

Graph and Web mining - Homework

Solve all questions. To be submitted by individuals. Homework due Thursday May 27th.

- 1) Use the three graphs in Slide #6 in Chp2a.ppt. Show all frequent patterns which have support of 66% (2 out of 3). Show the details of the computation for both gSpan and FSG algorithms. You may change colors with labels. You may change all edges to solid ones without labels. To save time you may draw the sub-graphs by hand and scan them.
- 2) Solve Q9.1 in Han
- 3) Solve Q9.3 in Han

Below is a copy of Han's questions

- 9.1 Given two predefined sets of graphs, *contrast patterns* are substructures that are frequent in one set but infrequent in the other. Discuss how to mine contrast patterns efficiently in large graph data sets.
- 9.2 Multidimensional information can be associated with the vertices and edges of each graph. Study how to develop efficient methods for mining *multidimensional graph patterns*.
- 9.3 *Constraints* often play an important role in efficient graph mining. There are many potential constraints based on users' requests in graph mining. For example, one may want graph patterns containing or excluding certain vertices (or edges), with minimal or maximal size, containing certain subgraphs, with certain summation values, and so on. Based on how a constraint behaves in graph mining, give a systematic classification of constraints and work out rules on how to maximally use such constraints in efficient graph mining.