Formation of Modularity in a Model of Evolving Networks

Menghui Li,¹,² Shuguang Guan,³ and Choy-Heng Lai²,⁴

¹Temasek Laboratories, National University of Singapore, 117508, Singapore
²Beijing-Hong Kong-Singapore Joint Centre for Nonlinear & Complex Systems (Singapore), National University of Singapore, Kent Ridge, 119260, Singapore
³Institute of Theoretical Physics and Department of Physics, East China Normal University, Shanghai, 200062, P. R. China
⁴Department of Physics, National University of Singapore, Singapore 117542

(Dated: March 29, 2011)

Modularity structures are common in various social and biological networks. However, its dynamical origin remains an open question. In this work, we set up a toy dynamical model describing the evolution of a social network. Based on the observations of real social networks, we introduced a strategy of link-creating/deleting according to the local dynamics in the model. Thus the coevolution of the dynamics and topology naturally determines the network properties. It is found that for a small coupling strength, the networked system cannot reach any synchronization and the network topology is homogeneous. Interestingly, when the coupling strength is large enough, the networked system spontaneously forms communities with different dynamical states. Meanwhile, the network topology becomes heterogeneous with modular structures. It is further shown that in certain parameter regime, both the degree and the community size in the formed network follow power-law distribution. These results are consistent with the characteristics of many empirical networks, and are helpful to understand the mechanism of formation of modularity in complex networks.

Real-world complex networks usually have certain universal properties, such as small-world, scale-free, and modularity [1–4]. Scale-free means that the degree of a network follows a power-law distribution, and modularity refers to that networks typically consist of communities or clusters in which the nodes are more highly connected to each other than that to the rest of the network [1]. In the past decade, many evolving network models have been proposed from the topological perspective, aiming at constructing networks with power-law distribution of degree [3] and modular structure [5, 6]. However, from functional perspective, taking biological networks for instance, modularity is generally believed to correspond to certain functional groups, where nodes within the same group may share similar dynamical characteristics [3]. This implies that in principle the network topology and dynamics are strongly dependent on each other. In fact, any formed network structures and dynamical patterns are actually the result of the coevolution of both network dynamics and topology [8]. For example, in various biological and social networked systems, such as the email networks [9] and the mobile communication networks [4], individuals are more likely to interact with others “similar” to themselves [10], which usually leads to networks consisting of communities driven by shared activities, attributes, and affiliations, and so on.

The formation mechanism of modularity and scale-free property is crucial to the understanding of the structural and functional/dynamical properties of complex networks. Recently increasing attentions have been paid to the adaptive coevolutionary networks [8]. However, so far, it has not been well understood how modular structure and power-law degree distribution concurrently emerge in an evolving network. Motivated by the above idea, in the present work, we set up a toy model of dynamical network whose nodes are represented by phase oscillators. Basically, the model is an evolving network which grows from a few nodes at the very beginning. The most important characteristic of the model is that it describes the interplay between the topological structure and the dynamics on the network. On the one hand, the node dynamics are coupled to each other according to the network topology; on the other hand, the connections among nodes can be created or deleted according to the local dynamical states. Our particular interest is focused on what types of network structures can be formed as a result of the coevolution of network dynamics and topology. Mainly, our study presented the following results: (i) The communities, within which the nodes have similar dynamical states and the connections are denser than outside, can be naturally formed during the network evolution. (ii) In certain parameter regime, the degree distribution and the distribution of community size follow power law. These properties exhibited by the model concurrently emerge as the network evolves, and are consistent with the observations in many real social and biological networks.

In our model, the node dynamics are simplified as phase oscillators, which are coupled as in the following dynamical equations:

$$\dot{\theta}_m = \omega_m + \frac{\gamma}{k_{mn}} \sum_{n=1}^{N} a_{mn} \sin(\theta_m - \theta_n + \phi_{mn}).$$

Here, $m, n = 1, 2, \ldots, N$ are the oscillator (node) indices, and $\gamma$ is the uniform coupling strength. $A = \{a_{mn}\}$ is the adjacency matrix, where $a_{mn} = 1$ if nodes $m$ and $n$ are connected, and $a_{mn} = 0$ otherwise. $\{\omega_m\}$ are the intrinsic frequencies of oscillators, and $k_{mn}$ is the degree of oscillator $m$. We noticed that in many social networks individuals tend to contact with others with similar attributes. Furthermore, if two individuals are in the same environment, they are also more likely to make friends with each other. Considering the
influence of environment, in our model we particularly introduced an extra phase coupling term $\phi_{mn}$ as

$$
\phi_{mn} = (\psi_m - \psi_n) \mod (\pi).
$$

Here $\psi_m$ is the average phase of the local order parameter which is defined as

$$
r_m e^{i\psi_m} = \frac{1}{k_m} \sum_{n=1}^{N} a_{mn} e^{i\theta_n}.
$$

In an adaptive networked system, the network topology usually changes according to the dynamical interaction among nodes during the evolution. Our model is basically a growing network which starts from a few seed nodes at the very beginning. Particularly, the evolution rule of the network incorporates two main manipulations: one is the node-adding, and the other is link-adjusting, including adding and removing links. It has been shown that in empirical networks [9, 21], most new connections are likely to build up between one node and its second neighbors, i.e., the neighbor’s neighbor. Based on this idea, in our model we proposed a link-adjusting strategy according to the local dynamical states. Depending on the local order parameter $r_m$, an active individual has two options to adjust its link. One option is that if his neighbors do not reach a consensus, the individual is free and can link to any other individuals. Otherwise, if his neighbors reach a consensus, he is “frozen” in its neighborhood and can only make new link to his second neighbors. Moreover, existing links may be removed from the existing network for various reasons [22]. This effect has also been considered in our model. Specifically, the dynamics and topology of the networked system coevolve according to the following rules.

1. At the very beginning, the network consists of $N_0$ isolated oscillators, whose dynamical states evolve according to Eq. (1). The initial states of oscillators are randomly selected from $[-\pi, \pi]$, and their intrinsic frequencies are chosen from the distribution of $g(\omega) = 0.75(1 - \omega^2)$.

2. At every $T_a$ time steps, one new node is added into the network by connecting it to an arbitrarily selected node in the existing network.

3. At every $T_r$ time steps, one node $m$, randomly selected from the existing network, will be activated to adjust its links. Suppose the local order parameter of node $m$ is $r_m$, which characterizes the local coherence around node $m$. Then with total probability $r_m$, node $m$ first chooses one of its nearest neighbors $k$ according to the following probability partition function

$$
\prod select = f(m, k) = \frac{L_{mk}(\Delta \theta_{mk})}{\sum_{j \in \partial_m} L_{mj}(\Delta \theta_{mj})},
$$

where $\partial_m$ denotes the set of the nearest neighbors of the node $m$. $L_{mk}(\Delta \theta_{mk}) = [1 + \cos(\Delta \theta_{mk})]/2$ is the similarity distance [23][24]. According to the definition, the larger the similarity distance is, the closer the dynamical states between these two connecting oscillators are. Then node $k$ introduces one of its nearest neighbor $i$ to node $m$ according to probability partition function

$$
\prod add = f(k, i)_{i \in \partial_k, i \in \partial_m}.
$$

In this way, a new link between $m$ and one of its second neighbor $i$ is established. Meanwhile, with the total probability $1 - r_m$, node $m$ connects to an existing node in the network according to probability partition function $f(m, j)_{j \notin \partial_m}$.

4. Parallel to the above step, node $m$ will remove one of its existing connections with the total probability $p$ according to the following probability partition function

$$
\prod cut = \frac{1}{\sum_{j \in \partial_m} 1/L_{mj}(\Delta \theta_{mj})},
$$

This implies that the probability removing the link between $m$ and $k$ is inversely proportional to their similarity distance. According to the above rules, given the parameters, the formed network will approximately has a constant average degree

$$
\langle k \rangle \approx 2[1 + \frac{T_r}{T_a}(1 - p)],
$$

which is independent on the network size. However, the degree distribution can be rather heterogeneous. In order to investigate the possible communities of the formed network, we
employed the $k$-clique percolation method, in which communities can be defined in terms of adjacent cliques [25]. Specifically, we calculated the following quantities to characterize the community structure in network: the average cluster size $⟨s⟩$ (except the largest one), the relative largest cluster size $R_L$, and the average clustering coefficient $⟨C⟩$. The average community size $⟨s⟩$ is defined as

$$⟨s⟩ = \frac{\sum s^2 n_s s}{\sum s n_s s},$$

where $s$ is the size of cluster and $n_s$ is the number of $s$-size cluster. The sums run over all possible values of $s$ but the largest cluster. The relative largest cluster size $R_L$ is defined as

$$R_L = \frac{s_L}{N},$$

where $s_L$ is the size of the largest cluster, and $N$ is the size of system. In our model, the active node can connect to one of its second neighbors with certain probability, which generates at least one triangle. Therefore, we focused on 4-clique and 5-clique in our computation.

Now we report the main results of our numerical simulations. We first study how the network evolves with the increase of the coupling strength. Figure 1 shows various properties of the final networks with respect to the coupling strength $\gamma$. It is found that when $\gamma \to 0$, both $⟨s⟩$ and $R_L$ are very small. This means that the network structure is basically homogeneous and no obvious clusters are formed at this stage. However, the network structure significantly changes when $\gamma$ is large enough. For instance, when $\gamma > 2$, both $⟨s⟩$ and $R_L$ are significantly large indicating that the network structure is characterized by limited number of communities. This can be further verified by the average clustering coefficient $⟨C⟩$ as shown in Fig. 1(c). A careful examination of the dynamical states on the network shows that when the coupling strength $\gamma$ is small, the system can reach neither synchronization nor clustering. In this case, the links are almost randomly generated, so the formed network is a random one. There is no obvious communities as shown in Fig. 2(a). On the contrary, when $\gamma$ is sufficiently large, clustering occurs in the networked system with the growth of the network size. Physically, there are two factors affecting the formation of dynamical groups. On the one hand, according to the network-growing rule, a node is more likely to connect to its second neighbors with similar states. This will gradually generate a core of oscillators which are partially synchronized. On the other hand, when two oscillators are in two different dynamical groups, their “environment” are different, i.e., $\phi_m$ between them is large. Since large $\phi_m$ does not in favor of synchronization [26], these two nodes have less chance to build a connection between them. In addition, once the phase difference between two oscillators is large, it is likely that the existing link between them will disconnected, while new link is hardly generated between them. Therefore, once a dynamical group, i.e.,

![FIG. 2: Typical examples of the final network: adjacent matrix (top) and the final phase states of oscillators (bottom). (a) $\gamma = 0.2$. Approximately, the network is homogeneous and the phases of oscillators follow a uniform distribution. (b) $\gamma = 6$. Communities is formed, and the oscillators within the same community have close phase states. The indices of the oscillators have been rearranged according to the phase. Other parameters are the same as those in Fig. 1.](image1)

![FIG. 3: (Color online) The topological properties of final networks of $N = 10000$. (a) The distribution of degree. Note that the inset plot is on log-normal coordinate. (b) The distribution of community size for 4-clique, where $\gamma = 6$. Other parameters are the same as those in Fig. 1.](image2)
community size, which has been observed in many empirical networks. In Fig. 3 (a), it is also observed that after a long time evolution, the oscillators self-organize into many communities, both dynamically and topologically as shown in Fig. 2 (a). This is because that when the coupling strength is small, the interactions among nodes are weak. Thus it is difficult to form dynamically coherent state on the network. In this case, most new links are generated randomly. As a result, the network is of random nature and the degree approximately follows an exponential distribution as shown in Fig. 3 (a). Nevertheless, when the coupling strength becomes large enough, e.g., \( \gamma > 2 \), the formation of communities is greatly enhanced by the dynamical interactions among nodes on the network. In this case, we observe approximate power-law distribution of degree as shown in Fig. 3 (a) in certain parameter regime. In Fig. 3 (a), it is also seen that the the power-law distribution has the “droop head” shape, which has been observed in many empirical networks [28]. In our study, it is found that the distribution of community sizes, also approximately follow a power law, as shown in Fig. 3(b), a phenomenon which has been observed in many empirical networks [28].

In this work, we have further investigated how the formed network structures depend on the main control parameters, i.e., \( T_a/T_r \), \( p \), and the final network size \( N \). The results are presented in Fig. 4. As shown in Fig. 4(a), the parameter \( T_a/T_r \) has remarkable effect on the properties of the formed networks. The main network quantities, such as \( \langle s \rangle \), \( R_L \), and \( C \), all increase with the increase of \( T_a/T_r \). In our model, the ratio \( T_a/T_r \) represents the interaction between two opposite regulations. \( T_a \) is the time period to “purposely” build a link, while \( T_r \) is the time period to “purposely” build a link. The former manipulation is in favor of increasing randomness in the network, while the latter manipulation help generate triadic closure inside the network. Larger \( T_a/T_r \) means that the second factor is dominant, and this explains why network quantities \( \langle s \rangle \), \( R_L \), and \( C \) increase when \( T_a/T_r \) becomes larger. In our model, parameter \( p \) is the probability deleting links with small similarity distance. When \( p \) is relatively small, this manipulation helps forms dynamical groups, where oscillators within the same group have similar dynamical states, and this process in turn enhance the formation of topological communities in the network. However, with further increase of \( p \), according to Eq. 7, the network gradually becomes too sparse to form any distinct communities. Therefore, there exists an optimal parameter range \( 0.2 < p < 0.4 \), where communities with various sizes can be formed in the network. In our model, the final network size \( N \) is proportional to the total integration time. As shown in Fig. 4(c), all three network quantities \( \langle s \rangle \), \( R_L \), and \( C \) become approximately stationary when the network evolves for a long time.

To summarize, we have studied a network-growing model of phase oscillators, in which the dynamics and the network topology interact with each other and concurrently
evolve. Following simple rules of node-adding, link-adding/removing, the model exhibits several interesting behaviors. Within certain parameter range, the dynamical communities and the topological modules can spontaneously emerge in the network. It is found that in the formed networks, the degree and the community size approximately satisfy power-law distributions. Our study also reveal that the interaction strength among nodes on network can essentially determine the formation of network structures, both dynamically and topologically. This is an important point which has been ignored or over simplified by many previous network-growing models. The findings in this work captured the typical properties of many realistic networks. Thus it is helpful for us to further understand the complicated interaction between network topology and the dynamics.

This work is supported by Temasek Laboratories at National University of Singapore through the DSTA Project No. POD0613356. SGG is sponsored by the Science and Technology Commission of Shanghai Municipality under grant no. 10PJ1403300, and also by the NNSF of China under grant no. 11075056.