HW3: Markov Clustering Algorithms

**Group #: 1**

Tianle Ma, [tianlema@buffalo.edu](mailto:tianlema@buffalo.edu), 50096690

Other members

1. Understanding and implementation of MCL algorithm
2. In your own words, describe what MCL algorithm is, how it works, why it works, and when it works (advantages and disadvantages). (20 points)
3. Pseudo code for **your implementation**. (10 points)
4. Briefly describe how you implemented the algorithms using your preferred language, e.g., pattern design, data structures. (10 points)
5. Results

There are majorly two tuning parameters in MCL algorithms: inflation rate , and expansion rate , (both are default 2). Tune the two parameters, visualize clustering result using Pajek or other software.

1. Calculate modularity for each parametric configuration (inflation, expansion). The modularity is defined Equation (45) on Page 44, <http://dollar.biz.uiowa.edu/~street/graphClustering.pdf>

You should choose a reasonable range of parameters **based on your own experiment**. For example, .

The parameters could be different for each dataset. Some parameters may not even converge.

Show the results in a table for each datasets. (10 \* 3 = 30 points)

1. Present the best clustering result (**specify the “optimal” parameters**, and copy the corresponding **Pajek network partition output** in this report) for each dataset based on network visualization using Pajek or other software. (5 \* 3 = 15 points)
2. Discuss how the two parameters, inflation rate and expansion rate, influence the clustering results, and how to choose the parameters. (15 points)

(Note: modularity is only one metric for assessing your clustering result, it is not comprehensive at all. You can try to implement other metrics if you have time and are interested in doing so. Since it is hard to assess your cluster results quantitatively, you can choose the “best” clustering result based on visualization result).

1. Put all any other considerations or extra work here. Depending on how significant it is, we may or may not consider to give you extra points (at most 10 points) to boost your score. The maximal score is 100 for this homework.