Ceci est un titre de document

Antonin DUREY, Jihad EL FAKAWI

Novembre 2016

Table des matières

Introduction

De nos jours, les crashs de logiciels ou de programmes sont quotidiens et les grosses applications doivent faire face à des milliers de rapport de crashs. L'envoi de ces rapports aux développeurs est néanmoins essentiel pour permettre à ceux-ci de corriger les problèmes et failles. Pour simplifier l'analyse de ces rapports, ils sont triés par bucket. Un bucket rassemble ainsi une série de rapports dont la source du problème est très certainement la même. Grâce à cela, lorsque les développeurs analyserons les rapports contenus dans chacun des bucket, ils disposeront d'une grande quantité d'information pour corriger le problème.

Dans le cadre du module OPL de notre master 2 IAGL, nous avons été amenés à réaliser un projet qui porte sur l'analyse de crash. Sur ce projet, nous avons implémenté le tri de rapports d'erreurs dans des buckets.

Pour trier ces rapports, nous avons implémenté plusieurs algorithmes. Dans un premier temps, nous les avons triés de manière aléatoire. Dans un second temps, nous avons attribué un certain nombre de points aux buckets en fonction des points communs entre le rapport à trier et les rapports contenus dans le bucket. Le bucket choisi est alors celui dont le score est le plus élevé. Le troisième algorithme est une variante du second, il calcule la moyenne des points par nombre de lignes d'erreurs et par nombre de rapports dans un bucket.

L'évaluation de nos résultats a été facilité par l'infrastructure mise en place. En effet, nous pouvions soumettre nos résultat, c'est-à-dire l'attribution des rapports d'erreurs à un bucket précis. Grâce à cela, nous avons constaté que notre troisième et dernier algorithme est plus performant que le second, lui même plus performant que le premier. Néanmoins, ces 2 dernières solutions présentent l'inconvénient de prendre un temps conséquent lors du tri - plusieurs dizaines de secondes avec les données fournies.

Ce rapport présente ces travaux en 3 parties. La première est consacré au travail technique qui a été effectué. La seconde se concentre sur l'évaluation des résultats et des algorithmes. Enfin, la troisième et dernière partie traite des améliorations et des évolutions possibles du projet.