Offre n°2024-07078

Post-Doctoral Research Visit F/M Automatic porting of vulnerability-fixing patches using Coccinelle

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Post-Doctorant

Niveau d'expérience souhaité : Jeune diplômé

Contexte et atouts du poste

The position is part of the project "SWHSec: Leveraging Software Heritage to Enhance Cybersecurity", funded by the Programme de Transfert au CampusCyber

Mission confiée

Objectives: Software projects today commonly co-exist in multiple versions. Some end-users may prefer to stay with older versions for stability, while others adopt the latest versions to take advantage of the newest features. These older versions of the software may be forked, and then customized for specific requirements. The co-existence of these many forks, however, means that when a vulnerability is detected and fixed, many variants of the fix are needed, to protect all known affected forks. The objective of this task is to develop and evaluate techniques for automating the creation of such fix variants. We will use Software Heritage as a source of case studies, as it contains many co-existing forks of software projects, including their complete development history.

Work description: Over the last 20 years, the Whisper team has been developing tools for automating the processing of C (and to a lesser extent C++) software projects. These tools include: (i) Coccinelle, which provides a patch-like domain-specific language SmPL for matching and transforming code, (ii) Prequel, which adapts SmPL to the task of searching for changes in a git history, and (iii) Spinfer], which generalizes change examples (as could be identified using Prequel) into SmPL transformation rules. In this project, we will investigate how to bring together and extend Coccinelle, Prequel, and Spinfer to automate the porting of vulnerability-fixing patches across software forks. Specifically, the main challenge is how to adapt the fix to the specific APIs, data structures, etc. of the target fork. We envision that it will be possible to use Prequel to collect information about the changes that have led to the differences between forks, Spinfer to translate these collected changes into transformation rules, and Coccinelle to apply these rules to the vulnerability
fix, to adapt it to the target fork. At the same time, we would like to exploit the collected information to create an explanation for the adaptations, to provide to the user, to help give confidence that the adapted fix will not itself introduce new vulnerabilities. We will initially focus on the Linux kernel, the traditional target of Coccinelle, and then scale up the approach to the wide variety of software available on Software Heritage.

References:

Coccinelle: Julia Lawall, Gilles Muller: Coccinelle: 10 Years of Automated Evolution in the Linux Kernel. USENIX Annual Technical Conference 2018: 601-614

Prequel: Julia Lawall, Derek Palinski, Lukas Gnirke, Gilles Muller: Fast and Precise Retrieval of Forward and Back Porting Information for Linux Device Drivers. USENIX Annual Technical Conference 2017: 15-26

Spinfer: Lucas Serrano, Van-Anh Nguyen, Ferdian Thung, Lingxiao Jiang, David Lo, Julia Lawall, Gilles Muller: SPINFER: Inferring Semantic Patches for the Linux Kernel. USENIX Annual Technical Conference 2020: 235-248

Software Heritage: https://www.softwareheritage.org/

Principales activités

- Studying recent backports in the Linux kernel to identify the issues that arise, and in particular any errors that have occurred in the backporting process.
- Manually simulating the collection of the information required to perform backporting.
- Automating the identified information collection strategies.

- Designing strategies for converting the collected information into transformation rules.
- Developing tools to support the evaluation of the proposed approach at a large scale.
- Assessing the precision and recall of the proposed approach, on the Linux kernel and on several projects available in Software Heritage
- Writing papers and giving presentations describing the approach and its results.

Compétences

Technical skills and level required : The project requires a strong background in program analysis. Some familiarity with algorithms for processing source code, such as clone detection, differencing, clustering, etc. would be appreciated.

Languages : The position requires reading a large amount of C code. Some reading familiarity with C++ could also be beneficial. All of the existing software infrastructure related to the project is written in OCaml.

Other appreciated skills : Good communication skills (spoken and written). The ability to plan a project and work independently based on the plan.

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Informations générales

- **Thème/Domaine :** Systèmes distribués et intergiciels
  Ingénierie logicielle (BAP E)
- **Ville :** Paris
- **Centre Inria :** Centre Inria de Paris
- **Date de prise de fonction souhaitée :** 2024-07-01
- **Durée de contrat :** 2 ans
- **Date limite pour postuler :** 2024-08-31

Contacts

- **Équipe Inria :** WHISPER
- **Recruteur :**
  Lawall Julia / Julia.Lawall@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

The position is well suited to someone who enjoys looking at a lot of code and code changes, and who is passionate about improving software development and maintenance processes, with a potential impact on the real world.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Sécurité défense :
Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un

poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.