Extraction de structure de code bas niveau
Sujet de stage de Master 2 recherche

Encadrants : Olivier Ly <ly@labri.fr>
Aymeric Vincent <vincent@labri.fr>
Igor Walukiewicz <igw@labri.fr>
Thème : Modélisation et Vérification
Equipe : Méthodes Formelles
Financement : ANR BINCOA

Description du sujet

Le but de ce stage est d'explorer les informations de structure que l'on peut extraire de code exécutable. Deux voies complémentaires sont à explorer : la reconstruction du graphe de flot de contrôle [BGRT05, KV08] qui structure le code, et la découverte des structures de données manipulées par le code [BR07].

Il s'agira de comprendre l'état de l'art sur ces deux aspects, et d'étudier comment la reconstruction de ces deux structures peuvent s'aider mutuellement.

Nous attendons une étude théorique et algorithmique du problème, en prenant en compte la difficulté supplémentaire amenée par le volume de code important puisqu'il correspond à un code directement exécutable. La communauté vérification utilise plutôt habituellement des modèles de haut niveau ou des programmes écrits en langage plus évolué comme le C.

References

[BGRT05] Gogul Balakrishnan, Radu Gruian, Thomas W. Reps, and Tim Teitelbaum. Codesurfer/x86-a platform for analyzing x86 executables. In *CC*, pages 250–254, 2005.

[BR07] Gogul Balakrishnan and Thomas W. Reps. Divine: Discovering variables in executables. In *VMCAI*, pages 1–28, 2007.

[KV08] Johannes Kinder and Helmut Veith. Jakstab: A static analysis platform for binaries. In *CAV*, pages 423–427, 2008.