NED: An Inter-Graph Node Metric Based On Edit Distance

Haohan Zhu, Xianrui Meng, George Kollios

Abstract: Node similarity is a fundamental problem in graph analytics. However, node similarity becomes more important in a new graph based on the knowledge of an existing graph (transfer learning on graphs) and has applications in biological, communication, and social networks. In this paper, we propose a novel distance function for measuring inter-graph node similarity with edit distance, called NED. In NED, two nodes are compared according to their local neighborhood structures which are represented as unordered k-adjacent trees, without relying on labels or other assumptions. Since the computation problem of tree edit distance on unordered trees is NP-Complete, we propose a modified tree edit distance, called TED*, for comparison among neighborhood trees. TED* is a metric distance, as the original tree edit distance, but more importantly, TED* is polynomially computable. As a metric distance, NED admits efficient indexing, provides interpretable results, and shows to perform better than existing approaches on a number of data analysis tasks, including graph de-anonymization. Finally, the efficiency and effectiveness of NED are empirically demonstrated using real-world graphs.

TED* vs TED & GED
A Practical and Good Approximation to Tree Edit Distance

Nearest Neighbor & Top-K Queries
Better Performance and Query Quality

Case Study: Graph De-Anonymization
Higher Precision than Feature-Based Similarities