Multiscale community geometry in a network and its application

Chen Chen
University of California, Davis, California 95616, USA

Hsieh Fushing*
Department of Statistics, MSB 4232, University of California at Davis, California 95616, USA
(Received 26 April 2012; revised manuscript received 17 August 2012; published 12 October 2012)

We introduce a between-ness-based distance metric to extract local and global information for each pair of nodes (or "vertices" used interchangeably) located in a binary network. Since this distance then superimposes a weighted graph upon such a binary network, a multiscale clustering mechanism, called data cloud geometry, is applicable to discover hierarchical communities within a binary network. This approach resolves many shortcomings of community finding approaches, which are primarily based on modularity optimization. Using several contrived and real binary networks, our community hierarchies compare favorably with results derived from a recently proposed approach based on time-scale differences of random walks and has already demonstrated significant improvements over module-based approaches, especially on the multiscale and the determination of the number of communities.

DOI: 10.1103/PhysRevE.86.041120 PACS number(s): 02.50.Ga, 36.40.Ei, 02.70.Hm, 64.60.aq

I. INTRODUCTION

In modern research, networks are collected and are studied in nearly all disciplines of sciences, and many previously unthinkable domains in the real world [1,2], including social dynamic relational networks [3], genetic pathway networks [4–6], ecological food-web and competition networks [7,8], financial payment and banking networks [9–11], and many others. These networks attempt to manifest relationships ranging from macroscopic to microscopic levels. Although they are diverse in format, these networks often successfully reveal intriguing structures of the system under study. The most intriguing structures, when a network is viewed as an approximation of a complex system, are primarily brought out through identifying communities [12]. This is why the investigation of community structures in networks has been a very intensive research area. Well-studied examples include Zachary's karate club [13], the scientific coauthorship network [14], the bottlenose dolphin network [15], and the protein-protein interaction network [16]. So far, the key idea underlying most approaches is the minimum-cut-maximum flow from graph theory. One of the most well-known approaches is the modularity optimization [17], which is popularized in a series approach along with many of its variants. These methods work well in partitioning the whole network into several separate communities with high intraconnectivity but with low interconnectivity [18]. Many of these identified communities are indeed capable of extracting meaningful structural information hidden in the underlying complex system.

However, as advances in data acquisition techniques have been revolutionized on a surprisingly rapid pace, much research in this area has raised a common issue that computational approaches for identifying communities need to be revised for adapting more conceptual and realistic features relevant to the complex system under study, such as hierarchical, multiscale, and overlapping structures [19–21]. In this paper, we particularly focus on the issue of the multiscale structure, which is significantly related to the hierarchical issue. From our point of view, the overlapping issue also heavily involves the multiscale one from the aspect of similarities among links or edges.

The resolution limit of communities—identified via modularity optimization and its variants—has been pointed out as a significant drawback by Fortunato and Barthelemy [22]. The importance of multiscale community detection has become better recognized, and resolving attempts are proposed [23–26] in order to completely depict the multilevel community structure in networks. Among these approaches, the one proposed by Delvenne *et al*. [27] is a step toward the direction of trying to unify the modularity optimization and the clustering idea via intrinsic time scale differences as random walk traveling within large and small communities.

In this paper, we will study the geometric sense of the multiscale structure among network communities. In addition to extracting information regarding whether a community, indeed, contains many denser but smaller communities, we also compute the information about the proximity between the communities. Basically, a pertinent geometry is constructed as the network community structure.

Two pieces of information are essential for any multiscale structure: (1) what and how many scales are relevant; and (2) given a relevant focal scale, how can we extract the right structural configuration. The two pieces of information will be rigorously extracted in this paper. The approach begins with defining a distance measure between any pair of nodes based on their global and local positions in the network. After identifying the outliers, data cloud geometry is applied as a nonparametric clustering algorithm to explicitly derive the two pieces of essential information mentioned [28].

Here, we link the community detection with the classic clustering mechanism. This approach is in even sharper contrast to existing optimization approaches than that proposed in Delvenne *et al*. [27]. A brief view of our approach is as

*Author to whom correspondence should be addressed: fushing@wald.ucdavis.edu

follows. We start with offering a new "distance" perspective for all possible pairs of nodes given its empirical edge connectivity in the network. This distance is calculated based on the concept of edge between-ness and defined along a shortest path between two nodes. The underlying idea is that, if a shortest path contains one or several large edge between-nesses, then these two nodes are likely to be apart. With such a distance, a distance matrix can be derived as an empirical one measured in a classic clustering setting.

How far apart two nodes are globally is not only decided by the distance between them, but is also critically determined by the focal scale used. To introduce the concept of scale, a sequence of different temperatures, taking values from very small to extremely large, is employed to transform the empirical distance matrix to a sequence of scale-sensitive similarity matrices, which are, therefore, manipulated into a Laplacian or Markovian transition matrix. A regulated random walk is then devised to effectively and exhaustively explore the whole collection of nodes to simultaneously reveal the clustering structure. In an ensemble fashion, such a piece of information is summarized into a clustering sharing probability matrix pertaining to the focal scale. The eigenvalue plot of such a matrix reveals the number of clusters involved, whereas, the properly constructed hierarchical clustering tree reveals the clustering memberships. Finally, phase transitions are searched along the evolution of the sequence of eigenvalue plots corresponding to the sequence of employed scales to determine the set of relevant scales. The geometry sense among communities is seen through a community merging process. Two communities being closer to each other merge earlier than two communities being far apart in the process of relevant scale changing from small to large. As a network is taken as an approximation of a complex system of interest, it is believed that this geometric perspective of multiscale network structure can offer a new hierarchical insight into the system under study. An equally important implication is that such structural information is potentially useful for comparing among different networks and, likewise, for comparing among different complex systems.

II. ESTABLISHING DISTANCES IN BINARY NETWORKS

We start with the description of a binary network $\mathcal{G} = \{\mathcal{N},\mathcal{E}\}$ with node collection $\mathcal{N} = \{N_1, N_2, \ldots, N_K\}$ and edge collection $\mathcal{E} = \{e_{k,k'}: 1 \leqslant k,k' \leqslant K\}$. Here, $e_{k,k'} = 1$ when an edge exists between two nodes N_k and $N_{k'}$, otherwise 0. The edge between-ness of an existing edge $e_{k,k'}$ is denoted as $b_{k,k'}$, and its number of common nearest neighbors of both ending nodes N_k and $N_{k'}$ is denoted as $c_{k,k'}$.

A monotonically increasing kernel function $H(\cdot)$ is applied to derive the distance. It can be simply $H(b) = b$ or $H(b) = e^b$. Another proper choice of $H(\cdot)$ is data driven, which can be derived as the reciprocal of the fitted smooth right tail of an edge between-ness histogram. The data-driven version is recommended when the number of node K is not small. The bottom line is that the kernel function keeps the order of the connections according to the between-ness. For convenience, we always normalize $H(b)$ by the maximal fitted value so that the kernel function eventually takes values between 0 and 1.

For an arbitrary pair of nodes (N_A, N_B), the corresponding path length l_{AB} is computed as the length of the shortest paths, and its full collection of shortest paths is computed as $\mathcal{P}(N_A, N_B)$. Any one of the shortest paths between N_A and N_B, say $(N_0, N_1, \ldots, N_{l_{AB}})$, with $N_0 = N_A$ and $N_{l_{AB}} = N_B$, corresponds to a vector of edge between-ness $(b_{0,1}, b_{1,2}, \ldots, b_{l_{AB}-1,l_{AB}})$.

With all ingredients for a pair of nodes (N_A, N_B), we define the distance as follows:

$$d(N_A, N_B) = \min_{\mathcal{P}(N_A,N_B)} \sum_{i=1}^{l_{AB}} H(b_{i-1,i}).$$

The distance aggregates the kernel function over all the segments in which intercommunity edges contribute much more than intracommunity edges.

In some cases, the component $H(b_{i-1,i})$ is replaced by the power transform $H(b_{i-1,i})^{1/(1+c_{i-1,i})}$, which is a device that we use to couple the global information of edge between-ness with the local information of the number of sharing nearest neighbors $c_{i-1,i}$. The piece of local information, as illustrated in the following example, is not necessarily included in the distance when the size of the network is large. However, the piece of local information does help us obtain a more reasonable community structure.

As we briefly examine this distance, it becomes clear that it yields a relatively small-scale distance for a pair of nodes within the same community and a relatively large distance for two nodes falling into two different communities. This fact effectively realizes the classic idea in clustering.

III. DATA CLOUD GEOMETRY AND COMMUNITY DETECTION

Once the distance matrix is obtained, an algorithm named data cloud geometry is applied to detect the multiscale community structure over the binary network. The algorithm proposed in Ref. [28] is to display the multiscale clustering structure in a given dataset. With an appropriate distance matrix defined, the topology of the data is shown on a series of scales, represented by the temperature T. Here, we briefly introduce the procedure of the algorithm at one temperature T:

(1) Identify the potential outliers, which are the nodes far from the others in the distance defined. The distance to the nearest node is chosen to be the proxy of the distance from each node. By calculating the interquartile range of the short distance distribution, we mark the nodes outside the upper fence as potential outliers to guarantee that no isolated nodes would be identified as communities.

(2) We calculate the similarity between any pair (N_A, N_B) at the temperature T by

$$s_{AB}(T) = \exp\left\{-\frac{d(N_A, N_B)}{T}\right\},$$

and, therefore, construct the similarity matrix under temperature T by

$$S(T) = [s_{AB}(T)]_{K\times K}.$$

As mentioned above, T is the scale parameter with which we view the network. Under a high temperature, the differences

between the distances are shrunk to be close, whereas, under a low temperature, the gaps between the distances tend to be magnified.

(3) A regulated random walk is established based on the Markovian transition matrix,

$$L(T) = D^{-1}(T)S(T),$$

where D is the diagonal matrix with

$$d_{ii} = \sum_{j=1}^{K} s_{ij}.$$

Compared to the traditional random walk, a modification has been made in this regulated one. A node is removed when it has been visited for a certain number of times. The number is always preassigned, for example, five times. By this modification, it is observed that the algorithm tends to remove all the vertices in one community one by one, and the random walk is forced to jump to another community after all the vertices in the previous community are removed. This pattern is more evident when the intracommunity distances are much smaller than the intercommunity distances. The successive time of each removal is recorded to make a profile in which a "spike" signals that the algorithm enters a new cluster. By detecting the spikes, we create a series of removal segments and conclude that, if two vertices are in the same removal segment, they are likely to be in the same community.

(4) The regulated random walk is repeated a large number of times. An ensemble connectivity matrix is then constructed with each element being the proportion of times in which the pair belongs to the same removal segments in the last step. It is actually an empirical estimate of the probabilities that any pair of nodes belongs to the same community. The ensemble connectivity matrix is more informative and reliable than the original similarity matrix or distance matrix.

(5) At a working temperature, a hierarchical clustering tree is built from the ensemble matrix and is then cut into a set of subtrees with a number of pieces determined from the eigenvalue plot of the ensemble matrix. The choice of the number of communities is empirical, usually to be selected as the number of eigenvalues which are significantly larger than zero [14,29]. In the following simulated and real network examples, we checked the drops in eigenvalue plots. All the eigenvalues larger than the last drop will be counted as nonzero eigenvalues.

(6) To confirm the outliers, the communities are marked to see whether the potential outliers connect between two different communities. A potential outlier would not be confirmed when it is actually linking two disconnected parts of the same community.

(7) The community structure will be detected on a series of temperature scales, which usually vary from very small to large. Plotting the number of communities against the temperature enables us to track the merging process and, therefore, to easily identify the stable phases. Later, in the simulation studies, the phase transition will be shown. It is noted that the outlier is related to the data geometry of the majority of data and is temperature dependent.

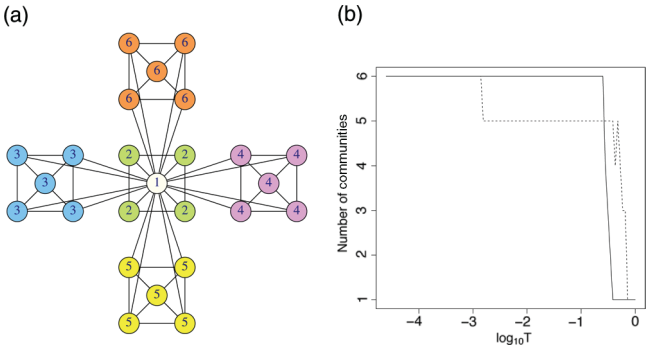

FIG. 1. (Color online) The fine level geometry with six communities is given in (a). Another stable phase is the trivial single-community structure. Both phases can be identified from the solid line in (b), which plots the numbers of communities against temperature scales. The dashed line is the the number of communities when the visit times to burn a node is set as three.

IV. SIMULATION STUDIES

A. Scale-free hierarchical networks

A small hierarchical network is analyzed to illustrate how the multiscale community structure is detected by the geometry we proposed. The network, shown in Fig. 1(a), was first raised in Ref. [30] as a deterministic network with a scale-free characteristic. For simplicity, only the network with one level of replicas is discussed here. By varying the temperature scale in a large range, only two stable phases: six community and a single community are detected through the merging process shown in Fig. 1(b). This process is very distinct from the type reported in Fig. 3 of Ref. [27] in which the stability function $r(t)$ smoothly goes through many phases: from 125 (total number of nodes) to 1. This result indicates that our multiscale community geometry is rather stable within each of its phases. The stability implies the intrinsic structural information being discovered in the network.

The label in Fig. 1(a) represents the community structure on the fine scale. (The trivial structure on the coarse scale is not shown.) Compared to the similar multiscale community structure given in Ref. [27], an extra community, which consists of the central node is given from our approach. The detection of this central vertex shows that the established geometry is even capable of perceiving a hierarchical structure, which is missing in Ref. [27]. Another notable aspect is that our multiscale community geometry does not reveal a phase in which each single node is a community as seen in Ref. [27]. We believe that this property should be avoided by any community detection approach.

This network also allows us to check the effect of changing the visit time for removing a node in the algorithm. An additional phase with five communities is detected when the visit time is set as three. In this phase, the central node (in red) is clustered with the four adjacent nodes (in yellow), which actually conforms to the results reported in Ref. [27]. In this case, the smaller visit time provides more stable phases, which complete the multiscale structure. However, it is not always the case in real network examples. For practitioners, selecting a series of visit times is one of the best choices.

B. Network with no hierarchical structure

In this simulation, a network with no hierarchical structure is constructed and is analyzed to perform a comparison between multiscale community geometries with the result obtained via modularity optimization. The illustrative example is a binary network as shown in Fig. 2(a), consisting of 21 motifs, and each of them is a fully connected network of four nodes. This network with the symmetric structure has no hierarchical structure since none of the motifs has a dominating number of connections, such as the central vertex in the last simulation study. Another significant distinction is that the one we show here is not scale free.

By varying the temperature scale, we found that the evolution of the community merging process has only passed through three stable phases: 21, 4, and 1 communities as shown in Fig. 2(b). Again, only short transition phases are present between any two stable phases, which provide a clear geometry structure on this simulated network.

To compare the multiscale community geometry with the community detection from modularity optimization, very briefly we review the modularity approach. The modularity approach optimizes the following distancelike quantity [17]:

$$Q = \frac{1}{4m}\sum_{i,j}\left[e_{ij} - \frac{d_i d_j}{2m}\right] s_i s_j,$$

where the edge e_{ij} and degrees d_i and d_j are pieces of local information and the total degree $m = \sum_i d_i$ is the global information and so are the prospective community indicators s_i, which equal 1 if N_i is in the first community and 0 is in the other community. By mixing the two pieces of information, the optimization targets finding a partition of two on the node's index set $(1, \ldots, K)$.

It is clear that the modularity optimization provides only a single scale of the community structure and has been shown to have the tendency to oversplit the network [22]. This tendency is also observed in this illustrative network as shown in Fig. 3. This result reveals an undesirable community structure.

In contrast, the multiscale community geometry is computed at temperature T ranging from 0.01 to 1. Before converging into the single-community phase, along the evolution of the community merging process, two apparent stable phases are found: a 21-community phase in Fig. 2(c) and a 4-community phase in Fig. 2(d). The former one is much more stable than the latter. It is also interesting to note that the five-community phase is even more transient than the four-community phase.

V. MULTISCALE COMMUNITY GEOMETRIES IN REAL WORLD NETWORKS

Three real world networks, from very small sizes to large sizes, are analyzed in this section for their multiscale community geometry. The karate club is a classic example, whereas, net-science coauthorship is popularly studied in recent network analysis literatures. The third example is one of a series of networks constructed from Lewis Carroll's English word game called Doublets [31]. This network has the unique feature of having many long dendrites. The presence of dendrites is expected to cause some computational difficulties in most existing community detection approaches.

A. Karate club

As a famous example in the community detection area, the karate club network was first introduced by Zachary [13]

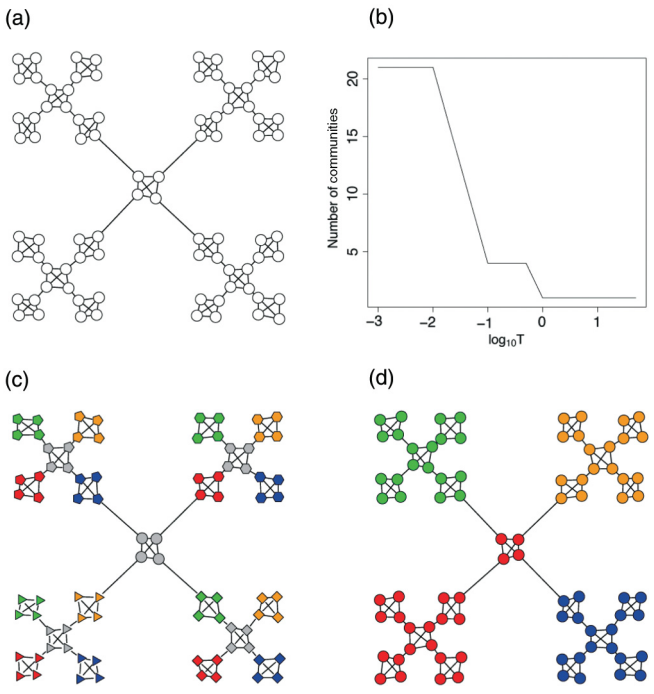

FIG. 2. (Color) The network, shown in (a), is simulated to illustrate the procedure. A series of temperatures is set to determine the number of communities shown in (b). The two stable phases, 21 community and 4 community, are provided in (c) and (d).

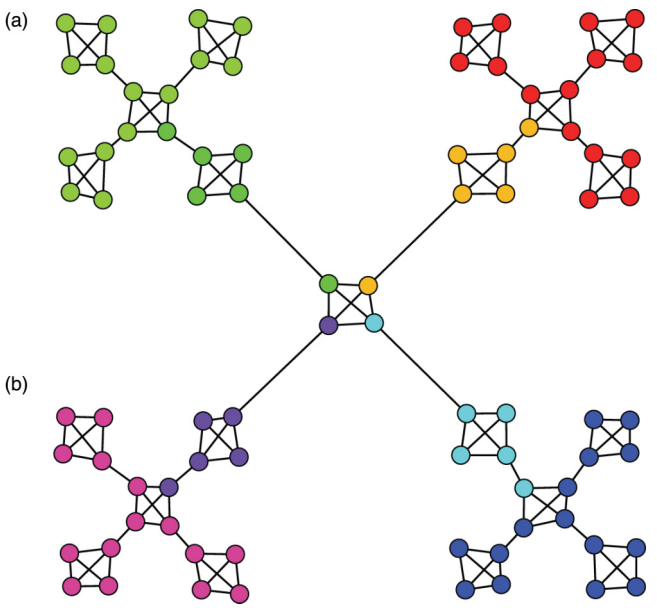

FIG. 3. (Color) The communities detected by the modularity optimization approach.

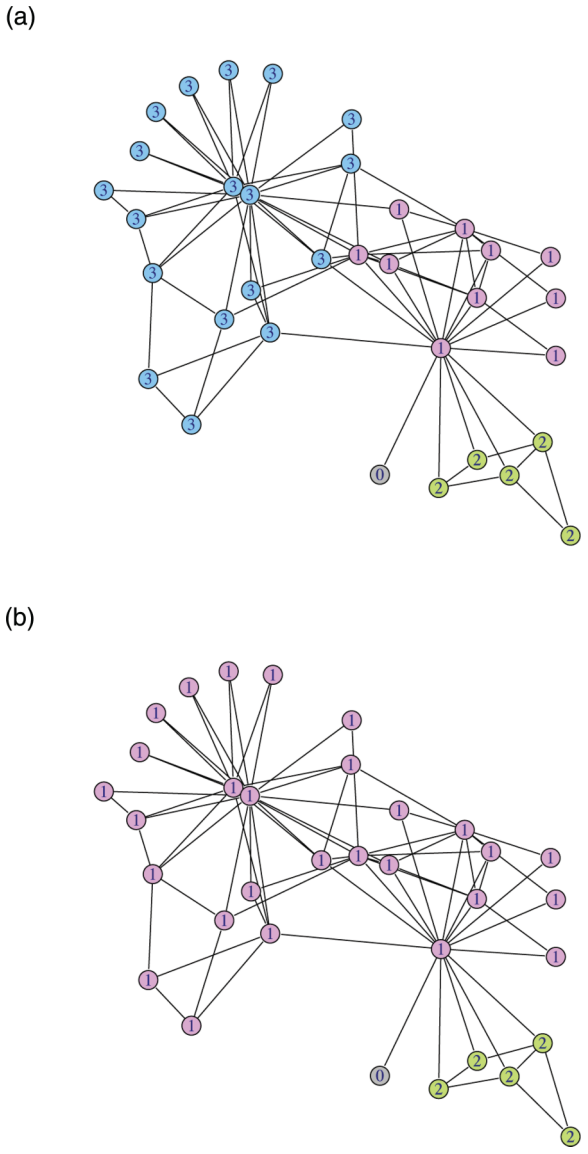

FIG. 4. (Color online) Two scales of geometry are detected on the karate club network, which are (a) the three-community structure and (b) the two-community structure. The node labeled 0 is an outlier.

and, to some extent, has been regarded as a benchmark for evaluating the effectiveness of the community detection algorithm. Our computed multiscale community geometry provides an intrinsically different view on the how the club could be split. Two levels of the geometric structure are reported in Fig. 4 to display the differences from most other community detection algorithms.

At the comparatively lower temperature $T = 0.02$, three communities are detected as shown in Fig. 4(a), whereas, two communities are presented in Fig. 4(b) at temperature $T = 0.05$. Only one outlier node is identified and is labeled 0. Without this outlier, the merging process shown in the figures gives rise to the third community (labeled 2) that has never been studied. As we can see, instead of merging communities 1 and 2 in Fig. 4(a) as in much literature, our approach leaves the second community alone and combines the others. It indicates

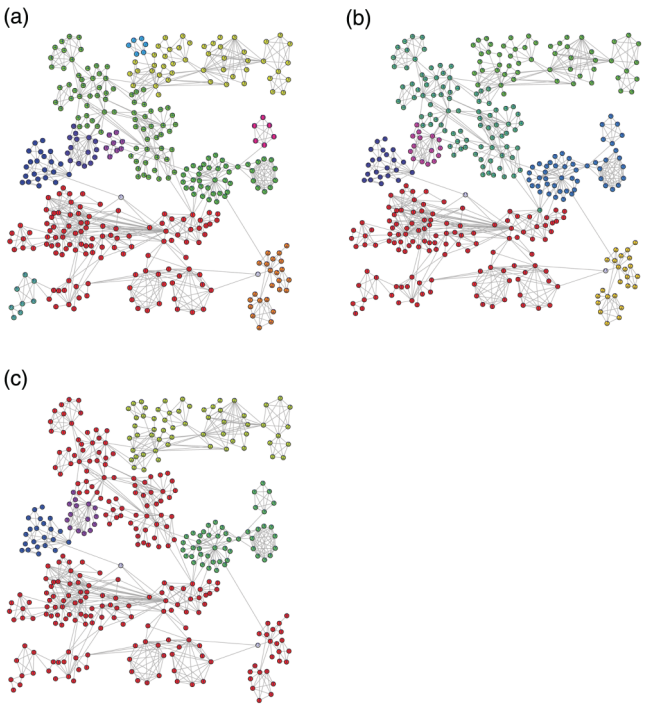

FIG. 5. (Color) Three scales of community geometry are detected on the coauthorship network, which are (a) the ten-community structure, (b) the seven-community structure, and (c) the five-community structure.

that these members in community 2 are even further away from the other members in the distance we have defined.

B. Net-science coauthorship

In this example, the network of collaboration in the net-science field is examined. It consists of 379 authors with the edges representing if the two authors have ever shared a publication in the net-science area. A community structure may depict the common research interest these authors share. Here, we report three levels of the multiscale community geometry in Fig. 5 with ten, seven, and five communities. The composition of the communities in Fig. 5(a) represents a fine partition over the network in which each community nearly represents a specific research topic. In Fig. 5(b), the communities of closer research topics are merged, whereas, the communities that are distant from the others remain unchanged.

In comparison, the multiscale community structure of this network, which was computed and was reported in Ref. [27], showed three stable levels with 21, 5, and 2 communities, respectively. The community structures at the five-community scale, which is the most stable one in both analyses, are coherent. The Rand index (with no potential outliers) is calculated as 0.874, which indicates the coherence between them.

We have to note that, in Fig. 5(c), only one of the potential outliers is confirmed at this temperature. The bottom right one connects the disjoint parts from community 1, which suggests that it is embraced into this community at this working

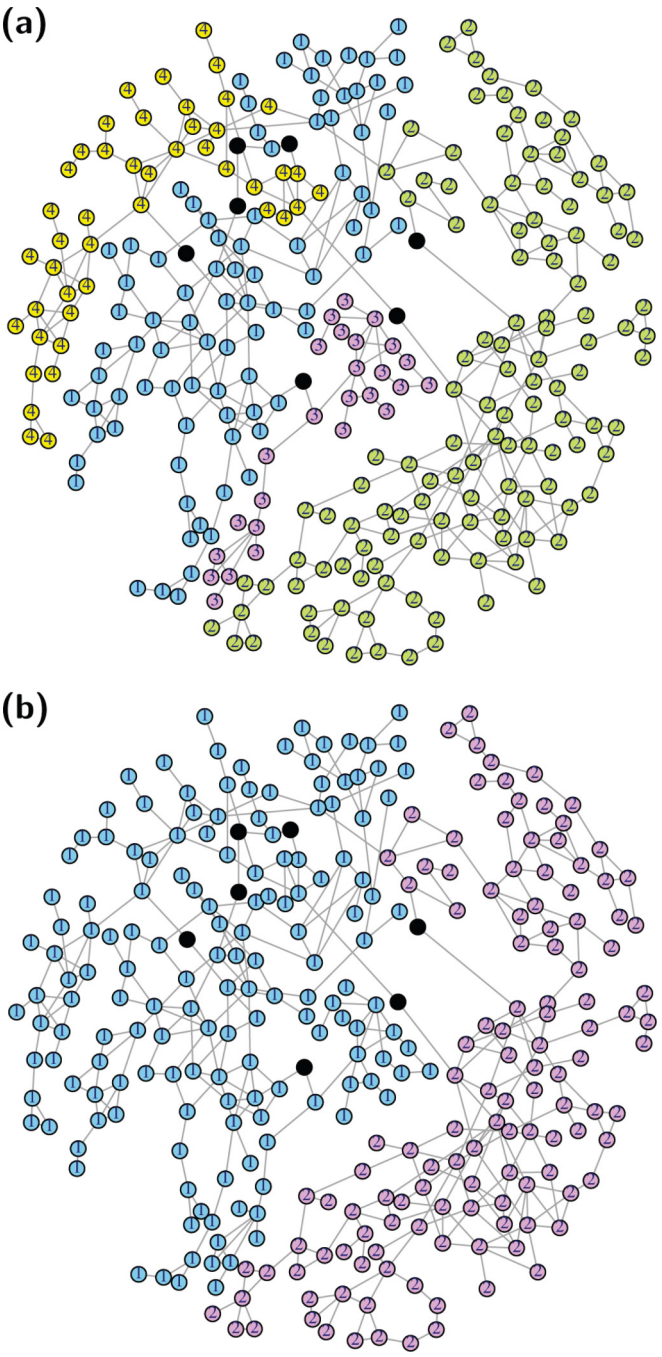

FIG. 6. (Color online) Two scales of geometry are detected on the eight-letter Doublets network, which are (a) the four-community structure and (b) the two-community structure.

temperature. Under the other two structures shown in Figs. 5(a) and 5(b), both potential outliers are confirmed.

C. Eight-letter Doublets

The Doublets network has recently been derived and has been constructed based on the word game Doublets, which was first created by Lewis Carroll and was studied in Ref. [31]. With all English words as the whole collection of nodes, a link is wired between two English words if they share the same alphabetic letters except one (obviously they are of the same length). In this example, only the largest connected clique of eight-letter words in which each word has eight letters is studied. The eight-letter Doublets network, shown in Fig. 6, consists of 291 vertices. Its multiscale structure is illustrated in the two panels of Fig. 6, corresponding to two temperature scales. The upper panel consists of four communities, whereas, the bottom one, under a higher temperature, has only two clusters. The composition of the detected communities usually reveals distinct English word structures with regard to linguistic constraints and phonological rules or even redundancy, see details in Ref. [31].

It is interesting to see that seven potential outliers are detected in which most of them are the nodes lying between the clusters illustrated in the upper panel. When the temperature is raised, shown in the bottom panel, many of the potential outliers should be merged into identified communities since they are merely intracommunity nodes.

VI. CONCLUSION AND DISCUSSION

In this article, we proposed a computational approach to derive the process of community evolution through the computed multiscale community geometry. Not only the formation of any conglomerate community, but also a distance metric among communities is recorded. This new in-depth perspective of a binary network is likely to offer a potential insight and a better understanding of the complex system to which a binary network attempts to approximate. At the same time, we unify the computations for community geometry with the classic clustering mechanism. This unification would broaden network analysis and, more importantly, would place it upon the same solid and rich mathematical foundations on which clustering mechanisms are based [28].

One of the most attractive advantages of this approach, compared to the optimization approaches and modeling methods, is its computational efficiency. The algorithm avoids nondeterministic-polynomial-time-hard computation by taking only the $O(K)$ computation time, which makes it capable to be applied to the community detection on large networks.

Another significant feature in this approach is that the number of communities could be naturally determined at each temperature scale. By introducing the distance into the network analysis, we provide a way to compute the geometry of the underlying complex system and, therefore, the number of communities. In contrast, a preassigned number of communities has been a persistent flaw suffered by many model-based approaches.

Although the distance among the nodes provides significant improvement in understanding the system, its definition could be quite empirical. Some suggestions have been given and have been illustrated in this article. However, a more subject-knowledge-based definition should be expected to be more effective in most of the real world studies.

ACKNOWLEDGMENT

This research is supported, in part, by the NSF under Grant No. DMS-1007219.

[1] M. E. J. Newman, A.-L. Barabási, and D. J. Watts, *The Structure and Dynamics of Networks* (Princeton University Press, Princeton, New Jersey, 2006).
[2] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, Science **323**, 892 (2009).
[3] A. Cho, Science **325**, 406 (2009).
[4] I. Farkas, H. Jeong, T. Vicsek, A.-L. Barabási, and Z. N. Oltvai, Physica A **318**, 601 (2003).
[5] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, Nat. Genet. **31**, 64 (2002).
[6] N. Guelzim, S. Bottani, P. Bourgine, and F. Képès, Nat. Genet. **31**, 60 (2002).
[7] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[8] J. M. Montoya and R. V. Sole, Oikos **102**, 614 (2003).
[9] G. Iori, S. Jafarey, and F. G. Padilla, J. Econ. Behav. Organ. **61**, 525 (2006).
[10] G. De Masi, G. Iori, and G. Caldarelli, Phys. Rev. E **74**, 066112 (2006).
[11] R. M. May and N. Arinaminpathy, J. R. Soc., Interface **7**, 823 (2010).
[12] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).
[13] W. W. Zachary, J. Anthropol. Res. **33**, 452 (1977).
[14] M. E. J. Newman, Phys. Rev. E **74**, 036104 (2006).
[15] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, Behav. Ecol. Sociobiol. **54**, 396 (2003).
[16] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, Nature (London) **411**, 41 (2001).
[17] M. E. J. Newman, Proc. Natl. Acad. Sci. USA **103**, 8577 (2006).
[18] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, Proc. Natl. Acad. Sci. USA **101**, 2658 (2004).
[19] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).
[20] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[21] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Nature (London) **435**, 814 (2005).
[22] S. Fortunato and M. Barthélemy, Proc. Natl. Acad. Sci. USA **104**, 36 (2007).
[23] P. Pons and M. Latapy, Theor. Comput. Sci. **412**, 892 (2011).
[24] A. Arenas, A. Fernandez, and S. Gomez, New J. Phys. **10**, 053039 (2008).
[25] A. Lancichinetti and S. Fortunato, Phys. Rev. E **80**, 056117 (2009).
[26] P. Ronhovde and Z. Nussinov, Phys. Rev. E **81**, 046114 (2010).
[27] J. C. Delvenne, S. N. Yaliraki, and M. Barahona, Proc. Natl. Acad. Sci. USA **107**, 12755 (2010).
[28] H. Fushing and M. P. McAssey, Phys. Rev. E **82**, 061110 (2010).
[29] M. Fiedler, Czech. Math. J. **25**, 619 (1975).
[30] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[31] H. Fushing, C. Chen, P. Scully, and P. Farrell (unpublished).