2019-01548 - PhD Position F/M [Campagne CORDI-S] Program analysis applied to the Linux kernel

Type de contrat : CDD de la fonction publique
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonction : Doctorant

Contexte et atouts du poste

The Linux kernel is used pervasively across the computing spectrum today, on hardware from smartphones to supercomputers, and in environments from battleships to stock markets. The Linux kernel is also one of the largest open source projects, with around 17 million lines of C code, around 13 000 commits per release (every 2-3 months), and over 1800 developers contributing to each release. These properties raise a number of challenges for software maintenance, which in turn raise a number of interesting research directions. The Linux kernel development community is quite open to new contributors and solutions involving tool support, implying that research in this area can have an immediate and high practical impact.

The Whisper team has over 10 years of experience in research around tools for improving the reliability of operating systems code. The foundation of this work is Coccinelle [1], for matching and transformation of C code.
Coccinelle orignially targeted the Linux kernel, and today has been used in the development of thousands of Linux kernel commits. In recognition of this achievement, the original paper on Coccinelle [2] received the 2018
EuroSys test of time award. Recent and ongoing activity in this area include a tool for finding resource release omissions [3], a tool for querying git histories [4], and a tool for inferring transformation rules from examples.

[1] http://coccinelle.lip6.fr/

[2] Yoann Padioleau, Julia L. Lawall, René Rydhof Hansen, Gilles Muller:
Documenting and automating collateral evolutions in linux device drivers. EuroSys 2008: 247-260

[3] Suman Saha, Jean-Pierre Lozi, Gaël Thomas, Julia L. Lawall, Gilles Muller:
Hector: Detecting Resource-Release Omission Faults in error-handling code for systems software. DSN 2013: 1-12

[4] 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

Mission confiée

Possible PhD projects include the following:

* **Preventing regressions in Linux kernel stable versions:** The Linux kernel development community maintains a number of stable versions, which are previous releases to which recent bug-fixing patches are applied. A patch may apply cleanly to older code, but introduce errors, when the invariants found in the recent kernel version for which the patch was developed are different from the invariants found in the older stable version to which the patch is applied. The goal of this project is to develop methodologies to detect such changes in invarants, to be able to warn the stable-kernel maintainer of potential problems in applying a patch to an older kernel version. This project requires a substantial background in program analysis.

* **Linux kernel memory management:** Correctly managing memory is critical to a long-running system such as the Linux kernel, but the C language provides no native support for doing so. In response, the Linux kernel has developed a number of strategies for managing different types of memory usage. The goal of this project is to collect and categorize these strategies, detect and understand the bugs in their usage, and propose more robust coding strategies. This project requires a substantial background in program analysis.

* **Correctness of Linux kernel data structures:** While the Linux kernel provides a number of libraries for data structure manipulation, it also contains a number of ad hoc data structures that are specific to particular services. The goal of this project is to develop methods for automatically identifying these data structures and their properties in Linux kernel code, and formally verifying that the Linux kernel code maintains their invariants. This projects requires a background in program analysis and also in proving properties of source code.

* **Inference of Linux kernel interfaces:** Fuzzing tools, such as syzkaller, bombard a software with inputs that are designed to try to find bugs. A key to the success of such approaches is understand the expected form of the inputs, so that the test cases constructed by the fuzzer can reach code deep in the software. In the case of the Linux kernel, descriptions of valid inputs have are available for system calls, but are available to a much lesser degree for the IOCTL interfaces exposed directly by device drivers. The goal of this project is to infer such interfaces from the source code, including both type constraints and expected semantic relationships between input values, such as that one argument should represent the length of another argument. This project requires a substantial background in program analysis. Some familiarity with devices would also be helpful.

Principales activités

Main activities: Research on program analysis methodologies. Presentation of results in research conferences. Writing papers about the results.

Compétences

Technical skills and level required : Strong experience in program analysis, compilation, or related areas is required. Experience in reading and understanding large C code bases would be helpful. Software development associated with the project is expected to be carried out in OCaml. Previous experience in writing either a research paper or a masters thesis, preferably in English, would be helpful.

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 (after 6 months of employment) 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

Rémunération

1982 € la première et la deuxième année, 2085 € la troisième année.

1982 € during the first and second years, 2085 € the last year.

Informations générales

- **Thème/Domaine :** Architecture, langages et compilation
- **Ville :** Paris
- **Centre Inria :** CRI de Paris
- **Date de prise de fonction souhaitée :** 2019-09-01
- **Durée de contrat :** 3 ans
- **Date limite pour postuler :** 2019-05-26

Contacts

- **Equipe Inria :** WHISPER
- **Directeur de thèse :**
  Lawall Julia / julia.lawall@inria.fr

A propos d'Inria

Inria, l'institut national de recherche dédié aux sciences du numérique, promeut l'excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l'interface d'autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 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.

Consignes pour postuler

La candidature doit contenir :

CV
lettre de motivation
notes de master
Des lettres de recommandation peuvent être envoyées directement à la personne au recruteur.

The application must contain:

CV
cover letter
master's notes
Letters of recommendation can be sent directly to the recruiter.

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.

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.