# Graph and Web mining – Project / Paper

## 1. Project description

In the final project the students (1 or 2 students) will implement one of studied graph mining algorithms and will test it on some public available data. In addition to the software, a report detailing the problem, algorithm, software structure and test results is expected.

For testing the project we will provide two types of datasets:

a) Synthetic datasets (both transactions and single graph – use one of them according to the implemented algorithm)

b) Links to real datasets

It is best to test the implementation on both types of datasets.

As for real datasets, you do not have to use the suggested links but can also provide your own datasets

Topics for projects:

You may implement one of the studied algorithms (e.g. FSG, gSpan, Path, gIndex, Subsea ) or another relevant algorithm mentioned in the lecture. If you like to implement another algorithm not mentioned in the lecture (e.g. one from the graph mining seminar ) you should get my approval.

## 2. Paper description

In the final paper the student (1) will review at least two recent papers in graph mining not presented in class in detail and explain them in detail. It

is best that the two papers will be related, so use Google scholar to find related papers. For example assume I like to do a paper on algorithms for closed graphs, then I found in Google scholar the original and the followup paper:

a) CloseGraph :mining closed frequent graph patterns, Yan, J Han - Proceedings of the ninth ACM SIGKD, 2003

b) Mining closed relational graphs with connectivity constraints Yan, X Zhou, J Han - Proceedings of the eleventh ACM SIGKDD, 2005

Note that if the base paper was not covered in detail in class, its ok to include it in the survey (this is the case above), however, if the base paper was covered in detail in class you should only include a brief description of it and use another two papers.

Comment:

In case you reviewed some papers for the graph mining seminar, you cannot use them again for this course.