

인메모리 기반 치안정보 데이터뱅크 설계에 관한 연구

이상윤, 염경록*, 이영준*, 박원주, 이용태
한국전자통신연구원, *㈜아이와즈

sylllee@etri.re.kr, *yumgr@iwaz.co.kr, *yjlee@iwaz.co.kr, wjpark@etri.re.kr, ytleee@etri.re.kr

A study on the design of in-memory-based police information data bank

Sang-Yun Lee, Gyung-Rok Yeom*, Yeon-Jun Lee, PARK,WON-JOO, Lee Yong Tae

ETRI, *IWAZ

요 약

데이터뱅크에 모이는 치안 정보들이 방대하게 많아지면, 데이터를 처리하는 속도가 급격히 떨어진다. 본 논문에서는 이러한 방대한 빅데이터 처리 속도를 향상시키기 위해 인메모리 기반으로 새로 설계한 치안 정보 데이터뱅크를 제안한다. 또한, 유사한 범죄 정보를 최대한 많이 찾을 수 있도록 순서쌍을 이용한 새로운 검색 방법을 제안한다

I. 서 론

치안 정보는 민감한 정보들을 많이 보유하고 있어서 공개되거나 획득하기가 쉽지 않은 데이터들이다. 하지만, 최근에 인공지능 기술이 급속도로 발전됨에 따라 치안 빅데이터를 구축하여 이를 범죄 분석이나 범죄 예방에 활용하려는 시도들이 빈번히 일어나고 있다[1]-[3]. 특히, 치안 분야에서의 검색 정확도나 분석의 성능을 높이기 위해 치안 전문 용어에 대한 중요성이 부각되어 이를 구축하기 위한 도구 연구 사례도 있다[4]. 이 도구는 치안 정보 빅데이터 구축을 위한 치안 정보 데이터뱅크 구축의 한 기능으로 개발되었다.

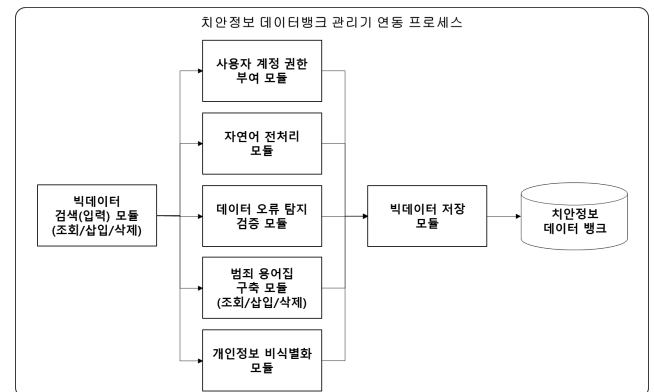
데이터뱅크에 모이는 치안 정보들이 방대하게 많아지다 보면, 데이터를 처리하는 속도가 급격히 떨어진다. 본 논문에서는 이러한 방대한 빅데이터 처리 속도를 향상시키기 위해 인메모리 기반으로 [5], [6] 새로 설계한 치안 정보 데이터뱅크를 제안하고 치안정보 데이터뱅크 연동 프로세스를 기술한다. 또한, 유사한 범죄 정보를 최대한 많이 찾을 수 있도록 순서쌍을 이용한 새로운 검색 방법을 제안하고 사용자가 편리하게 검색 기능을 이용할 수 있는 사용자 인터페이스 설계를 제시한다.

II. 인메모리 기반 치안정보 데이터뱅크

치안정보 데이터뱅크 관리기는 빅데이터 검색 모듈, 사용자 계정권한 부여 모듈과 개인정보 비식별화 모듈, 데이터 오류 탐지 및 검증 모듈, 범죄 용어집 구축 모듈, 빅데이터 저장 모듈의 연계를 통해 치안정보 통합 데이터뱅크 DB를 관리한다[1].

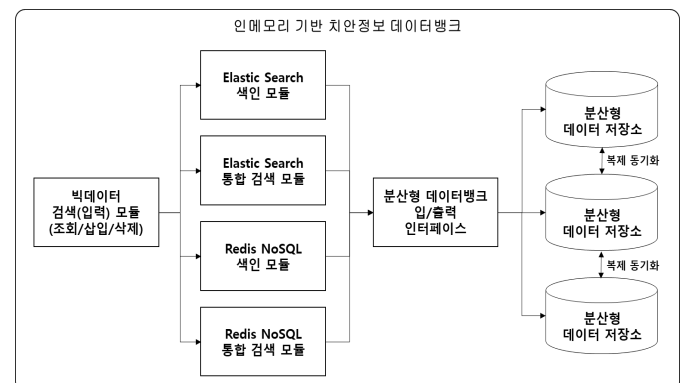
각 모듈은 빅데이터 검색 모듈과 빅데이터 저장 모듈을 통하여 데이터의 조회 및 수정/삭제를 수행하고 외부에서 활용할 수 있도록 인터페이스를 제공한다. 이에 대한 연동 프로세스가 [그림 1]에 나타나 있다.

인메모리 기반 치안정보 데이터뱅크는 빅데이터 검색 모듈, Elastic Search 색인 모듈과 Elastic Search 통합 검색 모듈, Redis NoSQL 색인 모듈, Redis NoSQL 통합 검색 모듈, 분산형 데이터뱅크 입/출력 인터페이스의 연계를 통해 인메모리 기반 치안정보 데이터뱅크의 검색을 효율적으로 지원한다.



[그림 1] 치안정보 데이터뱅크 관리기 연동 프로세스

각 모듈은 빅데이터 검색 모듈과 분산형 데이터뱅크 입/출력 인터페이스 모듈을 통하여 대용량의 치안정보 데이터에서의 빠른 조회(검색)을 수행할 수 있으며 완전일치, 부분일치, 연산자 검색 등 다양한 검색 기법 및 색인 기법을 활용할 수 있다. 이에 대한 프로세스는 [그림 2]에 나타나 있다. 여기에 사용되는 DBMS 는 오픈소스인 Redis NoSQL 을 통합 검색 모듈로 사용하였다.



[그림 2] 인메모리 기반 치안정보 데이터뱅크

III. 치안정보 유사 범죄 검색 시스템

오픈소스인 분산형 MariaDB 로 관리되는 치안정보 통합 데이터뱅크 DB 는 경찰청이 보유한 치안정보를

저장하기 위한 속성을 추가하고, 이를 기반으로 치안정보의 특징 정보들을 저장하기 위한 테이블을 새로 추가하여 치안정보 통합 데이터뱅크를 다양한 분석 및 기능에서 활용할 수 있도록 한다. 이 중 빅데이터 검색 모듈은 치안정보 통합 데이터뱅크 DB 를 검색하여 원하는 치안정보들을 전달받고, 빅데이터 저장 모듈은 데이터의 수정 및 추가 대상인 치안정보를 데이터뱅크 DB 에 저장한다.

치안정보 유사 범죄 검색은 동일한 수법의 범인을 색출하거나, 수사의 단서를 신속하게 찾기 위해 필요한 기능이다. 본 논문에서는 유사 범죄 검색의 정확도와 가능한 한 많은 연관된 유사 범죄 검색을 위해 순서쌍을 이용한다.

유사 범죄 검색 프로세스는 사용자가 수사중인 문서를 입력하면 자연어 전처리 모듈을 통하여 다양한 전처리를 수행한 결과를 이용하여 인물 중심 또는 사건 중심의 검색 알고리즘을 통해 수사 중인 사건과 가장 유사한 기준 범죄 정보를 제공한다. 이에 대한 전체 UI 는 [그림 3]에 나타나 있다.

[그림 3] 치안정보 유사 범죄 검색 시스템

유사 범죄 검색 결과는 사용자가 수사중인 원문 문서와 가장 유한 결과 리스트를 제공한다. 각 문서의 순위, 문서 ID, 데이터 출처(KICS, 112, 보도자료 등), 유사도 점수, 유사 사건 정보를 제공한다. 검색결과 중 특정 문서 클릭 시 상세 정보를 확인할 수 있다. 이에 대한 결과 UI 는 [그림 4]에 나타나 있다.

[그림 4] 유사 범죄 검색 결과

IV. 결론

본 논문에서는 치안정보 데이터뱅크의 검색 성능을 향상시키기 위한 인메모리 기반 치안 정보 데이터뱅크에 대해서 제안하였다. 대용량의 데이터를 저장하고 관리하기 위한 분산형 DB 구축과 저장소의 데이터를 In-Memory 기반에서 빠른 검색을 위한 Redis NoSQL 색인/검색 모듈, 다양한 언어처리를 적용 및 연산자 검색을 위한 Elastic Search 색인/검색 모듈로 구성하여 치안정보 데이터뱅크의 최적의 데이터 검색 환경을 제공한다.

또한, 순서쌍을 이용한 치안 정보 유사 검색 시스템을 제안하였다. 유사한 범죄를 검색하기 위한 자연어 전처리 모듈, 색인 및 검색 모듈, 사용자 친화적 UI 등으로 구성하여 수사 중인 사건과 유사한 범죄를 검색할 수 있는 시스템을 제공한다.

향후에는 인메모리 기반 데이터뱅크를 이용하여 분산 DB 환경에서 치안 정보를 검색하기 위한 DB 구축 및 검색 성능 향상에 대한 연구를 지속히 추진할 계획이다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2018-0-00440, 위험 상황 초기 인지를 위한 ICT 기반의 범죄 위험도 예측 및 대응 기술 개발)

참 고 문 헌

- [1] S. Y. Lee, W. J. Park, Y. T. Lee, G. R. Yeom, J. W. Hwang, K. H. Jang, "Police Information Databank for Big Data Analysis", International Conference on Electronics, Information and Communication (ICEIC) 2020, pp. 773-776, 2020.01.
- [2] 김승우, 정유진, "스마트치안에서의 자율주행차 및 드론을 활용한 여성 범죄 예방 연구", 한국정보통신학회논문지, 제 22 권 제 10 호, pp. 773-776, 2020.01.
- [3] 백영민, 김경운, 박현호, "CCTV 설치현황에 대한 인식과 CCTV 치안효과 인식이 거주민의 체감안전도에 미치는 효과", 경찰대학 경찰학연구, 제 18 권 제 3 호(통권 제 55 호), pp. 9-38, 2018.09.
- [4] 이상운, 염경록, 정인규, 장광호, 박원주, 이용태, "치안정보 통합 데이터뱅크용 용어집 구축에 관한 연구", 통신학회 하계학술대회 2019, ISSN: 2383-8302(Online) Vol. 69, 2019.06.
- [5] Secil Yuzuk, Murat G. Aktas, Mehmet S. Aktas, "On the Performance Analysis of Map-Reduce Programming Model on In-Memory NoSQL Storage Platforms: A Case Study", 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), pp. 45-50, 2018.
- [6] Hai Thanh Mai, Kyoung Hyun Park, Hun Soon Lee, Chang Soo Kim, Miyoung Lee, Sung Jin Hur, "Dynamic Data Migration in Hybrid Main Memories for In-Memory Big Data Storage", ETRI Journal, Volume 36, Issue 6, pp. 988-998, 2014.