Description-Oriented Community Detection using Exhaustive Subgroup Discovery

Martin Atzmueller, Stephan Doerfel, Folke Mitzlaff

Summarised by Laura Rodríguez Navas

January 9, 2020

The main contribution of the article

This article is focuses on the description-oriented community detection, using exhaustive subgroup discovery to provide an structurally valid and interpretable description for communities (i. e., a set of nodes), with a graph's structures and descriptive features of a graph's nodes. Also, presents the COMODO algorithm that use an efficient branch and bound method with appropriate pruning techniques to provide the exhaustive subgroup discovery. The article also shows an implementation of COMODO's algorithm that is evaluated by a selectable community quality measure using five real world datasets, and analysed the results.

As a personal opinion, I consider the approach of this article so interesting and introduces an innovation because can solve problems that in contrast to global approaches is usually not achieved by classical community mining methods, that consider the nodes of a network as mere strings or ids. Also the discovered algorithm it is very relevant since can easily be incorporated into practical applications and is not limited to any systems and can be applied to any kind of graph-structured data.

The article structures

The article has the structure that an article should have as a general scheme: the introduction, the methodology, the results of the investigation and the conclusion, as well as a list of references.

First introduce description-oriented community detection and present the COMODO algorithm. Also, consider a number of standard community quality measures. Specifically, the article consider a three standard community quality measures: The segregation index, the inverse average ODF (out degree fraction) and the local modularity. BLA BLA

The remainder of the article summarizes basics of subgroup discovery, and provides general notions of graphs and community mining measures. The preliminaries BLA BLA

Next, introduces the proposed approach for description-oriented community detection and presents a number of optimistic estimates for standard community evaluation functions. ???

After that, discusses related work.

For demonstrating the effectiveness and validity of the presented approach, provides experiments using five data sets and discusses their results in the context of the three real-world applications.

Finally, concludes the paper with a summary and directions for future research.

Table of contents for a presentation

Quantes slides. Quin seria l'índex de continguts. Per a cada slide que hi posaria.