

# Chapter 9

## A Discrete Black Hole Optimization Algorithm for Efficient Community Detection in Social Networks

**Mohamed Guendouz**

*Dr. Moulay Tahar University of Saida, Algeria*

### ABSTRACT

*In recent years, social networks analysis has attracted the attention of many researchers. Community detection is one of the highly studied problems in this field. It is considered an NP-hard problem, and several algorithms have been proposed to solve this problem. In this chapter, the authors present a new algorithm for community detection in social networks based on the Black Hole optimization algorithm. The authors use the modularity density evaluation measure as a function to maximize. They also propose the enhancement of the algorithm by using two new strategies: initialization and evolution. The proposed algorithm has been tested on famous synthetic and real-world networks; experimental results compared with three known algorithms show the effectiveness of using this algorithm for community detection in social networks.*

### INTRODUCTION

In last years, the use of online social networks has increased significantly due to the increasing use of Internet in general and the large availability of various kinds of connected devices such as: desktop computers, laptops and smartphones. Every day millions of people connect to social networking websites to communicate with each other and to share information, which made these websites very popular among others, Facebook and Twitter are the most famous ones. Facebook which was created in 2008 has now more than one billion active users from around the world, this huge number of users makes it the most popular social networking website in the world. Twitter known as the first microblogging website created earlier in 2006 has now more than 200 million active user. The huge amount of data and information shared between people on these websites has made social networks analysis a very important research field.

DOI: 10.4018/978-1-5225-3004-6.ch009

Social Networks are generally represented by graphs where nodes represent individuals and edges represent the relation between them, for example, in Facebook individuals are people and the relation between them is friendship. A social network is usually found to be grouped into a set of interconnected groups which means that individuals who are very similar between them have strong possibility to be in a same group, these groups are generally known as communities, one can classify communities into two types: Explicit and Implicit ones (Kadushin, 2012). In an explicit community members know that they belong to it and nonmembers know who its members are, Facebook groups are famous example of this type of communities where members can post messages and images, comment on other's posts and see activities of other people. In the contrary, an implicit community is a community where individuals interact with each other in a form of unknown connections which means that members of such communities do not know that there is a community and that they are members of it, for example people who use same hashtags in their tweets or follow the same celebrities in Twitter form implicit communities. In social networks, researchers are typically interested in finding implicit communities since explicit ones are already known and identified; this task is known as community detection or community discovery.

Community detection is defined as the process of identifying interconnected groups also known as clusters or partitions in a network, it assigns each individual in the network to his/her suitable group. Community detection algorithms are often provided with a graph  $G(V, E)$  where nodes  $(V)$  represent individuals and edges  $(E)$  represent relations between individual, Formally, the task of community detection is to find a set of communities  $\{C_i\}_{i=1}^n$  in  $G$  such that  $\bigcup_{i=1}^n C_i \subseteq V$ . Community detection algorithms are not specific only to detect communities inside social networks, they have many applications in different domains such us: Biology (Girvan & Newman, 2002), Social (Lozano et al., 2007) and Communications (Faloutsos et al., 1999). In fact, they can be applied to any complex problem that takes the form of a complex network; social networks are a specific type of complex networks. However, these algorithms are extensively used on social networks because of the important applications that can be done, for instance, community detection algorithms are widely used by marketing agencies to find potential clients among millions of people or to detect group of people who have the same food taste to recommend for them a special food product or may be a specific restaurant.

The problem of community detection in complex networks and especially in social networks has known in the last years an enormous interest by many researchers from different domains, several methods for community detection have been proposed and applied on many real world problems. One can classify these methods into different types (Fortunato, 2010); traditional methods can be classified into two groups: Graph Partitioning and Hierarchical Clustering. Graph Partitioning methods cluster a network into a predetermined number of communities, usually with equal size, these methods require the number and the size of communities before partitioning, the Kernighan-Lin algorithm proposed in (Suaris & Kedem, 1988) is one of the earliest methods for graph partitioning. Hierarchical Clustering methods cluster a network into groups of nodes based on their similarity which means that similar nodes are grouped into communities according to a similarity measure. Cosine similarity, Hamming distance and Jaccard (Singhal, 2001) index are frequently used measures. These methods are classified into two categories:

1. **Agglomerative Algorithms:** Initially each node represents a partition of its own, then partitions are successively merged until the desired network partition structure is obtained.

19 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/chapter/a-discrete-black-hole-optimization-algorithm-for-efficient-community-detection-in-social-networks/197700](http://www.igi-global.com/chapter/a-discrete-black-hole-optimization-algorithm-for-efficient-community-detection-in-social-networks/197700)

## Related Content

---

### Query Recommendation for Improving Search Engine Results

Hamada M. Zahera, Gamal F. El-Hady and W. F. Abd El-Wahed (2011). *International Journal of Information Retrieval Research* (pp. 45-52).

[www.irma-international.org/article/query-recommendation-improving-search-engine/53126](http://www.irma-international.org/article/query-recommendation-improving-search-engine/53126)

### Proximity-Based Good Turing Discounting and Kernel Functions for Pseudo-Relevance Feedback

Ilyes Khennak and Habiba Drias (2017). *International Journal of Information Retrieval Research* (pp. 1-21).

[www.irma-international.org/article/proximity-based-good-turing-discounting/181723](http://www.irma-international.org/article/proximity-based-good-turing-discounting/181723)

### Query Sense Discovery Approach to Realize the User's Search Intent

Tarek Chenaina, Sameh Nejia and Abdullah Shueb (2022). *International Journal of Information Retrieval Research* (pp. 1-18).

[www.irma-international.org/article/query-sense-discovery-approach-to-realize-the-users-search-intent/289609](http://www.irma-international.org/article/query-sense-discovery-approach-to-realize-the-users-search-intent/289609)

### A Study on Models and Methods of Information Retrieval System

Manisha Malhotra and Aarti Singh (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications* (pp. 246-265).

[www.irma-international.org/chapter/a-study-on-models-and-methods-of-information-retrieval-system/198553](http://www.irma-international.org/chapter/a-study-on-models-and-methods-of-information-retrieval-system/198553)

### Bates' Berrypicking Model (1989, 2002, 2005)

Linda L. Lillard and YooJin Ha (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications* (pp. 65-76).

[www.irma-international.org/chapter/bates-berrypicking-model-1989-2002-2005/198545](http://www.irma-international.org/chapter/bates-berrypicking-model-1989-2002-2005/198545)