

Relatório T4

“Simulando reconhecimento de padrões através de caminhos ótimos em grafos (classificação supervisionada)”

Teoria dos Grafos 2014/02

Prof. Dr. Alexandre Luis Magalhães Levada

Daniel Nobusada	344443
Thales Eduardo Adair Menato	407976
Jorge Augusto Bernardo	407844

“Nossa maior fraqueza está em desistir. O caminho mais certo de vencer é tentar mais uma vez.” - Thomas Edison

Resumo

A partir dos grafos fornecidos pelo professor para este experimento, utilizando o algoritmo de Dijkstra encontraremos os melhores caminhos de todos os vértices para, não apenas uma origem, mas várias “origens” as quais chamaremos de sementes. A partir da lista de predecessores obtida, é possível gerar um novo grafo conectando cada vértice com seu predecessor até chegar na origem (semente) a qual o mesmo foi determinado. Dessa forma podemos visualizar grupos (comunidades) que se formam num grafo desconexo.

Metodologia

Os grafos são carregados utilizando a estrutura da biblioteca *NetworkX* e então para cada um deles temos a seleção aleatória das sementes.

Para o caso “a” – através da lista da periferia do grafo (obtida através da função *periphery*) e então são selecionados k vértices distintos desta lista.

Para o caso “b” – é selecionado um vértice aleatório dentro da lista de vértices do grafo, a partir dele obtemos a lista de vizinhos do mesmo, que então preenche os $k-1$ vértices restantes.

Com as sementes já obtidas, executamos o algoritmo do Dijkstra utilizando-as e nos é retornado:

λ - dicionário com os vértices e seus respectivos pesos até a semente a qual está conectado

π - dicionário com os vértices e seus respectivos predecessores

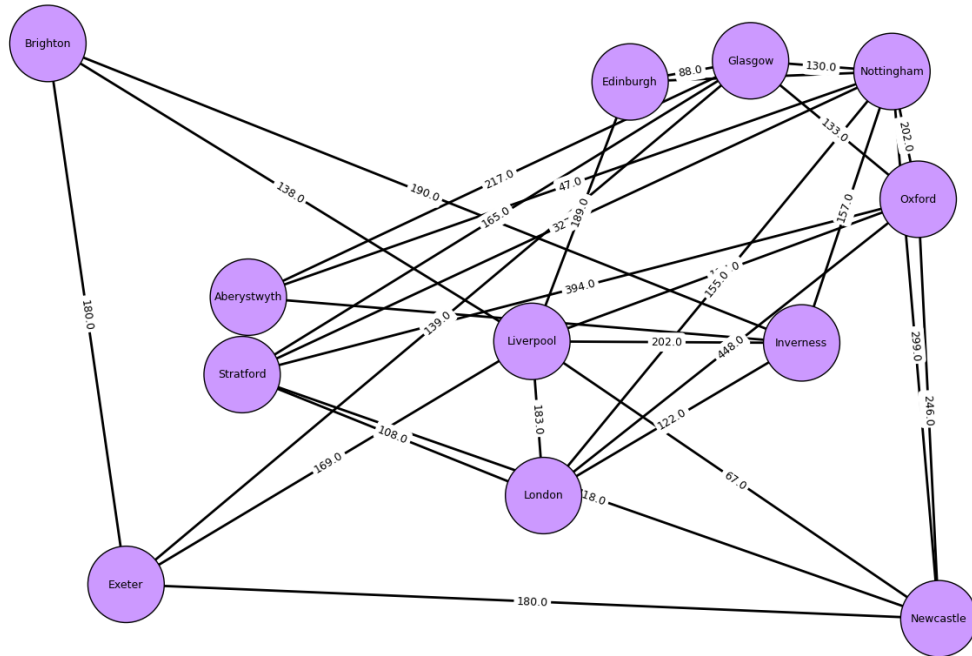
H – grafo gerado à partir de π , contendo apenas as arestas dos vértices de cada grupo e seus respectivos predecessores.

A partir desses dados obtidos, é gerado uma imagem com o plot do grafo H e um arquivo de texto com todas as informações contidas em λ e π , além de quem foram as sementes e as extensões dos grupos formados.

Questionamentos

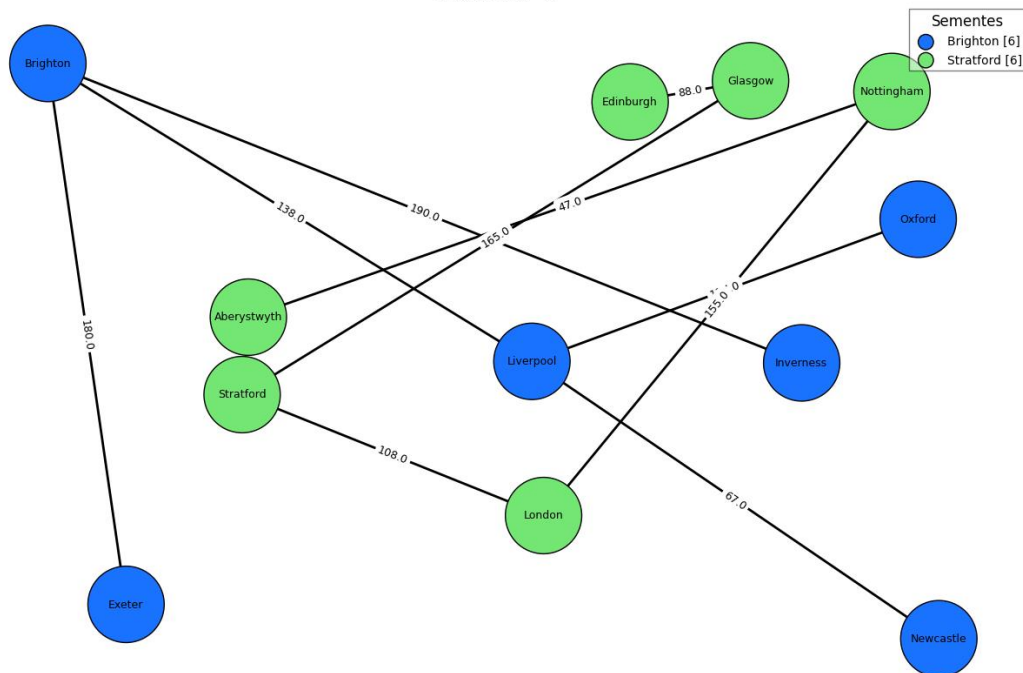
Considerando o grafo a seguir de 12 vértices: “uk12”

Grafo uk12 - Original

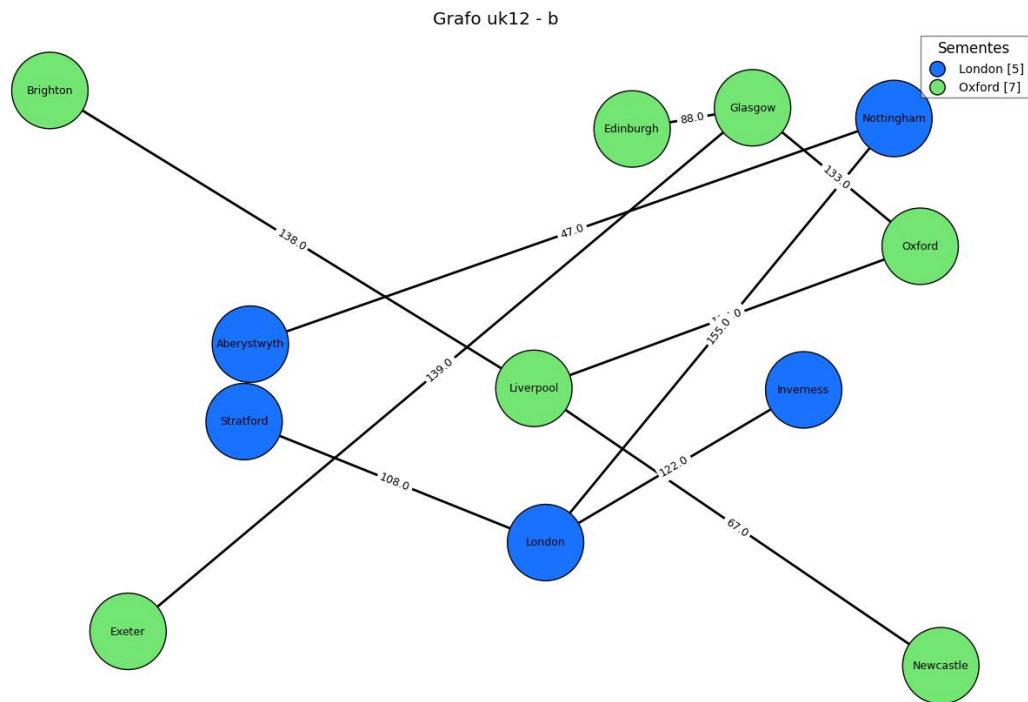


a) 2 agrupamentos (k=2) com sementes espalhadas

Grafo uk12 - a

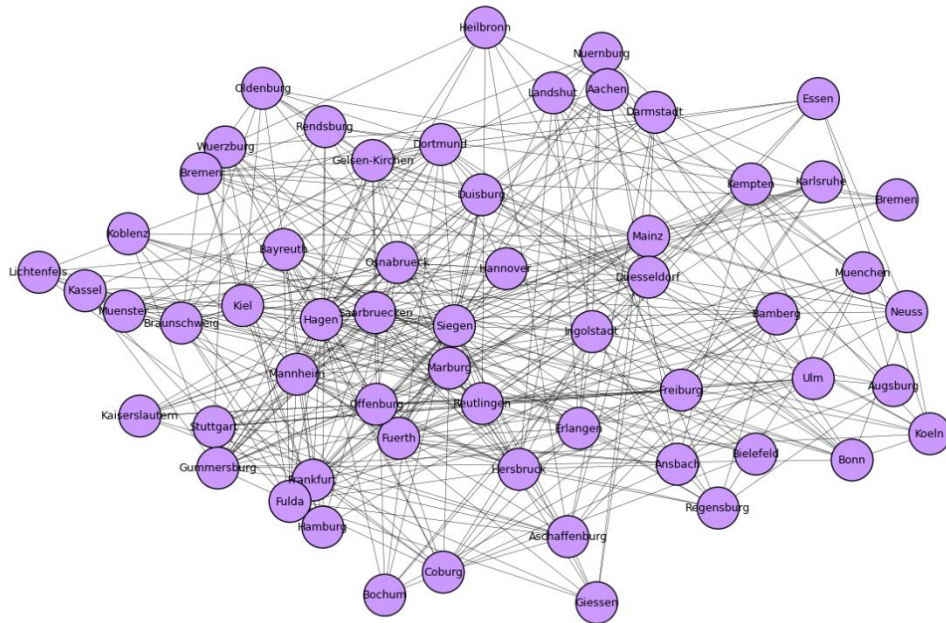


b) 2 agrupamentos (k=2) com sementes próximas uma da outra



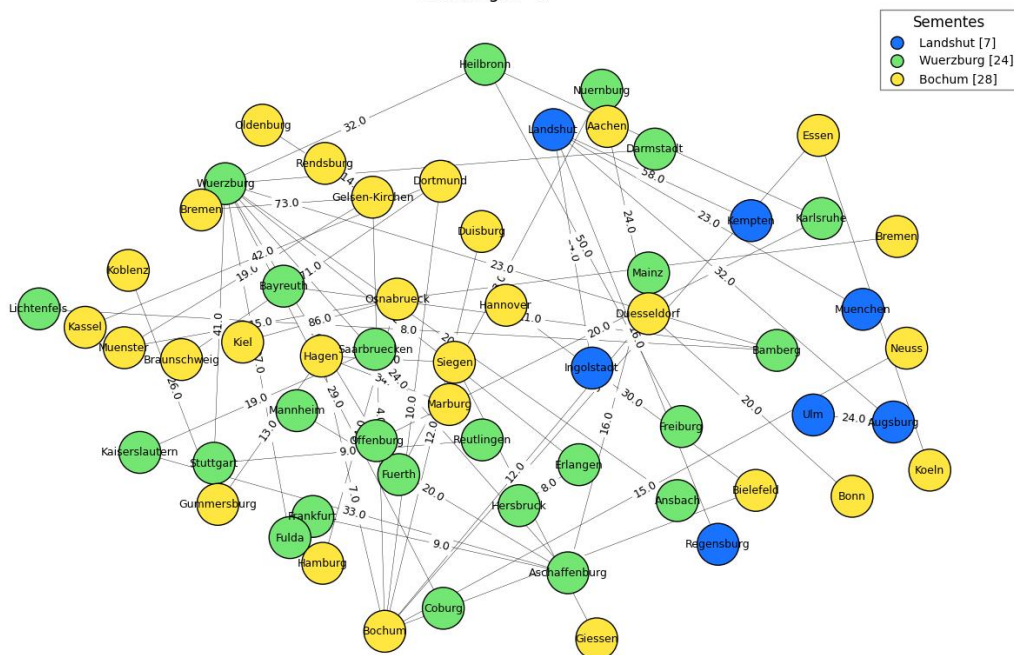
Considerando o grafo a seguir de 59 vértices: “wg59”

Grafo wg59 - Original

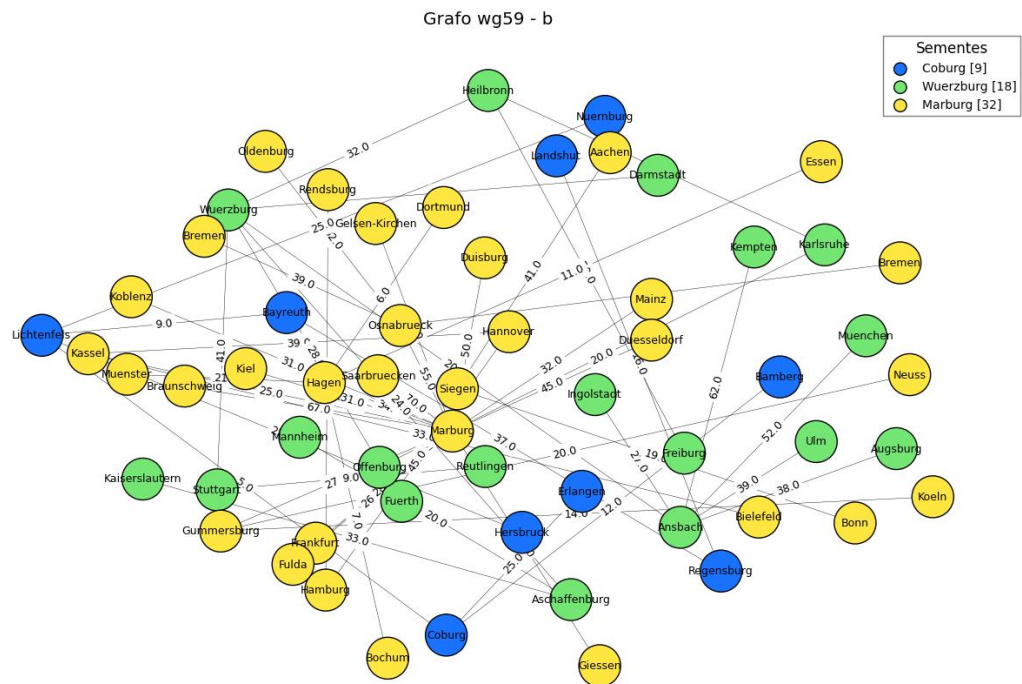


a) 3 agrupamentos (k=3) com sementes espalhadas

Grafo wg59 - a

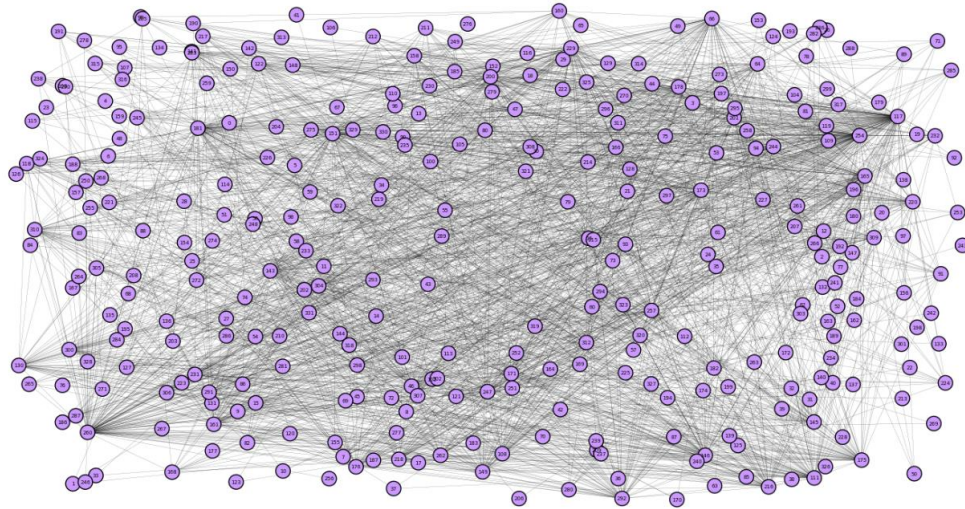


b) 3 agrupamentos (k=3) com sementes próximas uma da outra



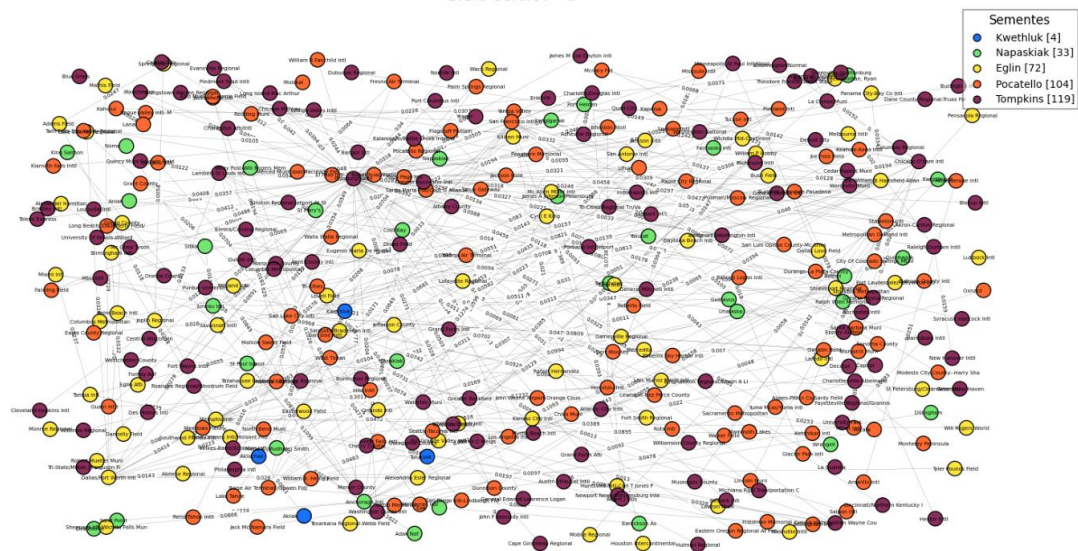
Considerando o grafo a seguir de 332 vértices (rede de aeroportos dos EUA):
“USAir97”

Grafo USAir97 - Original

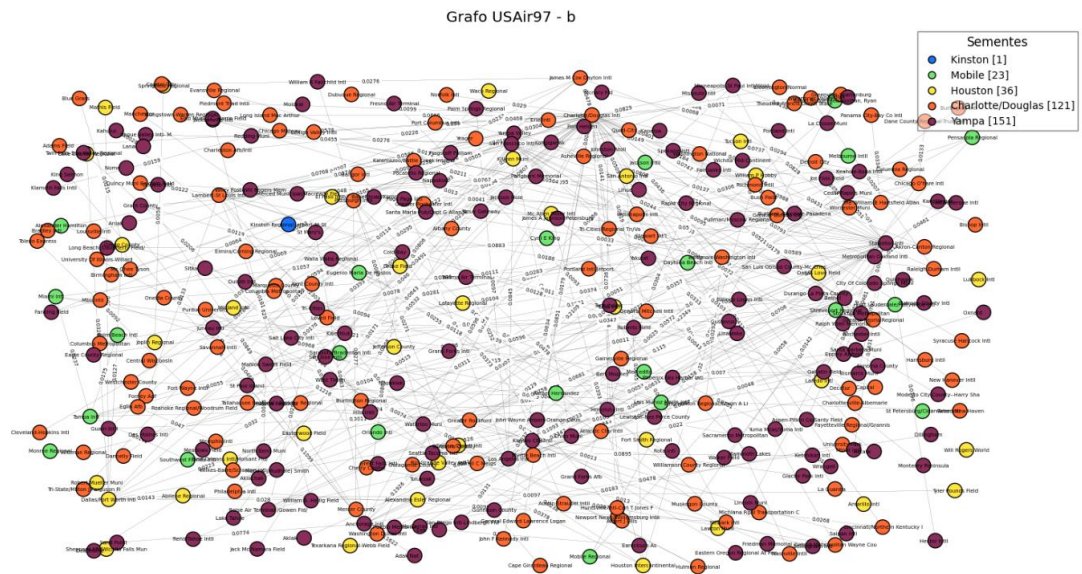


a) 5 agrupamentos (k=5) com sementes espalhadas

Grafo USAir97 - a



b) 3 agrupamentos (k=3) com sementes próximas uma da outra

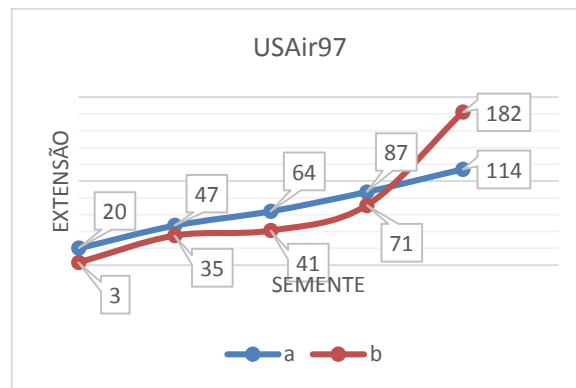
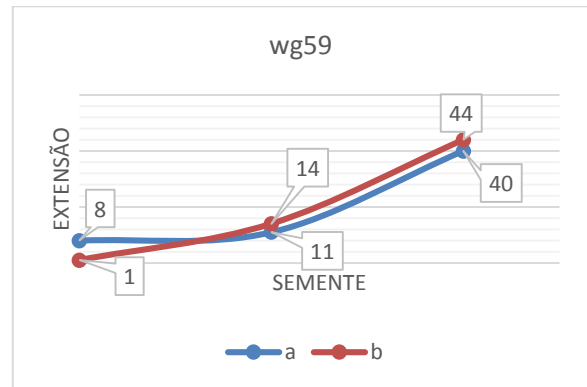
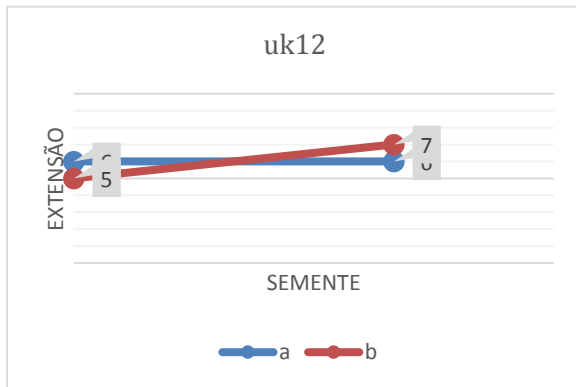


Conclusão

Analisando os resultados obtidos, observa-se que quando as sementes são escolhidas espalhadas uma das outras há uma distância (diferença) menor entre as extensões dos grupos obtidos do que nos casos em que as sementes foram escolhidas próximas umas das outras. Percebe-se que a competição é maior nos casos “b” e que nos casos “a” e se plotarmos as extensões em ordem crescendo, vê-se que quando as sementes são próximas, a curva de crescimento (ordenando as extensões crescentemente) é maior do que no outro caso.

Portando é possível imaginar que caso deseje-se grupos com extensões mais homogêneas, é recomendado que as sementes utilizadas estejam distantes umas das outras.

Os grupos são formados pelos melhores caminhos entre os vértices e a semente os quais estão conectados, sendo assim cada grupo está otimizado para aquela semente a qual está conectado.



Gráficos mostrando a curva obtida a partir da plotagem das sementes utilizadas em cada um dos casos “a” e “b” em função das extensões dos grupos obtidos.

Notas finais

O código fonte pode ser encontrando no GitHub:

<https://github.com/Nobusada/ufscar-2014-graph-theory-task4>

Todos os dados obtidos das simulações e utilizados neste relatório podem ser encontrados em:

<https://github.com/Nobusada/ufscar-2014-graph-theory-task4/tree/master/report-generated-data>

Os arquivos de texto gerados para cada grafo podem ser importados para um editor de planilhas, foi testado com sucesso no *Microsoft Office Excel*, e *OpenOffice Calc* utilizando separação por **tabulação**, abaixo uma tabela com os links dos arquivos já gerados:

<i>Nome Arquivo</i>	<i>extensão</i>	
UK12 Simulation Results	txt	ods
WG59 Simulation Results	txt	ods
USair97 Simulation Results	txt	ods

Post com explicações detalhadas do algoritmo no blog do grupo:

<http://ufscargrafos14a.blogspot.com.br/2014/11/simulando-reconhecimento-de-padroes.html>