Understanding the Effectiveness of State-of-the-art Deep Learning Models on Vulnerability Detection

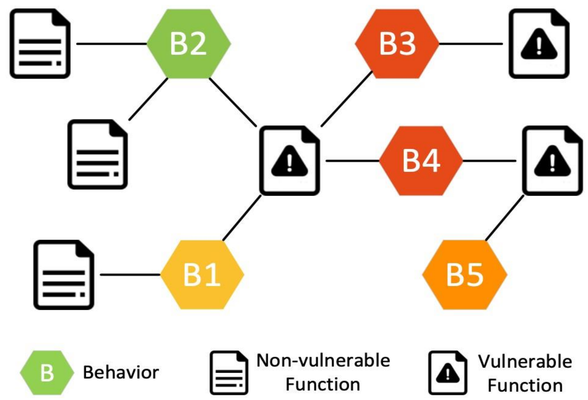
Ansh Agrawal[1], Xiaokuan Zhang[1]

[1]Department of Computer Science, College of Engineering and Computing, George Mason University

Aspiring Scientists Summer Internship Program 2025

Background

Deep learning models often miss vulnerabilities—like **buffer overflows** that let attackers crash programs or run malicious code—because they treat functions in isolation. VulBG improves detection by slicing functions into semantic "behaviors" and linking them across a **Behavior Graph**. This reveals patterns shared by vulnerable code and boosts model performance.



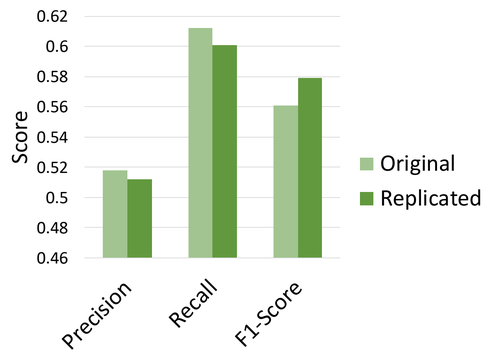
Results

The replicated model achieved strong performance and closely matched the original VulBG study. Behavior Graph features improved recall, confirming that inter-function relationships help detect more real vulnerabilities.

Key Metrics:

- **Accuracy:** 0.7104
- **Precision:** 0.5119
- **Recall:** 0.6010
- **F1-Score:** 0.5791

Replicated vs. Original Model Performance on Vulnerability Detection



Conclusions and Future Work

Key Conclusions:

- Successfully replicated VulBG with similar results.
- Behavior Graphs improved recall by capturing inter-function patterns.
- Confirms that connecting code across functions helps detect more vulnerabilities.

Future Work:

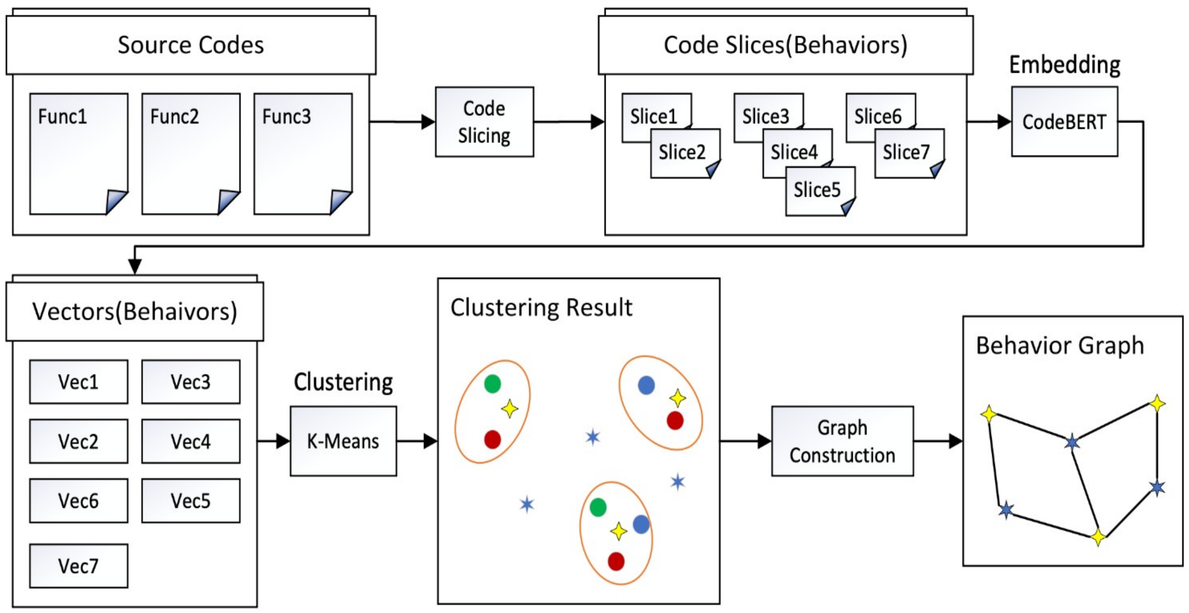
- Better slicing methods
- Combine with pretrained models (e.g., CodeBERT)
- Add explainability and slice-level insights
- Test on larger or different datasets

Materials and Methods

1. Preprocessing
Cleaned and loaded real-world C function data.

2. Feature Extraction
Embedded code slices using **CodeBERT**. Clustered slices with K-means to form behaviors.

3. Training and Evaluation
Trained a neural network classifier using both baseline and graph features. Evaluated with accuracy, precision, recall, and F1-score.



Acknowledgements

This research was made possible through the support of George Mason University's College of Science, which supports the ASSIP Program.

Major Citations

Yuan, B., Lu, Y., Fang, Y., Wu, Y., Zou, D., Li, Z., Li, Z., & Jin, H. (2023). Enhancing Deep Learning-based Vulnerability Detection by Building Behavior Graph Model. *Proceedings of the 2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE),* 2262-2274. https://doi.org/10.1109/ICSE48619.2023.00190