



NoSQL 기반 교통사고 및 교통 환경 빅 데이터 활용기법에 대한 연구

A Study of the Method of Traffic Accident and Traffic environment's Big Data based on NoSQL

저자 (Authors)	박세한, 이종혁, 김응모 SeiHan Park, JongHyeok Lee, Ung-Mo Kim
출처 (Source)	한국정보과학회 학술발표논문집 , 2015.12, 1436-1438 (3 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/Article/NODE06602754
APA Style	박세한, 이종혁, 김응모 (2015). NoSQL 기반 교통사고 및 교통 환경 빅 데이터 활용기법에 대한 연구. 한국정보과학회 학술발표논문집, 1436-1438.
이용정보 (Accessed)	성균관대학교 자연과학캠퍼스 115.***.170.150 2018/03/05 15:35 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

NoSQL 기반 교통사고 및 교통 환경 빅 데이터 활용기법에 대한 연구

박세한, 이종혁, 김응모
성균관대학교 정보통신대학

e-mail : scarletface@nate.com, jhl0866@skku.edu, ukim@skku.edu

A Study of the Method of Traffic Accident and Traffic environment's Big Data based on NoSQL

SeiHan Park, JongHyeok Lee, Ung-Mo Kim
College of Information & Communication Engineering
Sungkyunkwan University

요 약

자동차가 발명되면서 교통이 매우 편리해졌지만 그와 동시에 자동차 혹은 유사 교통수단에 의해 많은 사건이 일어나고 있다. 이러한 교통과 관련된 사건 사고들은 꾸준히 발생되고 있기 때문에 많은 기관에서 예방과 대책마련에 주의를 기울이고 있다. 교통사고는 단순한 경제적 손실뿐만 아니라 인명 피해와 사회적 위기감을 불러일으킨다. 교통사고를 예방하기 위해 운전자 교육과 교통안전법 등의 대안이 있지만 그럼에도 불구하고 매 해 교통사고는 무수히 많이 일어난다. 교통사고의 발생빈도를 줄이기 위해서는 교통사고의 발생원인 및 운전자 주의뿐만 아니라 사고 당시의 환경과 도로 상황 등 특징을 분석하고 규명하는 작업이 필요하다고 할 수 있다. 본 논문에서는 교통사고가 임의적으로 나타나는 것이 아니라 이러한 특징들이 중첩되어 규모와 빈도수가 결정된다고 판단하고, 교통 사고의 빅 데이터를 이용하여 유사하고 반복적인 사고들을 예방할 수 있는 방법을 제시한다. 따라서 대용량의 교통과 관련된 자료들을 신속하고 정확하게 분석하고 NoSQL기반, Map-Reduce 프로그래밍 모델을 이용하여 위치, 시기 및 시간대, 규모, 날씨, 도로 현황, 운전자 특성 등의 관점에서 효율적으로 예방할 수 있도록 도로교통 사고를 최소화할 수 있는 솔루션 제시를 목적으로 한다.

1. 서 론

2003년에서부터 2013년까지의 10여 년간 도로교통여건에서 도로연장거리의 연평균 0.9%, 자동차는 2.9%, 운전면허 소지자는 2.7% 증가하였다. 이 통계는 연평균 인구증가율 0.5%를 능가한 자동차와 운전면허 소지자의 증가로 자동차운전면허의 대중화와 자동차문화시대를 도래시켰다. 하지만 이로 인해 국민 경제의 생산 현장인 도로의 기능과 효용의 감소를 가져왔고, 도로에서 각종 교통사고가 빈번하게 발생하였다.

도로 교통사고는 국가의 소중한 자원인 인명과 재화의 손실, 좌절, 슬픔 등의 정신적 손해와 더불어 도로시설의 파괴, 차량 지, 정체로 인한 도로효율성의 저하 등의 피해를 동반한다. 이러한 교통사고를 도로교통공단에서는 국민의 경제적인 측면에서 사회적 비용으로 추계하고 있다. 2013년 1,119,280건의 교통사고로 5,092명이 사망, 1,782,594명이 부상하였고 도로교통사고의 비용은 인적, 물적 피해비용, 사회기관비용을 포함하여 약 24조 444억 원으로 추계되었다. 이 비용은 2013년도 GDP의 약 1.7%, 국가예산의 약 10.2% 수준이다[1].

1일 평균 교통사고는 590건, 사망 13.9명, 부상자 900명에 달하여 2~3분에 한 번의 형태로 교통사고가 계속해서 일어난다고 할 수 있다. 또한 2015년 최근에도 인천 영종대교에서 100중 추돌 사고가 일어나 2명 사망 65명 중경상 등 교통사고로 인한 피해와 위험은 생활과 밀접

하다고 생각된다. 최근 10년간 교통사고 사망자는 점차 줄어들고 있지만 아직까지도 많은 교통사고가 발생한다. 이러한 교통사고들은 통계를 이용했을 때, 휴가철 혹은 연말에 많이 발생하고, 지방별로도 인구수에 비례하지 않게 빈도수가 발생, 요일이나 시간으로 구분지어도 매번 다르다고 나타난다[2].

때문에 오늘날에도 교통사고로 인한 피해를 줄이기 위해 정부와 서울시에서도 “빅 데이터를 활용 교통사고 사망 20% 줄인다.”의 발표를 하고 민원, SNS정보를 분석하여 위험지역을 분석하는 등 많은 노력을 하고 있다 [3]. 따라서 교통사고와 관련된 기존 데이터를 신속하고 정확하게 분석하여 잠재적인 교통사고의 원인을 규명하고 제거하는 것이 중요해지고 있다.

본 논문에서는 교통사고통계 등의 공공 데이터와 NoSQL 데이터베이스를 이용하여 지역, 날씨, 상황, 월과 요일의 날씨, 시간대, 사고유형, 운전자의 연령대, 법규위반, 음주운전 등의 가능한 많은 자료 등을 토대로 도로 교통 사고를 데이터화하여 Map-Reduce 프로그래밍 모델을 사용, 교통사고의 데이터를 분석하고 처리하는 방법을 제시할 것이다.

2. 관련연구

2.1 NoSQL

NoSQL Database는 디자인의 단순화, 수평적 확장성,

세세한 통계를 포함하여 관계형 데이터베이스보다 제한이 적고 일관성 모델을 이용하는 데이터의 저장 및 검색을 위한 구조를 제공한다. NoSQL Database는 단순 검색 및 추가 작업에 매우 최적화된 Key Value 저장 공간으로 성능 이익을 내는 것이 목적이다. NoSQL Database는 빅데이터와 실시간 웹 애플리케이션의 상업적 이용에 널리 쓰이고, SQL계열 Query언어를 사용할 수 있다는 특징을 가지고 있기 때문에 “Not only SQL”로 불리기도 한다[4].

2.2 Map-Reduce

Map-Reduce는 Map과 Reduce가 합쳐진 용어로, Map 함수와 Reduce라는 함수의 조합을 통해서 분산/병렬 시스템운용을 지원하는 프레임워크 기술이다. Map-Reduce는 데이터를 {Key, Value}의 쌍으로 만들고 이를 처리하는 프로세스를 가진다. Map 함수는 사용자 정의 자료구조이며, 입력데이터로부터 Key/Value 쌍으로 이루어진 중간 데이터 형태의 데이터를 만들어낸다. Map-Reduce의 라이브러리에서는 Value에 대해서 Key를 함께 관리하고, 이로부터 나오는 중간데이터를 Reduce함수로 넘겨주어 처리한다. Reduce함수 역시 사용자에게 의해서 작성되며, Map함수로부터 나온 Key와 Key에 대한 Value를 받아들이는 후에 각 중간데이터에서 Key값을 이용해 Value값을 합쳐 대규모 데이터를 처리하게 된다.

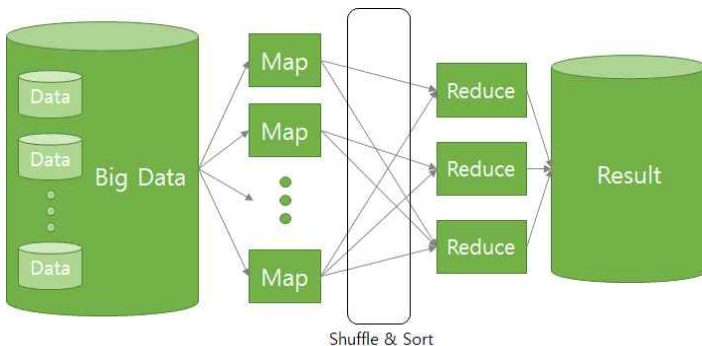


그림 1 Map-Reduce의 알고리즘

Shuffle-Sort는 Map과 Reduce 중간의 과정이다. Shuffle-Sort는 메모리에 저장되어 있는 출력 데이터를 분할과 정렬을 통해 로컬 디스크에 저장한 후, 네트워크를 통해 Reducer의 입력 데이터로 전달하는 과정을 의미한다.

2.3 중요 교통지표의 추이

도로교통공단(KoROAD)에서 조사한 2010년부터 2013년까지의 교통지표의 추이 자료이다. 교통사고에 대한 세부적인 데이터 이전의 전체적인 연도 별 통계자료다[1].

표 1 중요 교통지표의 추이

	인구 (천명)	운전면허 소지자 (천명)	자동차 (천대)	교통경찰 (명)
2010	49,410	26,402	17,941	9,715
2011	49,779	27,251	18,437	9,916
2012	50,004	28,263	18,871	9,615
2013	50,220	28,848	19,401	9,377
10년 평균 증감률(%)	0.5	2.7	2.9	-1.0

표 2 교통사고 통계 - 통합DB

	교통사고 발생 (통합DB)	교통사고 사망자 (통합DB)	교통사고 부상자 (통합DB)
2010	979,307	5,505	1,533,609
2011	897,271	5,229	1,434,786
2012	1,133,145	5,392	1,777,604
2013	1,119,280	5,092	1,782,594
10년 평균 증감률(%)	2.4	-3.1	2.8

3. 제안 시스템 설계 및 구현

3.1 교통사고 발생 시 피드백의 구조

3.1.1 현재 교통사고 발생 시 예방 피드백 예상 구조

현재 운전 시에 교통사고 위험 지역 혹은 이전 사고 발생 지역 등을 내비게이션이나 도로의 전광판을 통해서 운전자에게 정보를 전달한다. 운전자는 이 때 경각심을 가지고 운전 집중하게 되는 경우가 많아지고 통계적으로도 이러한 노력으로 교통사고가 점차 감소하고 있는 것을 확인할 수 있다.

하지만 수집되는 데이터에 비해 제공되는 정보는 단순 사고 발생 지역에 대한 경고와 운전자 사전 교육, 음주, 졸음운전 금지 캠페인 정도로 활용되고 있다.

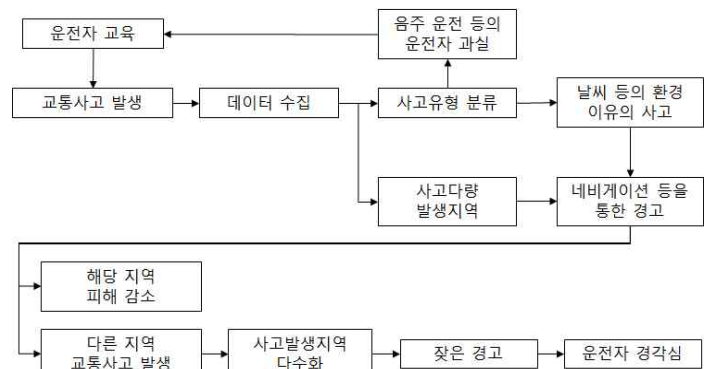


그림 2 현재 교통사고 발생 시 예방 피드백 예상도

3.1.2 빅 데이터 활용 후 예방 피드백 예상 구조

경찰DB에 들어가는 사건은 ‘교통사고보고 현황조사서’에 따라 많은 데이터를 얻을 수 있지만 일반적인 사고의 경우는 그렇지 않기 때문에 데이터 수집에서도 분류가 되었다. 본 논문에서 사용하게 되는 Map-Reduce와

