# Global Report

# 재난 극복력 있는 스마트 도시를 목표로:사물 인터넷과 빅데이터 분석이 게임체인저가 될 수 있는가?



국토교통과학기술진흥원

작성자 | 이현재

## 1. 서언

재난 (자연적 또는 인위적)의 강도는 지난 수십 년 동안 증가했다. 2018년 세계 재난 보고서(IFRC)는 지난 10년 동안 전 세계적으로 발생한 지진, 홍수, 쓰나미 등 3,751건의 자연재해가 총 1조 6,580억 달러의 피해를 입히고 20억 명이 넘는 사람들에게 영향을 미쳤다고 확인하였다. 재해 관리는 모든 재해 단계 (예 : 완화, 구조, 대응 및 복구)에서 활동의 계획 및 관리를 통합하는 일련의 조직화 된 프로세스로 간주할 수 있다. 재난 관리의 주요 목표는 재난을 분석, 모니터링 및 예측하는 효율적인 수단을 제공할 수 있는 상호 관련된 프로세스를 통합하는 것이다. 재난 관리의 핵심 기능은 조기 경보를 발동하고, 실시간으로 정보를 수집하고, 피해를 정확하게 추정하고, 대피 경로를 신속하게 파악하고, 비상 자원을 효과적으로 관리하는 것이다. 현대의 재난 관리 시스템은 이기종 소스(heterogeneous source)에서 생성된 다양한 유형의 데이터를 지원해야 하므로, 귀중한 정보를 추출하기 위해 효과적인 데이터 통합 및 다중 모드 데이터 분석 방법을 배치해야 한다. 상황 인식의 범위를 넓히고 효과적인 의사 결정을 위한 새로운 통찰력을 얻으려면 다양한 잠재적 데이터 소스 (예 : 센서, 소셜 미디어, 위성, 스마트 폰, 권위 있는 / 역사적 데이터 저장소 등)에서 관련 정보를 수집해야 한다. 다행히도 사물인터넷(IoT), 빅데이터 분석 (BDA), 클라우드 컴퓨팅, 포그 컴퓨팅 등과 같은 새로운 기술의 출현으로 재난 관리 프로세스 자동화는 시기적절하고 정확한 서비스를 제공하고 있다.

도시 및 인근 지역에 거주하는 세계 인구는 2050년까지 약 68%까지 증가할 것으로 예상된다. 도시들의 인구 밀도의 증가는 더 나은 서비스와 거주민들을 위한 적절한 기반시설을 필요로 한다. 스마트도시 인센티브는 인구 증가, 환경 오염, 에너지원 부족 등으로 인해 발생하는 문제에 답하기 위한 이상적인 해결책으로 학계와 산업계 전문가들에 의해 고려되고 있다. 다양한 전자기기와 네트워크 인프라가 한데 어우러져 고품질의 양방향 협업 멀티미디어 서비스를 제공하는 스마트도시 개념이 인기를 끌고 있다. 따라서, 재난의 조기 경고, 모니터링 및 예측 기능을 갖춘 스마트도시는 관련 당국이 재난 시나리오를 지능적으로 관리하는 데 필요한 정보와 통찰력을 생성함으로써 사망자를 최소화하는데 있어 게임 체인저가 될 수 있다. 모든 스마트도시의 중요한 구성요소는 사물인터넷(IoT)이다. 사물인터넷은 기기들이 인터넷을 통해 서로 통신할 수 있게 해주는 인프라이다. 이러한 시스템에서 다양한 장치 간의 상호 통신으로 인해 빅데이터로 알려진 상당한 양의 데이터가 생성된다. 빅데이터 애플리케이션의 경우 가장 중요한 작업은 다양한 유형의 데이터 (예: 구조화, 반 구조화 및 비구조화)를 보유할 수 있는 막대한 크기의 데이터 세트에서 숨겨진 값을 빠르게 발견하는 것이다. BDA(빅데이터 분석)는 여러 소스의 대규모 데이터 세트를 검사하여 조직이 정보에 입각한 결정을 내리는 데 도움이 되는 귀중한 정보와 통찰력을 추출한다.

재해 관리 자동화를 위한 IoT 및 BDA 기술을 구현하는 아키텍처 모델을 개발하고 이와 관련된 잠재적인 설계 과제를 해결하는 것은 최근 문헌에서 간과되고 있는 측면이 있다. 여러 소스에서 대량으로 분산된 데이터를 처리할 때 직면한 주요 문제는 다중 소스 이기종 데이터의 데이터 집계, 통합 및 처리이다. 전통적인 재난 관리 시스템의 데이터 관리 문제를 해결하려면 다중 소스 데이터의 통합을 지원하고, 효과적인 통신 및 빠른 액세스를 제공하고, 업데이트되고 적합한 데이터를 제공하고, 정보의 표준화를 지원하는 시스템 아키텍처를 개발하여야 한다. 그림1은 BDA 및 IoT 기반 DRSC(Disaster Resilient Smart City) 개념의 전체 시나리오를 보여준다.

## 2. 동기

재해 관리 커뮤니티와 관련 당국으로부터 가능한 모든 데이터 소스를 사용하여 재해 관리 프로세스를 위한 최신의 정확한 정보를 제공해야 한다는 요구가 증가하고 있다. 전통적인 재난 관리 시스템이 여러 소스에서 실시간으로 대량의 데이터를 수집, 통합 및 처리하는 것은 어려운 일이다. 재난의 초기 단계에서 책임 당국은 정확하고 신속하게 결정할 필요가 있다. 이러한 결정은 여러 차원을 포괄하는 다양한 출처의 품질이 좋은 정보가 제공되는 경우에만 성공적으로 구현될 수 있다. 재해 관리를 위한 잠재적

데이터 소스 중 하나로 IoT 기반 센서가 포함된다. IoT 기반 센서는 재난 발생 시 필요한 정보(온도, 방사능, 유독 가스 등)를 수집하는 데 도움이 되는 다차원 데이터를 제공한다. IoT 기반 플랫폼은 시간이 중요하고 확장 및 상호 운용이 가능한 서비스를 제공하는 조기 경보 시스템과 같은 재난 관리 시스템을 제공할 수 있다. 재난 관리를 위해 새롭게 떠오르고 있는 또 다른 빅데이터 소스는 소셜 미디어이다. 스마트도시는 시민과의 커뮤니케이션을 강화하고 피드백을 얻고 장려하기 위해 소셜 미디어를 고려해야 한다.

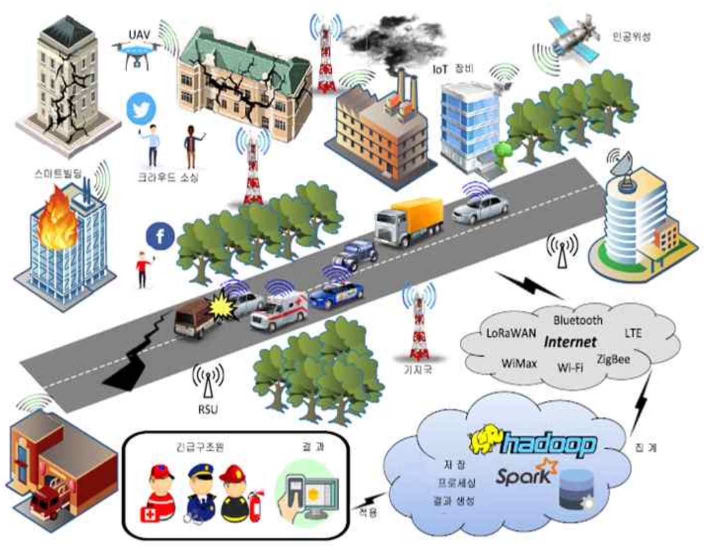


그림 1. BDA와 IOT 기반의 재난 극복형 스마트 도시의 예시 시나리오 (출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

스마트도시가 "재난 극복력을 갖춘 스마트 도시(DRSC)"가 되려면 방대한 양의 데이터를 효과적으로 집계 및 저장하고, 이기종 데이터 세트를 통합하고, 실시간 및 오프라인 분석을 수행하여 필요한 정보를 추출해야 한다. 이러한 과제는 BDA 및 IoT 발전의 첨단 기술을 의미하며, 문제의 긴급성을 총체적으로 처리할 수 있다. 따라서 본 연구는 DRSC용으로 설계 및 평가되고 최첨단 빅데이터 분석 툴에서 지원하는 다양한 데이터 소스와 함께 작동할 수 있는 아키텍처 솔루션을 제시하고자 한다. 우리의 노력의 동기는 보다 즉각적이고 통찰력있는 결과를 위해 이기종 데이터 소스와 실시간 처리를 고려하는 혁신적이고 효과적인 BDA 및 IoT 기반 DRSC 아키텍처를 제공하는 것이다.

## 3. BDA 및 IOT 기반의 재난 극복력 있는 스마트 도시

### A. BDA 및 IOT 기반 DRSC에 대해 제안된 참조 아키텍처

본 연구에서 DRSC에 대해 제안된 IoT 및 BDA 아키텍처는 i) DRSC를 위한 IoT 및 빅데이터 엔티티 간의 관계를 정의하는 데이터 추상화를 위한 참조 모델 및; ii) 전체 DRSC시스템 엔티티를 조립하기 위해 표준화된 프레임 워크로 간주될 수 있다. 그림 2와 같이 제안된 아키텍처는 5개의 계층, 즉 1) 데이터 자원 2) 데이터 전송; 3) 데이터 집계; 4) 데이터 분석 및 관리, 5) 적용 및 지원 서비스 계층으로 분할된다.

1) 데이터 자원 계층

이 계층은 모든 잠재적 데이터 소스의 인식과 생성된 데이터 수집을 처리한다. 여기에는 현장 센서, RFID 태그 감지, GPS, 감시 카메라, 스마트 폰, 위성 원격 감지 등과 같은 DRSC에 대한 모든 잠재적 IoT 기반 데이터 소스가 포함되어 있다. DRSC는 정부 또는 기타 재해 관리 조직이 보유하고 있는 소셜 미디어 스트림 및 권위 있는/역사적 데이터와 같은 다양한 데이터 리소스에서 대규모로 활용될 수 있다.

2) 데이터 전송 계층

데이터 전송 계층은 데이터 소스를 데이터 집계 계층에 연결하고 안전하고 효율적인 방식으로 주변 환경 전반에 통신 채널을 제공한다. "절대 네트워크 환경"을 보장하기 위해 유선, 무선 또는 위성 네트워크의 형태로 이용 가능한 모든 옵션을 고려하여 재난 정보 네트워크를 구축할 것을 권고한다. 전송은 LAN(Local Area Network), WAN(Wide Area Network) 및 PAN(Personal Area Network)으로 분류된 유무선 매체에서 수행할 수 있다. 전송 계층은 ZigBee, Bluetooth, Wi-Fi, 이더넷, WiMax, NFC 및 RFID와 같은 접속 전송 통신 기술과 CDMA, GPRS, 3G/LTE, 5G와 같은 네트워크 전송 통신 기술의 결합으로 지원된다.

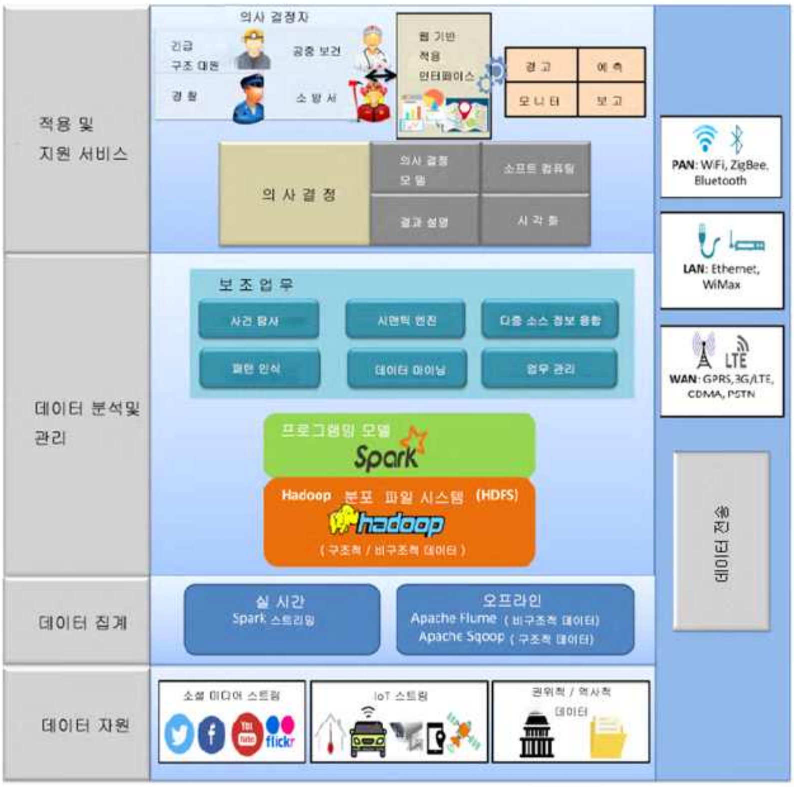


그림 2. **BDA와 IOT 기반의 재난 극복형 스마트 도시를 위해 제안된 참조 아키텍처** (출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

3) 데이터 집계 계층

데이터 집계자는 서로 다른 전송 매체를 통해 하나의 다중 소스 데이터베이스에 있는 모든 데이터를 수집할 책임이 있다. Apache Flume 및 Apache Sqoop을 각각 사용하여 정형 및 비정형 데이터의 형태로 데이터를 별도로 수집할 수 있다. 또한, 스파크 스트리밍은 Twitter 및 IoT 기반 데이터 스트림과 같은 소스의 실시간 데이터 집계에 이상적이다. 이 단계에서는 데이터 소스, 콘텐츠, 타임스탬프, 위치 등과 같은 필수 메타 데이터 정보도 식별할 수 있다.

4) 데이터 분석 및 관리 계층

데이터 분석 및 관리를 위한 기본 계층에는 데이터를 집계, 저장, 처리, 쿼리 및 분석하기 위한 다양한 도구 세트가 포함되어 있다. 다양한 BDA 프레임워크(예: Hadoop Ecosystem 및 Spark)의 조합을 검토하여 재해 관리 프로세스를 위한 효율적인 실시간 시스템을 개발할 수 있다. HDFS(Hadoop Distributed File System)는 대용량 데이터를 처리하기 위해 더 높은 효율성으로 상용 하드웨어에서 작동하도록 설계된 분산 스토리지 파일 시스템이다. Spark는 대화형 쿼리에 이상적이며 실시간 데이터 스트림 처리를 지원한다. 이벤트 감지는 재난 관리에서 매우 중요하며 발생하는 재난 이벤트를 식별하기 위해 작동해야 한다. 패턴 인식 메커니즘은 재해 관리에 중요한 텍스트 또는 공간 데이터 세트에서 유용한 정보 패턴을 감지하는 기계 학습 기능을 제공한다. 시맨틱 엔진은 비정형 정보의 분류, 검색 및 추출과 같은 효과적인 정보 관리에 활용될 수 있다. 다양한 데이터 마이닝 기술을 활용하여 사용 가능한 정보에서 새롭고 효과적이며 숨겨진 통찰력 패턴을 발견할 수 있다. 다중 소스 정보 융합 기술은 이기종 데이터 소스에서 필요한 데이터를 통합하는 데 도움이 된다. 작업 관리는 시스템의 다른 엔티티에 대한 작업 부하를 식별하고 시스템 운영을 효과적으로 관리하는 데 도움이 된다.

5) 적용 및 지원 서비스

대화형 웹 기반 애플리케이션 인터페이스는 의사 결정자에게 필요한 결과를 제공할 수 있다. 웹 기반 API를 통해 액세스한 다양한 시각화 도구를 사용하여 결과를 조회하고 표시할 수 있다. 웹 기반 사용자 인터페이스를 제공하고 소프트웨어 솔루션을 의사 결정자를 위한 결과 생성 및 시각화에 활용할 수 있다. 또한, 의사 결정 프로세스는 특정 애플리케이션의 요건에 따라 의사 결정 모델, 소프트 컴퓨팅, 결과 해석 및 시각화 기술과 같은 다양한 서비스와 통합될 수 있다.

### B. 구현 모델

DRSC 범위 내에서 경고 생성에 중점을 두고 배치한 시스템에서 수행하는 모든 운영 단계의 세부 사항을 요약한 구현 모델이 그림 3에 나와 있다. 제안된 구현 모델은 4개의 계층, 즉 1) 데이터 수집; 2) 데이터 병합; 3) 데이터 사전처리; 4) 빅데이터 분석 및 서비스 플랫폼으로 구분된다.

1) 데이터 수집

처음에는 재난 관리의 맥락에서 중요한 정보를 제공하는 여러 잠재적 데이터 소스(예 : 날씨 센서, 스마

트 홈 생성 데이터, 차량 교통, 소셜 미디어 스트림)가 식별된다. 데이터는 주로 각 데이터 소스의 로컬 데이터 집계기를 통해 수집된다. 로컬 데이터 수집기는 아날로그 데이터를 기계가 읽을 수 있는 디지털 형식으로 변환한다. 데이터 수집 프로세스는 실제 상황을 측정하는 센서 환경에서 데이터를 수집하는 로컬 집계기에서 데이터를 전송한다.

2) 데이터 병합

필요한 값을 효과적으로 추출하기 위해 수집된 데이터를 분류하기 위해 데이터 병합 프로세스가 수행된다. 데이터 병합 프로세스는 사용 가능한 데이터 세트에서 필요한 데이터 값에 대한 접근성을 보장하고 추가 분석을 위해 이를 조합한다. 우리가 제안한 모델은 다양한 데이터 소스에 개방되어 있다. 서로 다른 데이터 소스 모음은 시스템에 필요한 데이터를 제공하는 데이터 리소스로 간주된다.

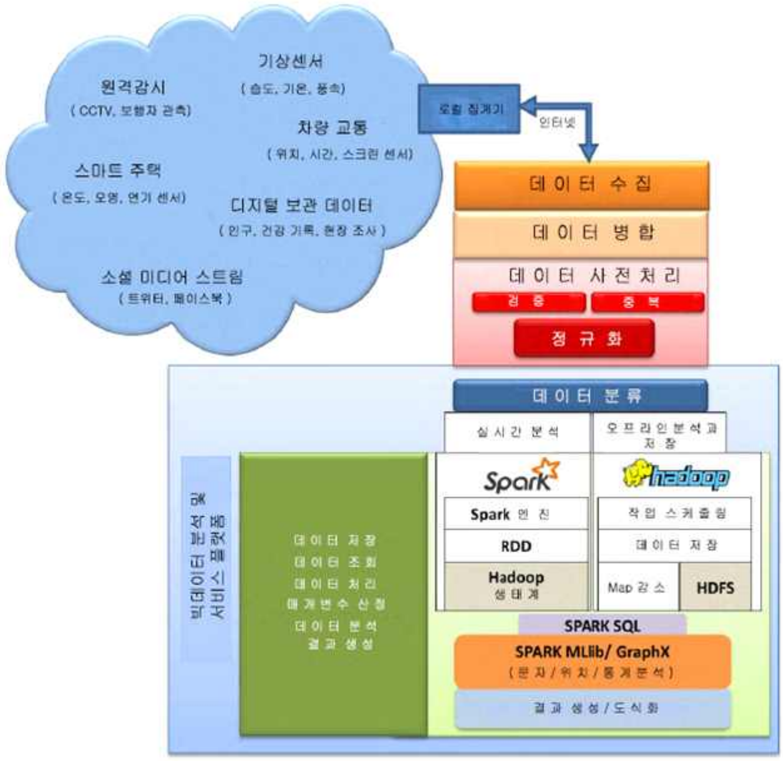


**그림 3. 배치된 시스템의 구현 모델** (출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

3) 데이터 사전 처리

데이터 세트는 초기에 불완전하고 모호하며 중복된 데이터를 제거하기 위해 사전 처리된다. 데이터 사전 처리 기술은 모든 종류의 데이터 분석에 앞서 적용되어야 한다. 정규화라고도 하는 데이터 사전처리는 다양한 데이터 변환 기술을 적용하여 데이터 값이 규정된 범위, 즉 [0, 1]에 속하도록 컴파일한다. 정규화 프로세스를 위한 의사 코드는 알고리즘 1에서 제안된다.

4) 빅데이터 분석 및 서비스 플랫폼

대량의 데이터에는 데이터 세트를 효율적으로 처리할 수 있는 최첨단 빅데이터 분석 도구의 조합이 필요하다. 제안된 아키텍처에서 볼 수 있듯이 이러한 요구 사항을 충족하기 위해 Hadoop 에코 시스템과 Spark 엔진의 조합이 사용된다. 처음에는 식별자(ID)와 메시지 유형의 도움을 받아 데이터가 분류된다. 분류된 데이터는 Hadoop 및 Spark가 이해할 수 있는 형식(예: 시퀀스 파일)으로 변환된다. Spark 엔진과 Hadoop 생태계가 탑재된 시스템 플랫폼은 규정된 알고리즘에 따라 데이터를 처리한다. Map 감소 메커니즘과 함께 Hadoop 에코 시스템을 사용하여 구현한다. Spark SQL은 빅데이터 분석 작업을 수행하는 SQL 기반 선언적 언어이다. Spark MLlib는 여러 최적화된 기계 학습 알고리즘을 제공하고 특수한 사용 사례를 위해 알고리즘을 수정하고 확장할 수 있는 범용 라이브러리이다. Spark GraphX는 그래프를 조작하고 그래프 및 데이터 병렬 시스템을 실행하는 대화형 그래프 계산 엔진을 구성한다.

**Algorithm 1** Data Normalization

**BEGIN**
Input: Datasets of each data values *(DS)*
Output: Dataset of normalized values *(NDS)*
Steps:
**FOR EACH** $i = 1$ to $n$ **LOOP** ▷ $i$ is ID of dataset
Calculate the Mean $(\bar{A})$ for each dataset
Calculate the Mean Absolute Deviation $(S_A)$ for each dataset
Find the Z-score normalization $(V_i')$ for each dataset value ▷ $V_i' = \frac{V_i - \bar{A}}{S_A}$
Divide all values by 5 ▷ to get 1-point range
Add 0.5 to all values ▷ to get values at scale of $[0 \sim 1]$
Return the normalized values in new datasets *(NDS)*
**CONTINUE**$(n+1)$;
**END LOOP**
**END**

## 4. 데이터 분석 : 결과 및 논의

### A. 임계 값 정의

임계 값(CT: Critical Threshold)은 초과할 경우 결과를 변경하거나 경고를 생성하는 특정 값 또는 경계 제한으로 정의할 수 있다. 이 연구의 애플리케이션 요구 사항에 따라 다양한 데이터 세트에 대해 다양한 CT가 설정된다. CT는 재감지 및 경보 생성을 위한 온도 CT, 오염 모니터링을 위한 독성 가스 수준 CT 등과 같이 각 데이터 세트에 대해 수동으로 정의된다.

### B. 분석 결과

IoT 생성 데이터 세트에 대해 제안된 아키텍처를 활용하기 위해 우리는 일상생활에서 일반적으로 발생하고 재난 관리를 위한 경보 생성 상황 내에서 적합한 세 가지 주요 사건을 고려했다. 이러한 사건의 적용은 1) 건물 내 화재 감지, 2) 도시 내 오염의 주요특성 모니터링, 3) 자연재해 또는 사고에 따른 긴급 대피를 돕기 위해 도로 막힘을 파악하는 것 등이다. 시스템이 이러한 이벤트를 감지하고 경보를 생성하는 방법을 자세히 설명했다. 건물의 온도 데이터는 건물의 화재 사고를 식별하기 위해 모든 방에 대해 모니터링된다. 화재 발생 여부를 확인하기 위해 온도 및 연기 수치의 상승률을 분석했다. 그런 다음 알고리즘 2에서 제안한 대로 화재 이벤트에 대한 온도 및 연기 판독 값에 대한 임계 값(CT)을 설정한다. 온도 및 연기 판독 값의 평균이 각각 할당된 CT 값을 초과하면 이벤트가 양수로 보고된다. 온도와 연기 값이 모두 양수의 보고로 결과가 나오면 알고리즘이 화재 경보를 생성한다.

표 1. IoT 적용에 정의된 주요 임계값

| Application | Critical Threshold (CT) |
|---|---|
| Temperature | 55°C |
| Smoke | $200\ g/m^3$ |
| Gases | $200\ g/m^3$ |
| Traffic | 125 Vehicles |

표 2. 경고 메시지 수준에 정의된 트위터 임계값.

| | Alert Message Status | Range |
|---|---|---|
| $T_{CT}$ | Negative | $[0-0]$ |
| | Informational | $[0-0.09]$ |
| | Warning | $[0.1-0.39]$ |
| | Critical/Emergency | $[0.4-1]$ |

**Algorithm 2** Fire Alert

```
BEGIN
Input: Temperature (T) and Smoke (S) Readings
Output: Fire Alert/No-Fire
Steps:
  FOREACH n Reading of Temperature (T) and Smoke (S) LOOP
    T_Avg = (Σ_{i=n-3}^{n} T_i) / 3
    S_Avg = (Σ_{i=n-3}^{n} S_i) / 3
    IF (T_Avg > CT)
      T_Report := TRUE
    ELSE
      GoTo(T_{n+1})
    ENDIF
    IF (S_Avg > CT)
      S_Report := TRUE
    ELSE
      GoTo(S_{n+1})
    ENDIF
  END LOOP
  IF (T_Report && S_Report = TRUE)
    GENERATE (Fire_Alert);
  ELSE
    CONTINUE(n+1);
  ENDIF
END
```

(출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

알고리즘 3은 오염 수준 경보 생성 프로세스에 대한 의사 코드를 보여준다. 그림 4.은 독성 가스, 즉 일산화탄소(CO), 이산화황($SO^2$) 및 이산화질소($NO^2$)가 심각한 임계 값을 초과하는 다양한 시간 간격을 보여준다. 따라서 시스템은 사람들에게 예방 조치를 취하거나 외부로 나가지 않도록 경고를 생성한다.

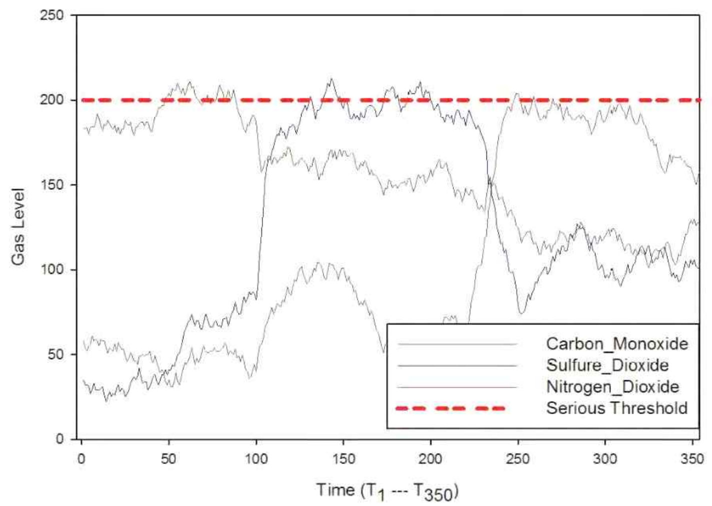

그림 4. 다양한 가스를 통해 오염 모니터링

**Algorithm 3** Pollution Level Alert

**BEGIN**

Input: Air Quality Metrics *(M)*

Output: Pollution Alert/not-polluted

Steps:

```
FOREACH Gas_Readings R (R_CO, R_SO2, R_NO2, R_O3) in (M) LOOP
IF (R_CO, R_SO2, R_NO2, R_O3) > CT)
    Rep =1
    Pollution_Alert();
ELSE
    Rep =0
    GoTo (Next_R)
ENDIF
END LOOP
```

**END**

(출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

또한, 비상 대피 경로 계획 및 실시간 교통 분석을 위해 도로 막힘 및 사고를 식별하기 위해 Volkhin 도로 교통 시뮬레이터의 수동 수정 버전을 사용했다. 우리는 이동하는 차량 수와 속도를 포함하여 차량의 위치와 교통량 정보를 개별 묶음으로 획득하였다. 도로 막힘은 차량 수가 임계 값을 초과하고 '도달 시간'이 기하급수적으로 증가할 때 식별된다. 알고리즘 4는 경로 막힘 경고 프로세스에 대한 의사 코드를 나타낸다. 도로 막힘 분석 결과가 그림5에 나와 있다.

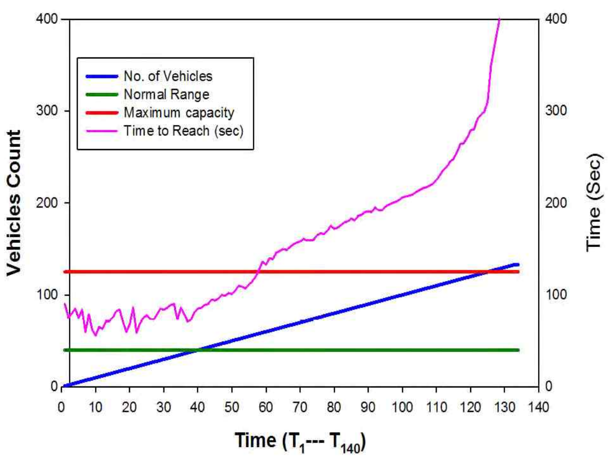

그림 5. 도로의 교통 막힘 분석

**Algorithm 4** Route Blockage Alert

**BEGIN**

Input: Traffic Data with *(Num_vehicles)* and Time interval *(T)*

Output: Route Status *(Blocked and New Route)*

Steps:

```
Identify Time interval (T)        ▷ (T) is time to reach
Identify (R)        ▷ (R) is Route towards destination
FOREACH Reading (Num_vehicles) at (T) on (R) LOOP
IF (Num_vehicles) > CT)
    GoTo(Next_Reading)
ELSE
    Blockage_Alert();
    Alternative_Route (Assign New_Route (R));
ENDIF
END LOOP
```

**END**

(출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

### C. 시스템 구현 및 평가

이 시스템은 3.2GHz 프로세서 및 8GB RAM이 지원하는 Core(TM)i5와 같은 시스템 사양을 사용하여

Ubuntu 14.04 LTS에서 작동하는 다양한 Spark 라이브러리의 지원을 받는 Hadoop 단일 노드 환경에서 구현된다. 제안된 시스템을 구현하는데 사용되는 주요 하드웨어 및 소프트웨어 구성은 표3과 같다. 이후 효율적인 실시간 처리가 필요한 대용량 데이터 세트 처리에 집중했기 때문에 데이터 크기 증가를 고려하여 데이터 처리 및 처리량 측면에서 시스템을 평가했다. 그림6은 증가하는 데이터 세트 크기에 따른 처리 시간 효율성 결과를 보여준다. 데이터 증가에 따라 처리 시간도 늘어날 것으로 예상된다. 그러나 우리 시스템을 사용하면 데이터 크기가 크게 증가함에 따른 처리 시간의 증가가 상당히 낮다. 그림 7은 시스템의 처리량 분석 결과를 보여준다. 처리량 결과는 지정된 기간 동안 시스템에서 처리한 MB 수를 보여준다. 시스템은 데이터 크기가 증가함에 따라 유망한 처리량 경향을 보여준다.

표 3. 시스템의 하드웨어 및 소프트 웨어의 사양

| Item | Version |
|---|---|
| Processor | Core(TM) i5-3470 3.2GHz |
| Hard Disk | SATA 7200RPM HDD |
| RAM | 8GB |
| OS | Ubuntu 14.04 LTS |
| Apache Spark | Spark 2.3.1 |
| Apache Hadoop | Hadoop 2.6.5 |

(출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

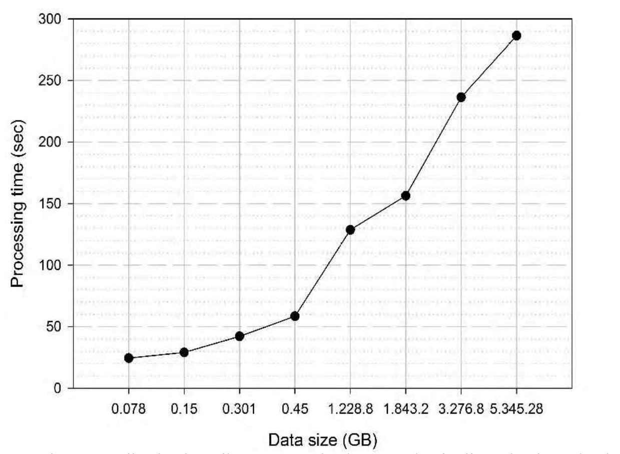


그림 6. 데이터 세트 크기가 증가함에 따라 처리 시간과 관련된 시스템의 효율성

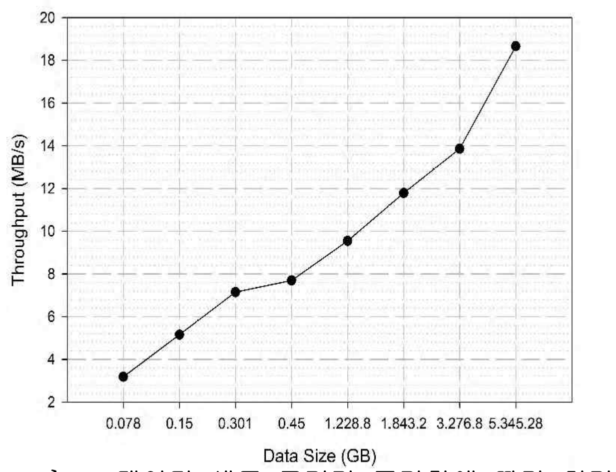


그림 7. 데이터 세트 크기가 증가함에 따라 처리량과 관련된 시스템의 효율성

(출처: S. A. Shah et al.: Towards Disaster Resilient Smart Cities: Can IoT and BDA Be the Game Changers?, https://ieeexplore.ieee.org/document/8759905)

## 5. 연구 과제

이 섹션에서는 DRSC 환경과 관련될 수 있는 주요 연구 과제에 대해 설명한다. 이 연구는 설계 및 구현 단계에서 직면할 수 있는 다양한 문제를 강조하고 환경의 효력을 줄일 수 있다.

### A. 내 결함성

데이터 소스가 여러 개가 있는 재난 상황에서는 물리적 손상, 배터리 소모 또는 통신 채널 장애로 인해 다양한 하드웨어 구성요소가 고장날 가능성이 높다. DRSC 환경에서 데이터 소스는 시스템 가용성을 유지하기 위해 정전 및 기반구조가 손상된 경우에도 데이터를 제공할 수 있어야 한다. 백업 전력 소비 메커니즘과 대체 통신 채널 설정이 보장되어야 한다.

### B. 상호 운용성

데이터 형식이 다른 다양한 실시간 및 정적 데이터 소스에서 데이터를 수집한다. 높은 데이터 중복성으로 인해 품질이 낮을 수 있는 대량의 이기종 데이터를 통합하는 것은 어렵다. 데이터 이질성 문제를 처리하려면 최고 수준의 의미적 상호 운용성과 데이터 품질을 획득할 수 있도록 샘플링 및 필터링 기술을 훈련하여야 한다.

C. 메타 데이터
재해 상황에서 이기종 빅데이터 소스에 대한 메타 데이터의 수집 및 관리는 중요한 과제이다. 빅데이터 패러다임에서 메타 데이터를 생성하고 유지하는 것은 여러 데이터 소스와 데이터 형식으로 인해 매우 어렵다. 메타 데이터를 통해, 여러 데이터 품질 문제 및 통합 관련 문제들을 제거하고, 분석을 위한 믿을 만한 데이터 세트를 제공할 수 있다.

D. 개인 정보 보호 및 보안
개인 정보 보호 문제는 빅데이터 분석에서 심각한 문제였다. 대부분 개인 정보를 사용하여 필요한 결과를 생성하기 때문이다. IoT의 최종 사용자는 IoT의 사용과 생산성을 제한하는 다양한 보안 및 개인 정보 문제에 직면해 있다. 악의적인 사용으로부터 보호하고 데이터 무결성, 가용성 및 기밀성을 유지하려면 재해 관련 데이터에 적합한 보안 메커니즘과 액세스 제어 검사를 적용하는 것이 중요하다.

E. 시간 제약
신속한 대응이 생명을 구할 수 있으므로 재난 관리에서 시간은 매우 중요하다. 긴급 대응을 위해 제한된 시간 내에 바람직한 결과를 추출하기 위해 방대한 양의 이기종 데이터를 사용하는 것은 매우 어렵다. 정해진 시간 내에 의사 결정자의 요구 사항에 따라 방대한 양의 이기종 데이터에서 양질의 데이터를 생성하는 것은 기존 기술과 도구에 있어 큰 도전이다.

F. 표준화
표준은 시스템 효율성을 보증하고, 기술 및 관리의 변경 사항을 채택하고, 용도, 정책 및 향후 연구에 대한 합법적인 지침을 제공하는 데 유용하다. BDA 및 IoT 기술의 사용이 증가함에 따라 통신 표준, 데이터 통합 표준 및 보안 표준을 재검토할 필요성과 범위가 커졌다.

G. GIS 기반 시각화
의사 결정자와 비상 대응자는 상호 작용하고 이해하기 쉬운 빠르고 정확한 예측, 통찰력 및 근거 사실을 필요로 하기 때문에 매핑 및 시각화는 DRSC 환경에서 가장 중요한 부분이다. 빅데이터 분석 및 시각화 도구는 실시간으로 효과적인 결과를 얻기 위해 완전하게 작동해야 한다. 지리정보 시스템(GIS)은 공간 데이터를 매핑하고 분석하기 위한 대화형 인터페이스를 제공한다.

## 6. 결론

최신 BDA와 IoT 기술의 협력은 이기종 데이터 소스가 재난 관리 애플리케이션에 사용되는 필수 정보를 추출하기 위한 효과적인 분석을 수행하는데 유용한 전문적인 환경을 제공한다. 이러한 접근 방식은 빠르고 효과적인 상황 인식을 가능하게 하여 재해의 영향을 줄이는 데 도움이 된다. 재난 관리와 같이 시간에 민감하고 성능이 필요한 애플리케이션을 위한 BDA 및 IoT 시스템 계획 및 설계에는 여전히 큰 연구 격차가 존재한다. 본 연구의 목적은 BDA 및 IoT 기반의 재난 극복력 있는 스마트 도시의 설계 및 구현에 관한 지식에 기여하고 향후 연구를 안내하는 것이다. 본 연구에서 BDA와 IoT를 통합하여 새로운 재난 극복력 있는 스마트 도시 개념에 대한 아키텍처를 제안했다. 제안된 일부 매개 변수와 결합된 BDA 및 IoT를 효과적으로 구현하여 데이터를 집계, 사전처리 및 분석하여 재해 관리자에게 업데이트되고 유용한 정보를 제공하는 방법에 대한 확실한 개요를 제공한다. BDA와 IoT가 공동으로 제공하는 모든 이점을 확보하여 두 기술의 장점을 모두 갖춘 개선된 재해 극복형 스마트 도시 개념을 설계 및 구현할 수 있도록 하는 것이 목표이다. 향후 작업을 위해 원격 감지, UAV 가상, 온라인 뉴스 미디어 및 감시 카메라와 같은 다양한 데이터 소스를 통합하여 재해의 대피, 모니터링 및 예측과 같은 다른 애플리케이션을 추가하여 보다 심층적인 분석과 더 나은 상황 인식을 제공할 것으로 예상한다. DRSC 환경은 시민에게 조기 경보를 생성하고 재난 상황으로부터의 예방, 모니터링 및 복구를 지원하기 위해 다중 소스 데이터로 뒷받침되는 빠르고 효과적인 분석을 가능하게 한다. 이 연구는 스마트 도시 및 재난 관리 영역에서 향후 정보 습득을 위한 연구원 및 산업에 대한 참고 자료를 제공할 수 있을 것이다.

| |
|---|
| **<원문제목>** Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?<br>**<원문출처>** https://ieeexplore.ieee.org/document/8759905 |