**Homework ll Community detection instructions**

1. The implementation of community detection, you can work on Python, C++, Java, R, or Other programming language
2. You can use library to implement this homework
3. The homework is due on 6/15/2018

**Evaluation Description**

The scoring metrics is Categorization Accuracy

**Submission Format**

The submission file should contain a header(as below example), id column represent the pair id in the question.txt,prediction column represent this pair is in the same community or not,You should predict by yourself，1=yes，0=no，The data looks like following:

Example:

id,prediction

0,1

1,0

2,0

**File descriptions**

**email-Eu-core.txt** is the historical data of email transmissions between a company for use in drawing

**question.txt** is a question that you need to answer. The id column represents the index. node\_1 and node\_2 represent a pair. After classifying the community, you need to determine whether the pair is in the same community or not.

**question.txt**

* id – pair’s index
* node\_1 – first node in the pair
* node\_2 – second node in the pair

**email-Eu-core.txt**

Each line represent the first node connected to the second node

**method:**

Step1. Use the email-Eu-core.txt to draw the graph(igraph, networkx)

Step2. Use community detection algorithm to classifying the community

Step3. Answer each pair in the question.txt is in the same community or not

Step4. Notice that the submission file only need the pair id and prediction result

Following is a link of kaggle competition:

<https://www.kaggle.com/t/f7b345bbec2c4cc98c237e54c712ba36>