Avaliação Guião PL04

Turma P5

Sara Gonçalves (98376) e Miguel Martins (103421)

Universidade de Aveiro – Métodos Probabilísticos para Engenharia Informática

Prof. Carlos Bastos

2021/2022



# **Introdução**

Este relatório serve de complemento para a realização da avaliação do guião PL04. A partir da matéria abordada nas aulas teóricas e consolidada nas aulas práticas devíamos ser capazes de desenvolver uma aplicação com algumas funcionalidades de uma rede social.

Esta tem em consideração um conjunto de utilizadores identificados por um ID e um conjunto de amigos também identificados por um ID.

Para resolver os desafios propostos usámos o MATLAB para implementar funções assentes principalmente em *MinHashes*, distância de Jaccard e Filtro de *Bloom*.

Como dados de entrada para que a aplicação funcione, são considerados dois ficheiros:

* *friends.txt:* Contém, na primeira coluna, o ID de um utilizador e na segunda coluna o ID de um amigo desse utilizador. Da coluna 3 à coluna 5 contém o dia, mês e ano do último contacto.
* *users.*txt: Contém a informação de cada um dos utilizadores.

Neste relatório iremos primeiramente apresentar os dois scripts desenvolvidos. De seguida será mostrado os métodos realizados para o restante funcionamento do programa e uma breve exemplificação dos resultados que foram obtidos. Por fim, apresentaremos as conclusões que conseguimos retirar deste projeto.

# **Scripts**

**Script 1** – data.m

O primeiro script desenvolvido tem como objetivo realizar a leitura dos ficheiros de entrada, mais concretamente dos ficheiros falados anteriormente. Além disso, ainda foi necessário guardar em ficheiro todas as estruturas de dados associadas aos utilizadores e aos amigos, incluindo a matriz com os vetores MinHash de cada utilizador, a matriz com os vetores MinHash de cada nome, tal como pedido no enunciado do trabalho.

Para isso foi desenvolvido o seguinte script:

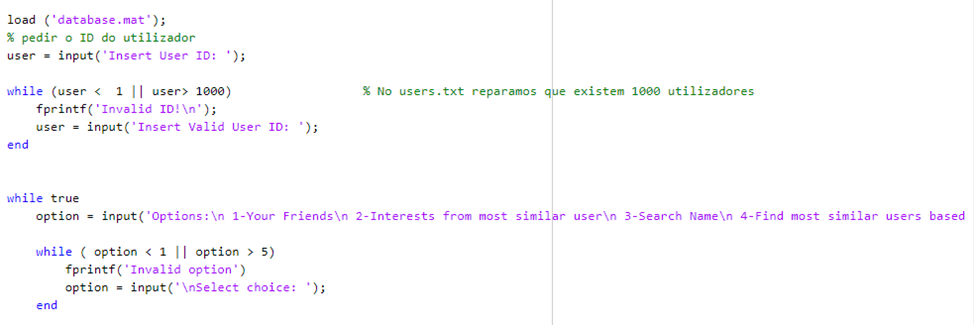


Figura 1 - Script 1

Uma imagem com texto

Descrição gerada automaticamente

Figura 2 - Script 1 (continuação)

Observando ambas as figuras, pode-se ver que foram guardadas todas as informações que serão necessárias ao longo do trabalho.

**Script 2** – app.m

O segundo script começa por ler do disco todas as estruturas previamente guardadas pelo primeiro script e depois implementa todas as interações com o utilizador.

Uma imagem com texto

Descrição gerada automaticamente

Figura 3 - Script 2

Uma imagem com texto

Descrição gerada automaticamente

Figura 4 - Script 2 (continuação)

Na figura 3 e 4 pode-se então ver que a aplicação será composta por cinco opções, tal que a última apenas tem a função de terminar o programa se o cliente desejar.

É de notar que também foi feita uma correta implementação de todas as interações com o utilizador e a validação de todos os dados inseridos pelo cliente.

As funções ainda desse script criadas como auxiliares das opções serão explicadas mais à frente.

# **Métodos**

A opção 1 tem como objetivo listar os amigos do utilizador atual, mostrando o seu id e nome completo. Nesta opção criámos uma matriz com os ids dos amigos do utilizador atual com a função friends, calculando também o seu tamanho. Tendo os ids dos amigos fizemos um ciclo for para percorrer a matriz e usando cada id da matriz criada fomos buscar o nome correspondente ao id com a função Name criada no script data.m que ia buscar as colunas com o nome e apelido de cada utilizador. Por fim fazer print do id e do nome completo por cada linha.Para esta opção foram criadas as seguintes funções:

* **função your\_friends**

Uma imagem com texto

Descrição gerada automaticamente

Figura 5 - Função yourFriends

A função *your\_friends* percorre todos os utilizadores e encontra através da função find do Matlab todos os amigos correspondentes a cada pessoa e retorna desse modo um cell array que representa a lista amigos para cada utilizador.

* **get\_your\_friends**

Uma imagem com texto

Descrição gerada automaticamente

Figura 6 - Função get\_your\_friends

Esta função tem como parâmetros o conjunto de utilizadores e os seus respetivos amigos (valor obtido da função getUserMovies), os nomes completos obtidos a partir do ficheiro “users.txt” lido no script 1 e ainda o número de identificação do utilizador atual. Serão percorridos todos os id’s dos amigos do user atual de forma a encontrar os nomes correspondentes a esses e listá-los, um por linha. Ainda neste método determinámos a quantidade de amigos que o user atual tem.

A *opção 2* tem como objetivo listar os interesses do utilizador mais “similar” ao utilizador atual e, após isso, listar os não partilhados pelo utilizador atuar, como sugestão de novos interesses. Para isso, começa por determinar qual dos outros users é mais similar ao atual em termos da lista de amigos.

Para esta opção foram criadas as funções:

* **getSimilarUsers**

Uma imagem com texto

Descrição gerada automaticamente

* **suggested\_interests**

Uma imagem com texto

Descrição gerada automaticamente

A *opção 3*  baseia-se no cliente introduzir uma palavra e através dela procurar os nomes completos de utilizadores mais similares à string introduzida.

Em conjuntos com esses nomes deve-se obter uma estimativa da distância de Jaccard entre essa string e cada nome obtido. A lista determinada deverá ter no máximo cinco nomes e têm de estar ordenados por ordem crescente de distância de Jaccard cuja distância seja menor ou igual a 0.99. Com base na utilização de um filtro de bloom, deve aparecer uma mensagem a indicar se a string introduzida pode responder exatamente a um dos nomes dos utilizadores. No entanto, esse último aspecto não foi feito devido a falta de tempo.

Para esta opção foram criadas a função:

* searchName

Uma imagem com texto

Descrição gerada automaticamente

Uma imagem com texto

Descrição gerada automaticamente

A opção 4 não foi trabalhada.

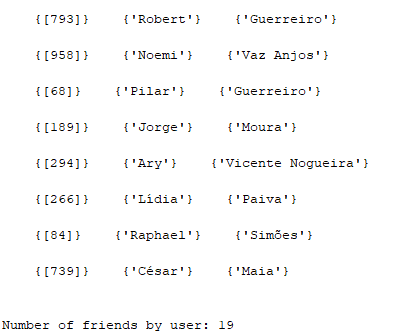
# Resultados obtidos

Nesta secção iremos simular através de um número de identificação do utilizador escolhido as diversas funcionalidades desta rede social. É de relembrar que a opção 5 é apenas uma funcionalidade de terminar o programa.

Uma imagem com texto

Descrição gerada automaticamente

* **Opção 1**



* **Opção 3**

Uma imagem com texto

Descrição gerada automaticamente