Ludwig-Maximilians-Universität München
Institut für Informatik
Dr. Eirini Ntoutsi
PD Dr. Matthias Schubert

Knowledge Discovery in Databases II
WS 2015/2016

Übungsblatt 12: Graph Comparisons and Frequent Subgraphs

Aufgabe 12-1   Graphlet Kernels

Graphlets are a means to transform the topology of a graph into a vector space. In this context, a graphlet is a sub-graph with exactly k vertices. k-graphlets differ from each other in topology, i.e. the configuration of vertices. The graph $G(V, E)$ is expressed as a vector through the number of individual k graphlets in $G(V, E)$.

(a) How many graphlets of size 4 does a graph with n vertices contain? What is the time complexity of extracting them?

(b) How expensive is comparing two graphs using graphlets of size 4?

(c) How can graphlet vector comparison be sped up/improved?

(d) What problems emerge with larger values of k? What are the consequences of using an vertex-labeled graph $G(V, E, L)$?

Aufgabe 12-2   Random Walk Kernel

In this excercise, we want to implement the random walk kernel of two graph having the adjacency matrixes A and B and the label vectors L_A and L_B for the nodes of A and B.

(a) Implement a method product_graph for computing the product graph.

(b) Implement the random walk kernel for A and B.

Aufgabe 12-3   Canonical Labeling

Calculate the Canonical Labeling of the following graphs: