Book Review

Graph Mining
Laws, Tools and Case Studies

Deepayan Chakrabarti
Christos Faloutsos
Synthesis Lectures on Data Mining and Knowledge Discovery
Morgan & Claypool Publishers, 2014

ISBN: 9781608451159

Graphs in millions appear in a wide range of computational domains.

The graph nodes and the links exceed several billions and their analyses thus become more pertinent in computational analysis. Realizing this potential the authors Deepayan Chakrabarti and Christos Faloutsos came up with an extensive synthesis note on mining graphs. The growing literature on it reflects such requirements.

Twenty one chapters in four parts constitute the compendium of the contributions by the authors. The first part on Patterns and Laws begins with a pesenatation about the ground reality on 'real nodes and links'. Despite the fact that existing networks have more numbers, the really existing and used ones are less which calls for a relevance with Zipf' law. The power laws related to the distributions with computational issues and real life examples are presented in a chapter in the first part. Citing is the basic premise in link and network building. In the chapter on Patterns in Evolving graphs, the authors have studied the patterns the systems connecting each others.

The connected graphs change over time and the weighing enables to measure distributions and behaviour. The chapter on Patterns on Weighted Graphs provide laws on datasets and weight power laws.

Using degree distribution as illustration, the authors in the next chapter, have estimated the slope of power law by computing the power-law exponent. They have shown accurately the measurement of deviations from power laws. The summary of patterns presented by them in the last chapter of the first part clearly documented the skewed distribution in power law estimation, with further observations of shrinking diameter and higher number of triangles.

The Part II starts with a basic background discussion on graph generators where the authors have viewed that the graph generators can lead to gain understanding and insight on graph creation by expressing the processes providing the development of the patterns. The taxonomy of graph generation models is detailed with literature support. In the chapter 9, the authors have discussed the growth of preferential attachment models extensively. The geographical constraints have impact on real graphs where the random graph and preferential models have neglected, the authors claim. These models are listed and discussed about the incorporation of the geographical information in the chapter on Incorporating Geographical Information. Besides the comparison studies are also addressed.

Through the help of the *RMat* generator the authors have balanced three issues, viz., model parameters, realism and efficiency. The RMat generated graphs are discussed by them with the different structures. In the twelfth chapter on Graph Generation by Kronecker Multiplication, the authors have provided a brief description to offset the unequal distribution of adjacency matrix followed by a summary in the next chapter.

The chapter 14 is a review of the tools and tensors. It is intended mainly to address the Singular value decomposition and the page rank arithmetic. The tensors are expressed in the later chapter. A specifically designed chapter on case studies on sub-graphs. The last part is occupied by related work and the bibliography of the literature forms the end of the book.

The book is technical in nature and explains many intricacies with adequate examples about the graph mining.

Hathairat Ketmaneechairat
King Mongkut's University of
Technology, North Bangkok
Bangkok, Thailand