THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY
Department of Computer Science and Engineering
COMP4641: Social Information Networks Analysis and Engineering
Spring 2017 Assignment 1
Due time and date: 9:30am, Mar 27th (Mon), 2017.

IMPORTANT NOTES

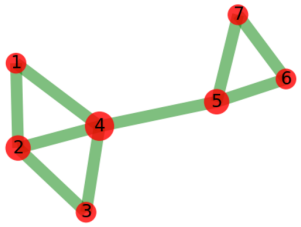
- **Your implementation should not use existing network analysis packages (such as snap or networkx).**
- **Please make sure that your program output exactly follows the format specified in this assignment.**
- **Your grade will be based on the correctness, efficiency and clarity.**
- **Late submission: 25 marks will be deducted for every 24 hours after the deadline.**

Q1. Given a adjacency matrix of some undirected graph, you are required to write a C++/python2.7/matlab program that performs the following two functions:

1. implements the Girvan-Newman algorithm, and outputs the resultant hierarchical decomposition of the network;
2. implements the calculation of modularity, and outputs the corresponding cluster structure.

Implementation Details

For example, given the following graph



- the input file is

```
7
0 1 0 1 0 0 0
1 0 1 1 0 0 0
0 1 0 1 0 0 0
```

```
1 1 1 0 1 0 0
0 0 0 1 0 1 1
0 0 0 0 1 0 1
0 0 0 0 1 1 0
```

  - the first line contains the number of nodes;
  - the remaining lines contain the adjacency matrix

- your program output should have two parts:

  1. the first part outputs the hierarchical decomposition produced by the Girvan-Newman algorithm, in the following format:

     ```
     network decomposition:
     ([1, 2, 3, 4], [5, 6, 7])
     ([1], [2, 3, 4], [5, 6, 7])
     ([1], [2, 3, 4], [5, 7], [6])
     ([1], [2, 3, 4], [5], [7], [6])
     ([1], [3], [2, 4], [5], [7], [6])
     ([1], [3], [2], [5], [4], [7], [6])
     ```

  2. the second part outputs the modularity results, in the following format:

     ```
     2 clusters: modularity 0.3642
     3 clusters: modularity 0.2531
     4 clusters: modularity 0.0926
     5 clusters: modularity 0.0186
     6 clusters: modularity -0.1173
     7 clusters: modularity -0.1543
     optimal structure: ([1, 2, 3, 4], [5, 6, 7])
     ```

Note that graphs different from the example graph above will be used in grading.

Submission Guidelines

The source code should be emailed to **comp4641@hotmail.com** before the deadline. All the submitted files should be contained in a single zip package named like **A1_awangab_12345678.zip** (replace awangab with your ust account and 12345678 your student id). Note that all the codes should be compilable and well-commented (provide enough comments for each key line of code), otherwise you may lose some points if the code is very difficult to understand. Plagiarism will lead to zero point on this assignment.