

Avaliação da qualidade das funções de similaridade no contexto de testes baseados em modelos

Gustavo L. Machado*, Eliane Martins

Resumo

Nesse trabalho analisamos funções de similaridade no contexto de agrupar casos de teste gerados a partir de modelos de estado, buscando a redução de custos com a realização desses testes.

Palavras-chave:

Geração de casos de teste – Diversidade – Funções de Similaridade

Introdução

Geração de casos de teste a partir de modelos de estado costumam produzir uma quantidade muito grande de dados, ao ponto que tentar utilizar todos eles para testar algum software pode se tornar uma tarefa inviável. Mas é intuitivo pensar que dentre todos os casos de teste gerados alguns serão semelhantes ou até iguais nos resultados que produzem, sendo assim redundante usar todos eles. Temos por objetivo então agrupar o conjunto de todos os casos de teste em conjuntos menores, de acordo com a similaridade entre os casos de teste, de forma também a não perdermos resultados e a otimizarmos a utilização dos testes. Para tal utilizaremos diferentes funções de similaridade, que produzem um escore quantificando a semelhança entre dois objetos comparados. A qualidade dos conjuntos gerados será avaliada em termos de diversidade. Quanto maior a similaridade dentro do conjunto e menor a similaridade entre conjuntos melhor.

Resultados e Discussão

Para a análise dos casos de testes os consideramos como sequências de caracteres, strings, que representavam as transições do modelo que cada caso de teste cobria. Portanto, após estudos e discussões, chegamos as seis conhecidas funções de similaridade para strings e textos que utilizamos: Jaro-Winkler, Damerau-Levenshtein, Levenshtein, Optimal String Alignment (OSA) e Normalized Compression Distance (NCD).

Com essas distâncias computamos as matrizes de similaridade de um grupo controle de casos de testes. E em seguida montamos os respectivos dendrogramas, utilizando o método "average" ("Unweighted Pair Group Method with Arithmetic Mean-UPGMA), para então para termos uma ideia inicial de possíveis grupos, pois independente do método utilizado é preciso que um especialista decida o número de grupos que serão formados. A figura 1 mostra o exemplo de um dendrograma, bem como possíveis divisões de grupos.

A princípio, como todas as meta-heurísticas que fazem as seleções de grupos pedem como entrada o número de grupos desejados, nós, após avaliação do dendrograma definimos este número e fizemos nossa primeira análise baseada neles.

Para análise dos grupos formados levamos em consideração duas coisas: primeiro, a razão entre

mutantes (bugs) em comum mortos pelos testes de grupo, pelo numero total de mutantes mortos por esse mesmo grupo; Em seguida analisamos outra razão, transições em comum dentre os casos de teste num mesmo grupo/total de transições cobertas pelo grupo. O que se espera é que essas taxas sejam próximas de um.

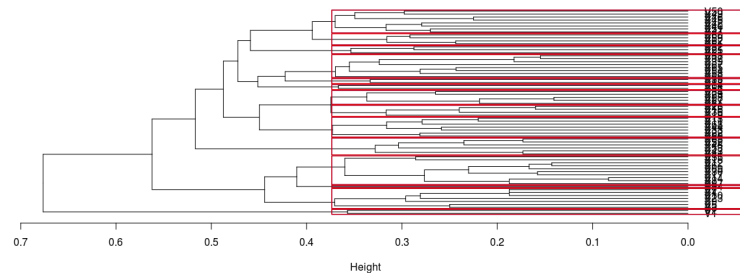


Figura 1. Dendrograma da distância NCD para o nosso grupo controle de casos de teste.

Conclusão

Este projeto iniciou em Abril/2017. Até o presente momento foram feitas as tarefas de tratamento dos dados, para análise, escolhas das funções de similaridade a serem estudadas e do método de agrupamento. O próximo passo agora é analisar os agrupamentos obtidos com as diferentes funções de similaridade e determinar qual(is) obté(ê)m melhor(es) resultado(s), i.e., que produzem grupos de casos de teste mais homogêneos em termos de detecção de defeitos. Estamos no caminho para conseguir estabelecer um método para a formação de grupos no contexto de agrupamento de casos de testes.

Agradecimentos

¹ Coutinho, A.E.V.B., Cartaxo, E.G. & Machado, P.D.L. Software Qual J, (2016)