JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 10, pp. 3105-3116, Oct. 2026

Check for updates

소방청 일일상황보고 분석을 활용한 동적 다차원 위험평가 프레임워크 기반 의사결정 지원 시스템 개발

김 지 호[1,2] · 심 준 석[1] · 임 언 택[1] · 정 민 수[1*]
[1]국립재난안전연구원 재난원인조사실 재난회복연구센터
[2]부산대학교 정보융합공학과 박사과정

Decision Support System Based on a Dynamic Multidimensional Risk Assessment Framework Using National Fire Agency Daily Situation Reports

Jiho Kim[1,2] · Junseok Sim[1] · Eonteak Lim[1] · Minsu Jung[1*]

[1]Disaster Prevention Research Division, National Disaster Management Research Institute, Ulsan 44538, Korea
[2]Department of Information Convergence Engineering, School of Computer Science and Engineering, Pusan National University, Busan 46241, Korea

[요 약]

본 연구는 소방청 일일상황보고에 축적된 비정형 사건정보를 활용하여 소방공무원의 도시 안전관리 의사결정을 지원하는 위험브리핑 시스템을 개발하는 것을 목적으로 한다. 2024년 1월부터 2026년 6월까지의 일일상황보고 875건에서 5,770건의 사건자료를 구축하고, Sentence-BERT, UMAP 및 HDBSCAN을 적용하여 의미적으로 유사한 사건군을 분류하였다. 이후 기준일을 중심으로 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험을 분석하고, 선정된 위험정보를 실제 대표사건과 연결하였다. 또한 생성형 언어모델을 활용하여 위험요약, 위험 인사이트 및 대응 참고사항으로 구성된 일일 위험브리핑을 생성하였다. 2026년 4월 25일을 기준일로 적용한 결과, 최근 증가 위험과 계절 집중 위험에서는 공장 및 창고화재 중심 사건군이, 과거 동일 시기 반복 위험에서는 교통사고 중심 사건군이 선정되었다. 본 연구는 장기간 축적된 비정형 소방 사건정보를 입력시점의 위험정보로 변환하고, 이를 실제 사건 근거와 연계한 동적 다차원 위험평가 프레임워크로 제시하였다는 점에서 의의가 있다.

[Abstract]

This study develops a risk briefing system to support urban safety management decision-making using unstructured incident information from the National Fire Agency's daily situation reports. A total of 5,770 incident records were extracted from 875 reports collected between January 2024 and June 2026. Sentence-BERT, UMAP, and HDBSCAN were applied to clutser semantically similar incidents. Based on a user-defined reference date, risks associated with recent increases, recurring historical periods, and seasonal concentrations were analyzed and linked to representative incidents. A Qwen-3-based generative language model was then used to generate daily risk briefings containing risk summaries, insights, and response guidance. Using April 25, 2026, as the reference date, factory- and warehouse-fire incident groups were identified for recent increase and seasonal concentration risks, while a traffic accident incident group was identified for recurring same-period risk. Overall, this paper presents a dynamic multidimensional risk assessment framework that transforms long-term unstructured firefighting incident records into actionable risk information and links the results to evidence from actual incidents.

색인어 : 일일상황보고 데이터, 소방 빅데이터, 위험브리핑, 의미 기반 군집화, 의사결정 지원

Keyword : Daily Fire Situation Reports, Fire Safety Big Data, Risk Briefing, Semantic Clustering, Decision Support System

http://dx.doi.org/10.9728/dcs.2026.27.10.3105

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 18 August 2026; **Revised** 04 September 2026
Accepted 02 October 2026

***Corresponding Author; Minsu Jung**
Tel: ※ 개인정보 표시제한
E-mail: msjung@korea.kr

Ⅰ. 서 론

도시의 과밀화와 시설의 복합화, 인구 및 교통활동의 증가는 도시에서 발생하는 안전사고의 유형과 대응환경을 다양하게 만들고 있다. 또한 기후와 계절적 환경의 변화는 화재, 구조, 구급 및 생활안전 사고의 발생 시기와 피해 양상에 영향을 미치며, 특정 사고가 단기간에 증가하거나 매년 유사한 시기에 반복되는 현상으로 이어질 수 있다. 이러한 도시 안전환경에서는 개별 사고에 대한 사후 대응과 함께, 축적된 사고정보를 활용하여 위험의 변화를 조기에 확인하고 대응 준비에 반영하는 과정이 필요하다[1].

도시 재난안전관리는 사고 발생 이후의 대응에 한정되지 않으며, 위험정보의 수집과 분석, 예방활동, 대응자원의 점검, 관계자 간 정보공유 등 사고 이전의 대비활동을 포함한다[2]. 특히 소방기관은 화재진압과 구조·구급 활동을 수행하는 동시에 계절별 안전대책, 소방대상물 점검, 장비와 인력의 사전 확인, 상황정보 공유 등 다양한 예방 및 대비 업무를 담당한다. 따라서 소방공무원이 현재 시점에서 주의해야 할 사고유형과 과거에 반복된 위험을 신속하게 확인하여 현장 대응 준비의 효율성을 높여야 한다.

최근 공공안전 분야에서는 행정과 현장 업무에서 축적되는 데이터를 활용하여 안전관리의 객관성과 일관성을 높이려는 시도가 확대되고 있다[3],[4]. 소방 분야에서도 화재, 구조, 구급 및 생활안전 활동과 관련된 다양한 자료가 수집되고 있으며, 이를 활용한 사고 발생 특성 분석, 출동수요 예측 및 위험지역 분석 등이 수행되고 있다[5],[6]. 이러한 데이터 기반 접근은 담당자의 경험과 직관을 보완하고, 반복적으로 발생하는 위험을 정량적으로 파악할 수 있다는 점에서 의미가 있다.

소방청 일일상황보고는 전국에서 발생한 주요 화재, 구조·구급 및 생활안전 사건의 발생시점, 장소, 사고내용, 인명 및 재산피해, 구조와 이송내용 등을 포함하는 대표적인 소방 사건정보이다. 일일상황보고는 개별 사건의 발생과 대응 내용을 확인하는 데 유용하며, 장기간 축적할 경우 사고유형별 발생 추세와 시기적 특성을 분석할 수 있는 자료로 활용할 수 있다. 그러나 사건내용이 서술형 문장과 비정형 문서 형태로 기록되므로, 여러 날짜와 연도의 자료를 동시에 비교하거나 표현이 서로 다른 유사 사건을 일관된 유형으로 분류하기 어렵다. 이로 인해 최근 증가한 사고, 과거 동일 시기에 반복된 사고, 특정 계절에 집중된 사고를 확인하려면 담당자가 다수의 보고서를 직접 검색하고 비교해야 한다.

기존 소방안전 데이터 연구는 화재 및 구조와 구급 사고의 발생 빈도, 공간적 분포, 출동수요와 피해 특성을 분석하는 데 기여해 왔다[7]-[9]. 또한 자연어처리와 텍스트마이닝 기술의 발전에 따라 사고보고서와 재난 관련 문서를 분석하여 주요 키워드, 주제 및 사고유형을 도출하는 연구도 진행되고 있다[10]-[12]. 그러나 기존 연구는 특정 사고의 발생 특성을 사후적으로 설명하거나 사건문서를 분류하는 데 집중하는 경향이 있다. 축적된 비정형 사건정보를 기준일에 따라 분석하고, 최근 변화와 반복성, 계절성을 함께 고려하여 소방공무원의 일일 업무에 활용할 수 있는 위험정보로 제공하는 연구는 상대적으로 제한적이다.

본 연구의 목적은 소방청 일일상황보고에 축적된 비정형 사건정보를 활용하여 소방공무원의 도시 안전관리 의사결정을 지원하는 위험브리핑 시스템을 개발하는 것이다. 이를 위해 첫째, 일일상황보고에서 화재, 구조·구급 및 생활안전 사건을 추출하고 사건 서술문을 의미기반으로 분류하였다. 둘째, 사용자가 입력한 기준일을 중심으로 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험을 분석하였다. 셋째, 선별된 위험정보를 발생 건수, 피해 특성 및 실제 대표사건과 연결하고, 생성형 인공지능을 활용하여 업무 점검과 정보공유에 활용할 수 있는 일일 위험브리핑으로 제공하였다. 결과적으로 본 연구는 축적된 비정형 소방 사건정보를 기준일에 따른 발생빈도, 단기 변화, 연도 간 반복성, 계절성 및 심각도를 종합적으로 반영하는 위험정보로 반영한다. 이를 실제 사건 근거와 브리핑으로 연결하는 동적 다차원 위험평가 프레임워크를 제시한다는 점에 의의가 있다.

Ⅱ. 관련연구

도시 안전관리는 다양한 위험요인을 사전에 파악하고 대응역량을 확보하는 방향으로 발전하고 있다. 도시의 고밀화와 시설의 복합화는 화재, 교통사고, 구조사고, 구급활동 및 생활안전 사고의 발생 양상을 다양하게 만들고 있다. 기후와 계절적 환경의 변화도 사고의 발생 시기와 피해 특성에 영향을 미친다. 이러한 도시환경에서는 사고정보를 지속적으로 수집하고 분석하여 위험요인을 확인하며, 분석결과를 예방활동과 대응 준비에 반영하는 과정이 중요하다[13],[14].

데이터 기반 분석은 도시 안전관리의 객관성과 효율성을 높이는 주요 수단이다. 공공기관과 지방자치단체는 행정자료와 사고자료를 활용하여 사고 다발지역, 안전취약요인, 시간적 발생특성, 공간적 분포 및 대응수요를 분석하고 있다[10]-[12]. GIS와 통계분석을 적용한 연구에서는 도시 내 위험지역과 안전취약지역을 도출하였으며, 머신러닝을 활용한 연구에서는 사고 발생 가능성과 피해 규모를 예측하였다[15],[16]. 이러한 분석결과는 안전관리 대상과 대응자원의 우선순위를 결정하는 기초자료로 활용될 수 있다.

소방 분야에서는 화재와 구조 및 구급 활동에서 축적된 데이터를 활용한 연구가 수행되고 있다. 화재 관련 연구에서는 사고 원인, 발생 빈도, 시간대, 계절적 특성, 공간적 분포 및 피해 규모를 분석하였다. 구조와 구급 관련 연구에서는 출동건수, 환자 유형, 지역별 수요, 출동시간 및 대응 특성을 중심으로 소방활동의 발생 양상을 분석하였다[17],[18]. 일부 연구에서는 GIS와 통계모형을 활용하여 화재위험지역과 출동

취약지역을 도출하고, 소방자원의 배치와 대응계획 수립을 지원하였다[18],[19].

비정형 안전정보의 활용을 위해 다양한 자연어처리 기법이 적용되어 왔다. 초기 연구에서는 키워드 빈도분석과 동시출현 분석을 활용하여 사고 원인과 주요 위험요인을 도출하였으며, 토픽모델링을 통해 대규모 사고기록에 포함된 잠재 주제를 분류하였다[20]. 이후에는 머신러닝 기법을 적용하여 사고유형과 피해수준을 분류하는 연구가 수행되었다[17]. 최근에는 언어모델과 문장 임베딩을 통해 문맥과 의미적 유사성을 반영하는 분석도 확대되고 있다[21].

기존 소방안전 연구는 사고 발생 특성의 분석과 사건유형의 분류에 기여하였다. 그러나 많은 연구가 특정 사고유형이나 제한된 기간의 자료를 대상으로 수행되었으며, 분석결과도 빈도, 공간분포, 예측 또는 문서분류 형태로 제시되는 경우가 많았다. 장기간 축적된 사건정보를 현재 시점의 안전관리 업무와 연결하고, 시간에 따른 변화와 반복성을 함께 고려한 연구는 충분하지 않다. 또한 분석결과를 실제 사건 근거와 연계하여 소방공무원의 일일 상황판단에 활용할 수 있는 정보로 제공한 사례도 제한적이다.

본 연구는 소방청 일일상황보고를 활용하여 기준일 중심의 위험정보를 제공한다. 화재, 구조, 구급 및 생활안전 사건의 서술문을 의미기반으로 분류하고, 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험을 분석한다. 선별된 위험정보는 발생 건수, 피해 특성 및 실제 대표사건과 연결하며, 생성형 인공지능을 활용하여 일일 위험브리핑으로 제공한다. 이를 통해 비정형 소방 사건정보의 활용범위를 확대하고, 도시 안전관리를 위한 소방공무원의 상황판단과 업무 의사결정을 지원하고자 한다.

Ⅲ. 연구방법

본 연구는 소방청 일일상황보고를 사건 단위 자료로 구축하고, 사건 서술문의 의미적 유사성과 시간적 발생 특성을 분석하여 일일 위험브리핑을 생성하였다. 연구절차는 데이터 구축, 유사사건 분류, 기준일 기반 위험정보 분석, 생성형 인공지능을 활용한 위험브리핑 작성, 시스템 구현의 순서로 진행하였다(그림 1). 결과적으로 본 논문은 장기간 축적된 비정형 소방 사건정보를 기준일에 따라 변화하는 위험정보로 변환하고 이를 실제 사건 근거 및 브리핑으로 연결하는 동적 다차원 위험평가 프레임워크를 제시한다.

3-1 연구자료 및 데이터 구축

본 연구는 소방청 일일상황보고에 수록된 화재, 구조, 구급 및 생활안전 사건을 분석자료로 활용하였다. 일일상황보고는 전국에서 발생한 주요 사고의 발생일시, 장소, 사고내용, 인명

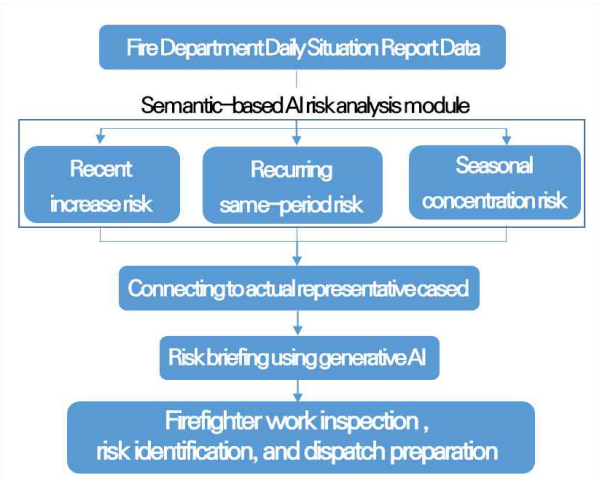


그림 1. 기준일 기반 소방 위험브리핑 생성 절차
Fig. 1. Procedure of generating fire risk briefings based on the reference data

피해, 재산피해, 구조활동 및 이송내용을 서술형으로 제공한다. 분석기간은 2024년 1월부터 2026년 6월 4일까지이며, 해당 기간의 일일상황보고 PDF는 소방안전 빅데이터 플랫폼을 활용하였다.

수집된 PDF는 보고일자와 사건내용을 기준으로 사건 단위 자료로 변환하였다. 먼저 보고서에 기재된 기준일을 확인하고 화재, 구조 및 구급, 생활안전 등으로 구분된 개별 사건을 추출하였다. 각 사건에서는 발생일시, 발생장소, 사건유형, 주요상황, 인명피해, 재산피해 및 대응내용을 추출하여 정리하였다. 이후 연도별 자료를 통합하고 중복 사건, 누락된 날짜, 문서 형식 차이 및 추출 오류를 정제하였다.

사건의 의미분석에는 사고유형과 주요 상황을 포함한 사건 설명문을 사용하였다. 사건 설명문은 보고서의 원문을 기초로 구성하였으며, 의미분석에 영향을 줄 수 있는 사건 구분어와 구분기호를 제거하고 연속된 공백을 삭제하였다. 이 과정에서 일일상황보고의 “화재”, “구조”, “구급”, “생활안전” 구분란의 텍스트는 삭제하고 주택, 공장, 교통사고 등 사건을 의미하는 사건텍스트를 추출하여 피해 및 대응 내용만을 추출하였다. 따라서 일일상황보고 875건에 대하여 최종 사건 분석자료는 5,770건으로 구성하였다. 표 1은 일일상황보고 사건 단위 전처리 기준 및 방법을 설명한다.

3-2 사건 의미분석 및 유사사건 분류

사건 서술문의 의미적 유사성을 분석하기 위해 Sentence-BERT를 적용하였다. 소방 사건은 동일한 사고유형이라도 보고서마다 표현이 다르게 기록될 수 있으므로, 특정 키워드의 일치 여부만으로 유사사건을 분류하기 어렵다. Sentence-BERT는 문장 전체의 의미를 고정된 차원의 벡터로 변환하므로 표현이 다른 사건 사이의 의미적 유사성을 분석하는 데 활용할 수 있다.

본 연구에서는 다국어 문장 임베딩 모델인 paraphrase-multilingual-MiniLM-L12-v2를 활용하여 각 사건 설명문을 384차원의 의미벡터로 변환하였다. 생성된 임베딩은 사건의 대상, 발생경위, 피해상황 등이 종합적으로 포함되어 있어 생성된 벡터는 이러한 내용을 반영한다. 이를 통해 공장과 창고 화재, 다중 차량 충돌, 산악사고 등과 같이 유사한 상황을 포함하는 사건을 의미적으로 비교하였다.

고차원 임베딩의 군집 특성을 효과적으로 분석하기 위해 Uniform Manifold Approximation and Projection을 적용하였다. 구체적인 파라미터는 n_neighbors=30, n_components=15, min_dist=0.0, metric=cosine으로 설정하였다. n_neighbors은 개별 사건 주변의 국소적 의미관계를 반영할 수 있도록 설정하였고, 384차원의 임베딩을 15차원으로 축소하여 의미적 유사관계를 유지하면서 계산효율과 해석가능성을 확보하였다.

HDBSCAN은 차원축소된 벡터를 대상으로 적용하여 밀도기반 군집화를 수행하였다. HDBSCAN은 군집 수를 사전에 지정할 필요가 없고 밀도가 낮아 특정 군집에 안정적으로 포함되기 어려운 사건을 노이즈로 처리할 수 있어 다양한 사고유형이 혼재된 사건자료에 적용이 용이하다. 최소 군집 크기인 min_cluster_size=20, 군집 형성을 위한 최소 이웃 수는 min_samples=8,로 설정하였으며, 군집 선택에는 EOM 방식을 적용하였다. Leaf 방식은 계층의 말단 군집을 중심으로 세분화된 군집 생성이 가능하나, 본 연구에서는 해석 가능한 규모의 군집을 확보하는 목적으로 EOM 방식을 선택하였다.

3-3 기준일 기반 위험정보 분석

위험신호 선정점수는 특정 지표의 절대적 크기로 편향되는 것을 고려하여 각 기준일의 후보 군집을 대상으로 평가항목별 Min-Max 정규화를 수행한 후 동일한 가중치로 평균하여 산정하였다. 계절 위험의 경우 과거 동일 계절의 합산 발생건수, 계절집중도, 두 연도 반복 여부, 평균 심각도를 평가항목으로 사용하였다. 각 항목에 대한 정규화 값은 다음과 같이 계산하였다.

$$z_j = \frac{x_j - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (1)$$

최종 선정점수 S는 사용 가능한 m개의 평가항목의 정규화 값을 동일하게 평균하여 산정하였다. 수식은 아래와 같다.

$$S = \frac{1}{m}\sum_{j=1}^{m} z_j \quad (2)$$

두 연도의 과거 자료가 모두 존재하는 경우 네 개의 평가항목을 사용하였으며, 한 연도의 자료만 존재하는 경우에는 두 연도 반복 여부를 제외하였다. 이를 통해 합산 발생건수와 같이 값의 범위가 큰 변수가 다른 평가항목을 과도하게 지배하는 영향을 완화하였다.

표 1. 일일상황보고의 사건 단위 전처리 기준 및 방법
Table 1. Criteria and methods for incident-level preprocessing of daily situation reports

Processing Item	Criteria and Method
PDF Text Extraction	Text was extracted from each page of the daily situation reports, and line breaks, redundant whitespace, and unnecessary formatting symbols were normalized.
Report Date Identification	Because the date in the PDF file name may differ from the date stated in the report, the date provided in the report body was preferentially used as the report date and as the basis for assigning the report ID. When the date could not be identified from the report body, the date in the file name was used as supplementary information.
Separation of Report Date and Activity Date	The report publication date was stored as report_date, while the date corresponding to the actual reporting period of incidents was stored separately as activity_date. When a report summarized incidents from the previous day, activity_date was assigned as the day preceding report_date.
Identification of Incident Sections	Incident sections, including fire, rescue and emergency medical services, daily safety activities, and major disasters, were identified to distinguish the areas containing individual incident descriptions.
Incident-Level Segmentation	Individual incidents were separated using markers indicating the beginning of each incident, and the text extending to the next incident or incident section was treated as a single incident.
Incident Description Preprocessing	The original incident description was preserved, while a separate text field was generated for semantic analysis. Incident-section labels, administrative prefixes, separators, and redundant whitespace at the beginning of the text were removed.
Preservation of Incident Information	Expressions describing the substantive characteristics of incidents, including residential facilities, factories, traffic accidents, occurrence conditions, damage, and response information, were retained for subsequent semantic analysis.
Review of Irregular Records	Preprocessing results such as empty incident descriptions, excessively short descriptions, and remaining incident-section expressions were separately identified and managed as records requiring review.

사건의 심각도는 사상자 유형별 상대적 영향을 반영하여 산정하였다. 사건 i에 대한 심각도 V_i는 사망자 수에 1.0의 가중치를, 중상자 수에 0.7의 가중치를, 경상자 수에 0.3의 가중치를 부여하여 평균 심각도를 계산하였다. 해당 가중치는 교통안전법 교통안전도평가지수 산출식의 가중치를 활용하여 산출하였다. 수식은 아래와 같다. 다만, 본 심각도 지표는 인명피해를 중심으로 산정되었으므로, 재산피해 규모가 큰 화재 등의 사고에 대한 의험특성을 충분히 반영하기 어렵다.

$$V_i = 1.0D_i + 0.7S_i + 0.3M_i \quad (3)$$

기준일 기반 위험정보 분석은 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험으로 구성하였다. 전체 분석기간의 사건을 의미기반으로 군집화한 후, 사용자가 입력한 기준일에 따라 분석기간을 설정하고 군집별 발생건수와 평균 심각도를 산출하였다. 위험유형별 선정기준은 발생건수, 시간적 변화 또는 반복성, 계절적 집중도를 중심으로 구성하였으며, 평균 심각도는 위험신호의 우선순위를 보완하는 변수로 활용하였다. 표 2는 분석기준을 나타낸다.

최근 증가 위험은 기준일을 포함한 최근 15일과 그 직전 15일의 사건 발생 변화를 이용하여 산출하였다. 최근 기간은 기준일 이전 14일부터 기준일까지이며, 직전 기간은 기준일 이전 29일부터 15일까지로 설정하였다. 각 군집에 대한 최근 기간 발생건수와 직전 기간 발생건수를 계산하고, 최근 기간에 3건 이상 발생하면서 직전 기간보다 1건 이상 증가한 군집을 위험 후보 군집으로 선정하였다. 후보군의 선정점수는 최근 기간 발생건수, 직전 기간 대비 증가건수, 최근 기간 사건의 평균 심각도를 평가 항목으로 구성하였다. 각 평가항목은 후보군 내에서 Min-Max 정규화한 후 평균으로 최종 선정점수를 산출하였으며, 점수가 가장 높은 군집을 최근 증가 위험으로 선정하였다. 예를 들어 특정 군집이 최근 15일간 8건, 직전 15일간 3건 발생했다면 최근 발생건수 8건, 증가건수 5건으로 평균 심각도를 각각 정규화하여 선정점수를 계산하는 것이다.

동일 시기 위험은 기준일과 동일한 시점을 중심으로 전년도와 전전년도의 사건 발생 특성을 비교하여 산출하였다. 각 과거 연도에 대해 기준일과 동일한 날짜를 중심으로 30일 간의 분석기간을 설정하여 군집별 발생 건수를 계산하였다. 선정점수는 두 연도의 합산 발생건수, 두 연도에서의 반복 발생 여부, 해당 기간 사건의 평균 심각도를 평가항목으로 구성하였다. 두 연도의 자료가 모두 존재하는 경우 반복 발생 여부를 포함하였으며, 한 개의 연도 자료만 이용이 가능한 경우 항목에서 제외하였다. 각 평가항목은 후보군 내에서 Min-Max 정규화한 후 평균값을 계산하여 최종 선정점수가 가장 높은 군집을 동일 시기 위험으로 도출하였다. 예를 들어 특정 군집 내 해당 기간 동안 연도마다 반복해서 발생하고 전년도 동일 시기에 16건, 전전년도에 13건 발생했다면 합산 발생건수는 29건과 평균 심각도를 정규화하여 선정한다.

계절 위험은 기준일이 포함된 계절과 동일한 과거 계절의 사건 발생 특성을 이용하여 산출하였다. 계절은 봄의 경우 3월부터 5월, 여름의 경우 6월부터 8월, 가을의 경우 9월부터 11월, 겨울의 경우 12월부터 다음 해 2월까지로 구분하였다. 전년도와 전전년도의 동일 계절을 대상으로 군집별 발생건수와 계절집중도를 계산하였으며, 선정점수는 동일 계절의 합산 발생건수, 계절집중도, 두 연도에서의 반복 발생 여부, 해당 계절 사건의 평균 심각도를 평가항목으로 구성하였다. 계쩔집중도는 해당 군집의 과거 전체 발생건수 중 동일 계절에 발생한 사건이 차지하는 비율로 산출하였다. 두 연도의 자료가 모두 존재하는 경우를 사용하고 각 평가항목은 후보군 내 Min-Max 정규화한 후 평균으로 산출하여 최종 선정점수가 높은 군집을 계절 위험으로 선정하였다. 예를 들어 특정 군집 내에 동일시기에 대하여 전년도 봄철 55건, 전전년도 봄철 37건이 발생했다면, 동일 계절에 대한 합산 발생건수는 92건이며 전체 발생 건수 중 봄철 사건이 차지하는 비율을 계절집중도로 계산하여 심각도와 함께 선정점수를 산출한다.

최종 위험신호는 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험에서 각각 선정점수가 가장 높은 군집 1개를 선택하였다. 각 군집에는 분석결과를 설명할 수 있는 실제 사건을 대표사건으로 연결하였다. 대표사건은 평균 심각도 점수, HDBSCAN 군집 소속 확률 및 발생일을 기준으로 정렬

표 2. 기준일 기반 위험정보의 분석기준
Table 2. Analysis criteria for risk information based on the reference data

Risk type	Analysis Period	Candidate Selection Criteria	Main Evaluation Factors	Final Selection
Recent Increase Risk	Recent 15 days including the reference data and the preceding 15 days	At least 3 cases in the recent period and an increase of at least 1 case compared with the preceding period	Recent case count, increase count, mean severity	Highest-scoring cluster
Recurring Same-Period Risk	30-day periods centered on the corresponding dates in the previous two years	At least 2 cases in total across the two years	Total case count, recurrence across both years, weighted mean severity	Highest-scoring cluster
Seasonal Concentration Risk	Same season in the previous two years	At least 2 cased in total during the corresponding seasons	Seasonal case count, seasonal concentration, recurrence across both years, weighted mean severity	Highest-scoring cluster

하였다. 심각도 점수가 높은 사건을 우선하였으며, 심각도 점수가 동일한 경우 군집 소속 확률이 높은 사건을 우선하였다. 두 값이 동일한 경우 최근 발생한 사건을 우선하였고, 동일한 사건 설명은 중복을 제거한 후 최대 5건까지 제시하였다.

위험신호 선정변수는 사건의 발생 규모, 시간적 변화와 반복성, 피해의 심각성을 함께 반영하도록 구성하였다. 발생건수는 해당 위험유형의 출현 규모를 나타내며, 증가건수와 연도 간 반복 여부는 시간에 따른 변화와 재현성을 평가한다. 계절 집중도는 특정 계절에 사건이 집중되는 정도를 나타낸다. 평균 심각도는 발생빈도가 낮더라도 피해 가능성이 큰 사건이 분석에서 제외되는 것을 방지하기 위한 보완변수로 활용하였다. 정리하자면, 기준일에 따라 발생규모, 단기 변화, 연도 간 반복성, 계절성 및 심각도를 종합적으로 평가하여 시점별로 위험신호를 도출하였다.

3-4 위험브리핑 생성 및 시스템 구현

본 연구는 통계분석을 통해 선정된 위험신호와 대표사건 정보를 소방공무원이 당일 상황판단에 활용할 수 있는 위험 인사이트와 대응 참고문장으로 변환하기 위해 생성형 인공지능을 적용하였다. 앞 단계에서 최근 증가 위험, 동일 시기 위험 및 계절 위험을 통계적으로 선정하고, 각 위험신호에 대한 발생특성과 대표사건을 확정한 후 이를 생성형 인공지능의 입력자료로 구성하였다. 생성형 인공지능은 위험신호 자체를 판단하거나 새로운 위험을 예측하는 역할을 수행하지 않으며, 확정된 통계정보와 대표사건의 근거 범위에서 사고 특성에 대한 인사이트, 당일 확인이 필요한 대응조치 및 출동대 공유사항을 생성하도록 하였다.

생성형 인공지능의 입력자료에는 위험신호 군집 내 대표사건을 입력한다. 대표사건은 해당 위험신호에서 실제로 확인된 사고 특성과 피해상황을 제시하기 위한 입력 데이터로 활용하였다. 생성 결과는 위험 인사이트, 당일 대응조치 및 출동대 공유사항으로 구성하였다. 위험 인사이트는 대표사건에서 확인된 사고 특성과 현장 위험요인을 요약하고, 대응조치는 해당 근거를 바탕으로 소방공무원이 사전에 점검하거나 확인할 사항을 제시하며, 출동대 공유사항은 출동 전에 공유할 필요가 있는 주요 사건 특성을 제시하도록 하였다.

대표사건 데이터는 상위 위험유형과 세부 근거태그의 두 단계로 구성하였다. 먼저 각 사건의 사건유형, 장소유형 및 사건 서술정보를 이용하여 하나의 상위 위험유형으로 부여하였다(표 3). 화재사건은 장소유형을 우선하여 주택화재, 공장 및 창고화재, 차량화재, 산불 및 임야화재 등으로 분류하였으며, 비화재사건은 교통사고, 추락 및 끼임, 붕괴 및 매몰, 수난 및 익수 등 사건 서술문에서 확인되는 사고 특성을 이용하여 분류하였다.

대표사건의 서술문에서는 사전에 정의한 주요 표현의 포함 여부를 확인하여 복수의 세부 근거태그를 부여하였다. 세부 근거태그는 사고유형, 충돌, 인명피해, 연소확대, 폭발, 화학물질, 대피, 병원이송 및 현장 접근과 같은 사건의 세부 특성을 나타내도록 구성하였다. 예를 들어 대표사건에 '전소', '중상', '병원이송'이 포함된 경우 각각 fire_spread, casualty, medical_transport 태그를 부여하였다. 하나의 대표사건에는 복수의 세부 근거태그를 동시에 부여하도록 하였으며, 세부 근거태그는 표 4와 같다.

표 3. 상위 위험유형의 분류 기준
Table 3. Classification criteria for major risk types

Major Risk Type	Key Classification Criteria
Residential Fire	house, apartment building, apartment, multi-family house, multiplex house, row house
Factory and Warehouse Fire	factory, warehouse industrial facility, workplace, junkyard, waste treatment facility, livestock facility, etc.
Vehicle Fire	vehicle, automobile, bus, truck, passenger car, cargo vehicle
Forest and Wildland Fire	wildand, forest fire, outdoor area, grass fire, forest
Commercial and Office Facility Fire	neighborhood living facility, commercial building, restaurant, lodging facility, market, office, etc.
Multi-use and Public Facility Fire	education facility, school, medical facility, hospital, welfare facility, religious facility, cultural facility, public facility, etc.
Collapse and Burial	collapse, burial, falling, structure, entrapment under debris
Fall and Entrapment	fall, entrapment, pinching, overturning, machinery accident, amputation
Traffic Accident	traffic accident, rear-end collision, collision, vehicle-to-vehicle, vehicle-to-pedestrian, etc.
Hazardous Material and Explosion	hazardous material, hazardous chemical, gas leak, hydrochloric acid, sulfuric acid, ammonia, explosion, etc.
Water Rescue and Drowning	water rescue, drowning, river valley, reservoir, offshore area, sea, rapid current
Mountain Accident	mountain accident, hiking, mountaineering trail, rock face
Facility safety Accident	elevator, lift, facility, manhole, ventilation opening, etc.

생성형 인공지능이 작성한 문장은 대표사건에서 확인된 정보의 범위를 벗어나지 않도록 검증하였다. 각 대표사건에는 상위 위험유형에 해당하는 기본 태그와 사건 서술문에서 추출한 세부 근거태그를 부여하였으며, 생성문에서도 동일한 기준으로 관련 태그를 다시 추출하였다. 생성문에서 확인된 태그가 해당 대표사건의 근거태그에 모두 포함되는 경우에만 문장을 채택하였고, 근거에서 확인되지 않은 위험요인이나 피해특성이 포함된 문장은 제외하였다. 또한 정량정보는 앞 단계의 통계분석 결과만 사용하도록 하여 생성형 인공지능이 확인된 근거 범위 안에서 위험 인사이트와 대응 참고정보를 작성하도록 제한하였다.

결과적으로 위험브리핑 시스템은 사용자가 기준일을 입력하면 해당 날짜를 중심으로 최근 증가 위험, 동일 시기 위험 및

표 4. 세부 근거태그의 분류 기준
Table 4. Classification criteria for detailed evidence tags

Detailed Evidence Tag	Classification Keywords
vehicle_fire	vehicle fire, engine compartment, automobile fire
factory_fire	factory, warehouse, factory fire, warehouse fire
residential_fire	house, apartment building, lodging facility, residential
traffic_accident	traffic accident, rear-end collision, collision
water_rescue	water rescue, drowning, offshore, vessel, beach
collision	collision, rear-end collision, impact
multi_collision	multiple collision, multi-vehicle collision, chain collision
pedestrian	pedestrian
slippery_road	snowy road, icy road, wet road, slippery
fire_spread	fire spread, partial burn, complete burn, destruction by fire
electrical	electrical, auxiliary battery, short circuit, battery
machinery	mechanical, machinery, equipment
welding	welding, hot work
dust	dust
flammable	flammable material, oil, fuel
explosion	explosion
hazardous_material	chemical substance, hazardous material, harmful substance, leakage
cargo	loading, cargo, wing-body truck, tanker truck
tunnel	tunnel
marine_isolation	isolation, loss of direction
vessel_overturn	overturning, flooding
casualty	death, serious injury, minor injury, injury, cardiac arrest, casualty
mass_casualty	multiple casualties, casualties, multiple injuries, missing persons
evacuation	evacuation, escape
medical_transport	first aid, hospital transport, transport
access_constraint	entry, access, passage
secondary_accident	secondary accident, subsequent accident, chain accident
structural_collapse	collapse, structural damage

계절 위험을 분석하고 결과를 화면에 제시하도록 구현하였다. 시스템 화면에서는 위험유형별 분석결과와 근거 수치, 대표사건, 위험 인사이트, 대응조치 및 출동대 공유사항을 확인할 수 있도록 구성하였다. 또한 제한된 생성문에 연결된 대표사건의 발생일, 위험유형 및 사건내용을 함께 확인할 수 있도록 하여 생성결과의 근거를 사용자가 직접 검토할 수 있도록 하였다.

Ⅳ. 분석결과

4-1 사건 의미 군집화 결과

전체 5,770건의 사건자료를 Sentence-BERT를 활용하여 384차원으로 임베딩하고, UMAP을 적용하여 15차원으로 축소한 후 HDBSCAN을 통해 의미적으로 유사한 사건을 군집화하였다. HDBSCAN의 최소 군집 크기는 20, min_sapples는 8로 설정하였고 군집 선택에는 EOM 방식을 적용하였다. 분석 결과 총 37개의 클러스터가 도출되었으며(표 5), 전체 5,770건 중 군집에 배정된 사건은 4,504건이며 노이즈는 1,266건이다.

그림 2는 전체 사건자료의 의미 기반 군집화 결과를 나타낸다. 각 점은 개별사건이며 동일한 번호로 표시된 점들은 의미적으로 유사한 그룹이다. 차량화재, 교통사고, 공망 및 창고화재 등 주요 사고 유형이 구분되었다. 특히 일부 군집은 독립적인 영역에 밀집되었으며, 유사한 사고특성을 포함하는 군집은 상대적으로 인접한 위치에 분포한다.

특히 군집 4에는 전체 사건 중 878건이 포함되어 있어 해당 군집을 대상으로 추가적인 UMAP 및 HDBSCAN 분석을 수행하였다. 군집 4에 대한 세부 군집 설정은 UMAP을 이용한 10차원 공간에서 최소 군집 크기는 동일하게 20으로 설정하였으며, min_samples는 5를 적용하고 군집 선택은 동일하게 EOM을 통해 선택하였다.

재군집화 결과 15개의 세부군집이 생성되었으며, 이 중 685건이 세부군집으로 분류되고 193건은 노이즈로 구분되었다(그림 3). 세부군집의 평균 군집 소속 강도는 0.925로 나타났으며, 0.5 미만인 사건은 확인되지 않았다.

재군집화된 세부군집에서는 수난 및 익수사고, 산악사고, 붕괴 및 매몰사고, 추락 및 끼임사고, 공장 및 창고화재 등 주요 사고특성을 중심으로 사건집단이 구분되었다. 이는 동일하거나 유사한 사고특성이 서로 다른 표현으로 기록된 경우에도 의미적 유사성을 기준으로 동일한 분석단위에 포함될 수 있음을 보여준다. 또한 대규모 혼합군집을 추가로 세분화함으로써 이후 군집별 발생빈도, 시간적 변화 및 심각도를 비교할 수 있는 분석단위를 구축하였다. 이러한 결과는 장기간 축적된 일일상황보고의 비정형 사건정보를 기준일 기반 위험분석에 활용할 수 있는 형태로 변환하였다는 점에서 의미가 있다.

4-2 기준일 기반 위험정보 분석결과

기준일 기반 위험정보 분석 결과를 확인하기 위해 서로 다른 시점과 계절을 포함하는 복수의 기준일에 대해 분석을 수행하였다. 기준일별로 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 집중 위험을 각각 산출하였으며, 이를 통해 동일한 사건자료에서도 기준일에 따라 서로 다른 사고유형과 시간적 위험 특성이 선정되는지를 확인하였다.

표 5. 전체 군집별 사고 특성
Table 5. Accident characteristics by total clusters

Cluster No.	Count	Interpretation of Major Accident Characteristics
0	62	Livestock facility fire
1	342	Vehicle fires
2	22	Tree-felling and crushing accidents
3	42	Mixed small-scale fires with unknown or negligent causes
4	878	Mixed rescue accidents including falls, water accidents, collapses, and facility safety incidents
5	118	Bus-related traffic collisions involving multiple casualties
6	421	General traffic accidents and multi-vehicle collisions
7	55	Railway and train-related collision accidents
8	132	Battery-related fires in residential, industrial, and transportation facilities
9	35	Vessel and fishing-boat fires
10	50	Refrigerator and electrical appliance fires
11	211	Fires caused by cooking and negligence
12	21	Kitchen fires in residential and commercial facilities
13	27	Restroom and ventilation-related electrical fires
14	44	Electrical fires in factories
15	314	Electrical fires in residential and commercial facilities
16	25	Fires in schools and educational facilities
17	370	Residential fires involving casualties and extensive preperty damage
18	412	Factory and warehouse fires
19	73	Fires caused by cigarettes and smoking-related negligence
20	24	LPG leakage, explosion, and related fires
21	27	Fires involving portable gas stoves and gas appliances
22	43	Gas leakage and hazardous gas exposure accidents
23	40	Hazardous chemical leakage and spill accidents
24	51	Fires in waste storage and recycling facilities
25	24	Warehouse and storage facility fires
26	39	Industrial and commercial fires involving multiple buildings or fire spread
27	104	Forest fires
28	30	Fires caused by waste burning with spread to wildland or nearby structures
29	126	Mixed fires with undetermined causes or ongoing
30	79	Detached residential house fires
31	34	Construction and welding-related fires
32	81	Indoor residential fires involving rooms and household contents
33	25	Fires in residential rooms and lodging facilities
34	44	Apartment and residential unit fires
35	23	Fires in multi-story residential and mixed-use buildings
36	56	General residential fires

최근 증가 위험은 기준일에 따라 서로 다른 사건군이 선정되었다. 2025년 4월 25일에는 상업 및 업무시설 화재 중심 사건군이 최근 15일간 4건 발생하여 직전 15일의 2건보다 2건 증가하였다(그림 4). 표 6의 2025년 7월 30일에는 수난 및 익수사고 중심 사건군이 최근 15일간 4건 발생한 반면에 직전 15일에는 발생하지 않아 여름철 수난사고의 단기 증가가 확인되었다.

또한 표 6의 2025년 11월 30일 브리핑 생성결과는 다음과 같다. 생성 결과는 차량화재 중심 사건군이 최근 15일간 9건으로 직전 기간의 4건보다 5건 증가한 것으로 나타났다. 2026년 1월 31일에는 주택화재 중심 사건군이 최근 15일간 12건으로 직전 기간의 5건보다 7건 증가하였다(표 6). 이러한 결과는 최근 증가 위험이 특정 사고유형에 고정되지 않고 기준일의 실제 발생 변화에 따라 서로 다른 사건군을 선택함을 보여준다.

과거 동일 시기 반복 위험에서 2026년 1월 31일은 교통사

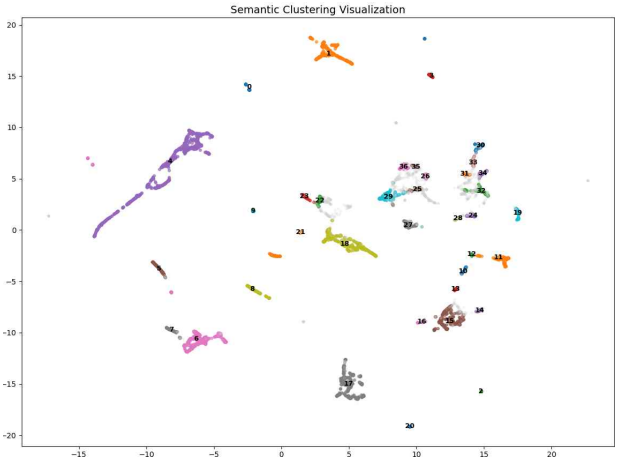


그림 2. 전체 사건자료의 의미 기반 군집화 결과
Fig. 2. Semantic clustering results of the entire incident dataset

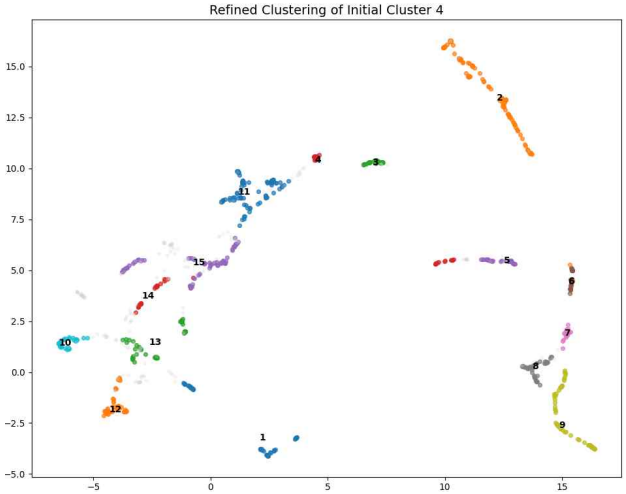
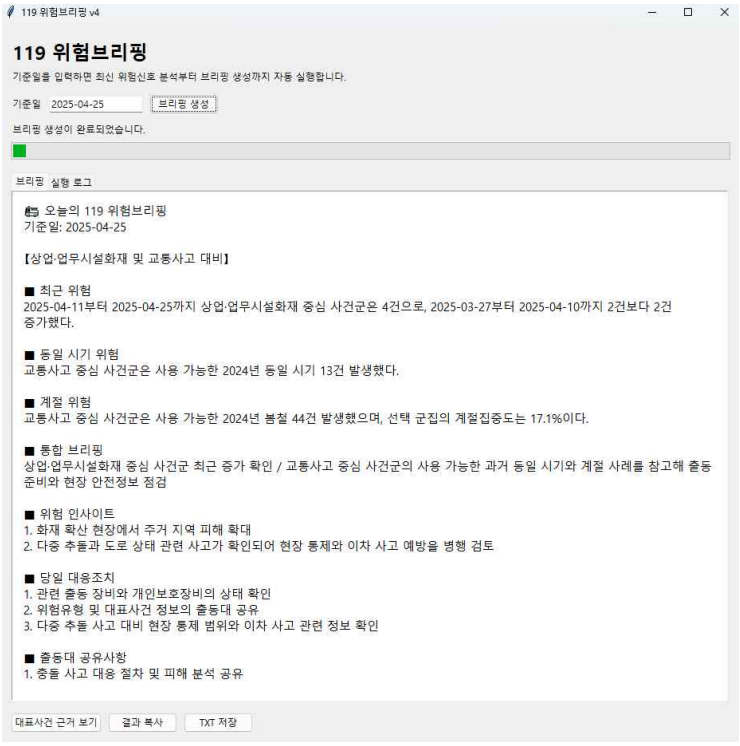


그림 3. 군집 4의 세부 군집 시각화
Fig. 3. Visualization of refined subclusters within cluster 4



*The above figure presents the interface in Korean, reflecting the original language used in the developed application

그림 4. 2025년 4월 25일의 위험브리핑 생성결과

Fig. 4. Risk briefing generated for the reference date of April 25, 2025

표 6. 기준일별 위험 군집 및 브리핑 요약

Table 6. Summary of risk clusters and briefings by reference date

Reference Date	Recent Increase Risk	Recurring Same-Period Risk	Seasonal Concentration Risk	Risk insight
2025.7.30.	수난 및 익수 중심 사건군 최근 15일 4건(7.16.~7.30.) 직전 15일 0건(7.1.~7.15.)	교통사고 중심 사건군 2024년 9건 발생	교통사고 중심 사건군 2024년 기준 여름 71건 계절집중도 27.6%	다중 추돌과 도로 상태 관련 사고가 확인되어 현장 통제와 이차 사고 예방을 병행 검토
2025.11.30.	차량화재 중심 사건군 최근 15일 9건(11.16.~11.30.) 직전 15일 4건(11.1.~11.15.)	교통사고 중심 사건군 2024년 19건 발생	교통사고 중심 사건군 2024년 가을철 78건 발생 계절집중도 31.3%	1. 교통사고 사례에서 다수 사상자와 현장 이송이 확인되어 응급처치와 병원이송 관련 정보를 병행 확인 2. 충돌 이후 화재와 주변 차량 또는 시설물로의 연소 확대가 함께 나타날 수 있어 교통 통제와 화재 대응을 병행 검토
2026.1.31.	주택화재 중심 사건군 최근 15일 12건(1.17.~1.31.) 직전 15일 5건(1.2.~1.16.)	교통사고 중심 사건군 2024년 1건 2025년 16건 발생	주택화재 중심 사건군 2025년 겨울철 77건 발생 계절집중도 37.2%	교통사고 사례에서 다수 사상자와 현장 이송이 확인되어 응급처치와 병원이송 관련 정보를 병행 확인 주택화재 사례에서 인명피해와 전소가 확인되어 피난과 인명검색 관련 정보를 우선 확인
2026.4.25.	공장 및 창고화재 중심 사건군 최근 15일 11건(4.11.~4.25.) 직전 15일 6건(3.27.~4.10.)	주택화재 중심 사건군 2024년 13건 2025년 16건	주택화재 중심 사건군 2024년 봄철 37건 2025년 봄철 55건 계절집중도 26.7%	공장과 창고 화재에서 연소 확대 사례가 확인되어 주변 시설로의 확산 가능성과 현장 진입 여건을 병행 확인 교통사고 사례에서 다수 사상자와 현장 이송이 확인되어 응급처치와 병원이송 관련 정보를 병행 확인

*The above figure presents the interface in Korean, reflecting the original language used in the developed application

고 중심 사건군이 2025년 동일 시기 16건, 2024년 동일 시기 1건 발생한 것으로 나타났다(표 6). 2026년 4월 25일에는 각각 16건과 13건이 발생하였다(표 6). 특히 2026년 4월 25일 사례에서는 동일한 교통사고 사건군이 두 연도 모두에서 다수 확인되어 동일 시기 위험에서 반복 발생 특성을 보다 확인할 수 있음을 나타낸다.

계절 집중 위험에서도 기준일이 속한 계절에 따라 서로 다른 발생 특성이 나타났다. 2025년 4월 25일에는 교통사고 중심 사건군이 2024년 봄철 44건 발생하였으며 계절집중도는 17.1%로 나타났다(그림 4). 2025년 7월 30일에는 교통사고 중심 사건군이 2024년 여름철 71건 발생하여 계절집중도 27.6%를 보였다.(표 6). 2025년 11월 30일에는 가을철 78건, 계절집중도 31.3%로 나타났다(표 6). 2026년 1월 31일에는 주택화재 중심 사건군이 77건 발생하였으며 계절집중도는 37.2%로 확인되었다(표 6). 계절집중도의 경우 기준일에 따라 31.3%와 37.2%와 같은 상대적으로 높은 계절 집중을 나타내는 사례를 확인할 수 있었다.

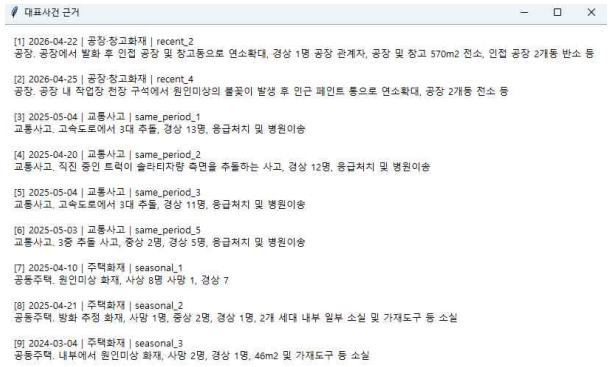
복수 기준일 분석 결과는 세 가지 위험 유형이 서로 다른 시간적 특성을 반영하고 있음을 확인하였다. 최근 증가 위험은 기준일 직전의 단기적인 발생 변화를 민감하게 반영하였으며, 동일 시기 위험은 과거 동일 시점에서의 발생 규모와 연도 간 반복성을 나타냈다. 계절 위험은 동일 계절의 발생 규모와 전체 발생 중 해당 계절이 차지하는 비중을 함께 반영하였다. 또한 2025년과 같이 과거 자료가 충분하지 않은 기준일에서는 활용 가능한 연도만을 이용하여 결과를 산출하고, 2026년 이후에는 두 연도의 자료를 함께 비교할 수 있었다. 이러한 결과는 본 분석방법이 특정 기준일에 한정되지 않고 데이터 가용범위에 따라 서로 다른 시점과 계절의 위험특성을 도출할 수 있다.

4-3 위험브리핑 생성 및 시스템 적용결과

기준일 기반 위험분석 결과는 동적 다차원 위험평가 프레임워크를 적용한 의사결정지원 정보로 변환하여 소방공무원이 일일 업무에서 활용할 수 있도록 위험브리핑 형태로 변환하였다. 시스템은 사용자가 기준일을 입력하면 최근 증가 위험, 과거 동일 시기 위험 및 계절 위험을 자동으로 분석하고, 선정된 사건군의 통계정보와 대표사건을 생성형 언어모델의 입력자료로 구성한다. 이후 Qwen3:8b 모델을 이용하여 위험 인사이트, 당일 대응조치 및 출동대 공유사항을 생성하도록 구현하였다.

생성된 브리핑은 통계 결과와 실제 사건 근거를 함께 확인할 수 있도록 구성하였다. 브리핑 화면에는 기준일, 위험유형별 분석결과, 통합 브리핑, 위험 인사이트, 당일 대응조치 및 출동대 공유사항을 순차적으로 제시하였다. 그림 4와 표 6은 서로 다른 기준일에 대해 생성된 위험브리핑의 적용 예시를 나타낸다. 기준일에 따라 선정되는 최근 위험과 과거 동일 시기 및 계절 위험이 달라지며, 이에 따라 생성되는 인사이트와 대응 정보 역시 서로 다르게 구성되었다. 이를 통해 하나의 고정된 결과를 제공하는 방식이 아니라 입력된 기준일의 분석 결과에 따라 브리핑 내용이 달라지는 것을 확인하였다.

각 브리핑 문장은 실제 대표 사건과 연결하여 근거를 확인할 수 있도록 하였다. 사용자는 '대표 사건 근거 보기' 기능을 통해 생성된 인사이트와 대응조치에 활용된 사건의 발생일, 위험 유형 및 주요 사고내용을 확인할 수 있다. 그림 5는 2026년 4월 25일 기준 연결된 대표 사건의 근거 확인 화면을 나타낸다. 예를 들어 교통사고 위험이 근거로 선정된 경우 브리핑에서는 다중 추돌, 다수 사상자 발생, 응급처치 및 병원 이송 사례가 인사이트로 생성되었으며, 공장 및 창고 화재가 근거로 선정된 경우에서는 연소확대가 포함되었으며, 이에 따라 위험 인사이트는 연소 확대가 제시되었다. 이와 같이 브리핑의 생성 결과를 원자료의 실제 사건과 연결함으로써 사용자가 생성 문장의 근거를 직접 확인할 수 있도록 하였다.



*The above figure presents the interface in Korean, reflecting the original language used in the developed application

그림 5. 2026년 4월 25일의 위험브리핑의 대표사건 근거 인터페이스

Fig. 5. Representative incident evidence interface for the risk briefing of April 25, 2026

Ⅴ. 결 론

본 연구는 소방청 일일상황보고에 축적된 비정형 사건정보를 활용하여 소방공무원의 도시 안전관리 의사결정을 지원하는 위험브리핑 시스템을 개발하였다. 2024년 1월부터 2026년 6월까지의 일일상황보고 875건에서 5,770건의 사건자료를 구축하고, Sentence-BERT, UMAP 및 HDBSCAN을 적용하여 의미적으로 유사한 사건군을 분류하였다. 이후 사용자가 입력한 기준일을 중심으로 최근 증가 위험, 과거 동일 시기 반복 위험 및 계절 위험을 분석하고, 선정된 사건군을 실제 대표사건과 연결하여 일일 위험브리핑으로 제공하였다.

위험신호 선정점수는 발생건수, 시간적 변화, 연도 간 반복성, 계절집중도 및 사건 심각도를 반영하도록 구성하였다. 각 평가항목은 후보군 내에서 Min-Max 정규화한 후 동일한 비중으로 평균하여 선정점수를 산출하였다. 사건 심각도는 사망, 중상, 경상에 각각 1.0, 0.7, 0.3의 상대 가중치를 적용하여 산정하였다. 이를 통해 서로 다른 척도의 평가요인을 동일한 기준에서 비교할 수 있도록 하였다.

본 연구에서는 복수의 기준일을 입력하여 사례를 검토하였다. 최근 증가 위험은 시점별 단기 발생변화를 반영하여 상업 및 업무시설 화재, 수난 및 익수사고, 차량화재, 주택화재, 공

장 및 창고화재 등 서로 다른 사건군을 선정하였다. 과거 동일 시기 위험에서는 자료 축적 수준에 따라 전년도 또는 전년도와 전전년도의 발생특성을 비교할 수 있었으며, 두 연도 자료가 모두 존재하는 경우 반복 발생 여부를 함께 평가하였다. 계절 위험에서는 봄, 여름, 가을, 겨울의 기준일을 각각 적용하여 계절별 발생 건수와 계절집중도의 차이가 나타나는 부분을 확인하였다. 이러한 결과는 기준일에 따라 단기 변화, 연도 간 반복성 및 계절적 특성이 서로 다른 위험 정보로 도출될 수 있음을 나타낸다.

본 연구는 장기간 축적된 비정형 소방 사건정보를 의미적으로 분류하고, 기준일에 따른 단기 변화, 연도 간 반복성 및 계절적 특성을 반영한 동적 다차원 위험평가 프레임워크를 제시하였다는 점에 의의가 있다. 또한 분석 결과와 생성형 인공지능의 역할을 분리하여 위험 판단은 통계분석에 의해 수행하고, 생성형 인공지능은 분석 결과를 현장 활용이 가능한 문장으로 변환하도록 제한하였다. 이를 통해 소방 사건정보의 활용 범위를 사고 특성 분석과 사건 분류에서 일일 상황판단, 대응 준비 및 정보 공유를 지원하는 방향으로 확장하였다.

본 연구는 분석기간, 사건 심각도 산정 및 현장검증 측면에서 향후 보완이 필요하다. 현재 자료는 약 2년 6개월의 기간을 대상으로 하므로 일부 기준일에서는 전년도와 전전년도의 동일 시기 또는 동일 계절 자료를 모두 활용하기 어렵다. 따라서 장기간의 일일상황보고를 추가로 입력하여 연도 간 반복성과 계절적 특성을 보다 안정적으로 평가할 필요가 있다. 또한 사건 심각도는 교통안전법 교통안전도평가지수를 활용하여 교통안전 분야의 공식 평가 체계에서 활용되는 사상자별 상대 가중치를 참고하여 사망 1.0, 중상 0.7, 경상 0.3으로 산정하였으나, 해당 가중치가 화재, 구조, 구급 및 생활안전 사건의 피해 특성을 충분히 반영하는지에 대해서는 실제 사용자 평가를 통한 추가적인 검토가 필요하다. 또한 해당 심각도 지표는 인명피해를 중심으로 산정되어, 재산피해 규모와 시설 특성 등을 함께 고려하는 복합 심각도 지표로 고도화할 필요가 있다. 향후에는 분석기간과 최소 발생건수 등의 기준을 사용자에 따라 설정가능한 변수로 확장하여 기관별, 업무별 의사결정 목적에 맞게 적용할 수 있도록 시스템을 고도화할 필요가 있다.

감사의 글

본 연구는 행정안전부 국립재난안전연구원의 지원(“도시단위 대규모 재난대피 시뮬레이션 기술 개발”, “MDMI-주요-2026-05-01”)에 의해 수행되었습니다. 이에 감사드립니다.

참고문헌

[1] S. Han, “A Comparative Study on the Safe Community Management System and Its Application,” *Journal of Regional Studies and Development*, Vol. 23, No. 2, pp. 465-493, 2014.

[2] J.-Y. Lee, M.-O. Choi, and I.-S. Moon, “Analysis of the Characteristics of the Urban Disaster Management Process - Fire Administration in Incheon Metropolitan City,” *Journal of Regional Studies*, Vol. 18, No. 1, pp. 103-125, 2010.

[3] S. Kanbara and R. Shaw, “Disaster Risk Reduction Regime in Japan: An Analysis in the Perspective of Open Data, Open Governance,” *Sustainability*, Vol. 14, No. 1, 19, 2021.

[4] S. He, X. Shu, J. Hu, L. Zhang, J. Zhang, J. Zhang, and Y., Zhou, “Prediction and Early-Warning Method of Electrical Fire Risk Based on Fire-Fighting Big Data,” *Journal of Tsinghua University (Science and Technology)*, Vol. 64, No. 3, pp. 478-491, 2024.

[5] T. Peng and W. Ke, “Urban Fire Emergency Management Based on Big Data Intelligent Processing System and Internet of Things,” *Optik*, Vol. 273, 170433, 2023.

[6] Sahebi, A., Sayyadi, H., Havasy, B., and Veisani, Y. “Predicting Firefighting Operation Time in Urban Areas Using Machine Learning: Identifying Key Determinants for Improved Emergency Response,” *Discover Applied Sciences*, Vol. 7, No. 4, 250, 2025.

[7] Z. Abohaia, A. Elkhouly, M. E. Barachi, and O. Al-Khatib, “Regional Prediction of Fire Characteristics Using Machine Learning in Australia,” *Fire*, Vol. 8, No. 8, 330, 2025.

[8] Y. Guo, R. Jena, D. Hughes, M. Lewis, and K. Sycara, K. “Transfer Learning for Human Navigation and Triage Strategies Prediction in a Simulated Urban Search and Rescue Task,” in *Proceedings of the 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN),* Vancouver, Canada, pp. 784-791, 2021.

[9] E.-H. Hwang, J. Choi, and D. Choi, “A Study on the Effective Methods of Securing the Golden Time of Fire Engine Move Out,” *Journal of the Korean Society of Hazard Mitigation*, Vol. 18, No. 5, pp. 119-126, 2018.

[10] M. M. Mirończuk, “Information Extraction System for Transforming Unstructured Text Data in Fire Reports into Structured Forms: A Polish Case Study,” *Fire Technology*, Vol. 56, No. 2, pp. 545-581, 2020.

[11] T. Buffington and O. A. Ezekoye, “Statistical Analysis of Fire Department Response Times and Effects on Fire Outcomes in the United States, *Fire Technology*, Vol. 55, No. 6, pp. 2369-2393, 2019.

[12] F. F. Balahadia, B. G. Dadiz, R. R. Ramirez, M. Luvett, J. A. P. Lalata, and A. C. Lagman, “Application of Data Mining Approach for Profiling Fire Incidents Reports of

Bureau of Fire and Protection," in *Proceedings of the International Conference on Computational Intelligence and Knowledge Economy (ICCIKE),* Dubai, United Arab Emirates, pp. 713-717, 2019.

[13] Z. Peng, B. Xiao, Y. Yao, J. Guan, and F. Yang, "U-Safety: Urban Safety Analysis in a Smart City," in *Proceedings of the IEEE International Conference on Communications (ICC),* Paris, France, pp. 1-6, 2017.

[14] A. Meijer and M. Thaens, "Quantified Street: Smart Governance of Urban Safety," *Information Polity*, Vol. 23, No. 1, pp. 29-41, 2018.

[15] J. Shi, J. Li, H. Zhang, B. Xie, Z. Xie, Q. Yu, and J. Yan, "Real-Time Gas Explosion Prediction at Urban Scale by GIS and Graph Neural Network," *Applied Energy*, Vol. 377, 124614, 2025.

[16] A. Doğan, M. Başeğmez, and C. C. Aydın, "Assessment of the Seismic Vulnerability in an Urban Area with the Integration of Machine Learning Methods and GIS," *Natural Hazards*, Vol. 121, No. 8, pp. 9613-9652, 2025.

[17] J. Lee, "Development of an AI-Based Algorithm for Early-Stage Fire Severity and Firefighting Resource Prediction," in *Proceedings of the 2025 KIIT Fall Conference,* Jeju, pp. 726-730, 2025.

[18] D. Im, C. Kim, and W. Kang, "Predicting Fire Risk Areas and Optimizing Firefighting Resources in Seoul," *Journal of the Korean Cadastre Information Association*, Vol. 26, No. 1, pp. 73-91, 2024.

[19] J.-I. Byun and I.-S. Park, "A Study on the Effect of Road Requirements at Fire Point on Fire Response: Based on the Case of Seoul Metropolitan Government," *Journal of the Korean Society of Disaster Information*, Vol. 22, No. 1, pp. 205-222, 2026.

[20] S. H. Jun and S. H. Myeong, "The Analysis of Safety Services in Local Smart City Plan: Focusing on Topic Modeling Analysis," *Journal of Korean Association for Regional Information Society*, Vol. 26, No. 1, pp. 1-35, 2023.

[21] E. Kwon, H. Park, S. Byon, and K. Lee, "Improvement of the Classification Model Performance in 119-Emergency Report Data," *Journal of KIISE*, Vol. 49, No. 1, pp. 89-96, 2022.



김지호(Jiho Kim)

2022년 : 중앙대학교 대학원 (공학석사)

2022년～현 재: 국립재난안전연구원
2023년～2025년: 부산대학교 정보융합공학과 컴퓨터공학전공 박사수료
※관심분야 : GIS(Geographic Information System), 그래프 이론(Graph Theory), 텍스트 마이닝(Text mining), 재난대피(Disaster Evacuation), 알고리즘(Algorithm) 등



심준석(Junseok Sim)

2011년 : 부산대학교 대학원(공학석사)
2018년 : 부산대학교 지형정보협동과정 박사수료

2018년～현 재: 국립재난안전연구원
※관심분야 : GIS(Geographic Information System), 재난대피(Disaster Evacuation), 국가재난관리(National Disaster Management), 재난 안전 정책(Disaster and Safety Policy)



임언택(Eontaek Lim)

2020년 : 국립한밭대학교 대학원 (공학석사)

2021년～현 재: 국립재난안전연구원
※관심분야 : 재난관리(Disaster Management), 재난대피(Disaster Evacuation), 공간정보(Geospatial Information), 드론맵핑(Drone Mapping) 등



정민수(Minsu Jung)

2008년 : Kobe University (공학박사)

2012년～현 재: 국립재난안전연구원
2011년～2012년: 한국건설기술연구원
2008년～2011년: Kobe University 학술추진연구원
※관심분야 : 재난관리(Disaster Management), 재난대피(Disaster Evacuation) 등