Summary on Mining Graph Data

July 2, 2015

1 Introduction

1. Much success has been achieved when the data to be mined represents a set of indepedent entities and their attributes. However, in most domains, there is interesting knowledge to be mined from the relationships between entities. This relational knowledge may take many forms from periodic patterns of transactions to complicated structural patterns of interrelated transactions. Extracting such knowledge requires the data to be represented in a form that not only captures the relational information but supports efficient and effective ining of this data and comprehensibility of the resulting knowledge. Relational databases and first-order logic are two popular representations for relational data, but neither has sufficiently supported the data mining process.

2. The graph representation is a collection of nodes and links betwen nodes, does support all aspects of the relational data mining process. The graph can easily repesents entities, their attributes, and their relationships to other entities. Using graphs for representing the data and the mined knowledge supports direct visualization and increased comprehensibility of the knowledge.

3. Mining graph data, sometimes called graph-based data mining, is the extraction of novel and useful knowledge from a graph representation of data. The most natural form of knowledge that can be extracted from graphs is also a graph.

2 Social Network Analysis

1.