Towards Explainable Hybrid Similarity for Cybersecurity Incident Retrieval

Marc Krüger[1]

[1]Stiftung Universität Hildesheim, Universitätsplatz 1, 31141 Hildesheim, Germany

Abstract

This doctoral research investigates explainable hybrid similarity for cybersecurity incident retrieval within a Case-Based Reasoning (CBR) framework. Cybersecurity analysts increasingly require retrieval systems that not only identify similar incidents but also provide transparent explanations of retrieval decisions. The proposed framework combines four complementary similarity perspectives: transformer-based semantic similarity, TF-IDF lexical similarity, structured feature similarity, and ontology-guided similarity reasoning. The research addresses four challenges: the integration of heterogeneous similarity methods, the improvement of retrieval explainability, the scalability of retrieval for large incident repositories, and the transferability of results from synthetic to real-world cybersecurity datasets. A preliminary evaluation on a synthetic repository containing approximately 100,000 cybersecurity incidents indicates that ontology-guided similarity improves retrieval quality while maintaining efficient retrieval performance. Future work will investigate adaptive similarity weighting, expert-centered explainability evaluation, and validation using real-world cybersecurity incident datasets.

Keywords

Case-Based Reasoning, Cybersecurity, Explainable AI, Information Retrieval, Hybrid Similarity

1. Introduction

Cybersecurity analysts increasingly face the challenge of investigating large volumes of incident reports, alerts, and threat intelligence data. Effective incident analysis often requires identifying previously observed incidents that share similar attack characteristics, affected systems, or operational contexts. Retrieval-based decision support is therefore becoming an important capability in modern cybersecurity operations.

Case-Based Reasoning (CBR) provides a natural paradigm for this task because new incidents can be analyzed by retrieving and reusing knowledge from previously observed cases [1]. However, cybersecurity incident retrieval remains challenging due to the heterogeneous nature of cyber incident information. Incidents typically contain structured attributes, textual descriptions, attack indicators, and conceptual relationships that must be considered simultaneously.

Existing retrieval approaches often focus either on lexical similarity measures, symbolic reasoning techniques, or neural embedding models. While embedding-based approaches capture semantic relationships between incidents, they often provide limited transparency regarding retrieval decisions. Conversely, symbolic approaches can offer interpretable reasoning but may fail to capture broader contextual similarities between incidents.

This research investigates how multiple similarity perspectives can be integrated within a unified retrieval framework for cybersecurity incident analysis. The proposed approach combines transformer-based semantic similarity, TF-IDF lexical similarity, structured feature similarity, and ontology-guided reasoning within an explainable Case-Based Reasoning framework. The objective is to support accurate, scalable, and transparent retrieval of cybersecurity incidents while enabling analysts to understand the factors that contribute to retrieval results.

ICCBR DC'26: Doctoral Consortium at ICCBR2026, August 16, 2026, Bremen, Germany
krueger.hannover@t-online.de (M. Krüger)
0009-0008-6943-4938 (M. Krüger)

2. Related Work

Case-Based Reasoning (CBR) supports problem solving through the reuse of previously observed cases [1]. Existing frameworks such as myCBR provide configurable similarity modeling mechanisms [2]. In cybersecurity, most machine learning approaches focus on detection and classification tasks rather than retrieval-oriented decision support [3].

Recent retrieval approaches employ transformer-based embeddings such as BERT to capture semantic relationships between incidents [4]. However, explainability remains a major challenge, particularly in cybersecurity environments where analysts must understand retrieval decisions [5]. Existing research rarely combines semantic embeddings, lexical similarity, structured attributes, and ontology-guided reasoning within a unified CBR framework. This doctoral research addresses this gap through an explainable hybrid similarity framework.

3. Proposed Research Framework

Figure 1 illustrates the proposed explainable hybrid retrieval architecture. The framework follows a multi-perspective retrieval approach that combines semantic, lexical, structured, and ontology-guided similarity within a unified Case-Based Reasoning (CBR) framework.

Cybersecurity incidents are stored in a structured case repository. Each incident contains attributes such as attack type, target system, industry sector, attack severity, security tools, and user roles. Before retrieval, both query incidents and stored cases are transformed into textual representations that enable semantic embedding computation while preserving the original structured attributes.

The framework combines four complementary similarity perspectives, each capturing a different aspect of incident relatedness:

1. Semantic similarity using transformer-based embeddings (captures contextual meaning),
2. Lexical similarity using TF-IDF (captures domain-specific terminology and exact matches),
3. Structured feature similarity (captures agreement on structured attributes),
4. Ontology-guided similarity (captures conceptual relationships between cybersecurity entities).

The overall similarity between a query incident q and a case incident c is computed as a weighted aggregation of the four component similarities:

$$S(q,c) = w_1 S_{bert}(q,c) + w_2 S_{tfidf}(q,c) + w_3 S_{feature}(q,c) + w_4 S_{ontology}(q,c)$$

where w_1, w_2, w_3, and w_4 are weighting factors that control the contribution of each similarity perspective.

The weighted aggregation was selected because the four similarity measures capture partially independent information about cybersecurity incidents. Semantic similarity captures contextual meaning, lexical similarity emphasizes exact terminology, feature similarity exploits structured incident attributes, and ontology similarity models conceptual relationships between cybersecurity concepts.

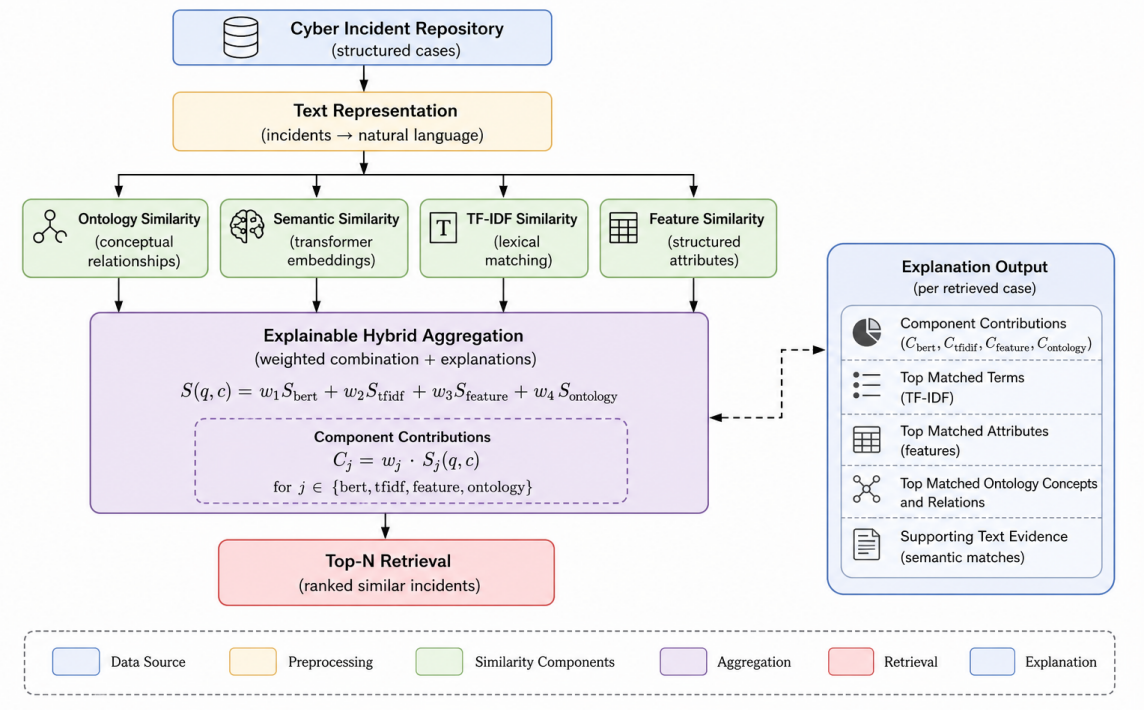

Figure 1: Explainable hybrid similarity framework.

4. Preliminary Results

To investigate the impact of ontology-guided similarity on cybersecurity incident retrieval, a preliminary evaluation of the proposed framework was conducted.

H1: Ontology-guided similarity improves retrieval quality compared to a hybrid retrieval model without ontology support.

The current implementation was evaluated using a synthetic cybersecurity incident dataset containing approximately 100,000 incidents. The dataset was selected because it allows controlled experimentation and large-scale scalability testing while avoiding confidentiality restrictions that commonly affect real-world cybersecurity incident repositories. The experiments included ontology ablation studies, retrieval evaluation, and runtime scalability analysis.

Retrieval effectiveness was assessed using Precision@10, Recall@10, Mean Reciprocal Rank (MRR), and normalized Discounted Cumulative Gain (nDCG). A 10-fold cross-validation procedure was used throughout the experiments.

Table 1 summarizes the preliminary retrieval results.

Table 1
Preliminary cross-validation results

Metric	Without Ontology	With Ontology
Precision@10	0.340	0.525
Recall@10	0.068	0.105
MRR	0.556	0.805
nDCG	0.340	0.565

The results indicate that ontology-guided similarity improves retrieval quality across all evaluation measures. The largest improvement was observed for MRR, which increased from 0.556 to 0.805. Similar improvements were observed for Precision@10, Recall@10, and nDCG, suggesting that ontology-guided reasoning contributes positively to ranking quality.

Runtime experiments demonstrated query latency below 50 milliseconds for repositories containing

approximately 100,000 incidents. These results indicate that the proposed framework can maintain efficient retrieval performance while operating on large incident repositories.

Overall, the preliminary findings support H1 and suggest that ontology-guided hybrid similarity is a promising direction for cybersecurity incident retrieval. However, the current evaluation is limited to synthetic data and focuses primarily on retrieval effectiveness. Future work will investigate real-world cybersecurity datasets, adaptive weighting strategies, and expert-centered explainability evaluation.

5. Conclusion, Open Challenges and Research Roadmap

This doctoral research investigates explainable hybrid similarity for cybersecurity incident retrieval within a Case-Based Reasoning framework. Preliminary results indicate that ontology-guided similarity improves retrieval effectiveness while maintaining efficient retrieval performance on large synthetic incident repositories.

Several research challenges remain open. First, the proposed framework must be validated using real-world cybersecurity incident datasets. Second, adaptive weighting strategies should be investigated to automatically balance semantic, lexical, feature-based, and ontology-guided similarity. Third, the explainability of retrieval results must be evaluated from the perspective of cybersecurity analysts through expert-centered studies.

The planned research consists of four phases.

- **Phase 1:** Development and evaluation of the hybrid similarity framework and ontology-guided retrieval.
- **Phase 2:** Investigation of adaptive similarity weighting and retrieval optimization techniques.
- **Phase 3:** Expert-centered explainability evaluation involving cybersecurity analysts.
- **Phase 4:** Validation on real-world cybersecurity incident repositories and assessment of transferability across cybersecurity domains.

The expected outcome is an explainable and scalable retrieval framework that supports cybersecurity analysts in identifying and understanding similar historical incidents while providing transparent retrieval explanations.

Implementation Availability

The implementation of the proposed framework is publicly available at:

- Hybrid CBR Framework for Cybersecurity Incident Retrieval (GitLab)

Declaration on Generative AI

AI-assisted language refinement tools were used during manuscript preparation. The author reviewed and validated the final content.

References

[1] A. Aamodt, E. Plaza, Case-based reasoning: Foundational issues, methodological variations, and system approaches, AI Communications 7 (1994) 39–59. doi:10.3233/AIC-1994-7104.

[2] K. Bach, K.-D. Althoff, Developing case-based reasoning applications using mycbr 3, in: Proceedings of the 20th International Conference on Case-Based Reasoning (ICCBR), volume 7466 of *Lecture Notes in Computer Science*, Springer, 2012, pp. 17–31. doi:10.1007/978-3-642-32986-9_2.

[3] A. L. Buczak, E. Guven, A survey of data mining and machine learning methods for cyber security intrusion detection, IEEE Communications Surveys & Tutorials 18 (2016) 1153–1176. doi:10.1109/COMST.2015.2494502.

[4] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of NAACL-HLT, 2019, pp. 4171–4186. doi:10.18653/v1/N19-1423.

[5] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, ACM Computing Surveys 51 (2018) 93:1–93:42. doi:10.1145/3236009.