

## Log da reunião realizada em 04/05/2018 sobre o projeto de TCC do aluno Gabriel Franco.

Foi feita análise dos resultados para o coeficiente de correlação Tau de Kendall, utilizando métricas obtidas a partir dos 10 melhores conjuntos de atributos selecionados pelo Algoritmo Genético (AG), acrescido da métrica de referência (KDA). As métricas foram obtidas usando dois tipos de cálculo. O primeiro deles segue o padrão:

$$M_{i,j} = \begin{cases} \frac{\sum_{k=1}^n A_k}{1+deaths}, \forall A_k \in \{C_j - \{deaths\}\} & \text{se } deaths \in C_j \\ \sum_{k=1}^n A_k, \forall A_k \in C_j & \text{caso contrário} \end{cases}$$

onde  $M_{i,j}$  é o valor da métrica para o indivíduo  $i$  da base de dados, para um conjunto  $C_j$  de atributos, sendo este o  $j$ -ésimo conjunto de atributos entre aqueles selecionados como melhores pelo Algoritmo Genético, e  $n$  é a quantidade de atributos presentes em  $C_j$  após a remoção de *deaths*, caso presente, ou o próprio tamanho de  $C_j$ , caso contrário.

O segundo tipo de cálculo foi baseado na média harmônica dos valores, ligeiramente adaptado pra evitar divisões por zero e para contemplar o fato de que para o caso de *deaths*, quanto menor o valor, melhor a avaliação do jogador, em oposição ao que acontece com os demais atributos. Sendo assim, temos:

$$M_{i,j} = \begin{cases} \frac{n+1}{\sum_{k=1}^n \frac{1}{1+A_k} + (1+deaths)} & \text{se } deaths \in C_j \\ \frac{n}{\sum_{k=1}^n \frac{1}{1+A_k}} & \text{caso contrário} \end{cases}$$

Os resultados encontrados não diferiram significativamente para as duas mecânicas de cálculo das métricas. Por ter uma análise mais intuitiva, adotamos o primeiro tipo.

Também foram selecionados 4 conjuntos de atributos para que se pudesse trabalhar, deste ponto da pesquisa em diante. O primeiro conjunto, contém os 5 melhores atributos rankeados no arquivo *top10.ods* no GitHub, os dois melhor classificados pelo Algoritmo Genético e o conjunto de referência, composto pelos atributos *assists*, *deaths* e *kills* (KDA). Desta forma, temos as seguintes métricas:

$$KDA = \frac{assists + kills}{deaths}, \quad (1)$$

a métrica de referência, utilizada pela comunidade,

$$ADG = \frac{assists + denies + gpm + hh}{deaths} \quad (2)$$

a principal métrica proposta, e mais completa, composta pelos 5 atributos com melhor avaliação individual na etapa anterior do estudo,

$$G = \frac{gpm + hh}{deaths}, \quad (3)$$

composta pelo conjunto com maior *fitness* obtido pelo AG, e

$$X = \frac{xpm + hh}{deaths}, \quad (4)$$

composta pelo conjunto com a segunda melhor avaliação na seleção de atributos. Esta foi também incluída por ser muito similar à anterior e também por ter *fitness* bastante próximo da melhor colocada.

Após estas conclusões, foram definidos os passos seguintes da pesquisa:

1. Computar o valor médio, desvio padrão e coeficiente de variação de cada métrica (equações 1, 2, 3 e 4) para cada cluster obtido anteriormente.
2. Tabular e/ou plotar os resultados encontrados.
3. Para fins de validação do valor-p encontrado na avaliação do coeficiente Tau, realizar um teste utilizando uma amostra menor, de cerca de 100 indivíduos da base e verificar se o valor-p passa a ser diferente de zero.
4. Fazer o *star plot* dos 10 melhores jogadores segundo cada uma das 4 métricas aqui produzidas.