

# Effects of Topology in Networked Systems: Stochastic Methods and Small Worlds

John S. Baras and Pedram Hovareshti

**Abstract**—The topology of a networked control system has critical consequences for its performance. We provide first substantial examples on the effects of topology. Then we proceed to develop a rigorous evaluation of topology effects using stochastic methods inspired from statistical physics and Markov chains. This analysis leads us to proofs on faster convergence of distributed algorithms in networked systems for certain topologies and especially small world topologies, which are given an ‘efficiency’ characterization. Finally, these results lead to the development of self-organization of such systems in hierarchies that provably improve performance and response.

## I. INTRODUCTION

In recent years the study of networked control systems has attracted substantial interest in the control community. An essential aspect of many of these systems is the lack of a central control authority: distributed control rules and algorithms are often utilized due to a host of reasons including energy considerations and reliability. Examples of different applications of local control rules and distributed algorithms in recent literature include flocking schemes and consensus in collaborative control of groups of robots [7], [23], [22], [18], synchronization of oscillators [19], trust establishment in a network of autonomous entities [2], [3], gossip algorithms and cooperation in sensor coverage [9], [27].

In all of these distributed schemes the agents are provided with simple sets of decision making algorithms or dynamics, such that each agent takes an action using its local information. The actions that each agent performs are also local, i.e. only are affected by and affect their neighbors. The goal of the system is to *emerge* a desired global behavior from the local interactions. The effectiveness of these schemes depends on three important factors: 1) The speed of convergence. 2) Robustness to agent/connection failures. 3) Energy/communication efficiency. There is usually a trade off in achieving these objectives. It is important to notice that the speed of convergence and robustness depend on both the structure of the network and the dynamics of the agents.

In this paper, we focus on the structural aspects of the network. We show that well-connectedness of agents significantly affects the convergence and robustness of the distributed algorithms running on the network. Since the decentralized nature of the algorithms require each agent

to be connected to only a few other agents, achieving high connectivity at the same time may seem counter-intuitive and even paradoxical. However, certain classes of graphs maintain this property. For example, *small world topologies* offer a favorable tradeoff between performance (convergence speed) versus cost of collaboration (connectivity cost) [4].

The subject of dynamic systems on graphs has also gained attention in other communities. Following the popular small world model of Watts and Strogatz [26] and the preferential attachment model of Barabasi and Albert [1], substantial research efforts have been dedicated towards constructing a scientific framework to study networks and the processes running on them. The main interest of the ongoing research in this community is to understand the formation and function of real world networks such as the World Wide Web, social networks, and biological systems.

The two research areas explained above have some overlap. Although the results in network science are usually asymptotic, they can provide the control society with useful insight. As an example, the study of mixing times of random walks on random and small world graphs [13], is in close connection to the convergence rate of consensus schemes considered in the control literature as mentioned in [24].

We address the small world effect in the context of perturbation of stochastic matrices. We also propose a hierarchical self-organization method and show that, by utilizing it, the network is capable of running distributed algorithms with high convergence speed. We generalize the concept of social leaders introduced by Blondel et al. [8] to classify the agents of more importance in a network. The scheme is based on a two stage algorithm which first finds the most effective choice of local leaders, and then provide nodes with information about their location with respect to other nodes and leaders and the choice of groups to form. This leads to self-organization of the systems in the hierarchy and improves performance and response.

The paper is organized as follows. We start by providing substantial examples on the effects of topology in the next section. In section III, we provide an overview of small world networks and address the convergence of consensus algorithms on them in the framework of perturbation of stochastic matrices. The two-level hierarchical self-organization scheme for faster convergence is developed in section IV.

## II. THE IMPORTANCE OF BEING WELL-CONNECTED

In this section we study some motivational examples that show the importance of being ‘well connected’ in distributed algorithms. The first two examples given in this section

Research supported by the U.S. Army Research Office under Award No DAAD 190110494, and by the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance program cooperative agreement DAAD19-01-2-001.

J.S. Baras and P. Hovareshti (baras@isr.umd.edu, hovaresp@isr.umd.edu) are with the Institute for Systems Research and the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA.

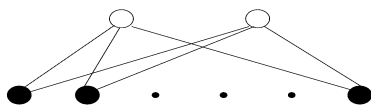


Fig. 1. Two “No-voters” control the local majority poll (from [16])

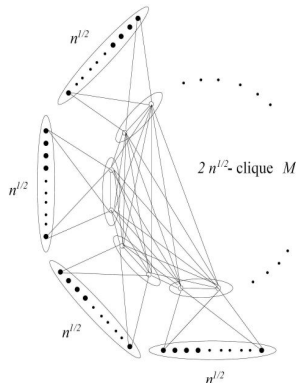


Fig. 2. A small coalition controls the local majority poll (from [16])

demonstrate how being well connected can serve agents’ selfish purposes. The next examples show the merits of being well connected for the performance of the overall system.

A. Local Majority voting

We start with a classic example by Peleg [16], which shows that in voting schemes a few well connected nodes can determine the outcome of the process. Consider  $n$  citizens each living on a vertex of a graph. Each citizen has an opinion about voting “Yes” or “No” on a subject. However, citizens observe a rule by which they first ask privately their neighbors’ opinions. Each person then casts “Yes”( resp. “No”) if the majority of their neighbors are “Yes-voters”(resp. “No-voters”). An important question is: what is the minimum number of “No-voters” that can guarantee a “No” outcome. As shown in Figure 1 the answer is 2. Every one of the  $n-2$  “Yes- voters” should change its vote, because of having two “No-voters” in their neighborhood.

It is worthwhile to notice that all the “Yes-voters” observe a 2 to 1 majority of “No-voters” in their neighborhood. However, each “No-voter” observes a huge majority of “Yes-voters” in their neighborhood. If the nodes follow the polling rule iteratively, each node will oscillate between Yes and No. However, if the “No-Voters” fail to observe the rule, the iteration will converge and all the nodes will vote No after the first iteration. Peleg also shows that for large  $n$ , a negligible minority of  $2\sqrt{n}$  “No-voters” can force all the voters to decide to vote No in just one iteration. This can be achieved by a clique of well connected “No-voters” who are attached to groups of badly connected “Yes-voters” as in Figure 2. In this case, by following the rules the “No-voters” can force the “Yes-voters” to change their vote while maintaining their own No votes in a single iteration.

B. Consensus schemes

Consider a set of  $n$  agents and the interconnection between them modelled by a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . The nodes of the graph,  $\mathcal{V} = \{1, 2, \dots, n\}$  represent the agents. If node  $i$  is able to communicate with node  $j$  a directed edge exists between the two nodes. Given an  $n$  by  $n$  matrix  $F$ , the graph  $\mathcal{G}(F)$  is defined to be a graph on  $n$  nodes  $\mathcal{V} = \{1, 2, \dots, n\}$  with an edge from node  $i$  to node  $j$  if and only if  $F_{ij} \neq 0$ . In the context of this paper, we consider the case of symmetric interconnections. Therefore, the graphs considered here are undirected. By a discrete time consensus equation, we mean a linear iteration of the form:

$$x(k + 1) = F(k)x(k) \tag{1}$$

in which  $F(k)$  is a nonnegative stochastic matrix. The entry  $F_{ij}$  denotes the weight that node  $i$  applies to the values that it gets from node  $j$  and its value may be time dependent. If there is no link between nodes  $i$  and  $j$ ,  $F_{ij} = 0$ .

A natural set of weights come from Vicsek’s model for leaderless coordination [18], [25]. Vicsek’s model assigns weights to neighbors of a node in a way in which each node performs a local averaging in its neighborhood. Therefore,

$$F(k) = (I + D(k))^{-1}(A(k) + I), \tag{2}$$

where  $A$  is the adjacency matrix of the graph  $\mathcal{G}$  and  $D$  is the diagonal matrix with each node’s degree on the corresponding diagonal. Consider the fixed topology case. Vicsek’s model in this case is related to a random walk on a fixed graph with self loops. If the graph is connected, then all the agents’ state variables will converge to  $x_\infty = \sum_{i=1}^n \pi_i x_i(0)$ . It can be easily verified that  $\pi_i = \frac{n_i + 1}{2l + n}$ , where  $l$  equals the number of edges in the graph. Therefore each agent’s contribution to the consensus is proportional to the number of neighbors it has. Therefore, for example in Figure 1, the ratio of the contribution of white agents to the contribution of black agents is of order  $O(n)$ .

C. Robustness

In order to be functional, distributed algorithms need to be robust to agent and link failures. In [14], Gupta *et al* address the robustness issue in distributed algorithms using a cost function approach. [14] defines three failure modes 1,2, and 3, respectively according to the faulty agent’s ceasing functionality, propagation of constant arbitrary values, and propagation of varying values. It is shown that consensus problems are robust to  $l - 1$  node failures of mode 1, in an  $l$ -connected graph. A classic theorem in graph theory states that the minimum number of vertices separating two independent set of nodes is equal to the maximum number of disjoint paths between them [12]. Thus being well-connected improves the robustness of distributed algorithms.

D. Fast convergence in consensus schemes

Since the iteration matrices  $F$  are non-negative and stochastic, the study of convergence of consensus schemes can be achieved by the study of the corresponding Markov chains. For fixed topology, consensus will be reached if and

only if  $F$  is primitive, i.e. the graph should be strongly connected with at least one self loop. Convergence of consensus algorithms in the presence of topology changes has been extensively studied and different sufficient conditions have been proposed based on different assumptions on the connectivity as well as synchrony ([15] and the references therein). Under our symmetric communication assumption, a sufficient condition is existence of infinite nonempty, bounded contiguous time intervals, across which the union of the communication graphs is strongly connected [18].

If  $F$  is a primitive stochastic matrix, according to the Perron-Frobenius theorem [10],  $\lambda_1 = 1$  is a simple eigenvalue with a right eigenvector  $\mathbf{1}$  and a left eigenvector  $\pi$  such that  $\mathbf{1}^T \pi = 1$ ,  $F^\infty = \mathbf{1}\pi^T$  and if  $\lambda_2, \lambda_3, \dots, \lambda_r$  are the other eigenvalues of  $F$  ordered in a way such that  $\lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_r|$ , and  $m_2$  is the algebraic multiplicity of  $\lambda_2$ , then

$$F^t = F^\infty + O(t^{m_2-1}|\lambda_2|^t) = \mathbf{1}\pi^T + O(t^{m_2-1}|\lambda_2|^t)$$

This shows that the convergence of the consensus protocol is geometric, with relative speed equal to SLEM.

We denote by  $\mu = 1 - SLEM(G)$  the spectral gap of a graph, so graphs with higher spectral gaps converge more quickly. If the matrix  $F$  is symmetric, its SLEM can be written as the norm of its restriction to the subspace orthogonal to  $\mathbf{1} = [111\dots 1]^T$ . In general the SLEM of  $F$  matrices are not easily computable.

For the general case where topology changes are also included, Blondel *et al* [7] showed that the joint spectral radius of a set of matrices derived from  $F$  matrices determines the convergence speed. For  $\Sigma$  a finite set of  $n \times n$  matrices, their joint spectral radius is defined as:

$$\rho = \limsup_{t \rightarrow \infty} \max_{A_1, \dots, A_t \in \Sigma} \|A_t \dots A_1\|^{1/t}$$

Calculation of the joint spectral radius of a set of matrices is mathematically hard. It is worthwhile to notice that graphs with well-connected nodes guarantee fast convergence. This is a direct result of the Cheeger inequality which relates the spectral gap of an  $F$  matrix to the *conductance* of the corresponding graph [10]. Switching over such topologies will also result in good convergence speed.

Since agents usually have energy constraints, the number of agents with which they communicate is limited. An important design issue is to find topologies which satisfy certain performance provided that the number of the links each agent can establish is less than an upper bound. In the next section we study the small world topologies and show their advantages from a perspective of favorable trade off between communication cost and speed of convergence [4].

### III. CONSENSUS PROBLEMS ON SMALL WORLD GRAPHS

Watts and Strogatz [26] introduced and studied a simple tunable model that can explain behavior of many real world complex networks. Their “Small World” model takes a regular lattice and replaces the original edges by random ones with some probability  $0 \leq \phi \leq 1$ . It is conjectured

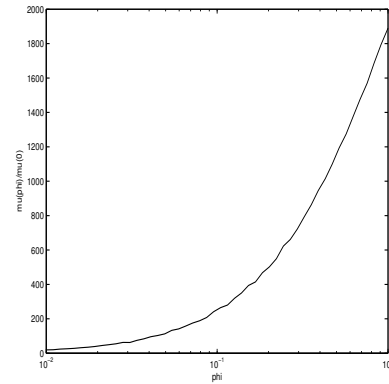


Fig. 3. Spectral gap gain for  $(n, k) = (1000, 5)$

that dynamical systems coupled in this way would display enhanced signal propagation and global coordination, compared to regular lattices of the same size. The intuition is that the shortcuts between distant parts of the network cause high speed spreading of information which may result in fast global coordination. Olfati-Saber [21] studied consensus protocols on small world networks and proposed some conjectures. We have performed similar studies in two areas [3], [5]. We have used a variant of the  $\phi$ -model originally proposed by Watts and Strogatz [26], which starts with a ring of  $n$  nodes, each connected by undirected nodes to its nearest neighbors to a range  $k$ . Shortcut links are added -rather than rewired- between randomly selected pairs of nodes, with probability  $\phi$  per link on the underlying lattice; thus there are typically  $nk\phi$  shortcuts. In reference [5] we actually forced the number of shortcuts to be equal to  $nk\phi$  (comparable to the Watts  $\phi$ -model) and used Vicsek’s weights. We studied different choices of initial ring lattices  $C(n, k)$ . Figure 3 shows the effect of added shortcuts to a base ring  $C(1000, 5)$ . Our Simulation results in [5] show that adding a small number of links to a ring-structured graph should result in high convergence rate. However analytical verification of this result is difficult for general graphs. Here we justify our results using a “mean field” approach and perturbation analysis. In the present analysis, following [17], we reflect the effect of shortcuts by adding “small” nonzero positive numbers to the entries of  $F$  corresponding to non-adjacent nodes of the lattice. We show the results for the cases of one and two dimensional lattices. Generalization of the results to higher dimensions is achievable at the cost of increasing notation. The small perturbation corresponds to using lots of shortcuts with negligible weights on them. Although by adding a uniform perturbation the topology of the graph is not respected, the analysis gives insight on random communication patterns for Small World networks. We state the result for the case where the base lattice is a ring but the result can be extended to  $C(n, k)$  for other  $ks$ . We show the results for Vicsek’s weights.

We follow the perturbation approach to Small World networks proposed by Higham [17]. We start with the one dimensional case. Consider the base lattice to have a ring

topology on  $n$  nodes,  $G(n, 2)$  and the corresponding  $F$  matrix  $F_0$ . This can be also viewed as a random walk with self loops. This is similar to a particular case of our base circulant matrix  $F_0$ . Therefore the base matrix is:

$$F_0(i, j) = \begin{cases} \frac{1}{3} & |i - j| = 0, 1, n - 1 \\ 0 & \text{otherwise.} \end{cases}$$

We know that:

*Lemma 3.1:*  $SLEM(F_0) = \frac{1}{3}[1 + 2\cos(\frac{2\pi}{n})]$ . Furthermore it has multiplicity at least 2.

Now we perturb the nonzero entries of the matrix  $F_0$  by  $\epsilon = \frac{K}{n^\alpha}$  for fixed  $K > 0$  and  $\alpha > 1$  in the limit  $N \rightarrow \infty$ , to get the perturbed matrix  $F_\epsilon$  which is given by:

$$F_\epsilon(i, j) = \begin{cases} \frac{1}{3} - \frac{(n-3)\epsilon}{3} & |i - j| = 0, 1, n - 1 \\ \epsilon & \text{otherwise.} \end{cases}$$

We call the ‘‘shortcuts’’ created this way  $\epsilon$ -shortcuts.  $F_\epsilon$  is also a circulant matrix. It can be easily shown that the SLEM is equal to:

$$\lambda_2(F_\epsilon) = \left(\frac{1}{3} - \frac{n}{3}\epsilon\right)(1 + 2\cos\frac{2\pi}{n}) \quad (3)$$

Thus we can state the following proposition:

*Proposition 3.2:* Let  $\epsilon = \frac{K}{n^\alpha}$ ,  $\alpha \geq 1$ .

- For  $\alpha > 3$ , the effect of  $\epsilon$ -shortcuts on convergence rate is negligible.  $\alpha = 3$  is the onset of the effectiveness.
- For  $\alpha = 2$ , the shortcuts dominantly decrease SLEM.
- For  $\alpha = 1$ , almost all of the nodes communicate effectively and thus the SLEM is very small.

*Proof:* (Sketch)

For large  $n$  we can write:

$$\lambda_2(F_\epsilon) = 1 - \frac{4\pi^2}{3n^2} + o\left(\frac{1}{n^3}\right) - n\epsilon + \frac{4\pi^2\epsilon}{3n} + o\left(\frac{1}{n^3}\right)$$

The first three terms are the contributions of the base lattice and the rest are the contributions of the perturbation. Comparing this to the SLEM of the base lattice yields the desired results. ■

For the base lattice, the spectral gap decreases as fast as  $n^2$ . If  $\epsilon$  is  $O(n^\alpha)$ ,  $\alpha > 3$ , then terms coming from the lattice are dominant, and therefore the shortcuts do not affect the spectral gap. For  $\alpha = 3$  the terms regarding the shortcuts will be of the same degree as the terms from the base and for  $k$  large enough, the SLEM starts decreasing from the corresponding lattice SLEM. For  $\alpha = 2$  the terms regarding the shortcuts are dominant and the SLEM has considerably decreased compared to the base lattice. Only for the case of  $\alpha = 1$  the spectral gap does not vanish as  $n \rightarrow \infty$ .

Similar statements can be established for the case of a two-dimensional grid. We consider a two-dimensional  $m \times m$  grid with periodic boundary, in which each node is connected to 4 nodes. Let  $n = m^2$ . Using Vicsek’s weights, the corresponding  $F_0$  is a block circulant matrix, in which each of its blocks is circulant. The matrix  $F_0$  can be written as:

$$F_0 = \text{BlockCirc}[C_1, C_2, \dots, C_m],$$

where  $C_1 = \text{Circ}[\frac{1}{5}, \frac{1}{5}, 0, \dots, 0, \frac{1}{5}]$ ,  $C_2 = C_m = \frac{1}{5}I_m$ , and  $C_3 = C_4 \dots = C_{m-1} = \mathbf{0}$ . The matrices  $C_i$  are either in diagonal form or can be diagonalized. Let  $\Lambda_1$  denote the diagonalization of  $C_1$ , and  $\Lambda_2 = \Lambda_n = \frac{1}{5}I_m$ .

The SLEM of  $F_0$  then can be found using Lemma 3.3 :

*Lemma 3.3:*  $SLEM(F_0) = \frac{1}{5}[3 + 2\cos(\frac{2\pi}{\sqrt{n}})]$ .

*Proof:* (Sketch) The matrix  $F_0$  is  $m^2 \times m^2$  block circulant. Its blocks are  $m$  dimensional circulants, so it can be diagonalized using the Kronecker product of  $m$  dimensional Fourier matrices ([11], Theorem 5.8). The diagonal matrix  $\Lambda$  is given by:

$$\Lambda = I_m \otimes \Lambda_1 + \Omega_m \otimes \Lambda_2 + \Omega_m^{m-1} \otimes \Lambda_m,$$

where  $\Omega_m = \text{diag}(1, \omega, \dots, \omega^{m-1})$  and  $\omega = \exp(2\pi i/m)$ . Comparison of the eigenvalues of the  $m$  diagonal blocks of  $\Lambda$  yields the result. ■

Now, consider the perturbed matrix  $F_\epsilon$  in which the zero elements of  $F_0$  are perturbed by  $\epsilon$ , and the previously nonzero elements are also perturbed to retain the stochastic form of the matrix. It can be verified that,  $F_\epsilon$  is again a block circulant matrix with circulant blocks  $C_1, C_2, \dots, C_m$ , where  $C_1 = \text{Circ}[\frac{1}{5} - (\frac{n-5}{5}\epsilon), \frac{1}{5} - (\frac{n-5}{5}\epsilon), \epsilon, \dots, \epsilon, \frac{1}{5} - (\frac{n-5}{5}\epsilon)]$ ,  $C_2 = C_m = (\frac{1}{5} - \frac{n}{5}\epsilon)I_m + \epsilon \mathbf{1}_m \mathbf{1}_m^T$ , and  $C_3 = C_4 \dots = C_{m-1} = \epsilon \mathbf{1}_m \mathbf{1}_m^T$ .

Therefore  $F_\epsilon$  can be diagonalized to  $\Lambda$ :

$$\Lambda = I_m \otimes \Lambda_1 + (\Omega + \Omega^{m-1}) \otimes \Lambda_2 + (\Omega^2 + \Omega^3 + \dots + \Omega^{m-2}) \Lambda_3$$

where again  $\Lambda_1, \Lambda_2$ , and  $\Lambda_3$  are diagonalizations of  $C_1, C_2$ , and  $C_3$ . Using the same arguments as in Lemma 3.3, the SLEM of  $F_\epsilon$  is calculated:

$$SLEM(F_\epsilon) = \left[\frac{3}{5} + \frac{2}{5} \cos \frac{2\pi}{\sqrt{n}}\right] - n\epsilon \left[\frac{3}{5} + \frac{2}{5} \cos \frac{2\pi}{\sqrt{n}}\right].$$

For large values of  $n$ , the spectral gap for the unperturbed lattice is approximately  $\frac{4}{5} \frac{\pi^2}{n}$ , whereas the spectral gap of the perturbed matrix is approximately  $\frac{4}{5} \frac{\pi^2}{n} + n\epsilon(1 + \frac{2\pi^2}{n})$ . Thus, we get the results as in the 1-D case.

As observed above  $\epsilon$ -shortcuts are loosely analogous to the shortcuts in the  $\phi$ -model. Since the Small World model is a probabilistic model, we anticipate that adding small weights is analogous to choosing graphs with low probability shortcuts. This idea has been further developed in [6].

#### IV. A HIERARCHICAL SELF ORGANIZING METHOD

In this section we show that a very efficient communication pattern with substantial improvement in performance is possible by a two level hierarchical scheme. The idea here is that selecting a few well connected and controlled agents which are well protected should enhance the speed of convergence of distributed schemes like consensus algorithms.

Given  $N$  agents, suppose we can divide them into  $K$  groups each having  $M$  members, so  $N = K.M$ . It is assumed that  $K \leq M \ll n$ . The exact sizes of  $K$  and  $M$  are problem specific and influence the performance of the algorithm. For each group suppose that we can select a ‘‘leader’’. The leaders should be able to have two properties: they should be well connected to the members of their group,

and they should also be able to communicate with other leaders when necessary. If the distributed algorithm is carried out at each group separately and the leaders communicate on a higher level, the agents can enjoy faster convergence rates; the reduction of the size of each group from  $N$  to  $M \leq \sqrt{N}$  results in faster intergroup convergence whereas the ease of communication between the leaders upon demand results in overall fast convergence. This can be illustrated for example, by considering the consensus problem of Figure 2 with the agents known as “no-voters” selected as leaders.

We now provide a semi-distributed method which can categorize the agents as “leader” or “regular”. Further, the method assigns each regular agent with an influence vector which indicates which leader has more influence on it. This provides the nodes with some global picture of the network.

#### A. Distributed exploration of the graph structure

The structure of a graph plays a crucial role in properties of a distributed algorithm that is running on it. Given a graph, individual nodes have only local knowledge about its structure, which includes information about their neighboring nodes. If any node wants to either improve its own performance or a global performance measure it needs to know more about the global picture of the network. This information can be used by the node to refine its choice of neighbors in order to improve its performance.

The most complete measure of global graph structure is the adjacency matrix. Since each node has limited memory, energy, and computational capacity, they cannot store and process the adjacency matrix. Our goal is to devise a scheme to provide each node with a small vector that includes compact global information on how the node is located with respect to the other nodes. It is desired that the scheme can be disseminated via an implementable distributed manner.

We propose a two stage algorithm for this purpose. Apart from a single data transmission and reception at a central authority by each node, the algorithm is carried out in a decentralized manner. In the first stage nodes will collaborate to find their *social degree* [8]. This is a local measure of how ‘well connected’ each node is. Once the nodes find out their social degree, they will transmit it to a central authority which determines  $K$  “social leaders” of the graph—the better connected nodes among all. The central authority then broadcasts the list of  $K$  social leaders to all of the nodes. In the second stage, each node uses a simple iterative scheme to maintain its *influence vector*, a vector of size  $K$  which determines the influence of each social leader on it. We acknowledge that while developed independently, the idea is similar to recent community evaluation schemes developed in the network science literature [20].

#### B. Social degrees and leader nodes

To find the leaders or the agents with the highest influence we use a generalization of a framework proposed by Blondel *et al* [8]. They define the social degree of a node as the number of the cycles of length 3 passing through that node.

They also define a social leader as a node with the highest degree in its neighborhood. This can be generalized as:

*Definition 4.1:* *Social degree of order  $k$*  of a node (denoted by  $SD^{(k)}(v_i)$ ) is defined to be the number of cycles of length  $k$  passing through the node, if  $k > 2$  and the number of its neighbors, if  $k = 2$ .

*Definition 4.2:* A *leader node of order  $k$*  is the node with the highest social degree of order  $k$  among its neighbors.

*Definition 4.3:* For given  $0 < \alpha < 1$  and  $\beta = 1 - \alpha$ , a node’s *social score* is defined as  $SC(v_i) = \alpha \cdot SD^{(2)}(V_i) + \beta \cdot SD^{(3)}(v_i)$ .

Notice that each node can determine its social degree of orders 2 and 3 by a simple query from its neighbors. Since determining higher order degree requires more effort, we use the orders 2 and 3 for our present application.

In the first stage of the algorithm, each node computes its social degrees of order 2 and 3. It also queries the social degrees of its neighbors. Upon comparing its social degrees with its neighbors, if a node is found to be a leader of order 2 or 3, it transmits its degrees to the central authority. Upon receiving these data from the leader nodes, the central authority selects  $K$  nodes  $l_i, i = 1, \dots, k$  with the highest social scores,  $SC(l_i)$ , gives an arbitrary order to them and transmits their assigned order to them. Notice that the choice of  $\alpha$  and  $\beta$  determines the preference between leaders in a “star-like” neighborhood versus leaders of better connected neighborhoods. Once a selected leader is assigned its order  $1 \leq i \leq K$  it will maintain the constant vector  $e_i \in R^K$ . This is the unit vector with 1 in its  $i$ th entry.

#### C. Determination of the influence vector

Our objective in this part is to associate with each of the regular nodes a vector that determines how well it is related to each of the leaders and how it is influenced by them. The amount of influence that a leader has on a local node is not only determined by their distance but also by the number of paths between them. We provide a definition for the influence vector based on the properties of random walks on graphs.

*Definition 4.4:* Consider a graph with  $K$  leaders and  $N - K$  regular nodes. The influence of leader nodes  $l_k$  ( $k = 1, \dots, K$ ) on any regular node  $i$  is the probability that a random walk that starts from  $i$  hits  $l_k$  before it hits any other leader node.

Given the leaders and the arbitrary order assigned to them, we first describe the algorithm to determine the influence vectors for each regular node. Then we will show why it converges and why it outputs valid vectors as influence vectors. We denote the influence vector of node  $i$  by  $x_i \in R^K$ . By  $x_i^k(t)$  we mean the  $k$ th entry of the influence vector of node  $i$  evaluated at time  $t$ .

The algorithm operates as follows. The influence vector of leader  $l_i$  is first assigned to be the unit vector  $x_i = e_i$ . These  $K$  vectors do not vary. For all regular nodes  $i$ ,  $x_i$  is initialized randomly, distributed uniformly on  $[0, 1]^K$ . At each iteration time  $t + 1$ , each regular node  $i$  updates its influence vector

entry-wise ( $k = 1, 2, \dots, K$ ) using the following rule

$$x_i^k(t+1) = \frac{1}{1+n_i(t)} [x_i^k(t) + \sum_{j \in N_i(t)} x_j^k(t)] \quad (4)$$

Theorem 4.1 shows the effectiveness of the above scheme.

*Theorem 4.1:* If the underlying graph is connected, the iteration (4) converges to a set of unique vectors. Furthermore,  $\lim_{t \rightarrow \infty} x_i^k(t)$  is equal to the probability that a random walk starting at node  $i$  hits the leader node  $l_k$  before any other leader node.

*Proof:* The proof of convergence is standard. The particular form of the solution arises because the procedure solves a discrete version of the Dirichlet problem on the graph. We follow the proof of Bremaud [10]. Relabel the nodes, such that  $D = \{1, 2, \dots, N - K\}$  denote the regular nodes and  $\partial D = \{N - K + 1, \dots, N\}$  denote the leader nodes, where  $l_i = N - K + i$ . For all  $k = l_1, \dots, l_K$ , define a function  $\phi^i$  on the graph such that  $\phi^k(l_j) = \delta(k, l_j)$ , where  $\delta$  is the Dirac Delta function.

Let  $P = (I + D)^{-1}(A + D)$ , where  $A$  is the adjacency matrix of the graph  $G$  and  $D$  is the diagonal matrix with  $i^{\text{th}}$  diagonal element equal to the degree of node  $i$ . Fix  $k$ , i.e. consider the influence of leader  $l_k$ . Then  $x_i^k$  converges to a value that satisfies the following equation

$$x_i^k = \begin{cases} (Px^k)_i & i = 1, 2, \dots, N - K \\ \phi^k(i) & i = N - K + 1, \dots, N \end{cases} \quad (5)$$

Note that  $P$  is a stochastic matrix and the first equation is valid for the regular nodes. Let  $\{Z_n^k\}_{n \geq 0}$  be a homogeneous Markov chain with state space  $V = \{1, 2, \dots, n\}$ . Let  $T$  be the hitting time of  $\partial D$ . For each state  $i \in V$  define:

$$h_i^k = E[\phi(Z_T^k) | Z_0^k = i] \quad (6)$$

Since the underlying graph  $G$  is connected,  $P$  is irreducible. Also  $\forall i \in V$ ,  $p_{ii} > 0$ , which means the chain is aperiodic. The number of states is finite and therefore the chain is positive recurrent and  $P(T < \infty | Z_0^k = i) = 1$ .

By definition  $h^k = \phi^k$  on  $\partial D$  and  $x^k = \phi^k$  on  $D$ . By first step analysis:  $h_i^k = \sum_{j \in V} p_{ij} h_j^k$  on  $D$ . So:

$$h^k = \begin{cases} Px^k & \text{on } D \\ \phi^k & \text{on } \partial D \end{cases} \quad (7)$$

Therefore  $h = x$  on the graph  $G$ . The proof of uniqueness of the solutions also follows from [10]. Notice that  $\phi_i^k$  is defined such that  $h_i^k$  is equal to the probability of hitting the leader node  $l_k$  before the other leader nodes. ■

## V. CONCLUSION

We showed in different contexts that the performance of distributed collaborative schemes can depend dramatically on the structure of the underlying topology. We showed that small world topologies offer several advantages from a perspective of a favorable tradeoff between performance of collaborative behaviors vs. costs of collaborative behaviors. A two level hierarchy of self-organization was shown to provide a very efficient communication pattern with substantial improvement of performance. Finally, a semi-distributed

method of finding proper ‘‘leader’’ nodes and measuring the influence of them on the regular ones in large networks was proposed and analyzed.

## REFERENCES

- [1] A. L. Barabasi, and R. Albert. Emergence of scaling in random networks. *Science*, 286:509-512, 1999.
- [2] J.S. Baras and T. Jiang. Cooperative games, phase transitions on graphs and distributed trust in MANET. *Proc. of the 43rd IEEE Conference on Decision and Control*, 1:93-98, 2004.
- [3] T. Jiang and J. S. Baras. Autonomous trust establishment. *Proc. of the 2005 International Network Optimization Conference*, Lisbon, Portugal, March 20-23, 2005.
- [4] T. Jiang and J. S. Baras. Fundamental tradeoffs and constrained coalitional games in autonomic wireless networks. *Proc. of 5th International Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Limassol, Cyprus, April 16-20, 2007.
- [5] P.Hovareshti and J.S. Baras. Consensus problems on small world networks: A structural study. *Proc. of the 2006 International Conference on Complex Systems*, Boston, Mass., May 28-31, 2006.
- [6] P. Hovareshti, J.S. Baras and V. Gupta. Average consensus over small world networks: A probabilistic framework. *Submitted to 47th IEEE Conference on Decision and Control*, 2008.
- [7] V. Blondel, J. Hendrickx, A.Olshevsky, and J. Tsitsiklis. Convergence in multiagent coordination, consensus and flocking. *Proc. of the 44th IEEE Conference on Decision and Control*, pages 2996-3000, 2005.
- [8] V. Blondel and C. de Kerchove and E. Huens and P. Van Dooren. Social leaders in graphs. *Submitted to POSTA06*, 2006.
- [9] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Trans. on Inf. Theory*, 52(6):2508-2530, 2005.
- [10] P. Bremaud. *Markov chains, Gibbs Fields, monte Carlo Simulations and Queues*. Springer Verlag, 2nd edition, 1999.
- [11] P. J. Davis. *Circulant Matrices*. Wiley, 1979.
- [12] R. Diestel. *Graph Theory*. Springer-Verlag, 2000.
- [13] R. Durrett. *Random Graph Dynamics*. Cambridge Univ. Press, 2007.
- [14] V. Gupta, C. Langbort, and R. M. Murray. On the Robustness of Distributed Algorithms. *Proceedings of the 45th IEEE Conference on Decision and Control*, pages:3473-3478, 2006.
- [15] L. Fang and P. Antsaklis. On communication requirements for multi-agents consensus seeking. *Lecture Notes in Control and Information Sciences (LNCIS)*, Springer, 331:53-68, 2006.
- [16] D. Peleg. Local majority voting, small coalitions and controlling monopolies in graphs: A review. *Proceedings 3rd Colloquium on Structural information and Communication Complexity*, 1996.
- [17] D.J. Higham. A matrix perturbation view of the small world phenomenon. *Siam Journal on Matrix Analysis and Applications*, 25(2):429-444, 2003.
- [18] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988-1001, 2003.
- [19] A. Jadbabaie, N. Motee, M. Barahona. On the stability of the Kuramoto model of coupled nonlinear oscillators. *Proceedings of the American Control Conference*, 5:4296- 4301, 2004.
- [20] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:02611, 2004.
- [21] R. Olfati Saber. Ultrafast consensus in small-world networks. *Proceedings of the American Control Conference*, 4:2371-2378, 2005.
- [22] R. Olfati Saber and R.M Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520-1533, 2004.
- [23] R. Olfati Saber. Flocking for Multi-agent Dynamical Systems: Algorithms and Theory. *IEEE Transactions on Automatic Control*, 51(3):401-4203, 2006.
- [24] A. Tahbaz-Salehi and A. Jadbabaie. Small world phenomenon, rapidly changing markov chains, and average consensus algorithm. *Proc. of 46th IEEE Conf. on Decision and Control*, Pages:276-281, 2007.
- [25] T. Vicsek, A. Czirok, E. Ben Jakob, I. cohen, and O.Schochet. Novel type of phase transitions in a system of self-driven particles. *Phys. Rev. Lett.*, 75:1226-1299, 1995.
- [26] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440-442, 1998.
- [27] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. *Proc. of International Conference on Information Processing in Sensor Networks*, pages 63-70, 2005.