International Journal of Hybrid Information Technology
Vol.7, No.2 (2014), pp.1-8
http://dx.doi.org/10.14257/ijhit.2014.7.2.01

A Novel Local Maximum Potential Point Search Algorithm for Topology Potential Field

Zhixiao Wang, Zhaotong Chen, Ya Zhao and Qiang Niu

China University of Mining and Technology, School of Computer Science and Technology, JiangSu China, 221116

softstone416@163.com

Abstract

Topology potential field is a novel model to describe interaction and association of network nodes, which has attracted plenty of attention in community detection, node importance evaluation and network hot topics detection. The local maximum potential point search is a critical step for this research. Hill-climbing is a traditional algorithm for local maximum point search, which may leave out some local maximum potential points, and search performance is greatly influenced by initial node sequence. Based on the detailed analysis of local maximum potential points' characteristics, this paper presents a novel local maximum potential point search algorithm. The results of simulation experiments showed that the new algorithm has better performance than the traditional hill-climbing method. It can find all local maximum potential points with high search efficiency.

Keywords: *Complex network; topology potential field; local maximum potential point*

1. Introduction

The classical field was used to describe non-contact interaction between matter particles. Later on, people abstracted it as a mathematical model to describe non-contact interaction between objects, and to depict distribution law of physical quantities.

In complex networks, nodes are not isolated, there exist relationship between nodes linked by edges. In this light, the topology potential field theory was introduced into complex networks to describe interaction and association among network nodes [1], and reveal the underlying characteristics of underlying importance distribution [2].

Topology potential is defined as the differential position of each node in the topology, that is to say, the potential of each node in its position. The potential of any node totally depends on the topological structure of its surroundings, the corresponding potential is known as topological potential. If a node in the complex network has more surrounding nodes, and then its potential is greater.

In a given network $G=(V,E)$, where $V=\{v_i \mid i=1,...,n\}$ is a set of nodes, n is the number of nodes in the complex network , $E=\{(v_i,v_j) \mid v_i,v_j \in V\}$ is a set of edges, and $|E|=m$. The topology potential of any node $v_i \in V$ can be computed as follows:

ISSN: 1738-9968 IJHIT

$$\varphi(v_i) = \sum_{j=1}^{n} [m_j \times e^{-(\frac{d_{ij}}{\sigma})^2}] \quad (1)$$

where, $\varphi(v_i)$ is the topological potential of node v_i; d_{ij} is distance between node v_i and v_j; m_j is the mass of node v_j; σ is an impact factor used to control the affecting hops of node in process of community expansion, which can be obtained by calculating the minimum entropy potential [1].

Just like other topology measurements in complex networks, such as degree, average path length, diameter etc, this index reflects the ability of each node to be influenced by other nodes in the network, and vice versa [2]. High potential node will attract nodes whose potentials are relatively low. There are some local maximum potential nodes in topology potential field. Moreover the potential distribution characterizes the structure and behavior of nodes in the topology space.

Recently years, topology potential field has attracted plenty of attention in community detection, node importance evaluation and network hot topics detection in complex networks area [1-10]. In these applications, the local maximum potential node search is a key step. On the one hand, this step requires high precision, for the number of local maximum potential nodes directly determine the final results. On the other hand, this step requires high efficiency. It is especially meaningful for large-scale complex networks.

Hill-climbing is a traditional method for local maximum potential node search, which may leave out some local maximum potential points, and search performance has close relationship with initial point selection. Based on the detailed analysis of local maximum potential points' characteristics, this paper presents a novel local maximum potential point search algorithm. The results of simulation experiments showed that the new algorithm has better performance than the traditional hill-climbing method. It can find all local maximum potential points with high search efficiency.

This paper is organized as follows: Section 2 describes related works; Section 3 shows the traditional hill-climbing algorithm; Section 4 analyzes the characteristics of local maximum potential node, and puts forward a novel local maximum potential node search algorithm named LMPS; Section 5 is simulation experiment and results; and section 6 comes to the conclusion of this paper.

2. Related works

Topology potential field has attracted plenty of attention in community detection, node importance evaluation and network hot topics detection in complex networks area.

Inspired from physical idea, reference [2] introduced topology potential field into complex network to describe the relationship among nodes being linked by edges and to reveal the general characteristic of underlying importance distribution. Topology potential can evaluate interaction ability of node and locality of complex networks structure.

Reference [3] put forward a community discovery algorithm based on topological potential. Each community is regarded as a local high potential area, the community structure in the complex network can be uncovered by detecting all local high potential areas margined by low potential nodes.

Reference [1] proposed a topology-potential-based method, which divides networks into separate communities by spreading outward from each local important element and extracting its neighbors within the same group in each spreading operation. The method can not only classify different types of nodes but also detect community structure more effectively.

Reference [4] and [5] proposed an overlapping community discovery method based on topology potential, The method not only considers the spread of the uncertainty of community identity of the overlapping nodes in the network，but also realizes a quantified representation, *i.e.*, uncertainty measure, of the community identity of the overlapping nodes.

Aiming at some inadequacies of the topology potential theory and its method, such as ambiguous application scope and excessively sparse overlapping nodes, reference [6] proposed a variable scale network overlapping community identification method based on identity uncertainty. The method identifies communities effectively by proposing an identity uncertainty measure of overlapping nodes and an idea of variable scale community.

Reference [7] proposed a novel approach to discover community membership in complex networks with node topology potential, along with the experiment on complex biological networks. Reference [8] and [9] presented node ranking algorithms for complex networks based on topology potential, which can reflect the importance of nodes more precisely. Reference [10] put forward an internet hot topics detection algorithm based on topology potential.

In above works, the local maximum potential node search is a key step. Based on the detailed analysis of local maximum potential points' characteristics, this paper presents a novel local maximum potential point search algorithm. The results of simulation experiments showed that the new algorithm has better performance than the traditional hill-climbing method. It can find all local maximum potential points with high search efficiency.

3. Hill-climbing algorithm

The traditional hill-climbing algorithm finds local maximum potential nodes by scanning the whole complex network, following the direction where topological potential values get highest [3].

The key steps of traditional hill-climbing algorithm are shown as follow.

(1) initialization. all nodes are set "unvisited".

(2) We randomly choose an "unvisited" node, climb up following the direction where topological potential values get highest, and mark along nodes "visited". The corresponding node of peak position is the local maximum potential node.

(3) In the climbing process, if we meet "visited" node, then jump to step (2).

(4) repeat step (2) and step (3), until the whole complex network is scanned.

(5) output all local maximum potential nodes.

Hill-climbing algorithm may leave out some local maximum potential points, and search performance has close relationship with initial point selection. Take the Zachary karate club network as an example. The network has 34 nodes and 78 edges, and it describes the social interactions between members of a karate club at American university [11]. The Figure 1 shows the structure of the Zachary karate club network.

Suppose node 9 is select as the initial search point, which has five neighbors: node 1, node 3, node 31, node 34 and node 33. The corresponding topology potential of these five nodes are 7.2343, 5.2138, 2.9868, 7.3534 and 5.8195. In these five nodes, the topology potential of node 34 is the biggest, according the Hill-climbing algorithm, we climb from node 9 to node 34, and all neighbors of node 9 are set "visited". Thus, we find a local maximum potential node, *i.e.*, node 34. Because node 1 has be set "visited", we will never revisited it again. Figure 1 shows us that node 1 has 15 neighbors, and the topology

potential of node 1 is higher than that of neighbors'. That is to say, node 1 is another local maximum potential node, but the hill-climbing algorithm leave this node out.

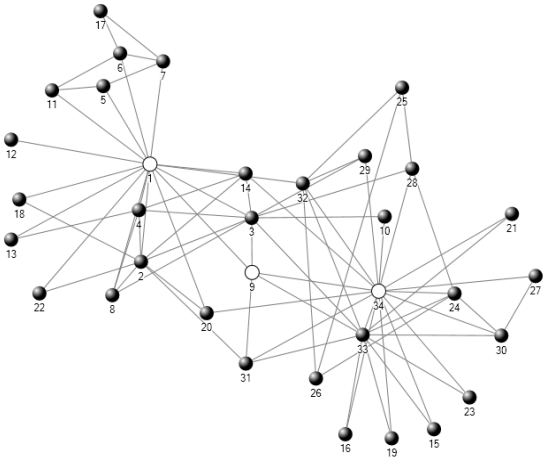

Figure 1. Zachary karate club network

4. LMPS algorithm

Based on the detailed analysis, we found some important characteristics of local maximum potential node.

Characteristic 1:

Suppose $v \in V$, V' represents the set of v's neighbor nodes, if $\forall v' \in V' \rightarrow \varphi(v') < \varphi(v)$, then v is local maximum potential node.

Characteristic 2:

Suppose $v \in V$, V' represents the set of v's neighbor nodes, if $\exists v' \in V'$, where $\varphi(v') > \varphi(v)$, then v is absolutely not a local maximum potential node.

Above characteristics show that, in complex networks, local maximum potential nodes are the nodes with higher potential value than all their neighbor nodes'. We can utilize these characteristics to directly determine whether a node is a local maximum potential node or not.

This paper puts forward a novel local maximum potential node search algorithm named LMPS. The main idea of this algorithm is: scanning all nodes in the complex network, choose the node with higher potential value than all its neighbor nodes' as local maximum potential node, and then add it in the queue of local maximum potential node.

The LMPS algorithm is shown as follows.

(1) initialization. all complex network nodes are set "unvisited".

(2) choose an "unvisited" node v_i, and compare the potential value of v_i with its neighbor nodes'. If the potential value of v_i is higher than all neighbor nodes', then jump to step (3), otherwise, jump to step(4).

(3) add v_i in the queue of local maximum potential node, mark v_i "visited", and mark all neighbors of v_i "visited".

(4) mark v_i "visited", and mark neighbors with lower potential value than v_i's "visited".

(5) repeat steps (2), (3), and (4), until all nodes in complex network are marked "visited".

(6) output all local maximum potential nodes.

5. Simulation Experiment

In this section, we empirically compared the LMPS algorithm with hill-climbing algorithm. Simulation program implemented with scientific computing software MATLAB in the Windows environments. All complex network data in the experiment come from www-personal.umich.edu/~mejn/netdata.

Firstly, we randomly selected six typical complex networks to evaluate the precision of local maximum potential node search. Table 1 shows the results of each complex network found by two algorithm.

Table 1. The results of local maximum potential node search

Complex network name	Local maximum potential nodes found by LEPS	Local maximum potential nodes found by Hill-climbing
Zachary karate club	1,34	1,34|34
Dolphin social network	15,18,21	15,18|15,21,18|15
Books about US politics	13,85	13,85
American College football	1,25,59,81	1,25,59,81|8,59,25|81,1, 59
Neural network	13,45	13,45|13
Les Miserables	12	12

Table 1 shows us that: the local maximum potential nodes of Dolphin social network and Les Miserables found by LMPS and Hill-climbing were the same. For the Zachary karate club, the Hill-climbing algorithm may miss node 1; for the Dolphin social network, the Hill-climbing algorithm may miss node 18 and 21; for the American College football, the Hill-climbing algorithm may miss node 1 and 25.

Secondly, we carried out 50 times local maximum potential node search for above six typical complex networks, and recorded the results of each search operation. The Figure 2 shows the precision of two algorithms.

From Figure 2, we can see that the precision of LMPS is apparently higher than Hill-climbing. Above results indicate that the Hill-climbing algorithm may leave out some local maximum potential points, and search performance has close relationship with initial point selection.

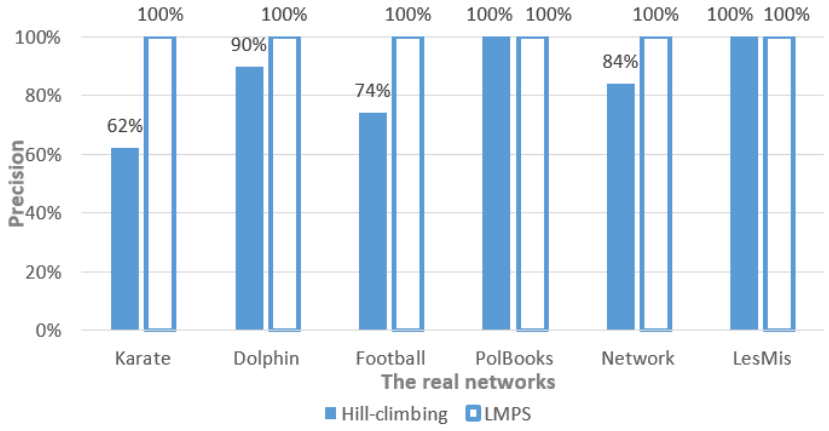

Figure 2. The precision of the two algorithms

Thirdly, we selected "Dolphin social network" complex network to evaluate the search speed of two algorithms. For each algorithm, we carried out 50 times local maximum potential node search, and recorded time consumed each time. Figure 3 shows the results.

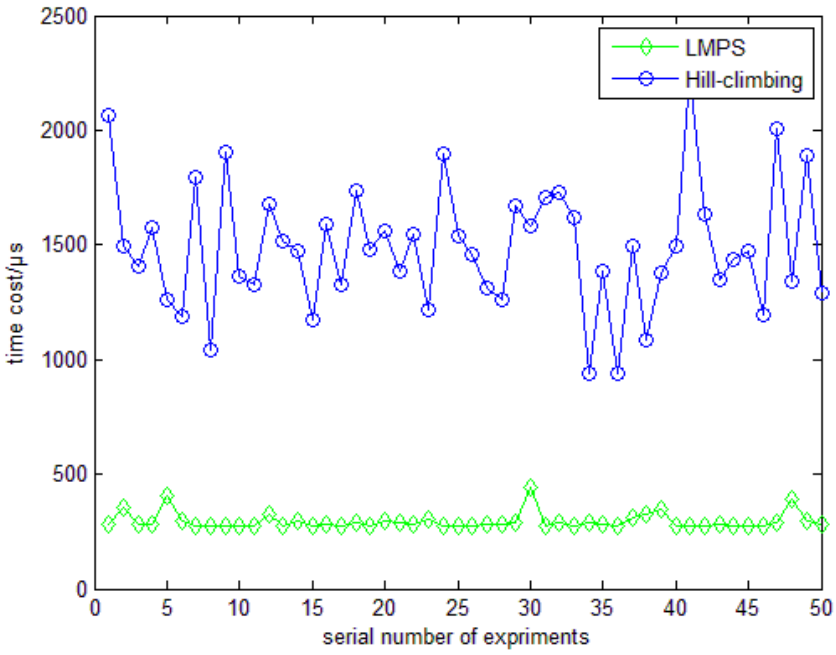

Figure 3. The time consumed by the two algorithms

In Figure 3, the time unit is microsecond. Because of the influence of many factors like computer system running state, the time consumed by the two algorithms does not remain constant, but fluctuates within a certain range. The fluctuation of hill-climbing is more significant than that of LMPS. The reason is: the hill-climbing algorithm uses random

strategy to select the initial node of each climbing, different initial node sequence results in different time cost. But, no matter how it fluctuates, the time consumed by LMPS is significantly less than that of hill-climbing.

6. Conclusions

Topology potential field is a novel model to describe network nodes interaction and association. In recent years, the topology potential theory attracts more and more attention in complex network area. The local maximum potential node search is a critical step for this kind of applications. The hill-climbing is a traditional method to search local maximum potential nodes which may leave out some local maximum potential points, and search performance has close relationship with initial point selection. Based on the detailed analysis of local maximum potential points' characteristics, this paper presents a novel local maximum potential point search algorithm. The results of simulation experiments showed that the new algorithm has better performance than the traditional hill-climbing method. It can find all local maximum potential points with high search efficiency.

At present, almost all research based on topology potential field ignore node difference and assume nodes have same mass. This hypothesis is debatable, and may lead to inaccuracy of topology potential calculation. In future, we will focus on node mass calculation. A node's mass should reflect its importance and influence in the complex network. The more importance a node is, the bigger its mass should be.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (2011QNB23)

References

[1] Y. Han, D. Li and T. Wang, "Identifying Different Community Members in Complex Networks based on Topological Potential", Fronters of Computer Science in China, vol. 5, no. 1, **(2011)**, pp.87-99.
[2] Y. Han and D. Li, "A novel measurement of structure properties in complex networks", Lecture Notes of the Institute for Computer Sciences, Social Information and Telecommunications Engineering vol. 5, **(2009)**, pp. 1292-1297.
[3] W. Gan, N. He, D. Li and J. Wang, "Community discovery method in networks based on topological potential", Journal of Software, vol. 20, no.8, **(2009)**, pp. 2241-2254.
[4] J. Zhang, H. Li, J. Yang, J. Bai and Y. Chu, "Network soft partition based on topological potential". The 6th International ICST Conference on Communications and Networking in China (CHINACOM), August 17-19, **(2011)**, Harbin, China, pp.725-729.
[5] J. Zhang, H. Li, J. Yang, J. Bai, Y. Chu and L. Zhang, "Community Discovery Method with Uncertainty Measure of Overlapping Nodes Based on Topological Potential", Journal of Harbin Institute of Technology (new series), vol. 19, no. 2, **(2012)**, pp. 16-22.
[6] J. Zhang, H. Li, J. Yang, J. Bai, L. Zhang and Y. Chu, "Variable Scale Network Overlapping Community Identification Based on Identity Uncertainty", ACTA ELECTRONICA SINICA, vol. 40, no. 12, **(2012)**, pp. 2512-2518.
[7] L. Xiao, S. Wang and J. Li, "Discovering Community Membership in Biological Networks with Node Topology Potential", Proceedings of IEEE International Conference on Granular Computing, **(2012)**, Hangzhou, China, pp. 231-237.
[8] J. Zhang, H. Li, J. Yang, J. Bai and L. Zhang, "An Importance-sorting Algorithm of Network Community Nodes Based on Topological Potential", Journal of Harbin Engineering University, vol. 33, **(2012)**, pp. 745-752.
[9] L. Xiao, H. Meng and D. Li, "Approach to Node Ranking in a Network Based on Topology Potential", Geomatics and Information Science of Wuhan University, vol. 33, no. 4, **(2008)**, pp. 379-383.
[10] W. Gan, C. Liu and D. Li, "Network Hot Topics Detection Based on Topology Potential", Military Operations Research and Systems Engineering, vol. 24, no. 3, **(2010)**, pp. 41-44.

[11] J. Leskovec, K. J. Lang and A. Dasgupta, "Statistical Properties of Community Structure in Large Social and Information Networks", Proceedings of the 17th international conference on World Wide Web. **(2008)**, pp. 695-704.

