

통합 지휘통제체계를 위한 벡터 데이터베이스 요구사항 분석

금형철, 문경렬, 백승호

LIGNEX

kyeongryeol.moon@lignex1.com, hyungcheol.geum@lignex1.com, seungcho.baek@lignex1.com

Requirements Analysis of Vector Databases for Integrated Command and Control Systems

Geum Hyeongcheol, Moon Kyeongryeol, Baek Seungcho

LIGNEX

요약

최근 대형 언어 모델(Large Language Model, LLM)의 발전과 함께, 인공지능 모델을 활용하는 각 조직들이 효율적으로 데이터를 참조하도록 데이터 저장 및 관리 기술들이 발전하고 있다. 특히 LLM을 포함한 AI 모델들의 입출력처럼, 요청 및 참조 데이터 또한 벡터 저장로 다루면 장점이 크다. 또한 이중 데이터에서도 의미론적으로 동일한 벡터로 사용 시, 데이터 관리 프로세스를 단순화하고 통합된 논리적 공간으로 사용할 수 있다. 이 개념은 국방에서 전 영역으로부터 수집한 정보를 통합 및 실시간 처리하고 인공지능 기반으로 지휘관 의사결정을 지원하는 지휘통제체계에도 응용할 수 있다. 이러한 체계에 적용하는 벡터 데이터베이스는 새로 들어오는 정보와 기존의 군 정보/자료들을 동일한 임베딩 공간의 지식베이스로 구축하여 지휘관 요청에 따른 데이터를 반환하고, 전장 상황을 분석하는 대형 모델 등 부가 기능에 활용할 수 있다. 따라서 본 논문은 한국형 통합 지휘통제체계를 위한 벡터 데이터베이스 채택 시, 장점과 JADC2(Joint-All Domain Command and Control)의 요구사항을 참고하여 중요사항을 분석한다.

I. 서론

지난 2년간 대형 모델, 특히 LLM 위주로 크게 발전하면서, 이를 활용하는 서비스들이 공개되고 있다. 특히 크게 주목받는 시장은 각 기업/조직만의 데이터에 특화된 온프레미스 솔루션이다. 이러한 솔루션들의 공통점은 내부 데이터들을 벡터로 embedding한 수학적 표현에 의존하는 것이다. 이 표현은 상호 의미론적인 정보를 공유한다. 그래서 소스가 서로 다른 데이터들이어도 벡터 공간에서 유사도 기반 검색 등의 알고리즘을 통해 사용자에게 데이터 간 연관성과 복잡한 분석 결과를 제공할 수 있다.

따라서 이러한 벡터 표현을 다루는 효율적인 방법을 고민하여야 한다. 이의 방안으로 최근 벡터 데이터베이스가 주목받고 있다. 벡터 데이터베이스는 종류에 상관없이 데이터를 임베딩 벡터로 저장하고, 자체적인 알고리즘을 통하여 출처에 무관하게 의미론적 관련성이 높은 벡터들을 인덱싱하여 사용자에게 제공한다. 벡터 데이터베이스는 수치 연산을 가속하는 GPU를 통해 효율성을 최대화할 수 있고, 시스템의 복잡도와 확장성에 유연하게 대처할 수 있는 장점이 있다.

다양한 벡터 데이터베이스가 공개되었지만, 아직 상용 AI 서비스에서도 서서히 도입되고, 최적 서비스 구성도를 설계하는 과정은 시행착오가 크다. 다만 광범위하고 대용량 데이터의 임베딩을 복잡한 추론에 사용하는 측면에서 미래 국방 영역에도 적용 가능한 요소가 많다. 특히 미 국방부가 개발하는 JADC2와 같은 통합 지휘통제체계가 고려 대상이다. 이 체계는 전 영역에서 정보를 수집하고, 실시간으로 데이터 통합 및 관리하는 것을 요구한다. 더 나아가 이 데이터들로 인공지능 기반의 전장 분석을 지원하여 지휘관 의사결정 지원을 목적으로 한다. 이러한 최근 방향에 착안하여 본 논문은 한국형 통합 지휘통제체계에 적용할 수 있는 데이터 관리 기법으로서 벡터 데이터베이스를 채택 시, JADC2의 요구사항과 관련하여 중요점을 분석한다.

II. 관련 연구

II-1. 벡터 데이터베이스를 응용한 대형 모델

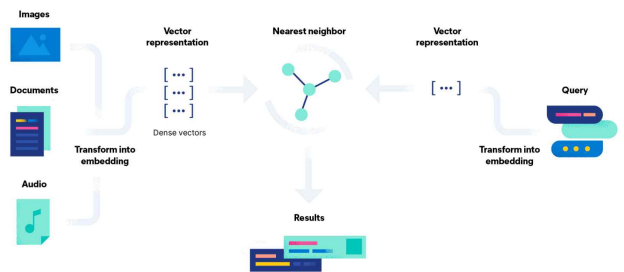


그림 1 벡터 데이터베이스 구조 [6]

최근 5년 간 Elasticsearch, Milvus, Pinecone, Weaviate, Chroma, MyScale, Marqo, MongoDB 등 수 많은 상용 벡터 데이터베이스 서비스가 공개되었다 [1]. 이러한 벡터 데이터베이스는 여러 데이터들을 벡터 임베딩한 수치 표현을 저장한다. 그리고 검색 및 관리를 위해 Graph, Quantization, Hashing 등 인덱싱 알고리즘으로 원래의 벡터를 간단한 구조로 변환한다. 그런 다음 Dot product, Cosine similarity 같은 유사도 알고리즘으로 외부 요청 벡터와 미리 인덱싱한 벡터들을 비교하여 의미적 관련성이 높은 벡터들을 랭킹하거나 반환한다. 이러한 벡터 데이터베이스는 차원이 큰 임베딩 벡터를 활용하는 LLM을 포함한 인공지능 모델과 결합하여 검색 증강 기법(Retrieval-Augmented Generation, RAG) 등 출력 품질을 향상하는 방법으로 발전하였다 [2, 3].

II-2. 통합 지휘통제체계

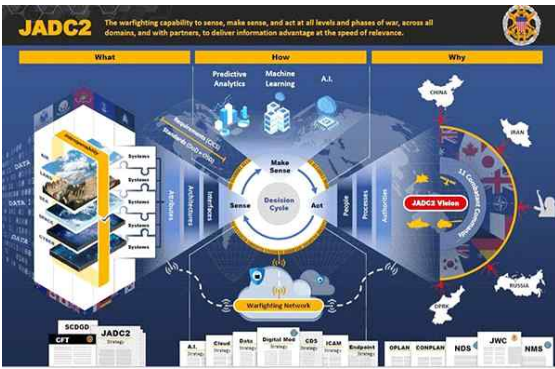


그림 2 JADC2 구조도[2]

전장 환경은 과거와는 달리 군중, 지역을 제한하지 않고, 전 영역 데이터를 통합하고 이로부터 분석 결과를 기반으로 지휘관이 의사결정을 하는 형태로 변하고 있다. 미국이 개발하는 JADC2는 대표적인 통합 지휘통제 체계이다. 육해공 및 우주·사이버 정보까지 포함하는 이 체계는 미군 의 현재 운용 중인 지휘통제체계를 하나의 네트워크로 연결하면서, 지휘관 의사결정 주기를 가속하는 것을 목표로 한다. [4]. 서방권의 또 다른 사례로 NATO는 Multi-Domain Operation(MDO)을 설계 중이다. 이 체계 또한 JADC2와 유사한 목적을 갖고 있으며, NATO 동맹국 간 효율적으로 전 영역 작전 능력의 동기화하는 것을 목적인다 [5].

	Elastic search	Milvus	Pinecone	Weaviate	Chroma
온-프레미스	✓	✓	✗	✓	✓
클라우드	✓	✓	✓	✓	✗
벡터 전용	✗	✓	✓	✓	✓
*QPS	100~700	2406	150 (1포드)	791	-
디스크 인덱싱	✗	✓	✓	✓	✓
접근 제어	✓	✓	✓	✗	✗
데이터 확장	정적 데이터 샤딩	동적 세그먼트 배치	-	정적 데이터 샤딩	동적 세그먼트 배치

표 1 주요 벡터 데이터베이스 비교 (*QPS: 초당 쿼리 처리 수)

III. 통합 지휘통제체계를 위한 벡터 데이터베이스 고려사항

JADC2에서 공개한 데이터 요구사항 중 벡터 데이터베이스 기반 체계 구축 시 특히 중요한 항목은 표준화된 데이터 인터페이스, 체계 확장성, 사용자 접근 제어이다. 각 벡터 데이터베이스들도 기업용 솔루션에 제공할 수 있도록 이러한 요구사항들과 관련된 기능을 제공하지만, 그 범위가 수준은 다양하다. 따라서 체계에 적합한 기능 및 성능을 보유한 벡터 데이터베이스를 선택하여야 한다. 표 1에서 JADC2 요구사항 기준으로 관련있는 주요 벡터 데이터베이스들의 특징을 볼 수 있다.

벡터 데이터베이스의 표준화된 데이터 인터페이스는 전 영역 데이터 간 일관된 임베딩을 위한 인코더의 성능을 보장하는 것이다. 서로 다른 데이터가 의미론적으로 같은 벡터 공간 표현으로 임베딩되지 않으면, 수학적 으로 유사도 연산 결과의 신뢰도가 낮아진다. 따라서 인공지능 모델의 추론 및 지휘관 의사결정의 불확실성이 커져 적합한 전장상황 판단을 하기 어렵다.

체계 확장성 또한 통합 지휘통제체계의 벡터 데이터베이스에서 중요한 항목이다. JADC2를 포함한 통합 지휘통제체계는 완전히 새로운 체계를

만들어 대체하는 것이 아닌, 모든 기존 체계와의 연동을 고려하여 설계되는 슈퍼셋 개념이기 때문이다. 따라서 벡터 인덱싱, 유사도 검색 뿐만 아니라 기존 체계의 데이터베이스와 연동되는 가능성을 고려하고, 추후 체계 복잡도가 커져도 응답 정확도와 대규모 데이터 스트리밍의 내구도를 보장하여야 한다. 표 1에서는 이러한 요구사항을 충족하는 벡터 데이터베이스로 Elasticsearch가 있다. 하지만 디스크 인덱싱, 동적 세그먼트를 지원하지 않는다. 이는 통합 지휘통제체계의 대규모 데이터를 처리하고 적절한 요청자에게 제공하는 측면에서 적절하지 않다. 한편 Milvus, Chroma 등은 디스크 인덱싱, 동적 세그먼트를 지원한다. 이는 통합 지휘통제체계의 복잡도와 확장성을 고려한 요구사항을 충족하는 특징이다.

사용자 접근 제어는 데이터 접근과 요청을 할 수 있는 권한과 관련있다. 이 기능은 많은 벡터 데이터베이스들이 엔터프라이즈에서 요구하는 보안 수준을 충족한다. 통상 벡터 데이터베이스에서 제공하는 방식은 역할 기반 액세스 제어(Role Based Access Control, RBAC)로 표 1의 접근 제어는 모두 RBAC를 의미한다. RBAC는 군 시스템에 적용하기는 어렵다. 따라서 지휘통제체계와 같이 보안 등급이 중요한 솔루션에서 벡터 데이터베이스를 적용하기 위해서는 강제적 접근 제어(Mandatory Access Control, MAC) 같이 중앙 집중으로 권한을 부여하는 접근 제어가 필요하다.

IV. 결론

본 논문은 최근 대형 모델의 발전과 함께 주목받는 벡터 데이터베이스의 장점을 군 통합 지휘통제체계에 도입할 시 기대되는 장점과 미 국방부 JADC2의 데이터 요구사항과 관련하여 주요한 벡터 데이터베이스 특징을 분석하였다.

미래의 지휘통제체계는 어떠한 형태로든 육해공, 우주 및 사이버, 민간 등 영역을 가리지 않고 대규모, 이종 간 데이터를 통합 관리하는 체계를 갖추게 될 것이다. 또한 JADC2에서도 계획하는 것처럼 인공지능 모델을 통한 지휘관 의사결정 지원과 그 주기성에 높은 비중으로 개입할 것이다. 이러한 체계의 내부 구성 간 데이터 플로우는 반드시 벡터 표현의 데이터 타입이 존재한다. 따라서 벡터 데이터 관리에 특화된 데이터베이스 구축이 필요하다. 이때 벡터 데이터베이스는 표준화, 확장성, 접근 제어 측면에서 요구사항을 충족할 수 있고, 본문에서 분석한 특징 외에도 실시간성, 의미론적 검색 등 고수준 데이터 분석을 위한 여러 장점들이 통합 지휘통제체계의 적절한 데이터베이스 솔루션으로서 도입 타당성이 크다. 물론 상용 벡터 데이터베이스를 그대로 사용할 수는 없다. 체계 구축 시 운용 비용, 속도, 정확도 측면에서의 균형 또한 고려하여야 한다. 그리고 기반 기술로서 국방 도메인 특성을 고려한 임베딩 모델, 유사도 연산 기술이 필요하다. 이러한 문제는 초기 단계에서 이론적 연구가 선행되어야 한다.

참고 문헌

- [1] James Jie Pan et al, Survey of 벡터 Database Management Systems, <https://arxiv.org/abs/2310.14021>
- [2] Yunfan Gao et al, Retrieval-Augmented Generation for Large Language Models: A Survey, <https://arxiv.org/abs/2312.10997>
- [3] Zhi Jing et al, When Large Language Models Meet 벡터 Databases: A Survey, <https://arxiv.org/abs/2402.01763>
- [4] Summary of the Joint-All-Domain Command & Control (JADC2), Department of Defence, March 2022
- [5] Multi-Domain Operations in NATO, NATO, October 2023
- [6] 벡터 데이터베이스란 무엇인가?
<https://www.elastic.co/kr/what-is/vector-database>