CYBERSECURITY

Large language models for automatic bug finding in source code analysis

Internship environment CEA-Leti's Information Technology Security Evaluation Facility (ITSEF) is a security evaluation laboratory certified by the Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) (France's national cybersecurity agency). It provides security evaluations for industrial products to obtain certification from the above certification bodies. It evaluates secure integrated microcircuits, HSMs and various smart cards such as bank cards, passports, health insurance cards, etc. It can also conduct "security tests" for pre-evaluations of electronic components and equipment. CEA-Leti ITSEF also performs security audits of design and production sites for secure products. CEA-Leti ITSEF designs, develops and advances internal hardware and software attack test benches to support security evaluations.

Objectives In the context of an ITSEF, the security evaluation of a software component usually requires a source code review (audit) performed by an evaluator who needs to be assisted by static analysis tools that can be configured and customized to help checking security requirements. The code analysis methodology applied at Leti ITSEF mainly consists in the following 2 operations: (1) extract a piece of source code to verify a particular property, (2) try to automatically prove the property, and in case of unknown status (the proof failed) search path conditions to violate the property. Such violations may reveal vulnerabilities to be exploited by malicious input data (software attack) combined with fault injection (hardware attack).

Main goals (1) Investigate how LLM can be used to assist evaluators in finding bug automatically in source code. For example, a research question is how IA could assist the user in generating formal specification, which is a long repetitive and complex process. (2) Assess how LLM perform and can be complementary to traditional tools used for evaluation (formal methods, using Frama-C and Lazart).

Internship tasks

- Literature review of LLMs solutions for automatic bug finding.
- Test of LLMs on open benchmarks of source code containing vulnerabilities ([3,4])
- Evaluation of a scope where LLM is relevant (i.e. where it performs better than traditional tools, where it can be complementary, to assist the evaluator)
- Proposition of a methodology to assist source code analysis with LLMs

Applicant profile We are looking for a motivated and curious candidate (BAC+5) in the field of cybersecurity to join our team. The candidate must have good programming skills (Python, C, assembly, ...) and some basic knowledge in artificial intelligence, embedded system security, vulnerability exploits. A prior technical knowledge in formal methods for static code analysis is highly valued. A proactive and autonomous profile, an enthusiasm for scientific research are encouraged.

References

[1] Sauze-Kadar Marine, Thomas. Loubier. (2025). A Multi-Model Approach to Enhance Automatic Matching of Vulnerabilities to Attack Patterns. Récupéré sur https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0013555900003979

[2] Lacombe, G., Feliot, D., Boespflug, E. et al. Combining static analysis and dynamic symbolic execution in a toolchain to detect fault injection vulnerabilities. J Cryptogr Eng 14, 147–164 (2024). https://doi.org/10.1007/s13389-023-00310-8

[3] WooKey challenge: https://wookey-project.github.io/

[4] ANSSI, Amossys, EDSI, LETI, Lexfo, Oppida, Quarkslab, SERMA, Synacktiv, Thales, Trusted Labs. (2020) Inter-CESTI: Methodological and Technical Feedbacks on Hardware Devices Evaluations. https://www.sstic.org/2020/presentation/inter-cesti_methodological_and_technical_feedbacks_on_hardware_devices_evaluations/

Keywords:
 # LLM
 # IA
 # static code analysis
 # Formal methods
 # cybersecurity

Contract Period:

- 6 months

Start date:

- Spring 2026

Workplace:

- CEA Grenoble, ITSEF

To apply, please contact:
Marine Sauze-Kadar
marine.sauze-kadar@cea.fr