

Étude de l'évolution des clones

7 octobre 2013

La réutilisation de fragments de code est une pratique courante dans le développement logiciel. Cette réutilisation peut être intentionnelle par l'action de copier-coller ou non intentionnelle quand les développeurs suivent les idiomes relatifs aux langages ou aux bibliothèques logicielles qu'ils utilisent. Ces fragments de code appelés « clones » peuvent représenter jusqu'à 10 à 15% du code total d'un logiciel [4].

Cette problématique a attiré de nombreux chercheurs et soulevé autant de questions de recherche. Le monde du génie logiciel s'intéresse aux conséquences de la présence de clones dans des logiciels et sur l'impact qu'ils peuvent avoir sur la qualité de ces derniers [1].

Votre travail consistera à étudier les articles cités ci-dessous afin de cerner et maîtriser les problématiques liées à la présence de clones dans un logiciel. Vous vous intéresserez plus particulièrement aux articles traitant de l'évolution des clones dans un logiciel [5].

Vous devrez ensuite comprendre le fonctionnement d'un outil de détection de clones ([2] [3]) et implémenter son support dans Merlion. Merlion est un outil permettant le suivi de l'évolution de clones par rapport à l'histoire d'un projet. Il utilise pour cela les informations calculées par un outil de détection de clones.

Références

- [1] N. Göde and R. Koschke. Frequency and risks of changes to clones. In *Proceedings of the 33rd International Conference on Software Engineering, ICSE '11*, pages 311–320, New York, NY, USA, 2011. ACM.
- [2] L. Jiang, G. Mishnerghi, Z. Su, and S. Glondu. Deckard : Scalable and accurate tree-based detection of code clones. In *Proceedings of the 29th international conference on Software Engineering, ICSE '07*, pages 96–105, Washington, DC, USA, 2007. IEEE Computer Society.
- [3] T. Kamiya, S. Kusumoto, and K. Inoue. Ccfinder : a multilinguistic token-based code clone detection system for large scale source code. *IEEE Trans. Softw. Eng.*, 28(7) :654–670, July 2002.
- [4] C. J. Kapser and M. W. Godfrey. "cloning considered harmful" considered harmful : patterns of cloning in software. *Empirical Softw. Engg.*, 13(6) :645–692, Dec. 2008.
- [5] M. Kim, V. Sazawal, D. Notkin, and G. Murphy. An empirical study of code clone genealogies. *SIGSOFT Softw. Eng. Notes*, 30(5) :187–196, Sept. 2005.