가 가장 최근 작성한 메시지 내용을 실시간으로 확인할 수도 있으며, 사용자들이 작성한 메 시지 데이터를 분석하여 기상청 예보 API를 이용한 기 상 정보 서비스와 비교하여 사용자에게 의해 제공되는 실 시간 기상 정보 서비스를 제공한다. 사용자 메시지 화 면에서 사용자가 작성한 메시지를 키워드 분석 모듈을 통해 메시지에서 가장 많이 사용되는 단어를 분석하고 해당 지역의 실제 기상 상태를 보여준다. 또한 화면 하 단에는 사용자들이 가장 최근 작성한 메시지를 보여줌 으로써 실시간 기상에 대한 체감과 가독성을 높였다.

[그림 6]은 현재 사용자가 위치한 지역이나 사용자 관 심 지역의 이벤트성 날씨에 대한 실시간 상황을 보여주 는 현재 기상정보 페이지를 나타낸다.



처음 비 많이 와요

그림 6. 현재 기상정보

2.2 주간 날씨

주간 기상 서비스 기능은 현재날씨와 마찬가지로, 접 속한 지역 사용자가 설정한 관심 지역의 향후 일주일간 의 날씨정보를 보여준다. 기상청 중기 예보 API를 활용 하여 향후 일주일간의 날씨정보를 제공하고, 오늘을 기 준으로 이틀 전의 날씨정보를 함께 제공하여 날씨의 변 화에 따른 체감온도의 변화를 사용자 스스로 판단하게 하고, 온도변화에 따라 사용자 체질에 맞는 대비를 할 수 있도록 기상정보를 표로써 제공한다. 또한 최저 온 도와 최고 온도를 일자별로 도식화하여 그래프로 제공 하는 등 시인성 향상을 위한 다양한 가시화 서비스를 제공한다. [그림 7]은 주간 기상정보 서비스를 보여준다.

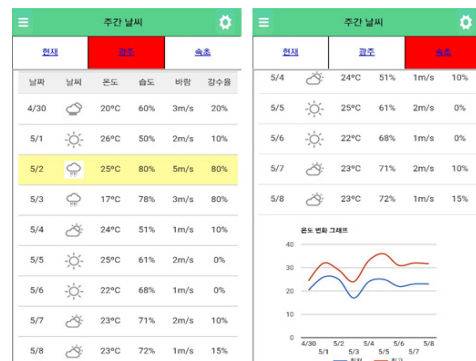


그림 7. 주간 기상정보

2.3 통계

사용자는 현재 자신이 위치한 지역 또는 사용자가 관심 있는 지역을 선택하여 그 지역의 통계 정보를 확인할 수 있다. 이 서비스를 통해 사용자는 일일 기상 또는 주간 기상 정보 서비스 기능에서 확인할 수 없는 특정한 날짜나 기간에 대해서 과거의 기상 이력을 통해 기상 변화를 예측하고 대비할 수 있다.

Hbase에 저장된 과거 10년의 데이터를 통계분석 모듈을 통해 지역별, 기상별 구분에 따라 분석하여 검색한 지역의 기온, 강수, 풍속과 같은 다양한 기상 정보를 최근 3년, 5년, 10년 단위로 확인할 수 있다. 예를 들어 사용자가 접속한 날짜가 2016년 3월이고, 검색한 지역이 서울시 강남구를 검색하여 결과를 조회 하면, 2016년 이전의 년도에 대해 3년, 5년, 10년 동안의 3월의 기온, 강수량, 풍속, 풍량 등을 확인할 수도 있다. [그림 8]은 다양한 기상 이력 정보들 중에서 강수 및 풍량에 대한 정보를 보여준다.

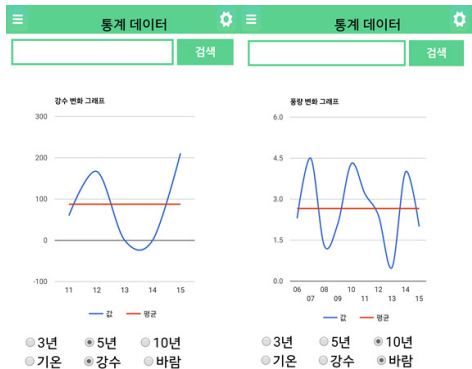


그림 8. 강수 및 풍량 통계 정보

2.4 사용자 메시지

사용자는 각 지역별 실시간 사용자 메시지를 최근에 입력된 메시지 순서로 확인하거나 직접 메시지를 작성할 수 있다. 이 서비스는 사용자의 직접적인 제보를 통해 이벤트성 날씨 혹은 기상예보와 다른 실제의 현재 날씨에 대한 정보를 공유하고, 공유된 메시지 데이터는 통계 분석하여 활용한다. 사용자의 관심 지역에서 입력된 메시지는 다른 사용자들의 메시지 데이터와 함께 1

시간 주기로 관심지역 내의 전체 메시지 데이터에서 가장 많이 언급된 기상관련 키워드를 추출한다. 이와 같이 분석된 정보는 오늘의 날씨 서비스에서 공공데이터를 활용한 실시간 기상 정보 서비스와 비교하여 [그림 6]에서와 같이 사용자 날씨 정보를 함께 보여준다. 제안하는 서비스를 통해 해당 지역의 기상 정보를 공공데이터만을 이용한 현재 날씨 정보 보다 정확하게 실시간으로 변화하는 기상 정보를 제공한다.

사용자는 사용자 메시지 화면 하단을 활용하여 메시지를 작성할 수 있다. 메시지 작성 기능은 사용자의 GPS 정보를 기반으로 현재 지역에 관한 날씨 정보를 작성하여 공유할 수 있는 기능이다. GPS 정보를 활용할 수 없는 상황이라면, 사용자가 직접 자신이 위치한 지역을 선택하여 사용할 수도 있다. [그림 9]는 사용자 메시지 서비스 기능을 보여준다.

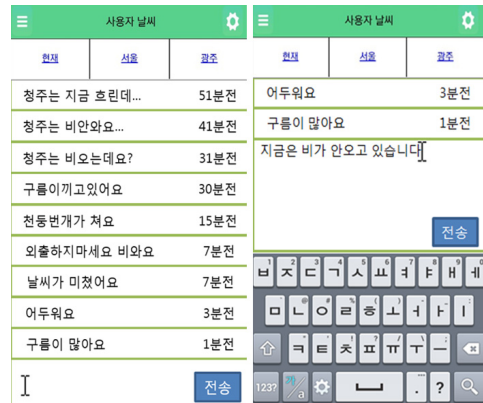


그림 9. 사용자 메시지

2.5 뉴스

사용자는 현재 위치한 지역 및 사용자가 설정한 관심 지역의 기상 또는 재난재해와 관련된 뉴스 정보를 볼 수 있다. 이를 통해 현재 위치한 지역이나 사용자 설정에 의한 관심 지역의 기상 및 재난 관련 기사를 확인하고 각종 상황에 대비할 수 있도록 도와준다. 각각의 기사를 선택하면 링크를 통해 해당 기사의 원본을 확인할 수 있다. 뉴스 데이터는 네이버 뉴스 검색 API를 통해 실시간으로 수집된 데이터이며 [그림 10]은 구현된 기상 및 재난관련 뉴스 페이지를 보여준다.



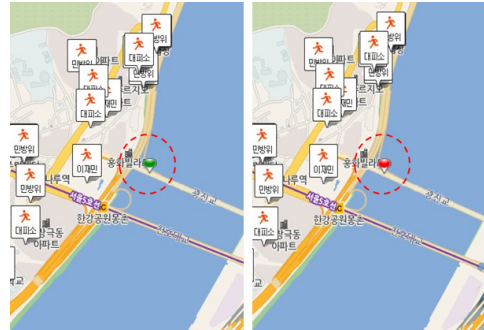
그림 10. 기상 및 재난 관련 뉴스

2.6 재난안전지도 서비스

재난안전지도 서비스는 GPS 센서를 통해 사용자의 현재 위치를 표시하고 주변의 과거 범람지역, 사용자 위치의 위험도, 대피소 위치를 보여준다. 과거 범람지역은 폭우 시 홍수 피해를 입었던 지역으로 통계분석을 활용하여 사용자가 현재 위치한 지역이 상습 침수 및 홍수 피해 지역이었음을 미리 파악할 수 있게 하며, 폭우 등과 기상 이변 발생 시 위험 지역으로부터 사전에 대피할 수 있도록 알림 서비스 등의 정보를 제공한다. 더불어 홍수 및 기타 재난상황에 대피할 수 있도록 사용자 위치 일대의 대피소를 확인할 수 있다.

서비스를 위한 지도는 공간정보 오픈플랫폼 지도서비스인 Vworld를 사용하여 개발하였다. 사용자 위험도 표시 기능은 사용자에게 현재 위치의 예측 위험성을 표시해 준다. 위험도는 사용자 위치를 표시하는 마커의 색상으로 표시되며, 색상에 따라 초록색은 안전, 노란색은 주의, 빨간색은 경계 상황임을 보여준다. 위험도의 계산은 지역 하천의 수문 데이터를 활용하여 하천의 범람 가능성과 각 지역별 피해 예상 정도를 고려한다. [그림 11]은 위험도에 따른 재난안전지도 서비스를 보여준다. 특히, [그림 11]의 (a)는 사용자의 현재 위치가 초록

색으로 안전함을 나타내고, (b)는 사용자의 현재 위치가 빨간색으로 위험도가 경계수준임을 나타낸다.



(a) 위험도 : 안전 (b) 위험도 : 경계
그림 11. 위험도에 따른 재난안전지도 서비스

V. 결론

본 논문에서는 실시간 기상 빅데이터를 활용한 재난 안전 시스템의 설계 및 구현을 제안하였다. 제안하는 시스템은 대용량의 빅데이터 처리를 위해 분산 저장소와 분산 처리 시스템을 기반으로 한다. 정형·비정형 형태로 수집된 데이터를 통해 실시간 기상 정보를 제공할 뿐만 아니라, 실시간 정보와 과거에 수집된 정보들을 결합하여 위험요인을 분석하고 예측 정보를 사용자에게 제공한다. 또한 뉴스 및 사용자 메시지를 활용하여 현재 위치한 지역의 주변 정보를 공유하고 향후 발생 가능성이 있는 재난안전 사고 위험에 대비할 수 있다.

제안하는 시스템은 태풍·폭우로 인한 홍수 피해와 같은 통계 정보와 기상 정보만을 고려한 빅데이터 분석 서비스를 제공하고 있다. 또한, 재난안전지도 서비스의 위험도 알림서비스를 위해서는 모바일 기기의 특성을 고려하여 사용자의 이동 속도, 위험지역과의 거리 등과 같은 요소들에 대한 분석이 요구된다. 향후 연구에서는 기상 관련 빅데이터 뿐만 아니라 더욱 다양한 자연 및 인적재난을 포괄할 수 있는 시스템으로 확장하고, 위험도 알림 재난안전지도 서비스는 다양한 요소들을 고려한 서비스로 개선하여 재난안전 사고 위험에 사전 대처 및 대응할 수 있도록 보완할 예정이다.

참고 문헌

- [1] 신현기, 박억중, 안성률, *경찰학사전*, 법문사, 2012
- [2] 오세연, 이재영, “IOT와 Big Data의 연계를 통한 범죄예방 활용방안,” *한국콘텐츠학회지*, 제13권, 제5호, pp.43-47, 2015.
- [3] 최선화, 김진영, 이종국, “재난관리의 새로운 해결 방안, 빅데이터,” *기상기술정책*, 제6권, 제2호, pp.77-87, 2013.
- [4] C. Pu and M. Kitsuregawa, “Big Data and Disaster Management: A Report from the JST/NSF Joint Workshop,” *Georgia Institute of Technology, CERCS*, pp.1-28, 2013.
- [5] 이동규, “빅데이터 기반의 재난관리 시스템 운용 방향에 대한 예비적 고찰 - 미국, 영국, 한국의 사례 비교를 중심으로,” *한국위기관리논집*, 제12권, 제1호, pp.17-32, 2016.
- [6] 이동규, “빅데이터를 활용한 재난관리 시스템 시론적 연구,” *한국행정학회 동계학술대회*, pp.1136-1155, 2014.
- [7] <http://www.mpss.go.kr/snskorea/snskoreainfo>
- [8] <http://www.safemap.go.kr/main/smap.do>
- [9] 임상규, “빅 데이터를 활용한 스마트 재난관리전략,” *한국위기관리논집*, 제10권, 제2호, pp.23-43, 2014.
- [10] 윤미영, “더 나은 미래를 위한 데이터 분석-Big Data 글로벌 선진사례 II,” *한국정보화진흥원*, 2013.
- [11] 배예나, “재난·안전 분야의 新ICT융합전략,” *한국정보화진흥원*, 2014.
- [12] 윤미영, 권정은, “빅데이터로 진화하는 세상. Big Data 글로벌 선진 사례,” *한국정보화진흥원*, 2012.
- [13] 이동훈, 김지윤, 강현숙, 이해림, “일본의 재난관리체계 및 재난심리지원체계 고찰과 시사점,” *한국콘텐츠학회논문지*, 제16권, 제7호, pp.73-90, 2016.
- [14] <http://semanticweb.kaist.ac.kr/hannanum>
- [15] <http://www.w3.org/standards/semanticweb/ontology>
- [16] <http://map.vworld.kr/map/maps.do>

저자 소개

김 연 우(Yeonwoo Kim) 정회원



- 2002년 2월 : 청주대학교 법학과 (법학사)
- 2012년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2012년 3월 ~ 현재 : 충북대학교 정보통신공학과 박사과정

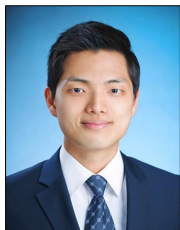
<관심분야> : 데이터베이스 시스템, 소셜 네트워크 서비스, 빅데이터 등

김 병 훈(Byounghoon Kim) 준회원



- 2015년 2월 : 청주대학교 컴퓨터정보공학과(공학사)
 - 2015년 3월 ~ 현재 : 충북대학교 빅데이터 협동과정(석사과정)
- <관심분야> : 빅데이터, RDF, 분산 질의 처리, 데이터베이스 시스템 등

고 건 식(Geonsik Ko) 준회원



- 2015년 2월 : 건국대학교 응용통계학과(경제학사)
 - 2015년 3월 ~ 현재 : 충북대학교 빅데이터 협동과정(석사과정)
- <관심분야> : 빅데이터, 데이터베이스 시스템, 소셜 네트워크 서비스, 추천 서비스

최 민 웅(Minwoong Choi) 준회원



- 2015년 2월 : 신라대학교 컴퓨터공학과, 경영학과(경영학 학사, 컴퓨터 공학사)
 - 2015년 3월 ~ 현재 : 충북대학교 빅데이터 협동과정(석사과정)
- <관심분야> : 데이터마이닝, 빅데이터, 사물인터넷, 경영전략 등

송 희 섭(Heesub Song)

준회원



- 2016년 2월 : 한림대학교 전자물리학과(이학사)
 - 2016년 3월 ~ 현재 : 충북대학교 빅데이터 협동과정(석사과정)
- <관심분야> : 소셜미디어, 소셜 네트워크, 빅데이터, 데이터베이스 시스템 등

김 기 훈(Gihoon Kim)

준회원



- 2016년 2월 : 충북대학교 정보통신공학과(공학사)
 - 2016년 3월 ~ 현재 : 충북대학교 정보통신공학과(석사과정)
- <관심분야> : 데이터베이스 시스템, 빅데이터, 고차원 인덱스 등

유 승 훈(Seunghun Yoo)

준회원



- 2016년 2월 : 충북대학교 정보통신공학과(공학사)
 - 2016년 3월 ~ 현재 : 충북대학교 정보통신공학과(석사과정)
- <관심분야> : 데이터베이스 시스템, 분산 컴퓨팅, 부하분산처리, 빅데이터 등

임 중 태(Jongtae Lim)

정회원



- 2009년 2월 : 충북대학교 정보통신공학과(공학사)
- 2011년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2015년 8월 : 충북대학교 정보통신공학과(공학박사)

<관심분야> : 데이터베이스 시스템, 시공간 데이터베이스, 위치기반 서비스, 모바일 P2P 네트워크, 빅데이터 등

복 경 수(Kyungsoo Bok)

중신회원



- 1998년 2월 : 충북대학교 수학과(이학사)
- 2000년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2005년 2월 : 충북대학교 정보통신공학과(공학박사)

- 2005년 3월 ~ 2008년 2월 : 한국과학기술원 전산학과 Postdoc
 - 2008년 3월 ~ 2011년 2월 : (주)가인정보기술 연구소
 - 2011년 3월 ~ 현재 : 충북대학교 정보통신공학과 초빙교수
- <관심분야> : 데이터베이스 시스템, 이동객체 데이터베이스, 소셜 네트워크, 빅데이터 등

유 재 수(Jaesoo Yoo)

중신회원



- 1989년 2월 : 전북대학교 컴퓨터공학과(공학사)
- 1991년 2월 : KAIST 전산학과(공학석사)
- 1995년 2월 : KAIST 전산학과(공학박사)

- 1995년 3월 ~ 1996년 8월 : 목포대학교 전산통계학과(전임강사)
 - 1996년 8월 ~ 현재 : 충북대학교 정보통신공학부 및 컴퓨터정보통신연구소 교수
 - 2009년 3월 ~ 2010년 2월 : 캘리포니아주립대학교 방문교수
- <관심분야> : 데이터베이스시스템, 빅데이터, 센서 네트워크 및 RFID, 소셜 네트워크 서비스, 분산 객체컴퓨팅, 바이오인포매틱스 등