AN ENHANCED ALGORITHM FOR COMMUNITY DETECTION IN LARGE SCALE COMPLEX NETWORKS

M. Hemalatha and N. Kamaraj

Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts and Science, India

Abstract

Complex networks have a large number of nodes and edges, which prevents the understanding of network structure and the discovery of valid information. Community detection is an important issue in studying network structure and network characteristics. It has received widespread attention in many fields. Most existing community detection algorithms obtain the final community structure by analyzing the relationship between each node and surrounding nodes. Starting from a portion of nodes in each community, the corresponding community for each node can be obtained through expansion operations, thereby obtaining the entire community structure. Such strategy can improve the accuracy of community detection algorithms. When solving large-scale combinatorial optimization problems, the traditional ant colony algorithm has a slow convergence rate and tends to fall into local optima. More and more scholars propose relevant optimization algorithms on the basis of classical ant colony algorithm. To overcome premature convergence, adaptively adjusted the pheromone on the path according to the existing solution, which enabled it to escape the local optimal value. To address these challenges, this research proposes an improved optimization method. This approach integrates community detection, multi-group cooperation, pheromone feedback mechanisms and Hybrid Dynamic Pheromone Updating Mechanism to improve exploration efficiency and convergence speed in large-scale TSP problems.

Keywords:

Community Network, Ant Colony Algorithm, Community Detection, Travelling Salesman Problem, Route Relation Network

1. INTRODUCTION

Complex networks consist of a great number of individuals and intricate relationships between them, they can be represented by graphs, in which individuals are abstracted as nodes, and relationships are abstracted as edges. Different types of complex network are included in reality, such as social networks, biological networks, technology networks, traffic networks, and so on. The community structure is one of the significant structural features in complex networks, and finding community structure can help us understand the network structure and discover valuable information [1]. The purpose of community detection is to partition the sample into internally tight and externally sparse communities by considering the similarity between nodes. An effective community detection technique can reveal potential community structures in many scenarios. Such as in social media, the community detection algorithm can find potential friends, recommend products for users and analyze social opinions; in the field of bioinformatics, the community detection algorithm can discover biomolecular with the same function or structure by partitioning biomolecular communities, etc. [2].

Recently, most such studies have been based on homogeneous networks, using network models with a single relationship between nodes and achieving certain results, but ignoring the information implied in the network structure. Furthermore, existing researches seldom discuss the attribute embedding of nodes and generally has weak global feature modeling capabilities. The fact that networks exist in the real world is inherently multiple, i.e., nodes have multiple types of relationships. In the citation network, papers have multiple types of relationships with each other, including authorship relationships between papers and citation relationships between papers. In movie networks, different movies have the same director or use the same actors, thus creating connections between different movies. At present, the research for multilayer network representation learning usually focuses on the fusion of network layers without considering the nodes' attribute features themselves [3]. The use of label information of nodes also poses a challenge in terms of algorithmic complexity. Furthermore, traditional community discovery methods in topology-based multilayer network analysis require complete alignment of nodes between layers, and the complexity of the algorithms is usually high.

Community structure exists in networks, and it gives more information about the network. For instance, we can understand very well the system, which is represented by a network, by finding its community structure and the relationship between communities. In addition, networks can represent many systems like social networks, electric networks, biological networks, etc. It is vital to develop new methods to find network communities [4]. When we analyze networks by studying relationships between nodes, we can get extra information about networks and systems. In general, nodes in the same community have common properties or insure similar tasks in network. A network has parts that are more densely connected than other parts. In other words, the nodes in these parts share many edges between them. These parts of nodes and edges are called communities (clusters). Finally, many studies have been done around networks and how to find community structure.

Although all these existing methods have applicability scenarios and good results in performing community detection in networks, they still have limitations. Most community detection methods use only the network topology to detect communities [5]. However, in many real networks, the distribution of node attributes is correlated with the community structure. Therefore, attribute networks need to consider node attributes to improve the quality of the discovered community structure. The practice of some existing community detection algorithms under attribute networks is to optimize some objective functions to obtain a partition result with the highest global score while ignoring other types of partition results, which tends to lose the global information in attribute networks and makes the results not globally generalizable.

In this paper, the author propose a representative community detection algorithm for attribute-oriented networks, the RCDAA

algorithm, which not only takes into account the connection information and attribute information of the network but also obtains multiple network division patterns by cooperating in the extraction of the representative community structure of the network division set, which more accurately characterizes the global structural information of the network, and also explores the relationship between the attributes of the network and the obtained representative division patterns.

2. RELATED WORK

Most existing local community detection algorithms expand from the given node to a community based on an expansion mechanism that can determine the membership of nodes. The expansion mechanism of previous algorithms is generally based on the relationship between the community under expansion and its neighboring nodes. However, the communities existing in the undetected area of the network also impact these neighboring nodes. It is arbitrary and inefficient to determine the membership of neighboring nodes based only on the relationship between the community and its neighboring nodes while ignoring the relationship between the neighboring nodes and these communities in the undetected area of the network.

Smith et al. [6] address the problem of identifying subgroups of highly connected individuals within a network, is an aspect of social network analysis that is relevant but potentially underutilized in prevention research [6]. Guidance on using community detection methods stresses aligning methods with specific research questions but lacks clear operationalization. Community detection using the Walktrap method best fit the hypothetical case study. The communities derived using the Walktrap method were quite different from communities derived through the other 5 methods in both the number of communities and individuals within communities. Compared to other methods of selecting high-risk groups, the Walktrap produced the most communities that met the hypothetical intervention requirements.

Li et al. [7] analyse structure and the development of communities in dynamic networks plays an important role in analyzing the evolution and development of the entire network [7]. The authors addressed the problem of the division of the community structure in social network big data, as well as ensure the continuity of the community between the current time and previous time period, are issues that need to be explored. This problem can be solved by fusing the three characteristics of temporal variability, stability, and continuity in dynamic social network communities, and by adopting the multi-objective optimization method to detect community structures in dynamic networks. The probability fusion method is added to the initial step of the algorithm to generate suitable network partitions and ensure fast convergence and high accuracy. Two neighboring fusion strategies are proposed that are suitable for communities: the neighbor diversity strategy and the neighbor crowd strategy. These two strategies make different changes to the candidate network partitions. A continuity metric for dynamic community evolution is formulated to compare the similarity of the dynamic network communities of two consecutive time steps. Experiments on synthetic datasets and actual datasets prove that the proposed method in this paper provides better performance than existing methods.

Heydariyan et al. [8] aimed to carried out in community detection for social network analysis because it plays a crucial role in social network systems in today's world [8]. However, most social networks in the real world have complex overlapping social structures, one of the NP-hard problems. The authors introduced a new model for overlapping community detection that uses a multi-objective approach based on a hybrid optimization algorithm. In this model, the Modified Selection Function (MSF) hybrids the algorithms and recovery mechanism, the Slime Mould Algorithm (SMA), the Sine Cosine Algorithm (SCA), and the association strategy. Also, considering that these algorithms have been presented to solve single-objective optimization problems, the Pareto dominance technique has been used to solve multi-objective problems. The proposed model has performed better than the other tested algorithms in comparing the tests conducted by us in all 16 data sets, in the comparisons made with the algorithms proposed in other works in 11 data sets out of 14 data. The set has performed better than competitors. As a conclusion, the findings show that this model performs better than other methods.

Gmati et al. [9] deals the community structure in social networks witnessed a determined effort. In this respect, a different category of social network can be handled, such as, dynamic social networks, social networks with node attributes, etc [9]. The authors introduced a new method to solve this thriving issue in the social network with node attributes. This latter can be represented by a bipartite graph, which consists of a two sets of nodes and edges connecting these nodes. The tendency of people with similar node attributes leads to the hidden information of clusters or communities. A wealthy number of community-detection algorithms have been proposed for bipartite graphs and applied to several domains in the literature. To palliate some of the highlighted shortcomings, we introduce a new approach, called Fast-Bi Community Detection (FBCD), that aims to an efficient community detection in social networks. The main idea of this approach is to explore the set of maximum matching in the bipartite graph in order to reduce the complexity of our algorithm.

Hu et al. [10] is aimed to study and validate an effective method to mine information from the complex network; community discovery has been widely used in social network, financial risk control and other fields [10]. However, the existing community discovery algorithms are not effective in dealing with complex network which always contains fuzzy community structure. The authors proposed an algorithm that defines the connectivity between any nodes in a network and constructs the symmetric doubly stochastic matrix. Then, the algorithm enhances the network by the nonlinear transformation of the eigenvalues of the symmetric doubly stochastic matrix and makes the original fuzzy community structure become clear. Experimental results show that this method can effectively sharpen the community structure of a network and improve the effect of community partitioning.

Jia et al. [11] finds the problem of the lack of user social attribute characteristics in the process of dividing overlapping communities in multilayer social networks [11]. The authors proposed a multilayer social network overlapping community detection algorithm based on trust relationship. By combining structural trust and social attribute trust, we transform a complex multilayer social network into a single-layer trust network. It

obtain the community structure according to the community discovery algorithm based on trust value and merge communities with higher overlap. The experimental comparison and analysis are carried out on the synthetic network and the real network, respectively. The experimental results show that the proposed algorithm has higher harmonic mean and modularity than other algorithms of the same type.

3. RESEARCH METHODOLOGY

Uncovering community structure is one of the most important problems in the field of complex networks. Several methods are developed to aim at detecting disjoint partition. However, in real systems vertices often belong to more than one community [12]. Forcing a node into a single community and not allowing for overlap could affect to detect the true underlying community structures. Therefore, the issue of detecting overlapping communities has become quite popular in the last few years [14]. Community detection is of great importance to understand the structures and functions of networks. Overlap is a significant feature of networks and overlapping community detection has attracted an increasing attention. Many algorithms have been presented to detect overlapping communities.

- Utilize route information from all ants rather than only elite ants to enhance diversity and exploration.
- Construct a route relationship network and use modularity-based community detection to divide routes into smaller subgroups.
- Identify high-quality route segments within communities and integrate them into the pheromone feedback system.
- Implement a multi-group cooperation strategy where superior and inferior ant colonies assist each other.

Compare the proposed method with existing state-of-the-art TSP solvers to evaluate improvements in accuracy, efficiency, and scalability.

3.1 PROBLEM STATEMENT

When solving large-scale combinatorial optimization problems, the traditional ant colony algorithm has a slow convergence rate and tends to fall into local optima [15]. More and more scholars propose relevant optimization algorithms on the basis of classical ant colony algorithm. To overcome premature convergence, adaptively adjusted the pheromone on the path according to the existing solution, which enabled it to escape the local optimal value.

- Existing Ant Colony Algorithms struggle with balancing accuracy and convergence speed in large-scale TSP.
- Traditional approaches prioritize elite ants while neglecting general ants' route information, leading to poor diversity in exploration.
- Conventional pheromone update mechanisms lack adaptability, often leading to stagnation and premature convergence.

Limited cooperation between multiple ant colonies reduces the algorithm's ability to explore optimal routes effectively.

4. PROPOSED SCHEME

When the ant colony algorithm uses multiple populations to solve the TSP problem, the communication and collaboration among multiple populations is a great advantage to improve the performance of the algorithm, but most scholars have achieved the improvement of the algorithm performance by exchanging the pheromones and optimal routes among different populations, and there is little research on the social relationship network when all ants of different populations explore the routes [13]. Therefore, this study proposes a community collaboration strategy to study the route structure relationship network among populations with the help of community detection methods from the route relationship network when different populations explore routes and apply it to solve the TSP problem.
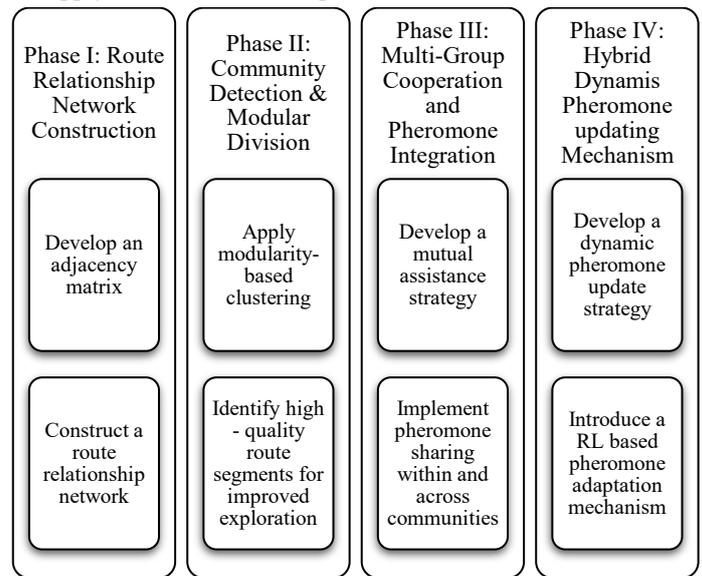

Fig.1. Proposed Work Structure

4.1 ROUTE RELATIONSHIP NETWORK CONSTRUCTION

In this phase, the research focuses on constructing a route relationship network by leveraging the path information collected from all ants. An adjacency matrix is developed to represent the connections between cities, forming the foundation for the optimization process. By considering the route choices of all ants, rather than only elite ones, a more comprehensive and diverse search space is created, allowing for better exploration and solution accuracy.

4.2 COMMUNITY DETECTION AND MODULAR DIVISION

To enhance route optimization, a modularity-based clustering approach is applied to divide the route network into smaller, more manageable communities. This step helps in identifying high-quality route segments, ensuring that ants can focus on the most promising paths. By detecting natural clusters within the route relationship network, the algorithm improves efficiency by reducing unnecessary computations while maintaining accuracy in large-scale TSP instances.

4.3 MULTI-GROUP COOPERATION AND PHEROMONE INTEGRATION

This phase introduces a mutual assistance strategy that allows superior and inferior ant colonies to share information dynamically. By integrating pheromones within and across communities, ants receive better guidance toward optimal solutions while avoiding stagnation. The cooperation mechanism ensures that less efficient colonies benefit from the experiences of stronger ones, improving overall exploration and exploitation balance.

4.4 HYBRID DYNAMIC PHEROMONE UPDATING MECHANISM

To further enhance the optimization process, a dynamic pheromone updating strategy is developed by combining both local and global pheromone updates. Additionally, a reinforcement learning-based pheromone adaptation mechanism is introduced, allowing the algorithm to adjust pheromone levels based on real-time performance. This adaptive approach ensures that the solution remains optimal even as the problem size increases, making it more effective for large-scale TSP problems.

5. ROUTE RELATIONSHIP NETWORK CONSTRUCTION

An enhanced hybrid Genetic Algorithm and Particle Swarm Optimization (GA+PSO) algorithm for community detection leverages the global exploration strength of the Genetic Algorithm (GA) with the local exploitation capability and fast convergence of Particle Swarm Optimization (PSO). This integration aims to overcome the limitations of standalone algorithms, such as PSO's tendency to get trapped in local optima and GA's potential for slow convergence or premature convergence.

5.1 CORE PRINCIPLES

- Objective Function: The algorithm typically optimizes a quality measure like modularity (Q), which quantifies the quality of a network division. A higher modularity value indicates better community structure.
- Hybridization Strategy: The key to an enhanced hybrid algorithm is a balanced approach to exploration (global search) and exploitation (local refinement). Common strategies include:
- Sequential Execution (Divide and Conquer): Applying GA for the initial iterations to explore the solution space broadly, then using the resulting solutions as input for PSO to fine-tune the results.
- Integrated Operations: Incorporating GA's crossover and mutation operations directly into the PSO's iterative update process to maintain population diversity and avoid local stagnation.

5.2. Enhancement Mechanisms

Researchers introduce several enhancements to the standard GA and PSO algorithms before or during hybridization:

- Dynamic Parameters: Adjusting PSO's inertia weight and learning factors dynamically (often non-linearly, e.g., using sinusoidal functions) allows the algorithm to favor global exploration in early stages and local exploitation in later stages.
- Improved Operators: Utilizing specific GA operators like Tournament Selection, Single-Point Crossover, and Bit-Flipping Mutation can improve diversity and prevent premature convergence.
- Local Search Integration: Embedding a local search strategy (like the Hooke-Jeeves algorithm or a chaotic local search) to refine the best solution found by the hybrid algorithm in each iteration, boosting precision and convergence speed.
- Efficient Encoding: Using an appropriate encoding scheme (e.g., real-valued or matrix encoding) that is compatible with both GA and PSO operations for community detection problems.

6. PERFORMANCE ANALYSIS

In the proposed work, the GN benchmark network data set is used to test the PSO-GA hybrid algorithm and the common community detection algorithm, and an accuracy metric based on information theory is adopted as the evaluation criterion of the clustering result [16]. The GN reference artificial network can be defined as GN (C, s, d, Z_{out}), where C represents the number of communities in the network, s represents the number of nodes in each community, and d is the node degree in the network.

Table.1. Performance Analysis

Parameter	Value / Description
Simulator	NS2 (Network Simulator 2)
Simulation Area	1000 m × 1000 m
Network Type	MANET/Large-scale Network
Number of Nodes	200, 400, 600, 800, 1000
Node Placement	Random / Uniform
Mobility Model	Random Waypoint
Node Speed	0 – 10 m/s
Pause Time	10 s
Transmission Range	250 m
MAC Protocol	IEEE 802.11
Routing Protocol	Ant-based Modular Routing with Community Detection
Simulation Time	500 s
Traffic Type	CBR (Constant Bit Rate)
Packet Size	512 bytes
Traffic Rate	4 packets/sec
Number of Connections	20 – 50
Queue Type	DropTail / FIFO
Queue Length	50 packets
Link Type	Wireless / Bidirectional

6.1 EXPERIMENTAL RESULTS

6.1.1 Packet Delivery Ratio (PDR):

Table.2. Comparison table of Packet Delivery Ratio (PDR)

Number of Nodes	M-ACO (%)	Louvain (%)	LPA-AB (%)	Proposed Method Community Detection and Modular Division (%)
100	82.4	85.7	91.2	92.5
200	80.6	84.3	90.5	91.8
300	77.9	82.1	89.3	90.7
400	75.5	79.6	88.0	89.5
500	73.2	77.8	87.4	88.7

6.1.2 End-to-End Delay:

Table.3. Comparison table of Packet Delivery Ratio (PDR)

Number of Nodes	M-ACO (%)	Louvain (%)	LPA-AB (%)	Proposed Method Community Detection and Modular Division (%)
100	50	48	45	43
200	70	65	60	57
300	95	88	82	78
400	120	110	102	97
500	150	138	128	122

7. CONCLUSION

In this work, we propose a methodology to enhance the performance of any community detection technique by providing more informative inputs, resulting in more realistic outcomes. Community detection is of great significance in studying the characteristics of network structure, analyzing user relations, exploring the way of message dissemination and mastering the trend of public opinion. The proposed work implements enhanced algorithm for community detection for large scale networks. The performance of the proposed algorithm is better than other existing ant colony based algorithms.

REFERENCES

[1] J. Zhao, X. You, Q. Duan and S. Liu, "Multiple Ant Colony Algorithm Combining Community Relationship Network", *Arabian Journal for Science and Engineering*, Vol. 47, No. 8, pp. 10531-10546, 2022.

[2] S. Bin, G. Sun, N. Cao, J. Qiu, Z. Zheng, G. Yang and L. Xu, "Collaborative Filtering Recommendation Algorithm based on Multi-Relationship Social Network", *Computers, Materials and Continua*, Vol. 60, No. 2, pp. 359-364, 2019.

[3] S. Rahimi, A. Abdollahpouri and P. Moradi, "A Multi-Objective Particle Swarm Optimization Algorithm for Community Detection in Complex Networks", *Swarm and Evolutionary Computation*, Vol. 39, pp. 297-309, 2018.

[4] S. Ding, Z. Yue, S. Yang, F. Niu and Y. Zhang, "A Novel Trust Model based Overlapping Community Detection Algorithm for Social Networks", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 32, No. 11, pp. 2101-2114, 2019.

[5] S. Pulipati, R. Somula and B.R. Parvathala, "Nature Inspired Link Prediction and Community Detection Algorithms for Social Networks: A Survey", *International Journal of System Assurance Engineering and Management*, pp. 1-18, 2012.

[6] N.R. Smith, P.N. Zivich, L.M. Frerichs, J. Moody and A.E. Aiello, "A Guide for Choosing Community Detection Algorithms in Social Network Studies: The Question Alignment Approach", *American Journal of Preventive Medicine*, Vol. 59, No. 4, pp. 597-605, 2020.

[7] W. Li, X. Zhou, C. Yang, Y. Fan, Z. Wang and Y. Liu, "Multi-Objective Optimization Algorithm based on Characteristics Fusion of Dynamic Social Networks for Community Discovery", *Information Fusion*, Vol. 79, pp. 110-123, 2022.

[8] A. Heydariyan, F.S. Gharehchopogh and M.R.E. Dishabi, "A Hybrid Multi-Objective Algorithm based on Slime Mould Algorithm and Sine Cosine Algorithm for Overlapping Community Detection in Social Networks", *Cluster Computing*, Vol. 27, No. 10, pp. 13897-13917, 2024.

[9] H. Gmati, A. Mouakher, A. Gonzalez-Pardo and D. Camacho, "A New Algorithm for Communities Detection in Social Networks with Node Attributes", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 15, No. 2, pp. 1779-1791, 2024.

[10] J. Hu, Z. Wang, J. Chen and Y. Dai, "A Community Partitioning Algorithm based on Network Enhancement", *Connection Science*, Vol. 33, No. 1, pp. 42-61, 2021.

[11] J. Jia, P. Liu, X. Du and Y. Zhang, "Multilayer Social Network Overlapping Community Detection Algorithm based on Trust Relationship", *Wireless Communications and Mobile Computing*, Vol. 2021, No. 1, pp. 1-8, 2021.

[12] Z. Li and J. Liu, "A Multi-Agent Genetic Algorithm for Community Detection in Complex Networks", *Physica A: Statistical Mechanics and its Applications*, Vol. 449, pp. 336-347, 2016.

[13] T. Yang, R. Jin, Y. Chi and S. Zhu, "Combining Link and Content for Community Detection: A Discriminative Approach", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pp. 927-936, 2009.

[14] X. Pan, G. Xu, B. Wang and T. Zhang, "A Novel Community Detection Algorithm based on Local Similarity of Clustering Coefficient in Social Networks", *IEEE Access*, Vol. 7, pp. 121586-121598, 2019.

[15] Z. Yuxin, L. Shenghong and J. Feng, "Overlapping Community Detection in Complex Networks using Multi-Objective Evolutionary Algorithm", *Computational and Applied Mathematics*, Vol. 36, No. 1, pp. 749-768, 2017.

[16] D. Lyu, B. Wang and W. Zhang, "Large-Scale Complex Network Community Detection Combined with Local search and Genetic Algorithm", *Applied Sciences*, Vol. 10, No. 9, pp. 1-8, 2020.