Unified Vulnerability Intelligence: An LLM and GNN Approach

Lihua Wang[1], Jiaojiao Jiang[1], Salil Kanhere[1], Sanjay Jha[1], Jiamou Sun[2], and Zhenchang Xing[2]

[1] The University of New South Wales, Sydney, Australia
[2] CSIRO Data61, Sydney Australia

1 Introduction

Effective cybersecurity depends on consistent and accurate vulnerability intelligence. Unfortunately, publicly available vulnerability databases (e.g., NVD, IBM X-Force, ExploitDB, and Openwall) often have incomplete, inconsistent, and heterogeneous reporting of the same vulnerability entry, hindering risk assessment and patch prioritization. Prior research has employed statistical analysis, traditional NLP techniques (e.g., CRF and BERT models) [2], and graph-based methods like CEAM [1] to address vulnerability data standardization. The seminal work by Sun et al. [2] highlights ongoing gaps in extracting key aspects of vulnerabilities. Meanwhile, graph-based alignment approaches, which often assume structural uniformity, face challenges with the irregularities of real-world data. This leads to unresolved semantic mismatches and difficulties in distinguishing the nuanced details of vulnerabilities. My PhD research explores mitigating vulnerability data discrepancies through large language models (LLMs) and graph neural network (GNNs) techniques to enhance alignment and generate actionable insights. My research addresses the following three core problems:

1. *How can the extraction of vulnerability aspects from heterogeneous, unstructured reports be optimised while addressing discrepancies in aspect-level information?*
2. *How can we resolve vulnerability discrepancies caused by missing or overly generic descriptors to ensure accurate cross-database vulnerability alignment?*
3. *How can external contextual information and causal reasoning improve vulnerability alignment where traditional deep learning fails due to nuanced, highly similar vulnerability descriptions?*

2 Methodology Design and Timeline

1) Vulnerability Aspects Extraction and Discrepancies Detection This project leverages advanced large language models (LLMs) such as GPT-3.5 and GPT-4 to extract and standardize critical vulnerability aspects from over 2,000 reports across heterogeneous threat intelligence sources. By employing both zero-shot and few-shot prompting strategies, the framework extracts seven key aspects (vulnerable products, versions, components, types, root causes, attack vectors, and impacts) while evaluating semantic similarity through direct prompting and standardized term-based comparisons. This approach overcomes the limitations of traditional NLP techniques with unstructured and variable vulnerability data. The experiments demonstrate an average F1-score of 0.96, significantly outperforming conventional models and reducing ambiguity in risk assessments.

2) DARA: Enhancing Vulnerability Alignment via Adaptive Reconstruction and Dual-Level Attention The DARA (Reconstruction and Dual Attention Alignment) model targets the precise alignment of vulnerable entities across multiple repositories by integrating different types of attributes into a unified graph representation. The model's Adaptive Reconstruction Branch employs random attribute masking and a GNN encoder-decoder pipeline to reconstruct missing information from cross-graph context, minimizing reconstruction loss and reducing false negatives. Additionally, DARA's dual-attention branch leverages intra- and inter-graph attention mechanisms with a contrastive loss framework to capture both local dependencies and cross-graph similarities, effectively reducing false positives. This integrated approach demonstrates a 2% improvement in alignment accuracy compared to state-of-the-art methods like CEAM [1].

3) VulRAG: RAG-Driven Context-Aware Vulnerability Intelligence (ongoing) VulRAG combines retrieval-augmented generation with graph reasoning for context-aware vulnerability analysis. The system employs a GNN reasoning module that processes vulnerability graphs through knowledge embedding, relation matching, and candidate scoring to identify matches, while its LLM-based reasoning generates additional contextual information for relevant vulnerabilities. By integrating structured graph data with unstructured vulnerability intelligence via LLM, VulRAG enhances vulnerability disambiguation and supports downstream tasks such as exploitability scoring and patch prioritisation.

Project	**Period**	**Status**
1st project: Vulnerability Aspects Discrepancies	2022.6 - 2023.12	Completed
2nd project: DARA	2023.12 - 2025.2	Completed
3rd project: VulRAG	2024.10 - 2025.6	Ongoing
4th project & thesis writing	2025.6 - 2025.12	Ongoing

Table 1. Timeline for the PhD Dissertation Projects

References

1. Qin, Y., Xiao, Y., Liao, X.: Vulnerability intelligence alignment via masked graph attention networks. In: Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security. pp. 2202–2216 (2023)
2. Sun, J., Xing, Z., Xia, X., Lu, Q., Xu, X., Zhu, L.: Aspect-level information discrepancies across heterogeneous vulnerability reports: Severity, types and detection methods. ACM Transactions on Software Engineering and Methodology **33**(2), 1–38 (2023)