* 1. Scope

This research will focus on evaluating the capability of large language models (LLMs), specifically GPT-3.5, GPT-4, and other advanced generative models, in detecting and addressing security vulnerabilities in network configurations. While the use of AI-powered tools in network management is ever increasing, their efficacy in identifying and mitigating security weaknesses, in specific the adherence to existing cybersecurity standards remains underexplored. This study aims to bridge the gap by investigating the use of LLMs to identify and offer solutions for vulnerabilities in network configurations, including issues such as access control flaws and configuration inconsistencies.

An important part of this research is the development of a vulnerability detection framework using prompt engineering, fine-tuning and evaluation mechanisms. Inspired by methodologies such as Netgen’s large dataset generation [1] and the systematic repair strategies shown in works like CirFix[2], The framework utilizes LLMs to analyse network configurations for weakness. The evaluation focuses on configurations following protocols commonly used in multi-vendor environments and benchmarks them against the security standards provided by the Centre for Internet Security (CIS)[4]

This study does not delve into the deeper architectural design of LLMs but instead focuses on their practical application in network security contexts. Using sample configurations and testing LLM outputs against known vulnerabilities as categorized by standards like MITRE’s CWE taxonomy[3] and CIS benchmarks, this research aims to offer useful insights for secure network configuration practices. By doing so, it seeks to demonstrate the potential of LLMs in improving cybersecurity frameworks for important infrastructure

[1]

I. Martín, J. A. Hernández, and Ó. González de Dios, “Netgen: A Fast and Scalable Tool for the Generation and Labeling of Networking Datasets,” in *2019 21st International Conference on Transparent Optical Networks (ICTON)*, Jul. 2019, pp. 1–4. doi: [10.1109/ICTON.2019.8840020](https://doi.org/10.1109/ICTON.2019.8840020).

[2]

B. Ahmad, S. Thakur, B. Tan, R. Karri, and H. Pearce, “On Hardware Security Bug Code Fixes by Prompting Large Language Models,” *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 4043–4057, 2024, doi: [10.1109/TIFS.2024.3374558](https://doi.org/10.1109/TIFS.2024.3374558).

[3]

“CWE - Common Weakness Enumeration.” Accessed: Nov. 27, 2024. [Online]. Available: <https://cwe.mitre.org/index.html>

[4]

“CIS BenchmarksTM,” CIS. Accessed: Nov. 26, 2024. [Online]. Available: <https://www.cisecurity.org/cis-benchmarks/>