

A Nature-Inspired Metaheuristic Cuckoo Search Algorithm for Community Detection in Social Networks

Ramadan Babers, Helwan University, Helwan, Egypt, and Scientific Research Group in Egypt (SRGE), Egypt
Aboul Ella Hassanien, Cairo University, Giza, Egypt, and Scientific Research Group in Egypt (SRGE), Egypt

ABSTRACT

In last few years many approaches have been proposed to detect communities in social networks using diverse ways. Community detection is one of the important researches in social networks and graph analysis. This paper presents a cuckoo search optimization algorithm with Lévy flight for community detection in social networks. Experimental on well-known benchmark data sets demonstrates that the proposed algorithm can define the structure and detect communities of complex networks with high accuracy and quality. In addition, the proposed algorithm is compared with some swarms algorithms including discrete bat algorithm, artificial fish swarm, discrete Krill Herd, ant lion algorithm and lion optimization algorithm and the results show that the proposed algorithm is competitive with these algorithms.

KEYWORDS

Community Detection, Cuckoo Search Algorithm, Social Networks, Social Networks Analysis

1. INTRODUCTION

Today, the fast way to contact person or organization became over the Internet by using social networks applications, social interactions over such applications became basic part of our daily activities. Social networks defined as set of nodes and edges, where nodes represent individuals or objects and edges represent the relationships or interactions between nodes (Xu et al. 2016; Rossi et al. 2015). Community is a term used to identify the tendency group of individuals to form condensed interactions and relations between them than individuals not belonging to that group (Bansal et al. 2011; Alzahrani & Horadam, 2016). Community structure used in networks analysis to achieve a specific target (Barabási et al. 2002), discover complex networks topology or discover valuable insights from the networks (Hassan et al. 2015). It used to detect groups and subgroups which have similar features (Xu et al. 2015; Kawamoto & Rosvall, 2015). There are several areas affected by community structure such as bio-informatics, sociology and information science (Babers et al. 2015). Several researches analysis complex networks and discover insights in it (Banati & Arora, 2015; Cao et al. 2015; Yu et al. 2015; Azar & Vaidyanathan, 2015; Emary et al. 2014a).

Traditional methods such as graph partitioning used to detect community within networks by dividing it based on predefined size. Spectral clustering method is based on similarity matrix and merging similar groups used by hierarchical clustering method (Choudhury & Paul). Community detection (CD) in social networks is one of the important researches in social networks and graph analysis (Wang et al. 2015). CD based on node centric which nodes in such community satisfy certain

properties as degree and reachability, also each group in community satisfy certain condition as density of interaction within nodes in the same group (Azar & Vaidyanathan, 2015; Hassanien et al. 2014).

The network defined as group of objects or entities interact with each other and linked by relationships called edges. Nodes represent anything as person, cell phone, PC's IP, car and product, the relationships between nodes are the connection linking them. There are different types of connections such that direct and indirect relation. The powerful tool for studying the structure of complex systems is network analysis (Albert & Barabási, 2002; Newman, 2003) which is useful in understanding the interactions and relationships established among nodes (Watts et al. 2002; Lusseau & Newman, 2004; Flack et al. 2006; Hassanien et al. 2015), also in describing and predicting the behavior of nodes.

Social Networks consist of two basics parts which are individuals (nodes) and relationships (edges) and can be defined formally using graph structure, $N = (O, V)$ where N is the set of nodes and relations or edges are V (Hassan et al. 2015). The number of nodes denoted by $n = |O|$ and the number of edges by $m = |V|$. Given a set of nodes $\{x_1, x_2, \dots, x_n\}$, the target is dividing those nodes into groups such that nodes in the same group are related and similar to each other and nodes in different group are dissimilar to each other. The challenge is to detect the number of communities within such complex networks, where each community is fulfilled objective quality function.

$$\min F(S), \quad S \in \Omega \quad (1)$$

Where $F(S)$ is an objective quality function and S is assuming that the quality measure minimized without lose in the context. Measures used for social networks analysis are different and depend on the network size. For example, the measures used: centrality, association index and node degree (Salama et al. 2012; Maharani & Gozali, 2015).

This paper presents a new community detection algorithm based on cuckoo search algorithm (Yildiz, 2013) which is a metaheuristic search algorithm. We used four social networks data sets as benchmark to apply proposed algorithm. The results are compared to other community detection algorithms. The rest of this paper is organized as: Section 2 states a literature review of the community detection problem and related works. Section 3 presents cuckoo search algorithm. Section 4 presents the proposed algorithm. Section 5 states the experimental results and a comparison analysis. The conclusion and future works of this work is states in section 6.

2. RELATED WORK

The previous work was dependent on graph partitioning, the problem was in dividing the nodes in graphs of predefined size. The number of clusters need to be specified. This problem can be solved by optimize the partitioning with choosing different measures as ratio cut (Wei & Cheng, 1989) and the normalized cut (Shi & Malik, 2000). As shown algorithms for graph partitioning are not suitable for community detection in social networks purpose as it is need to presented the input number of groups and its sizes.

The graph may have a hierarchical structure which displays several levels of clusters. Hierarchical clustering algorithm (Friedman et al. 2001) use multilevel structure of the graph and it is used in social networks analysis, engineering, biology, etc. The classification in hierarchical clustering depends on the similarity measure between nodes in the network. Hierarchical clustering techniques study the nodes identification within groups with high similarity, which can be classified in two categories:

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/a-nature-inspired-metaheuristic-cuckoo-search-algorithm-for-community-detection-in-social-networks/169751

Related Content

Adoption Process for VoIP: The UTAUT Model

Eduardo Esteva-Armida and Alberto Rubio-Sanchez (2012). *International Journal of E-Services and Mobile Applications* (pp. 15-31).

www.irma-international.org/article/adoption-process-voip/73714

Analysing Time-Related Properties of Service-Oriented Systems

Laura Bocchi, José Fiadeiro, Monika Solanki, Stephen Gilmore, João Abreu and Vishnu Vankayala (2012). *Handbook of Research on Service-Oriented Systems and Non-Functional Properties: Future Directions* (pp. 36-60).

www.irma-international.org/chapter/analysing-time-related-properties-service/60881

Factors Influencing the Adoption of Mobile Banking in India

Sindhu Singha and R. K. Srivastava (2014). *International Journal of E-Services and Mobile Applications* (pp. 1-15).

www.irma-international.org/article/factors-influencing-the-adoption-of-mobile-banking-in-india/119701

Value-Oriented Specification of Service Systems: Modeling the Contribution Perspective of Enterprise Networks

João Pombinho, David Aveiro and José Tribolet (2015). *International Journal of Information Systems in the Service Sector* (pp. 60-81).

www.irma-international.org/article/value-oriented-specification-of-service-systems/121616

Applying Dynamic Causal Mining in Health Service Management

Yi Wang (2010). *Intelligent Systems in Operations: Methods, Models and Applications in the Supply Chain* (pp. 233-252).

www.irma-international.org/chapter/applying-dynamic-causal-mining-health/42663