



Recent advance on detecting core-periphery structure: a survey

Wenli Tang¹ · Liutao Zhao² · Wei Liu³ · Yiping Liu¹ · Bo Yan¹

Received: 14 June 2019 / Accepted: 26 September 2019 / Published online: 17 October 2019
© China Computer Federation (CCF) 2019

Abstract

Recently, one type of mesoscale structure called core-periphery (CP) structure has received much attention in complex networks, as the algorithmic detection of such structures makes it possible to discover network features that are not apparent either at the local scale of nodes and edges or at the global scale of summary statistics. The core-periphery structure refers to that core nodes are densely interconnected, while periphery nodes are connected to core nodes to different extents, and periphery nodes are sparsely interconnected. Core-periphery structure containing a single core or multiple cores has been identified in various networks. However, investigation of the detection problems of the core-periphery has not been summarized in the literature. In this paper, we first introduce the definition of the core-periphery structure. The core-periphery structure has been paid more and more attention by researchers in various fields since its introduction, and it has been proved to be a powerful tool to analyze the theory of various topologies in our society, we briefly expounded the application of core-periphery structure in economics, sociology, medicine and other fields, and revealed the huge development potential of this theory. Then, we give a detailed overview of classical detection algorithms since the core-periphery structure theory was proposed. Finally, we give the development characteristics and the possible research directions of the core-periphery detection algorithm.

Keywords Core-periphery structure · Mesoscale structure · Detection method · Review

1 Introduction

A great quantity of the recent work on the structure of network systems, such as metabolic and social networks, has focused on how to measure the local structure of the network, such as degrees, correlations, clustering coefficients, etc. (Boccaletti et al. 2006; Newman 2010). However, researchers have also increasingly investigated a particular type of mesoscale structure known as community structure as well. Considerable effort has gone into algorithmic identification of community structure (Fortunato 2010; Girvan and Newman 2002; Noble 1992), and several of these methods focus on overlapping or fuzzy communities (Ahn et al. 2010; Airoidi et al. 2008; Palla et al. 2005), hierarchical structure (Clauset et al. 2008; Ravasz and Barabási 2003), and

ranking (Ball and Newman 2013), among others. Although the current research and identification methods for community structure have been very successful, another typical mesoscale structure called core-periphery in the network is equally important. However, the literature that focuses on this structure is not as much as the community structure. A global core-periphery structure with local community structure and a global community structure with local core-periphery structure is equivalent. Yang et al. (2018) use Twitter and Facebook data to prove that social groups always have a core structure of social communities always have core-periphery structures. We detail the differences and connections between CP structure and community structure in Sect. 2.4.

As early as the 1990s, the economist Paul Krugman has proposed the core-periphery structure, which Krugman 1991 increasing laid the theoretic foundation for new economic geography (Krugman 1991). Since then, the concept of core-periphery structure has also been widely developed and applied in various fields, such as in sociology (Doreian 1985; Burt 1976), international relations (Smith and White 1992; Steiber 1979), and economics (Hidalgo et al. 2007).

✉ Wei Liu
wei.liu@bnu.edu.cn

¹ Beijing Institute of Technology, Beijing, China

² Beijing Computing Center, Beijing, China

³ Beijing Normal University, Beijing, China

Recently, sociologists have conducted quantitative research on the core-periphery structure of a series of different types of networks (Holme 2005; Rombach et al. 2014). Increasingly, the research on the core-periphery structure is gradually maturing, however, there is no research summary about core-periphery structure except Csermely et al. (2013) summarized in 2013. In the review by Csermely et al. (2013), the detailed definitions of the core periphery structures and functions are introduced in different types of real-world networks, which focuses primarily on summarizing the development of *definitions* and *concepts*. However, there is currently no review that focuses on *detection methods* of core periphery structures. Therefore, in this paper, we mainly focus on summarizing the *detection methods* of this mesoscale structure and review the research status of core-periphery.

The rest of this paper is organized as follows. In Sect. 2, we review the core-periphery structure, including the definition of it, the application of it, and we briefly introduce the frequently used model, the stochastic block model, and relation of community structure and core-periphery structure. In Sect. 3, we give a detailed description of methods detecting core-periphery structure. We discuss these methods from the perspective of suitable network type, computation demand and other characteristics in Sect. 4, we summarize our work this paper in Sect. 5.

2 Core-periphery structures

2.1 Definition of core-periphery structure

The notion of core and periphery structure received attention in different fields from the late 1970s, as in social networks (Alba and Moore 1978; Laumann and Pappi 2013), in the context of international relation networks (Maslov and Sneppen 2002; Mullins et al. 1977), or in networks

related to the economy (Nemeth and Smith 1985; Snyder and Kick 1979). However, the most widely accepted definition of CP structure was proposed by Borgatti and Everett in the late 1990s (Borgatti and Everett 2000), who proposed two versions of structures in weighted, undirected graphs, the discrete version and the continuous version.

Here $G = (V, E)$ represents a graph with a set V of N nodes and a set E of M edges, with neither multiple edges nor self-loops. Let $A = (A_{ij})$ be the adjacency matrix of the original network G . $A_{ij} = 1$ if nodes i and j are adjacent by an edge, and $A_{ij} = 0$ otherwise. Note that $A_{ij} = A_{ji}$ and $A_{ii} = 0$ for all nodes i and j .

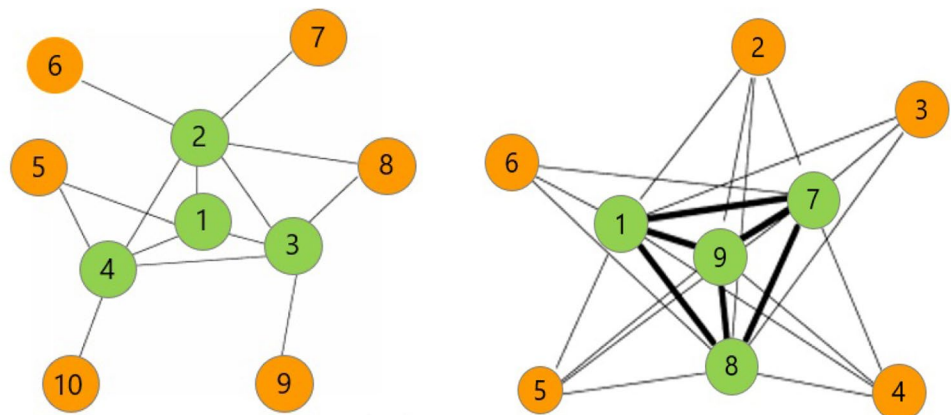
In their definition of the discrete version, they define a block model, which includes a fully connected core, and a periphery with no edges but fully connected to nodes of the core. They seek a vector C of length N with entries is in $\{1,0\}$, the i -th node belongs to periphery if the corresponding entry of C equals to 0, and belongs to core if the value is 1. Define:

$$pc = \sum_{i,j} A_{ij} C_{ij} \quad (1)$$

where $C_{ij} = 1$ if $C_i = 1$ or $C_j = 1$. If a random C makes pc to be higher compare to the former one, the algorithm think it's a better choice, if a random C maximizes pc , it will be the output.

In discrete version of core-periphery model, a node belongs to a core if and only if it is well connected both to other core nodes and to periphery nodes, and periphery nodes do not connect with other periphery nodes as in Fig. 1 (left) and (right), where the green node represents the core and the orange node represents the periphery. However, Fig. 1 (right) is an ideal core-periphery structure that is obtained by simply adding duplicates of the center of star network to the graph, and connects them to each other and to the periphery. It is more strict, that is, any core node is connected to other core nodes, and the

Fig. 1 A core/periphery structure and an ideal core/periphery structure



periphery node is connected to *every* core nodes, and there is no connection between the periphery nodes.

On the basis of discrete version, Borgatti and Everett (2000) proposed a variant notion of core-periphery structure, they define:

$$C_{ij} = \begin{cases} 1 & C_i = C_j = 1 \\ a \in [0, 1] & C_i = 1 \text{ xor } C_j = 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where “xor” denotes an “exclusive or” operation. Then a continuous version of CP is developed in which $C_{ij} = C_i * C_j = a$, every node is assigned to a value representing its “coreness”, if a node has a high value of “coreness”, it’s more likely to belong to a core.

Since the idealized core-periphery structure is too strict to be satisfied in the real networks. Hereafter, various notions and detection algorithms of CP structure have been developed (Borgatti and Everett 2000; Csermely et al. 2013; Da Silva et al. 2008; Rossa et al. 2013; Holme 2005; Shanahan and Wildie 2012; Xie et al. 2013), and a more relaxed concept of the CP structure is that the core nodes are often highly densely interconnected, the periphery nodes are connected to some core nodes, and the periphery nodes are sparsely connected or disconnected. This is a descriptive definition, and the strict definition of mathematics has not been well proposed yet.

2.2 Application of core-periphery structure

The core-periphery structure of the network consists of tightly connected core nodes and sparsely connected periphery nodes. Unlike the community structure, there is also good connectivity between the core nodes and the periphery nodes, which means that there is a certain information flow between the core and periphery. The core-periphery structure can better reflect the connection characteristics of the network. Just as a network may contain multiple communities, multiple cores can exist in the network. Research on the core edge of the network helps us to understand the deeper insight of large network topology.

Detection of the core-periphery structure can help improve the planning of traffic routes (Verma et al. 2016), and explore the deeper cooperation and development relationship of researchers (Rombach et al. 2014), improve the cooperation mode of trade networks (Doreian 1985; Fagiolo et al. 2010) and explain the characteristics of economic networks (Hidalgo et al. 2007; Kojaku et al. 2018). The purpose of detecting core-periphery structure is to identify which nodes in the network are in the core position and which nodes are in the marginal position. By analyzing the nodes in different positions in the network, it is helpful to infer the

influence of the behavior and evolution of individual nodes on the entire network. In this section, we use three real-world examples to demonstrate the significance of detecting CP structures in different filed and how CP structure can have substantial impact on discovering network features.

2.2.1 Core-periphery structure in the field of economy

In the field of economics, economics grow by upgrading the type of products they produce and export (Hidalgo et al. 2007). The core-periphery structure appears in various trade networks (Fricke and Lux 2015; Nemeth and Smith 1985) and financial networks (Hughes and Holland 1994; Malecki 1997; Anastasiou et al. 2019), which can help explain many economic mechanisms (Bailin 2017).

The product space (Hidalgo et al. 2007) can be described by a bipartite network, where a node indicates a country or a type of product, and a weighted edge between country and product indicate the ratio to the total amount of imported or exported products. It can be found that the product space has a clear CP structure by analyzing exports of goods between countries from 1998 to 2000, and most upscale products, such as metal products, machinery and chemicals, are located in a densely connected core, while lower income products such as agricultural products, handmade products, occupy a less connected periphery. Which shows that countries tend to turn to commodities that are close to the types of products currently specialized in this country. This may help explain why countries located in denser connected core of the product space can upgrade their export baskets more quickly, and poor countries have difficulty developing more competitive exports and cannot integrate into the income levels of rich countries.

2.2.2 Core-periphery detection and pervasive computing

In the concept of pervasive computing (Shneiderman and Plaisant 2010), the role played by computer devices is diluted, pervasive computing occurs on any device, anytime, anywhere, in which information flows through the virtual medium of the network, Devices are all regarded as nodes in the network. Individuals participating in pervasive computing are also regarded as nodes in the network (Szymanski and Yener 2006). Pervasive computing occurs in complex networks composed of nodes and edges. In fact, the core of pervasive computing lies in interaction. Interactions may occur between Individuals, between Individuals and devices, between devices and devices. Due to their different states, the frequency and scope of interactions are different. The core-periphery partitioning of nodes in this complex network helps us to make more accurate traffic prediction and resource allocation. In addition, in the application of social scenarios in the future, the peer-to-peer interaction on the

line and the offline perception information are highly integrated, which broadens the dimension of the social network. This high-dimensional data requires an effective node partitioning algorithm. The detection of the CP structure helps to understand the attributes of the nodes at the mesoscale level, and can play a good supporting role in social recommendation, group division, and information mediation perception. For example, MSRA (Tang 2012) proposes a mobile social network location service platform that combines mobile phone trajectory sensing data with background data integration analysis to recommend travel based on user behavior similarity. Yuan et al. (2012) use user-contributed data to improve data routing performance in the opportunistic network.

2.2.3 Core-periphery structure in the field of society

In social networks, there are many research literatures on core-periphery structures (Copus 2001; Craig and Von Peter 2014; Jeske 1999; Nocete et al. 2005; Forslid and Ottaviano 2003; Tickner 2013; Wasserman and Faust 1994). For example, the famous political blog network, which is a personal comment website that expresses the political views of the United States. The nodes represent 1225 blogs and represent hyperlinks between blogs. We can observe the political inclination of the people by identifying the core periphery structure. As shown in Fig. 2, we use the algorithm proposed by Zhang et al. (2015) to detect the core-periphery structure (the algorithm will be described in the Sect. 3) and visualize

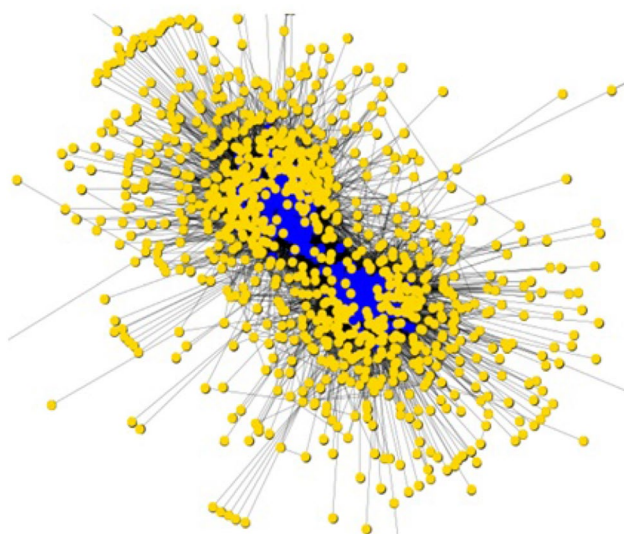


Fig. 2 Core-periphery division of a network of hyperlinks between political blogs. The network naturally separates into conservative and liberal communities, clearly visible as the two clusters in this picture. Within each group Zhang et al. algorithm (2015) finds a separate core and periphery indicated by the blue and yellow nodes respectively

the nodes in the network. The blue nodes in the network represent the core and the yellow represent the periphery. From Fig. 2, we can clearly distinguish two different political groups, and within each group, the algorithm can clearly distinguish the core nodes and periphery nodes, and the leaders located in the core part can communicate with the public more directly through the blog. Which is conducive to the dissemination of political and cultural ideas, and the core-periphery structure can clearly show the political opinion tendency of the people.

2.2.4 Core-periphery structure in the field of medicine

In the field of medicine, the human brain also exhibits core-periphery structures based on task (Bassett et al. 2013; Battiston et al. 2018). And core-periphery is commonly used as the processing and distribution department of tasks (Noble 1992; Virtanen et al. 2003; Waenerlund et al. 2011). The study of this structure in the brain can further promote the development of human medicine. Bassett et al. have studied patterns of correlated activity across a large set of brain regions (Bassett et al. 2013).

The results of experiments show that two typical mesoscale structures, community structure and core periphery structure exist in the human brains, and the emerge of the CP structure emphasizes the fact that different brain regions may play an inherently different role in information processing. Moreover, experiments have shown that one type of mesoscale structure can help characterize another. In the process of generating a sequence of motion, the brain consists of two parts, a set of time-stabilized, densely connected core regions and a set of temporally flexible and sparsely connected periphery regions.

More importantly, the core-periphery organization provides a profound way to understand how the assumed functional modules are linked. This, in turn, can predict basic human capabilities, including the generation of complex goal-oriented behavior. This has potential developmental value in the human medical field.

2.3 Stochastic block model

The stochastic block model is a generative model that has been well established and widely used in analysis of complex networks (Borgatti and Everett 2000; Cucuringu et al. 2016; Karrer and Newman 2011; Rombach et al. 2014; Xiang et al. 2018; Nowicki and Snijders 2001; Zhang et al. 2015). It can generate random networks that contain cluster structures, community structures, core-periphery structures by adjusting appropriate parameters. The different parameters can tell us the key information of the network. For example, we can know the best division of community structure or

find a clear-cut between core nodes and periphery nodes in networks. For this feature, the stochastic block model is commonly used in experiments to compare the effective of different methods.

The model is described as follows: Initially we generate n isolated nodes, without any edge connections among them, then divide them into some groups, here we assume that the nodes are divided into 2 groups (i.e., representing by group G_1 and group G_2), and the number of nodes are N_1 and N_2 respectively. According to the principle of generating random networks, we assume that the probability of connections between each pair of nodes is p_{rs} , where r represents the node belong to group r and s represents the node belong to group s . Here, we define three independent connection probabilities p_{11} , p_{22} , p_{12} for a pair of nodes in the group G_1 , group G_2 and between group G_1 and group G_2 , respectively. In this example, since the edges in the network are undirected, then the adjacency matrix corresponding to the network is symmetric, where, $p_{12} = p_{21}$ thus A connection probability matrix can be obtained as follows:

$$p_{rs} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (3)$$

$$p_{11} = 0.9, p_{12} = 0.2, p_{22} = 0.9, p_{11} = 1, p_{12} = 0.4, p_{22} = 0.15$$

$$p_{11} = 0.6, p_{12} = 0.4, p_{22} = 0.15.$$

By controlling the parameter of the stochastic block, let $p_{11} > p_{12}$ and $p_{12} < p_{22}$. It will generate traditional community structure as we expect, it is shown in the left of Fig. 3 (where $p_{11} = p_{22} = 0.9$, $p_{12} = 0.2$) from which we can clearly see two blocks that represent two community structures. The nodes in each group are closely connected, and the connections are sparse between different groups. When we set $p_{11} > p_{12} > p_{22}$, the network will generate a network with another mesoscale structure, core-periphery structure. This structure is shown in the middle of Fig. 3, where G_1 is the core part, corresponds to the block in the upper left corner of the adjacency matrix, G_2 is the periphery part, corresponds to the block in the lower right corner

of the adjacency matrix, and the rest is between the core and the periphery. It can be seen that the connections are more dense within group G_1 than connections between two groups than G_2 . And the core-periphery structure of the network is gradually weakened as the connection probability within group decreasing.

2.4 Relation of community structure and core-periphery structure

Community structure and core-periphery are two common and important mesoscale structures in social networks. The characteristics of the community structure are that the nodes in the same group are tightly connected, and the different groups are loosely connected (Chen et al. 2016; Newman and Girvan 2004). The core-periphery structure is characterized by tight connections between core nodes, tight connections between core nodes and periphery nodes, and loose connections between periphery nodes. These two structures characterize that the status of participants in social networks is unequal, and this inequality structure evolves into the characteristics of social networks (Zhang et al. 2008).

In the community structure detection, the basis of the partitioning algorithm is that the network contains cohesive community modules and weak cohesive inter-community nodes, which are often at the edge of the community and have weaker cohesiveness and core cohesion. In the community, it can often be divided into core and periphery sub-structures. For the core-periphery structure, the definition is two closely connected cores or a core node and a periphery node. From the above perspective, the two structures are very different, but they represent two views of the new community. The community structure is mainly used to define the boundaries of new communities (as shown in Fig. 4 left, community module in red circle), while the core edge structure is used to measure the interior of new communities and the interaction between nodes [as shown in Fig. 4 right), core nodes are in the dotted circle].

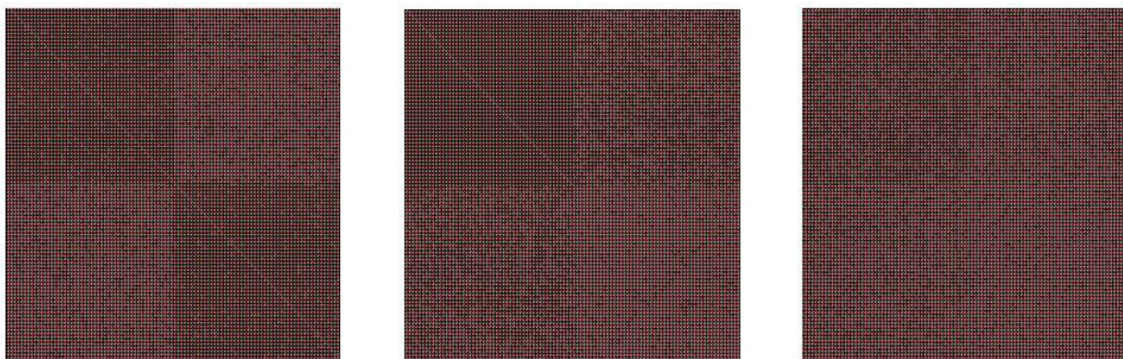
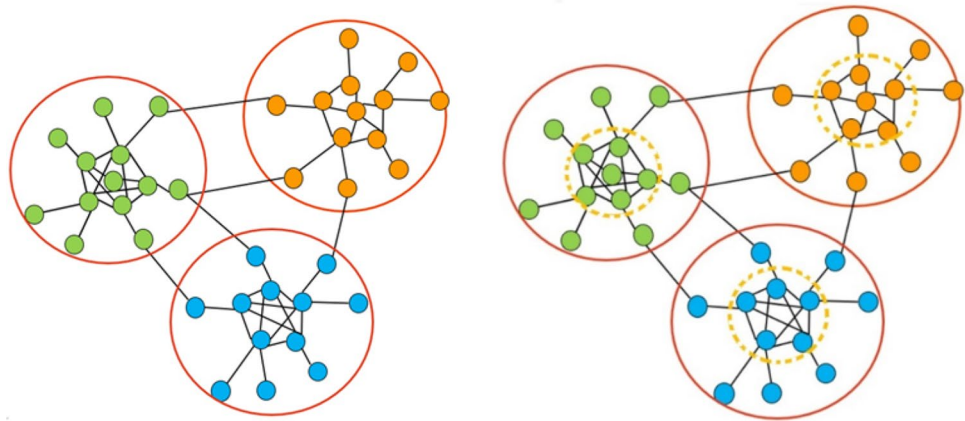


Fig. 3 (Left) Community structure; (middle) strong CP structure; (right) weak CP structure

Fig. 4 Community structures (boundary) and Multiple core-periphery structures (internal)



Through the analysis of these two mesoscale structures, we grasp the similarities and differences between the two structures from different angles. In the subsequent algorithm development process, it is expected that there will be a unified definition and detection method for the two from a new perspective. This will be significant for social network analysis.

3 Methods of detecting core-periphery structure

There are two main methods for CP detection. One type of methods distinguishing the core from the periphery based on classification. These methods regard the CP detection problem as a node binary classification problem. Any node is either belong to core or periphery. Another type attempt to assign a score to a graph or node from a quantitative perspective. This type of methods can be divided into two subclasses, the first class focus on analyzing the entire network to get the core significance of the holistic network, the other subclass of methods aims to assign a score to each node, which indicates the possibility of that the node is in the core. In the following two sections, we detail the cp detection method based on the above ideas.

3.1 Methods based on binary classification

Different from heuristic algorithm like exchange algorithm, simulated annealing and Genetic algorithm, branch-and-bound methods could obtain global optimal solution upon convergence, and they have been proven effective for small-scale numerous partitioning problems (Brusco 2011). Brusco and Cradit (2004), Cheng (1995), Brusco and Stahl (2001), Hansen and Delattre (1978), Klein and Aronson (1991) applies branch-and-bound programming to the detection problem of discrete core-periphery structure, their steps are as follows,

Step 0. Set $s = 0, P^* = \emptyset, Z^* = \infty, Z_1 = 0, Z_2 = 0$, and $S_k = \emptyset$, for $1 \leq k \leq 2$.

Step 1. Set $s = s + 1, k = 1$, and $S_k = S_k \cup \{s\}$.

where

$$\tau_j = \begin{cases} (2 - a_{js} - a_{sj}), & \text{if } k = 1 \text{ and } j \in S_k \\ (a_{js} + a_{sj}), & \text{if } k = 2 \text{ and } j \in S_k \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Step 3. Compute the following:

$$Z_2 = Z_2 + \sum_{j=s+1}^n \min(\eta_j, \omega_j) \quad (5)$$

where

$$\eta_j = \sum_{i=1}^s \alpha_{ji} \text{ for } s + 1 \leq j \leq n \quad (6)$$

$$\omega_j = \sum_{i=1}^s \beta_{ji} \text{ for } s + 1 \leq j \leq n \quad (7)$$

$$\alpha_{ji} = \begin{cases} (2 - a_{ji} - a_{ij}), & \text{if } i \in S_1 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$\beta_{ji} = \begin{cases} a_{ji} + a_{ij}, & \text{if } i \in S_1 \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Step 4. If $(Z_1 + Z_2) \geq Z^*$, then go to Step 7.

Step 5. If $s < n$, then go to Step 1.

Step 6. Set $Z^* = Z_1$ and $P^* = P = \{S_1, S_2\}$.

Step 7. If $k = 2$, then go to Step 9.

Step 8. Perform the following Sub-steps:

- Step 8a. Set $Z_1 = Z_1 - \sum_{j=1}^{s-1} \tau_j$, where τ_j is as defined in Step 2.
- Step 8b. Set $S_k = S_k - \{s\}, k = k + 1, S_k = S_k \cup \{s\}$.
- Step 8c. Go to Step 2.
- Step 9. Perform the following Sub-steps:
- Step 9a. $Z_1 = Z_1 - \sum_{j=1}^{s-1} \tau_j$, where τ_j is as defined in Step 2.
- Step 9b. Set $S_k = S_k - \{s\}$.
- Step 9c. Set $s = s - 1$.
- Step 9d. If $s = 0$, then STOP; otherwise, set $k = l : s \in S_l$ and go to Step 7.

where s is a position pointer, it records the currently considered assignment actor, p^* is the optimal bipartition, Z^* is the initial upper bound of the objective function, Z_1 means the direct contribution of actor s , Z_2 is the lower bound of the sum of contribution of all unhandled actors.

Brusco (2011) report that, networks within the node number of 60 will get good results in calculations, as the number of nodes continues to increase, the efficiency of the algorithm may be unacceptable due to the limitation of branch-and-bound methods, so this method can't replace the traditional heuristic algorithm, but they are a good complement in some special cases.

Zhang et al. (2015) proposed a method based on detection of core-periphery structure of statistical inference, they use the maximum-likelihood method for fitting generative network model, this method is similar to the popular first-principles methods in communication detection in theory. In fact, the basic model used by them is the same, except that the parameters are adjusted to be more suitable for detecting the core-periphery structure. The author uses the EM algorithm (Dempster et al. 1977) to find the parameters of the stochastic block model to be closest to a given network. The advantage of this approach is that it is very efficient and can be applied to large networks with millions of nodes, and theoretically can detect inconspicuous core-periphery structures (Karwa et al. 2017).

The EM algorithm solves equations by numerical iteration, at the start of the iteration, initial parameter p and γ should be given, after a round of iteration, the algorithm would output a pair of updated p and γ , It can be proved that after several round of iterations the algorithm will finally converge to some local maximum of the log-likelihood. Since the EM algorithm does not necessarily reach the global optimal solution, it may be necessary to try several sets of initial conditions, and then take the result of the group with the highest log-probability as the final output of the algorithm. The final output of the algorithm has the parameters of the random block model, and also includes the probability that each node belongs to a certain group, either core or periphery. For the algorithm proposed by the author, the latter is more valuable. Due to

too many calculations in the EM algorithm, the algorithm does not work within an acceptable range unless it is on a small network. Zhang et al. (2015) use belief propagation algorithm proposed by Decelle et al. (2011a, b) to improve this problem, which is faster than Monte Carlo sampling. Belief propagation is a message-passing technique for finding probability distributions on networks.

Xiang et al. (2018) proposed a joint algorithm that can simultaneously detect community structure and core-periphery structure. The algorithm is divided into three steps. The first step is to reorder the nodes in the network, initialize the node set U to store the new node sequence. Node set V' is used to store nodes that have not been reordered, $V' = V, V$ is nodes of the given network. At the beginning of the algorithm, the node with the largest closeness centrality is added to the set U , and then the node with the largest connection with the node in U , that is, the node with the most edges to nodes in set U , is selected from V' , and add the node to set U , if there are more than one node to be selected, choose the one with the highest degree, if there are still more than one node, randomly choose one. This process continues until the set V' is empty. At this time, there is a new rank of nodes in set U .

There is an example shown in Fig. 5, we suppose the blue node is in U , next step, choose one node in V' (outside of the circle) and add it into U (see Fig. 5 left). We find that node 1 and 2 has the same connections with U , but we add node 1 into U since node 1 has larger degree value (see Fig. 5 right).

The second step is to plot the region density curve. First, the author defines the connection density connection density (CD) to measure the tightness of the node connections in the sub-graph S . The CD is defined as follows:

$$CD(S) = \frac{2m'}{n'(n' - 1)} \quad (10)$$

where n' is the number of nodes in S and m' is the number of existing connections in S .

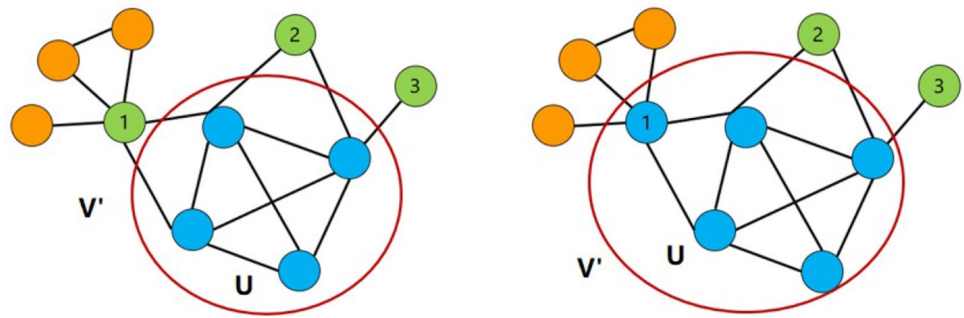
By defining a parameter α to control the minimum size of the core in a given network, define a node region density RD . For node i , its region density is the $CD(node_{i-\alpha+1}, \dots, node_i)$, the definition is as follows:

$$RD(u_i) = \begin{cases} CD(\{u_1, \dots, u_i\}), & i \leq \alpha \\ CD(\{u_{i-\alpha+1}, \dots, u_i\}), & i > \alpha \end{cases} \quad (11)$$

According to the reordered node sequence obtained in the first step, and the regional density RD of the node defined in the second step, a curve of the density of the node region on the new node sequence can be drawn, which will guide the core-periphery detection process.

The third step is the detection of core-periphery. If the connection density CD of a sub-graph S is greater than a

Fig. 5 An example of constructing a U set in detail, we will choose node 1 in the next step and add it into U



given threshold β , then S is considered to be a core, and the larger the β , the stricter the definition of the core. Look for the number of peak points on the region density curve that is greater than β . Note that the peak at the beginning of the curve is invalid. When the RD values of two consecutive sequences of nodes are greater than or equal to β , the two regions are combined into one. These peak node and the α nodes before the peak point form the core, which is defined as the class 0 group. In order to determine which periphery nodes are the affiliate nodes of these cores, the author set node which has connection with class 0 nodes as the class 1 node, and the class 2 node is nodes that has a direct connection with the class 1 nodes. This process continues until all edge nodes are assigned a class.

This method could detect multiple core-periphery pairs and different scale of cores by control the threshold value α , but it fails to assign a coreness value to each node to indicate the possibility of belonging to a core.

Kojaku and Masuda (2017) proposed a scalable algorithm capable of detecting non overlapping multiple cores, which automatically determines the number and size of core-periphery structures. The author extends the ideal core-periphery structure proposed by Borgatti and Everett (2000) to a situation suitable for multi-core detection. The author defines C as the number of cores, $c = (c_1, c_2, \dots, c_N)$, $c_i \in \{1, 2, \dots, C\}$, c_i is used to indicate which core the node i belongs to, and define the adjacency matrix.

$$B_{ij}(c, x) = \begin{cases} \delta_{c_i, c_j} & (x_i = 1 \text{ or } x_j = 1, \text{ and } i \neq j) \\ 0 & (\text{otherwise}) \end{cases} \quad (12)$$

where δ is Kronecker δ . In order to make the above adjacency matrix most similar to the given adjacency matrix A , the author looks for (c, x) by maximizing the following quality function Q^{cp} , which is defined as follows:

$$Q^{cp}(c, x) = \sum_{i=1}^N \sum_{j=1}^{i-1} A_{ij} B_{ij}(c, x) - \sum_{i=1}^N \sum_{j=1}^{i-1} p B_{ij}(c, x) \quad (13)$$

$$= \sum_{i=1}^N \sum_{j=1}^{i-1} (A_{ij} - p)(x_i + x_j - x_i x_j) \delta_{c_i, c_j}$$

where $p = M/[N(N - 1)/2]$ is the density of edges in the network, the term $\sum_{i=1}^N \sum_{j=1}^{i-1} A_{ij} B_{ij}(c, x)$ represents the number of edges that appear simultaneously in the network which contains ideal core-periphery and the given network, term $\sum_{i=1}^N \sum_{j=1}^{i-1} p B_{ij}(c, x)$ represents the number of edges that appear simultaneously in the network which contains ideal core-periphery structure and Erdős Rényi random graph (Erdős and Rényi 1960) in which the probability of edge between any two nodes is p . A large Q^{cp} indicates that a given network is more similar to a network with an ideal core-periphery structure.

The author uses the label switching heuristic method to maximize Q^{cp} . At the beginning, each node is set to belong to a different core, and then a node i is randomly selected, the switch the label of the neighbor of node i , either core or periphery, after performing the above process on all neighbors of node i , we select the tag that maximizes Q^{cp} and update it to node i . After scanning all nodes, if no label of any node in the current round changes, the algorithm converges. Otherwise, the nodes are reordered and all the above processes are executed again. After the algorithm converges, each node obtains its label and determines which core-periphery pair the node belongs to.

Ma et al. (2017) proposed an algorithm based on 3-tuple motif which is inspired by the idea of motif. Based on this fact that the degree of the core node is generally high, but the node with a high degree is not necessarily the core node. They define motifs by a 3-tuple (B, A, ϕ) . The parameter ϕ is used to emphasize that the core node is higher than the edge node and higher than the average of the network degree. The definition of motif is as follows:

$M(B, A, \phi) = \{(set(v), set(\mathcal{X}\mathcal{A}(v))) | v \in V^k, v_1, \dots, v_k \text{ distinct}, Av = B, f(\phi, ((set(v), set(\mathcal{X}\mathcal{A}(v)))) = 1\}$. where $f(\phi, ((set(v), set(\mathcal{X}\mathcal{A}(v)))) = 1$ indicates that the node of $set(v)$ and $set(\mathcal{X}\mathcal{A}(v))$ should satisfy the property of ϕ .

Let $M_1^C = (B1, A^c, \phi)$ and $M_1^P = (B1, A^P, \phi)$ be the motif of core and motif of the periphery, and $A^C = 1, 2, A^P = 3, 4$, one can form the adjacency matrix of M1 as follows:

$$(W_{M1})_{ij} = \sum_{(v, \mathcal{X}\mathcal{A}(v)) \in M_1^C \cup M_1^P} 1(\{i, j\} \in \mathcal{X}\mathcal{A}(v)) \quad (14)$$

Since $M1$ is too strict, they define a motif $M2$, and thus form a motif adjacency matrix W_{M2} , because $M2$ is too loose, they define a modest mode of motif $M3$, since it is not possible to determine whether the core-periphery structure contained in the network is obvious, the author Defines a mixed motif matrix M_W , which is a collection of $M1, M2$, which is defined as follows:

$$W_M = \alpha W_{M1} + \gamma W_{M2} + \beta W_{M3} \tag{15}$$

where $\gamma = 1.0, \beta = 2\gamma, \alpha = 2\beta$.

After obtaining the matrix W_M , the goal of a clustering algorithm for the node set S is as follows. First, the node set S should contain more elements in M On the other hand, the set S should avoid the elemental segmentation in M . Therefore, the optimization goal of clustering is

$$\phi_M(S) = \frac{cut_M(S, \bar{S})}{\min(vol_M(G)S, vol_M(G)\bar{S})} \tag{16}$$

where \bar{S} is the complement of $S, cut_M(S, \bar{S})$ means the number of motif M with at least one node in S and one node in \bar{S} , and $vol_M(G)S$ is the number of nodes in instance of M and belongs to S . One can use motif spectral clustering on the motif adjacency matrix W_M .

3.2 Methods based on quantization assignment

Yan and Luo (2019) focusing on the quantitative detection of multi-core CP structures, and for the first time officially defined the concept of multi-core periphery structure, that is, a sparsely dense core is surrounded by a sparse periphery, and proposes to find the best way to partition. The first step is to create a tree diagram of the network using a hierarchical clustering algorithm. The tree diagram defines a series of network partitions. Starting from the bottom, each vertex is independent, to the top, all the vertices belong to a separate cluster. When creating a tree diagram, the average linkage clustering algorithm is used to determine the distance between each pair of temporary vertex clusters (in the tree diagram) as the average of all pairwise distances between the vertices in the two clusters, guaranteeing the merger The cluster has the highest internal connection density, and the optimal kernel-peripheral partition is calculated by the connection density of vertices in all clusters (i.e., potential cores) and the connection density of vertices not assigned to any cluster (i.e., potential periphery) ratio density.

The core and peripheral connections are calculated as follows:

$$Density_{cores} = \frac{C}{\sum_{i=1}^k \frac{n_i(n_i-1)}{2}} \tag{17}$$

$$Density_{periphery} = \frac{P}{\frac{m(m-1)}{2} + m \sum_{i=1}^k n_i} \tag{18}$$

where n_i is the number of nodes in cluster i , and i belong to $[1, k]$ clusters; m is the number of nodes at the periphery; C is the sum of weights of all the connections of core nodes within cores; and P is the sum of weights of all the connections of peripheral nodes outside cores.

Measure the extent to which a partition implements a multicore peripheral structure by defining a core-peripheral ratio:

$$r = \frac{Density_{cores}}{Density_{periphery}} \tag{19}$$

In the research of core-periphery structure, a core is usually be defined as a sub-graph in which nodes are tightly interconnected (Borgatti and Everett 2000; Everett and Borgatti 2000), another notion is, the core of the graph should be central, Holme (2005) use a rather strong percept that a core should be both highly interconnected and central, they define the core as k -core which has the highest closeness centrality, to measure the tendency of a given network contains a core, they define normalized closeness centrality as a core-periphery coefficient c_{cp} , which is the core minus the same corresponding average value for our null-model, random networks of the same degree sequence as the original network.

Coefficient c_{cp} as

$$c_{cp} = \frac{C_C[V_{core}(G)]}{C_C[V(G)]} - \left\langle \frac{C_C[V_{core}(G')]}{C_C[V(G')]} \right\rangle_{G' \in \mathcal{G}(G)} \tag{20}$$

where $\mathcal{G}(G)$ is a null model (with same degree as G). Our goal is to maximize $C_C(U)$ on different sequences, and $C_C(U)$ is defined according to equation closeness centrality. The closeness centrality is defined as:

$$C_C(i) = (\langle d(i, j) \rangle_{j \in V \setminus \{i\}})^{-1} \tag{21}$$

$$C_C(U) = (\langle \langle d(i, j) \rangle_{j \in V \setminus \{i\}} \rangle_{i \in U})^{-1} \tag{22}$$

They found that c_{cp} is related to the characteristics of the network.

Da Silva et al. (2008) define a parameter called network capacity, which is defined as follows,

$$pc = \sum_{i=1}^n \frac{1}{PL_i} \tag{23}$$

where n is the total number of connected pairs in the network, PL_i is the shortest path between each pair. The

capacity is a parameter to measure the connectivity of a network. A network with more paths should have a higher capacity, if the paths in the network are shorter, the capacity will be higher too.

After defining the network capacity, the authors remove the nodes one by one in the order of decreasing centrality. Each time a node is removed, the network capacity C needs to be calculated once. After all nodes are removed, the network capacity becomes 0. The author proposed a parameter named core coefficient (cc) to evaluate the core-periphery structure of a metabolic network, the core coefficient is defined as follows,

$$cc = \frac{n}{N} \tag{24}$$

$$\int_{i=0}^n C_i = 0.9 \int_{j=0}^N C_j \tag{25}$$

where c_i is the network capacity of the network after remove i nodes.

In the obtained network capacity change curve, the author believes that the network capacity curve drop rate of the network containing core-periphery structure should slow down faster than that of the network without core-periphery structure, that is, the larger the core coefficient, the network have greater possibilities to contain core-periphery structure. The authors suggest that $cc = 0.5$ is a reasonable boundary. They proposed another algorithm to detect the core-periphery structure named network decomposition based on modularity. Its steps is as follows, the module with the most central nodes will be created as a core, and after the initial core module is selected, run the normal decomposition algorithm until all the nodes are clustered, then choose the module with the highest modularity, the closeness centrality is chosen to measure the centrality of a node. Although there is a certain degree of novelty, due to the use of closeness centrality, its calculation speed may be unacceptable on a large network.

When dealing with the adjacency matrix of the network, the MINRES method is very suitable for situations that the adjacency matrix is symmetrical and the diagonal elements are not considered, but when the matrix is asymmetrical, such as a directed network, SVD is needed, but this method requires diagonal elements, Boyd et al. (2010) proposed a minimum residual-based SVD method named MINRES/SVD to deal with asymmetric and diagonal missing networks. Each node gets two indicators: in-coreness and out-coreness, which indicates the outgoing and incoming trends of the node.

The MINRES/SVD algorithm aims at finding a best approximation to the origin matrix A , and the diagonal elements are excluded, just as MINRES, and where the vectors are allowed, their goal is to minimize the sum of the non-diagonal squared residuals,

$$f = \sum_i \sum_{j \neq i}^n (A_{ij} - u_i v_j)^2 \tag{26}$$

the main purpose of MINRES/SVD is to use the advantages both in MINRES and SVD, which means we can separate u and v to solve the problem of asymmetry and avoid the diagonal elements affecting the results. the calculation are as follows, replace Eq. (26) with the following equation:

$$f = \sum_i \sum_{j \neq i}^n (A_{ij} - u_i v_j)^2 \tag{27}$$

Let $d = \|u\| \cdot \|v\|$, where $\|u\| = \sqrt{\sum_i u_i^2}$ is the normal of u , and normalize u and v , differentiate with respect to u_i , and equals to zero, to minimize f in Eq. (27)

$$\frac{\partial f}{\partial u_i} = \sum_{j \neq i} 2u_i v_j^2 - \sum_{j \neq i} 2A_{ij} v_j = 0 \tag{28}$$

Simplified it to the following equation:

$$\sum_j A_{ij} v_j = u_i (\sum_j v_j^2 - v_i^2) \tag{29}$$

Then it can be expressed as the equation:

$$Av = u(v^T v - v^2) \quad \text{and} \quad \mathbf{u}^T A = \mathbf{v}^T (\mathbf{u}^T \mathbf{u} - \mathbf{u}^2)^T \tag{30}$$

Find MINRES/SVD using MATLAB or Mathematical software suite we could get

$$u.A = v(u.u - u^2), \{ \{u, u_0\}, \{v, v_0\} \} \tag{31}$$

where u_0 and v_0 are initial guess of u and v .

Lee et al. (2014) proposed two methods for detecting core-periphery structures, which are density-based and transport-based. In the density-based method, the author proposes a core-score indicator to indicate the probability that a node belongs to the core. It is calculated as follows: $\{C_{ij}\}$ into the core quality

$$R(\alpha, \beta) = \sum_{i,j} W_{ij} C_{ij}(\alpha, \beta) \tag{32}$$

where $\alpha \in [0, 1]$ determines the sharpness of the core-periphery division and $\beta \in [0, 1]$ determines the fraction of core nodes:

$$C_i(\alpha, \beta) = \begin{cases} \frac{i(1-\alpha)}{2[N\beta]}, & i \leq [N\beta] \\ \frac{(i-[N\beta])(1-\alpha)}{2(N-[N\beta])} + \frac{1+\alpha}{2}, & i > [N\beta] \end{cases} \tag{33}$$

$$CS(i) = Z \sum_{(\alpha, \beta)} C_i(\alpha, \beta) R(\alpha, \beta) \tag{34}$$

where Z is the normalization term, W_{ij} is the element of the i -th row and the j -th column of the adjacency matrix,

and represents the weight of the edge between the node i and the node j . When $W_{ij} = 0$, it indicates that there is no edge connection between the node i and the node j . i, j belongs to $\{1, \dots, N\}$, N is the number of network nodes. In the transport-based approach, the author believes that the core component of the network is more likely to appear in the path, so the author defines the path score, Path-Score, PS

$$PS(i) = \frac{1}{|\mathbb{E}|} \sum_{(j,k) \in \mathbb{E}} \sum_{\{p_{jk}\}} \sigma_{jik}[\mathbb{E} \setminus (j, k)] \quad (35)$$

where $\sigma_{jik} = 1/|\{p_{jk}\}|$ if node i in set p_{jk} , in which consists "backup paths" from node j to node k , and edge (j, k) is removed from \mathbb{E} , otherwise $\sigma_{jik}[\mathbb{E} \setminus (j, k)] = 0$.

The Path score of the network component can be measured for its importance after the other components of the graph are removed by calculating the centrality score. This paper calculates PS by calculating the shortest path and GCNP. This article calculates CS (Core-score), PS (PATH-score), BS (betweenness centrality) on the real network, and calculates the correlation (Pearson and Spearman correlation values) between them, and concludes that performance of CS and PS proposed in this paper is different on different networks, which highlights the fact that we need to propose different core-periphery structures for different networks.

Rossa et al. (2013) profile core-periphery structure by elaborating the behavior of a random walker. They proposed core-periphery profile curve, which is a non-decrease function with value $\alpha_1, \alpha_2, \dots, \alpha_n$, where n is the node number of a network, they also introduce a numerical indicator to measure the degree of centralization, and a coreness is assigned to each node, nodes with coreness value less than threshold α is included by the " α -periphery".

The core-periphery profile is obtained from standard random walk procedure. The step of obtaining core-periphery profile α_k is as follows, Define the core-periphery profile $\alpha_k, k = 1, 2, \dots, n$ of the network by the following algorithm:

Step 1. Randomly select a node i among those with minimal strength ($\sigma_i \leq \sigma_j$ for all $j \in N$), for generality, set the selected node is 1. Set $P_1 = \{1\}$, hence $\alpha_1 = 0$.

Step $k=2, 3, \dots, n$: Select the node attaining the minimum in:

$$\begin{aligned} \alpha_k &= \min_{h \in N \setminus P_{k-1}} \frac{\sum_{i \in P_{k-1} \cup \{h\}} \pi_i m_{ij}}{\sum_{i \in P_{k-1} \cup \{h\}} \pi_i} \\ &= \min_{h \in N \setminus P_{k-1}} \frac{\sum_{i \in P_{k-1}} \pi_i m_{ij} + \sum_{i \in P_{k-1}} (\pi_i m_{ih} + \pi_h m_{hi})}{\sum_{i \in P_{k-1}} \pi_i + \pi_h} \end{aligned} \quad (36)$$

The author defines the probability of persistence of the sub-graph S α_S as the probability that a random walker

in the sub-graph S still stays in S in the next step, and its expression is

$$\alpha_S = \frac{\sum_{i, j \in S} \pi_i m_{ij}}{\sum_{i \in S} \pi_i} \quad (37)$$

in which π_i is the asymptotic probability of visiting node i , and m_{ij} is:

$$m_{ij} = \frac{w_{ij}}{\sum_h w_{ih}} \quad (38)$$

where w_{ij} is the weight of the edge between node i and node j .

If a network contains an ideal core-periphery structure, the periphery nodes can only be linked to the core nodes, and there can be no links between the periphery nodes. This means that for sub-graph S , if it only contains periphery nodes, the α_S of the sub-graph is 0 because the random walker will jump out of the sub-graph at the next moment. Since the real-world network contains core-periphery structures that are often not ideal, there will still be weak connectivity between the periphery nodes, and the edge periphery is generalized to α -periphery containing a parameter α . That is, for a sub-graph S , if $\alpha_S < \alpha$, then the probability that the random walker in S jumps out at the next moment is $1 - \alpha$.

4 Discussion

In the previous section, we discussed two types of CP detection algorithms based on binary and quantization-based assignments. One type of methods regards the CP detection problem as a node binary classification problem. Any node is either belong to core or periphery. There are various classification methods. Some methods are based on label propagation to maximize the objective function (Kojaku and Masuda 2017), some based on graph cut Ma et al. (2017), some by fitting a stochastic block model (SBM) to empirical network data using EM algorithm (Zhang et al. 2015). The other type attempt to assign a score to a graph or node from a quantitative perspective. This type of methods can be divided into two subclasses, the first class focus on analyzing the entire network to get the core significance of the holistic network, Borgatti and Everett (Borgatti and Everett 2000) propose a method for detecting both discrete and continuous versions of core-periphery structure in weighted, undirected graphs, Da Silva and Marcio Rosa propose the core coefficient (Da Silva et al. 2008) that was defined based on the concept of closeness centrality and network capacity. Holme (2005) also propose a CP coefficient using the closeness centrality and k-cores deposition technique to determine core nodes. The other subclass of methods aims to assign a

score to each node, which indicates the possibility of that the node is in the core. This method is well developed by Rombach et al. (2014), who compute the node's core score by simulated annealing program. Cucuringu et al. proposed Path-Score (2016) that gives a score for each node based on transportation in networks (Lee et al. 2014).

Based on the binary method, the nodes can be clearly divided into two groups. The algorithm is simple and easy to design. But it can only distinguish whether a node belongs to the core or the periphery, it cannot determine its significance in their group. That is, if a node belongs to the core, we cannot determine its importance in the core cluster. Nodes within core are equivalent, which will undoubtedly bring some loss of information. The quantization assignment method solves this problem at the cost of using large time complexity, which may be unacceptable when applied to large networks.

The above classification method is based on the performance of the CP structure on the node attributes. From the perspective of the entry point of the algorithm, If the characteristics of the entire network are taken into consideration, methods can be divided into two types, densely-based (Brassil and Nodari 2018) and transport-based. The densely-based algorithm considers the core nodes to be closely connected, while the periphery nodes are loosely connected; and the transport-based algorithm consider the core node as a group, its path to other nodes is shorter and more, in addition, the remaining nodes in the network has a greater probability of passing through the group (Garas et al. 2012; Newman 2006; Ruggera et al. 2016; Yang and Leskovec 2014).

Reviewing the literature talked about in this paper, most of the core-periphery detection problems will eventually be transformed into an optimization problem. Some use heuristic algorithms to find near optimal solutions, such as simulated annealing (Lee et al. 2014), label switching (Rossa et al. 2013), etc., heuristic algorithm can almost guarantee a near optimal solution, but it is often not satisfactory in terms of computational efficiency; some literature transform the optimization problem into matrix computing domain Jia and Benson (2018), for instance, literature (Boyd et al. 2010) combined with MINRES algorithm and SVD. The MINRES/SVD is proposed to transform the problem into the singular value decomposition problem of the matrix; some literature construct the network feature matrix to transform the problem into the cut graph problem Ma et al. (2017); some also from the statistical point of view (Zhang et al. 2015), use the EM algorithm to calculate the random most similar to the original image, use nonlinear spectral method (Tudisco and Higham 2019). Researchers have considered using integer programming (Brusco 2011) to find the optimal solution, but this method is limited by the size of the network. This algorithm can only be obtained satisfactory speed when the network node is less than 60.

5 Conclusions and perspectives

In this paper, we give a detailed introduction to the current research status of core-periphery structure, including the definition of core-periphery structure, application fields, and identification methods. For the existing research methods, we roughly divide into two categories, one is densely-based, the other is transport-based, and we have reviewed and summarized the advantages and disadvantages of the two methods, from the above concluding observations. It can be clearly seen that there is still a lot of work to we need to do, including clarifying the definition of the core-periphery network, especially the core-periphery structure in dynamic networks. Here we list some research directions for the future network.

1. An important future study is to find a clear-cut point to separate core and periphery nodes. In other words, we can't know the size of the core of a network in advance in most realistic situations, we need to determine a clear-cut point to classify the nodes as core or periphery nodes, which has the great significance for detecting the effective of different algorithms in real networks. Besides, pursuing statistical methods for studying core-periphery structure and other mesoscale network structures are also important (Zhang et al. 2015).
2. Core-periphery structure is an important feature network in the real world. However, the efficiency and effectiveness of many current detection algorithms are not ideal, especially in large complex networks. Then it is necessary to continue to develop algorithms to identify it and compare the performance of these algorithms on a variety of networks.
3. The community structure has been studied in temporal and multi-layer networks, and has achieved good performance. However, the core-periphery structure has not received much attention that it deserved in these networks. And many literature have proved that core-periphery structure is as important as community structure. In this context, the study of core-periphery structure should also be considered in multi-layer networks in the future work.

References

- Ahn, Y.-Y., Bagrow, J.P., Lehmann, S.: Link communities reveal multiscale complexity in networks. *Nature* **466**(7307), 761 (2010)
- Airoldi, E.M., Blei, D.M., Fienberg, S.E., Xing, E.P.: Mixed membership stochastic block models. *J. Mach. Learn. Res.* **9**, 1981–2014 (2008)
- Alba, R.D., Moore, G.: Elite social circles. *Soc. Methods Res.* **7**(2), 167–188 (1978)

- Anastasiou, Dimitrios, Louri, Helen, Tsionas, Mike: Nonperforming loans in the euro area: a re core-periphery banking markets fragmented? *Int. J. Finance Econ.* **24**(1), 97–112 (2019)
- Bailin, A.: *From traditional to Group Hegemony: the G7, the Liberal Economic Order and the Core-Periphery Gap.* Routledge, Abingdon (2017)
- Ball, Brian, Newman, Mark E.J.: Friendship networks and social status. *Soc. Netw.* **1**(1), 16–30 (2013)
- Bassett, D.S., Wymbs, N.F., Rombach, M.P., Porter, M.A., Mucha, P.J., Grafton, S.T.: Task-based core-periphery organization of human brain dynamics. *PLoS Comput. Biol.* **9**(9), e1003171 (2013)
- Battiston, F., Guillon, J., Chavez, M., Latora, V., DeVicoFallani, F.: Multiplex core-periphery organization of the human connectome. *J. R. Soc. Interface* **15**(146), 20180514 (2018)
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: structure and dynamics. *Phys. Rep.* **424**(4–5), 175–308 (2006)
- Borgatti, S.P., Everett, M.G.: Models of core/periphery structures. *Soc. Netw.* **21**(4), 375–395 (2000)
- Boyd, J.P., Fitzgerald, W.J., Mahutga, M.C., Smith, D.A.: Computing continuous core/periphery structures for social relations data with minres/svd. *Soc. Netw.* **32**(2), 125–137 (2010)
- Brassil, A., Nodari, G.: A Density-based estimator of core/periphery network structures: analysing the Australian interbank market. No. rdp2018-01. Reserve Bank of Australia (2018)
- Brusco, M.: An exact algorithm for a core/periphery bipartitioning problem. *Soc. Netw.* **33**(1), 12–19 (2011)
- Brusco, M.J., Craidt, J.D.: Graph coloring, minimum-diameter partitioning, and the analysis of confusion matrices. *J. Math. Psychol.* **48**(5), 301–309 (2004)
- Brusco, M.J., Stahl, S.: An interactive multiobjective programming approach to combinatorial data analysis. *Psychometrika* **66**(1), 5–24 (2001)
- Burt, R. B. R. S.: Networks of collective action: a perspective on community influence systems. by Edward O. Laumann; Franz U. Pappi. *Contemp. Sociol.* **7**(2), 152–153 (1978)
- Chen, T., Tang, L.-A., Sun, Y., Chen, Z., Chen, H., Jiang, G.: Integrating community and role detection in information networks. In: Proceedings of the 2016 SIAM International Conference on Data Mining, pp. 72–80. SIAM (2016)
- Cheng, C.-H.: A branch and bound clustering algorithm. *IEEE Trans. Syst. Man Cybern.* **25**(5), 895–898 (1995)
- Clauset, A., Moore, C., Newman, M.E.J.: Hierarchical structure and the prediction of missing links in networks. *Nature* **453**(7191), 98 (2008)
- Copus, A.K.: From core-periphery to polycentric development: concepts of spatial and aspatial peripherality. *Eur. Plan. Stud.* **9**(4), 539–552 (2001)
- Craig, B., Von Peter, G.: Interbank tiering and money center banks. *J. Finan. Intermed.* **23**(3), 322–347 (2014)
- Csermely, P., London, A., Ling-Yun, W., Uzzi, B.: Structure and dynamics of core/periphery networks. *J. Complex Netw.* **1**(2), 93–123 (2013)
- Cucuringu, M., Rombach, P., Lee, S.H., Porter, M.A.: Detection of core-periphery structure in networks using spectral methods and geodesic paths. *Eur. J. Appl. Math.* **27**(6), 846–887 (2016)
- Da Silva, M.R., Ma, H., Zeng, A.-P.: Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. *Proc. IEEE* **96**(8), 1411–1420 (2008)
- Decelle, A., Krzakala, F., Moore, C., Zdeborová, L.: Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E* **84**(6), 066106 (2011a)
- Decelle, A., Krzakala, F., Moore, C., Zdeborová, L.: Inference and phase transitions in the detection of modules in sparse networks. *Phys. Rev. Lett.* **107**(6), 065701 (2011b)
- Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* **39**(1), 1–22 (1977)
- Doreian, P.: Structural equivalence in a psychology journal network. *J. Am. Soc. Inf. Sci.* **36**(6), 411–417 (1985)
- Erdős, P., Rényi, A.: On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**(1), 17–60 (1960)
- Everett, M.G., Borgatti, S.P.: Peripheries of cohesive subsets. *Soc. Netw.* **21**(4), 397–407 (2000)
- Fagiolo, G., Reyes, Javier, Schiavo, Stefano: The evolution of the world trade web: a weighted-network analysis. *J. Evol. Econ.* **20**(4), 479–514 (2010)
- Forslid, R., Ottaviano, G.I.P.: An analytically solvable core-periphery model. *J. Econ. Geogr.* **3**(3), 229–240 (2003)
- Fortunato, S.: Community detection in graphs. *Phys. Rep.* **486**(3–5), 75–174 (2010)
- Fricke, D., Lux, T.: Core-periphery structure in the overnight money market: evidence from the e-mid trading platform. *Comput. Econ.* **45**(3), 359–395 (2015)
- Garas, A., Schweitzer, F., Havlin, S.: A k-shell decomposition method for weighted networks. *New J. Phys.* **14**(8), 083030 (2012)
- Girvan, Mi, Newman, M.E.J.: Community structure in social and biological networks. *Proc. Natl. Acad. Sci.* **99**(12), 7821–7826 (2002)
- Hansen, P., Delattre, M.: Complete-link cluster analysis by graph coloring. *J. Am. Stat. Assoc.* **73**(362), 397–403 (1978)
- Hidalgo, C.A., Klinger, B., Barabási, A.-L., Hausmann, R.: The product space conditions the development of nations. *Science* **317**(5837), 482–487 (2007)
- Holme, P.: Core-periphery organization of complex networks. *Phys. Rev. E* **72**(4), 046111 (2005)
- Hughes, D.W., Holland, D.W.: (Core-periphery economic linkages: a measure of spread and possible backwash effects for the. *Land Econ.*, **70**(3), 1994
- Jeske, R.J.: World-systems theory, core-periphery interactions, and elite economic exchange in mississippian societies. *World-Systems Theory in Practice: Leadership, Production, and Exchange*, pp. 203–221 (1999)
- Jia, J., Benson, A.R.: Detecting core-periphery structure in spatial networks. arXiv preprint [arXiv:1808.06544](https://arxiv.org/abs/1808.06544) (2018)
- Karrer, B., Newman, M.E.J.: Stochastic blockmodels and community structure in networks. *Phys. Rev. E* **83**(1), 016107 (2011)
- Karwa, V., Pelsmajer, M.J., Petrović, S., Stasi, D., Wilburne, D., et al.: Statistical models for cores decomposition of an undirected random graph. *Electron. J. Stat.* **11**(1), 1949–1982 (2017)
- Klein, G., Aronson, J.E.: Optimal clustering: a model and method. *Naval Res. Log. (NRL)* **38**(3), 447–461 (1991)
- Kojaku, S., Masuda, N.: Finding multiple core-periphery pairs in networks. *Phys. Rev. E* **96**(5), 052313 (2017)
- Kojaku, S., Cimini, G., Caldarelli, G., Masuda, N.: Structural changes in the interbank market across the financial crisis from multiple core-periphery analysis. arXiv preprint [arXiv:1802.05139](https://arxiv.org/abs/1802.05139). (2018)
- Krugman, P.: Increasing returns and economic geography. *J. Polit. Econ.* **99**(3), 483–499 (1991)
- Laumann, E.O., Pappi, F.U.: *Networks of Collective Action: A Perspective on Community Influence Systems.* Elsevier, Amsterdam (2013)
- Lee, S.H., Cucuringu, M., Porter, M.A.: Density-based and transport-based core-periphery structures in networks. *Phys. Rev. E* **89**(3), 032810 (2014)
- Lu-An T., et al.: On discovery of traveling companions from streaming trajectories. In: 2012 IEEE 28th International Conference on Data Engineering. IEEE (2012)
- Ma, C., Xiang, B.-B., Zhang, H.-F., Chen, H.-S., Small M.: Detection of core-periphery structure in networks by 3-tuple motifs. arXiv preprint [arXiv:1705.04062](https://arxiv.org/abs/1705.04062) (2017)

- Malecki E.J.: Technology and economic development: the dynamics of local, regional, and national change. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship. (1997)
- Maslov, S., Sneppen, K.: Specificity and stability in topology of protein networks. *Science* **296**(5569), 910–913 (2002)
- Mullins, N. C., Hagens, L. L., Kick, H. E. L.: The group structure of cocitation clusters: a comparative study. *Am. Sociol. Rev.* **42**(4), 552–562 (1977)
- Nemeth, R.J., Smith, D.A.: International trade and world-system structure: a multiple network analysis. *Review (Fernand Braudel Center)* **8**(4), 517–560 (1985)
- Newman, M.E.J.: Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74**(3), 036104 (2006)
- Newman, M.: *Networks: An Introduction*. Oxford University Press, Oxford (2010)
- Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. *Phys. Rev. E* **69**(2), 026113 (2004)
- Noble, J.: General internal medicine in internal medicine: at the core or on the periphery. *Ann. Intern. Med.* **116**(12_Part_2), 1058–1060 (1992)
- Nocete, F., Sáez, R., Nieto, J.M., Cruz-Auñón, R., Cabrero, R., Alex, E., Bayona, M.R.: Circulation of silicified oolitic limestone blades in south-iberia (spain and portugal) during the third millennium bc: an expression of a core/periphery framework. *J. Anthropol. Archaeol.* **24**(1), 62–81 (2005)
- Nowicki, K., Snijders, T.A.B.: Estimation and prediction for stochastic blockstructures. *J. Am. Stat. Assoc.* **96**(455), 1077–1087 (2001)
- Palla, G., Derényi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435**(7043), 814 (2005)
- Ravasz, E., Barabási, A.-L.: Hierarchical organization in complex networks. *Phys. Rev. E* **67**(2), 026112 (2003)
- Rombach, M.P., Porter, M.A., Fowler, J.H., Mucha, P.J.: Core-periphery structure in networks. *SIAM J. Appl. Math.* **74**(1), 167–190 (2014)
- Rossa, F.D., Dercole, F., Piccardi, C.: Profiling core-periphery network structure by random walkers. *Sci. Rep.* **3**, 1467 (2013)
- Ruggera, R.A., Blendinger, P.G., Gomez, M.D., Marshak, C.: Linking structure and functionality in mutualistic networks: do core frugivores disperse more seeds than peripheral species? *Oikos* **125**(4), 541–555 (2016)
- Shanahan, M., Wildie, M.: Knotty-centrality: finding the connective core of a complex network. *PLoS One* **7**(5), e36579 (2012)
- Shneiderman, B., Plaisant, C.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Pearson Education India, New Delhi (2010)
- Smith, D.A., White, D.R.: Structure and dynamics of the global economy: network analysis of international trade 1965–1980. *Soc. Forces* **70**(4), 857–893 (1992)
- Snyder, D., Kick, E.L.: Structural position in the world system and economic growth, 1955–1970: a multiple-network analysis of transnational interactions. *Am. J. Sociol.* **84**(5), 1096–1126 (1979)
- Steiber, S.R.: The world system and world trade: an empirical exploration of conceptual conflicts. *Sociol. Q.* **20**(1), 23–36 (1979)
- Szymanski, B.K., Yener, B. (eds.): *Advances in Pervasive Computing and Networking*. Springer Science & Business Media, Berlin (2006)
- Tickner, A.B.: Core, periphery and (neo) imperialist international relations. *Eur. J. Int. Relat.* **19**(3), 627–646 (2013)
- Tudisco, F., Higham, D.J.: A nonlinear spectral method for core-periphery detection in networks. *SIAM J. Math. Data Sci.* **1**(2), 269–292 (2019)
- Verma, T., Russmann, F., Araújo, N.A.M., Nagler, J., Herrmann, H.J.: Emergence of core-peripheries in networks. *Nat. Commun.* **7**, 10441 (2016)
- Virtanen, P., Liukkonen, V., Vahtera, J., Kivimäki, M., Koskenvuo, M.: Health inequalities in the workforce: the labour market core-periphery structure. *Int. J. Epidemiol.* **32**(6), 1015–1021 (2003)
- Waenerlund, A.-K., Gustafsson, P.E., Virtanen, P., Hammarström, A.: Is the core-periphery labour market structure related to perceived health? findings of the northern swedish cohort. *BMC Public Health* **11**(1), 956 (2011)
- Wasserman, S., Faust, K.: *Social Network Analysis: Methods and Applications*, vol. 8. Cambridge University Press, Cambridge (1994)
- Xiang, B.-B., Bao, Z.-K., Ma, C., Zhang, X., Chen, H.-S., Zhang, H.-F.: A unified method of detecting core-periphery structure and community structure in networks. *Chaos Interdiscip. J. Nonlinear Sci.* **28**(1), 013122 (2018)
- Xie, J., Kelley, S., Szymanski, B.K.: Overlapping community detection in networks: the state-of-the-art and comparative study. *Acm Comput. Surv. (csur)* **45**(4), 43 (2013)
- Yan, B., Luo, J.: Multicores-periphery structure in networks. *Netw. Sci.* **7**(1), 70–87 (2019)
- Yang, J., et al.: Structural correlation between communities and core-periphery structures in social networks: evidence from Twitter data. *Expert Syst. Appl.* **111**, 91–99 (2018)
- Yang, J., Leskovec, J.: Overlapping communities explain core-periphery organization of networks. *Proc. IEEE* **102**(12), 1892–1902 (2014)
- Yuan, P., Ma, H.: Hug: Human gathering point based routing for opportunistic networks. In: 2012 IEEE Wireless Communications and Networking Conference (WCNC). IEEE (2012)
- Zhang, Y., Friend, A.J., Traud, A.L., Porter, M.A., Fowler, J.H., Mucha, P.J.: Community structure in congressional cosponsorship networks. *Phys. A Stat. Mech. Appl.* **387**(7), 1705–1712 (2008)
- Zhang, X., Martin, T., Newman, M.E.J.: Identification of core-periphery structure in networks. *Phys. Rev. E* **91**(3), 032803 (2015)



Wenli Tang received the bachelor's degree from Northeast Normal University in 2017. She is currently studying for the master's degree in computer science and technology at Beijing Institute of Technology. The research interests include social network, social computing, machine learning.



Liutao Zhao Research professorship, deputy director of Beijing Computing Center, Chairman of the board and general manager of Beijing Beike Rongzhi Cloud Computing Technology Co., Ltd.. Majoring in software engineering, big data, virtual reality, cloud computing as well as construction and operation of intelligent data center. Undertaking two tasks of national materials genome project, and having rich project management experience. Cocurrent instructor of postgraduates studying for master's

degree in Beijing University of Technology; concurrent instructor of postgraduates studying for master's degree in Beijing Institute of Technology.



Wei Liu PhD, associate professor, is director of User eXperience (UX) Master at Faculty of Psychology at Beijing Normal University, and board member of ICACHI. His research interests include developing methods, techniques, and tools to optimize design process for user centered interaction design. He holds a PhD on interaction design research from Faculty of Industrial Design Engineering at Delft University of Technology. LinkedIn profile: <http://www.linkedin.com/in/liuwei>.



Yiping Liu received the bachelor's and master's degrees in computer science from Beijing Institute of Technology in 2016 and 2017, respectively. His research interests include social computing, social networks, machine learning, and game theory. He will study for the Ph.D. at the University of Auckland.



Bo Yan Ph.D. candidate at Beijing University of Technology, is currently teaching at the Computer College of Beijing University of Technology. His research interests include data mining, cloud computing and big data applications. He has rich experience in project management and research and development. More than 30 academic papers have been published and 15 retrieved by EI.