A local average broadcast gossip algorithm for fast global consensus over graphs

Gang Wang*, Zhiyue Wang* and Jie Wu†

*School of Electrical and Information Engineering, Beihang University

Beijing, China, 100191

†Department of Computer and Information Sciences, Temple University

Philadelphia, PA, USA

Email: gwang@buaa.edu.cn, wangzy yaya@sina.com, jiewu@temple.edu

Abstract

Motivated by applications to wireless sensor, peer-to-peer, and social networks, the canonical average consensus problem is considered in random and regular graphs in this paper. A local average information exchange (LAIE) algorithm is developed to compute the global consensus of the initial measurements of the nodes at every node in the network. In the proposed algorithm, each node interacts with all of its neighboring nodes in each round of the diffusion process to compute and exchange the local average value, such that all nodes can asymptotically reach a global consensus in a distributed manner very quickly. This is in contrast to the conventional random gossip scheme, where each node only interacts with one of its neighboring nodes, leading to very long convergence time. Results show that in a random graph with n nodes, the convergence time of the LAIE algorithm is bounded below by $\mathit{\Omega}\left(\frac{(n-1)logn}{\Delta}\right)^1$, where the parameter Δ denotes the largest degree of the graphs. When a network has n nodes represented by d-regular topology graphs ($d > 2$), where each node has the same number of neighbors d, the convergence time of the LAIE algorithm is bounded below by $\mathit{\Theta}\left(\frac{n(d+1)logn}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right)$. This shows that the proposed algorithms can achieve quicker convergence to the global consensus than other schemes based on the classic random gossip algorithm. Finally, we assess and compare the communication cost of the local average algorithm to achieve consensus through numerical results.

Given two functions $f(n) > 0$, and $g(n) > 0$: $f(n) = o(g(n))$ means $\lim_{n\to\infty} f(n)/g(n) = 0$; $f(n) = O(g(n))$ means $\lim_{n\to\infty} sup(f(n)/g(n)) < \infty$; $f(n) = \omega(g(n))$ is equivalent to $g(n) = o(f(n))$; $f(n) = \Omega(g(n))$ is equivalent to $g(n) = O(f(n))$; $f(n) = \Theta(g(n))$ means $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

Index Terms

Broadcasting, Consensus, Local Average, d-regular Graph, Random Graph.

I. Introduction

Consider that a distributed system or network consists of a set of nodes where each node has an initial value that can share information with neighboring components via the connected edges, thus forming a general interconnection graph. The objective of a consensus problem is to have all nodes agree upon a certain quantity of interest, which is typically a function of some values that the nodes initially possess. When the nodes asymptotically reach an agreement on the same value, we say that the distributed system asymptotically reaches consensus. A special case of consensus is the case of average consensus, where the additional challenge is for the nodes to converge to the exact average of their initial values [1]. Consensus problems have attracted a considerable amount of interest in various fields including distributed detection [2], communication [3], control theory [4], distributed data fusion in sensor network [5], and biology [6]. The survey [7] provides the basic concepts and methods for consensus algorithms.

The essence of distributed consensus algorithms or gossip algorithms is that, in each round, one or more nodes can communicate with its immediate neighbors, following which each node updates its estimate with a quantity of interest, sometimes called its state, by combining the estimate with those of its neighbors. In this manner, every node bootstraps another until all of them agree on a common value. A common feature of a consensus problem is that agents can exchange information only locally, and there is no fusion center in the network. Distributed average consensus algorithms, which involve computations based only on local information, are advantageous because they obviate the need for global communication and complicated routing, and are robust against node and link failures.

In this paper, we propose a novel and simple distributed local average and information exchange algorithm. The proposed algorithm uses broadcast operation to exchange local information involving all neighboring nodes of each node at each computation to quickly obtain local agreement and achieve global consensus through an iterative process.

A. Related Work

A pioneering study on asynchronous distributed computation is [8], where the convergence of a gossip-based asynchronous distributed algorithm was investigated. Distributed consensus algorithms can be categorized in a few ways, one of which is on the basis of their deterministic or randomized operating protocol [9]. For deterministic consensus algorithms, the evolution matrix is largely fixed. Thus, in each round, almost every node exchanges information with its neighbors and updates its own state value. Yuan *et al.* propose distributed average consensus via gossip algorithm with real-valued and quantized data, and also prove the bounds with respect to the mixing parameter on the convergence rate [10]. Algorithms of this type include those described in [7], [11]. By contrast, the evolution matrix of randomized consensus algorithms changes randomly at every step of the iteration. Examples of this type of algorithm can be found in [9] and [12], [17].

Boyd *et al.* first proved that the bound of convergence time for a randomized gossip algorithm is $\Theta(\frac{\log n}{1-\lambda_2[EW]})$ in [12]. Convergence time is closely related to the second-largest eigenvalue $\lambda_2[EW]$ of the average weight matrix EW, depending on the design of the algorithm. It was shown that this algorithm converges to a consensus if the graph is strongly connected on average. Because the transmitting node must send a packet to the chosen neighbor and then wait for the neighbor's packet, this scheme is vulnerable to packet collisions, and yields a communication complexity of the order of $\Theta\left(n^2\right)$ over random geometric graphs, where n is the number of nodes. Moreover, the randomized gossip algorithm confines information iteratively among neighbors, and the choices of neighbors are random and repeated, which can lead to a considerable amount of wasted energy.

Dimakis *et al.* proposed the geographic gossip algorithm, which combined gossip with geographic routing, and showed that convergence time is $O(n^{1.5}\sqrt{\log n})$ [13] . By using geographic information, such as coordinates, the geographic gossip algorithm enabled any node to communicate with nodes far from it in the network, made computationally concentrated areas sparse, and hence remarkably enhanced convergence speed. However, the geographic gossip algorithm cannot be adapt to dynamic circumstances due to routing maintenance, and forwarding back the data renders packet loss inevitable, which increases complexity and reduces reliability.

On the contrary, enabling more computation is a useful approach to improving efficiency. Li *et al.* proposed a cluster-based gossip algorithm that reduced convergence time by a factor of $O(\log n)$ compared to the randomized gossip algorithm, but required considering the difficulty of cluster maintenance and the complexity of the theoretical mechanism [14].

Keren *et al.* studied the synchronized information dissemination problem and provided a relationship between the gossip model and the local model. However, in many realistic settings, synchronous algorithms require that the nodes agree on time, since the algorithm starts running and adds physical constraints for practical system design [15]. Furthermore, a deterministic gossip algorithm was proposed in [16], which only solved the k-local broadcast information-spreading problem using the local model in a synchronized manner.

Broadcasting is a simple and efficient network communication protocol to disseminate information to all nodes within a neighborhood, and can effectively adapt to unknown and dynamic network environments. Broadcast gossip algorithms are especially attractive for use in wireless networks [17]. In broadcast gossip algorithms, however, nodes asynchronously broadcast a message, and the contents of the message are immediately processed by all neighbors receiving it. Two broadcast gossip averaging algorithms are provided for distributed computation of averages in large Abelian Cayley networks in [18]. The results show that the robustness of broadcast gossip algorithms to interferences. The tight bounds on the largest perturbation parameter for which the system is provided, which is still guaranteed to converge to a consensus, and derived the value of the perturbation parameter that led to the fastest asymptotic rate of convergence on strongly connected digraphs [19]. The convergence of pairwise and broadcast gossip algorithms for consensus with intermittent links and mobile nodes are analyzed in wireless sensor networks in [20]. Reference [21] studies the problem of distributed parameter estimation in unreliable sensor networks by broadcast gossip algorithms.

B. Summary of Main Contributions

The main contributions of this paper are summarized as follows:

I. We propose a novel distributed local average and information exchange (LAIE) algorithm to accelerate the convergence of the global consensus. Each node still acts based only on local information, without a centralized controller. The proposed algorithm has very good scalability, and adapts well to a dynamic network environment, where the addition or deletion of nodes and links has no influence on performance.

II. The average weight matrix EW, according to the design of the algorithm, plays an important role in analyzing the convergence of the global consensus [12]. We first analyze an expectation matrix under random graphs, and then derive the convergence time $\Omega\left(\frac{(n-1)\log n}{\Delta}\right)$ and expected communication cost $\Omega\left(\frac{(\delta+2)(n-1)\log n}{\Delta}\right)$ of the LAIE algorithm.

III. Motivated by the properties of the algebraic connectivity of the Laplacian matrix, we use the mathematical tool of spectra of the Laplacian matrix to prove the convergence of the proposed algorithm, and provide a new bound of the convergence time under d-regular topology. A lower bound of the convergence time of the LAIE algorithm $\Theta\left(\frac{n(d+1)logn}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right)$ is developed, which is tighter than the bound derived by the randomized gossip algorithm $\Theta\left(n^2\log n\right)$ and the geographic gossip algorithms $\Theta\left(n\log n\right)$.

IV. The simulation results for different degrees d of the regular graph and the random graph are provided in this paper, and all results verify the efficiency of the LAIE algorithm. For different values of d, the proposed algorithm always outperformed the randomized gossip algorithm and geographic algorithms. For example, the LAIE algorithm was better than the randomized gossip algorithm and the geographic gossip algorithm by a factor of 0.986 n and 0.986 $n^{0.5}$ in a grid topology, respectively. The results show that both convergence time and expected communication cost decrease as the value of d increases, which means that our algorithm adapts well to graphs with large degrees.

C. Paper Organization

The rest of this paper is organized as follows: Section I-A contains a brief review of existing work. Section II id devoted to a description of the system model, the problem formulation, and the performance criterion. The LAIE algorithm and its convergence properties are presented in Section III. Section IV contains a discussion of the simulation results of the proposed algorithm and the convergence time property comparisons for different algorithms. The simulation results show that the proposed algorithm converges much more quickly than the randomized gossip algorithm, the geographic gossip algorithm, and the broadcast gossip algorithm. Section V contains our conclusions.

II. System Model

A. Network Model

A network is represented by a graph $G(V, E)$, where $V = \{1, 2, \ldots, n\}$ is the set of nodes and $E \subseteq V \times V$ is the set of edges. $(i, j) \in E$ if nodes i, j can communicate with each other. Note that we do not allow self-loops in the graph. We assume that communication between any two nodes is perfect. The neighbors of node i are denoted by $N(i) := \{j \in V : (i, j) \in E)\}$. The cardinality $N(i)$, denoted by $d_i = |N(i)|$, is called the degree of node i. Let Δ and δ denote the biggest and smallest degrees of graph G, respectively. Let Φ be the $n \times n$ adjacency matrix of G, where for $i \neq j$, $\Phi_{ij} = 1$ if $(i, j) \in E$ and $\Phi_{ij} = 0$ otherwise. The d-regular graph is a graph where each node has the same degree d. Here, we assume $d > 2$. For a d-regular graph G, we define an $n \times n$ diagonal matrix D where each diagonal entry is given by $D_{ii} = d$. The Laplacian matrix of graph G is defined as $L = D - \Phi$, and is positive semi-definite and singular [25]. Let $\lambda_1^L, \lambda_2^L, \ldots, \lambda_n^L$ be the eigenvalues of Laplacian matrix L. Here, without loss of generalization, we suppose $\lambda_1^L \leq \lambda_2^L \leq \cdots \leq \lambda_n^L$. It is well known that the second smallest eigenvalue of the Laplacian matrix is called the algebraic connectivity of a graph. In general, graphs that are more strongly connected have a larger λ_2^L [25].

B. Problem Formulation

We consider an n-dimensional vector $X(0) = [x_1(0), \ldots, x_n(0)]^T$, representing the initial states at n nodes, which are deterministic, at time slot $t = 0$. Let $x_{ave} = \frac{1}{n}\sum_{i=1}^n x_i(0)$, and let $X(t) = [x_1(t), \ldots, x_n(t)]^T$ be the vector of state values at the end of the t-th time slot. Suppose that the clock of node i ticks at the beginning of the $(t+1)$-th time slot. Node i then broadcasts its state value to all its neighboring nodes. The node states will be updated according to the following equations:

$$\begin{aligned} x_j(t+1) &= (x_j(t) + x_i(t))/2 \qquad && j \in N(i) \\ x_j(t+1) &= x_j(t) && j \notin N(i) \end{aligned} \tag{1}$$

In average consensus protocols, nodes communicate with each other in communication steps called rounds, and the amount of information exchanged in each round between two communicating nodes is limited. In this paper, we employ an asynchronous time model to choose the initiator node of information exchange [12], which is a good match with the distributed nature. Time is discretized, and the time interval $[Z_k, Z_{k+1})$ corresponds to the k-th round. In each round, we divide the time into slots (TSs). Every node is scheduled to transmit its message to one of its neighboring nodes in a TS. Our interest is in determining the time it takes for $X(t)$ to converge to the global average, i.e., $x_{ave}\cdot$ **1**, where vector **1** represents the column vector containing only 1s.

C. Performance Criterion

Definition **1** [12]: For any $0 < \varepsilon < 1$, the ε-averaging time $T_{ave}(n, \varepsilon, \eta)$ of an algorithm is defined as:

$$T_{ave}(n, \varepsilon) = \sup_{X(0)} \inf \left\{ t : Pr\left(\frac{||X(t) - x_{ave} \cdot \mathbf{1}||}{||X(0)||} \geq \varepsilon \right) \leq \varepsilon \right\} \tag{2}$$

where $||v||$ denotes the l_2 norm of the vector v. Thus $T_{ave}(n, \varepsilon)$ is the smallest time it takes for $X(t)$ to get within ε of $x_{ave} \cdot \mathbf{1}$ with high probability, regardless of the initial value $X(0)$.

Definition **2**: For any $0 < \varepsilon < 1$, the total communication cost, denoted by $C(n, \varepsilon)$, is defined as:

$$C(n, \varepsilon) = \sum_{t=1}^{T_{ave}(n,\varepsilon)} R(t) \tag{3}$$

where $R(t)$ denotes the transmission time required for a given node to communicate with other nodes at time slot t. Then, the expected communication cost is defined as

$$EC(n, \varepsilon) = E(R(t)) \cdot T_{ave}(n, \varepsilon, \eta) \tag{4}$$

where $E(R(t))$ denotes the expectation of random variable $R(t)$.

III. Main Work

A. Proposed Algorithm

In this section, we first propose a Local Average and Information Exchange algorithm for asymptotically achieving the global average consensus, which is shown in Table I. Suppose that, at time slot t, node i as an initiator broadcasts its state to all neighboring nodes. All its neighboring nodes compute and update their states according to Equation (1), and send the updated states to node i. Node i receives all states, computes its new average state, and broadcasts it to all its neighboring nodes once again. All neighboring nodes compute and update their states.

Node i and its neighbors then agree on the local consensus. This round of information exchange concludes. Obviously, the LAIE algorithm utilizes the broadcast feature operation twice, which will be shown to improve the convergence time of the proposed algorithm.

Algorithm 1 LAIE Algorithm.

1: **Initialization:** Let $G = (V, E)$, n, $X(0) = [x_1(0), \dots, x_n(0)]$
2: at time slot t, node i is chosen as the initiator node, its state is $x_i(t)$
3: node i broadcasts its state $x_i(t)$ to all neighboring nodes, which are denoted by the set $N(i)$
4: **if** $\forall j \in N(i)$ **then**
5: node j receives $x_i(t)$, and computes and updates its state $x_j(t+1) = \frac{x_j(t)+x_i(t)}{2}$
6: all nodes in set $N(i)$ send their updated states back to node i one by one
7: once node i receives all neighboring nodes' updated states, it computes the average of received states $x_i(t+1) = \frac{x_i(t)+\sum_{j\in N(i)} x_j(t)}{|N(i)|+1}$, and updates its state
8: node i sends its newest state $x_i(t+1)$ to its neighboring nodes again
9: all neighboring nodes $j \in N(i)$ update their states according to Step 5
10: **else**
11: $\forall k \notin N(i)$, only updates their states as follows: $x_j(t+1) = x_j(t)$
12: **end if**
13: repeat this process from Step 2

B. Analysis of Convergence Time Under Random Graph

In this paper, we consider two kinds of graphs: regular and random. Regular graphs have good topological properties that can provide fundamental insights into the proposed algorithm. Random graphs, where nodes are randomly deployed and have varying degrees, are very useful. For clear analysis, at each time slot t, the LAIE algorithm can also be formulated by a matrix $W(t)$ given by:

$$X(t) = W(t)X(t-1) \tag{5}$$

Matrix $W(t)$ represents a different algorithm that must satisfy the following necessary and sufficient conditions to correctly ensure asymptotic average consensus [12]:

$$W(t)\mathbf{1} = \mathbf{1},\ \ \mathbf{1}^T W(t) = \mathbf{1}^T,\ \ \rho(W(t) - J) < 1. \tag{6}$$

where $J = \frac{\mathbf{1}\cdot\mathbf{1}^T}{n}$ and $\rho(\cdot)$ denotes the spectral radius of the matrix, which is given by:

$$\rho(M) = max\{\left|\lambda_i^M\right| : i = 1, 2, \ldots, n\} \tag{7}$$

where λ_i^M denotes the eigenvalues of matrix M. We then analyze the properties of transfer matrix $W(t)$ of the LAIE algorithm.

Lemma **1**: Given graph $G = (V, E)$, $\forall i \in V$, at time slot $t \geq 1$, the degree of node i is d_i, $i = 1, 2, \cdots, n$. All weight matrices of the LAIE algorithm are given by $\{W^{(i)}(t) : \forall i \in V\}$. Then, vector $\mathbf{1}$ is both left and right eigenvector, i.e.,

$$W^{(i)}(t)\mathbf{1} = \mathbf{1}, \quad \mathbf{1}^T W^{(i)}(t) = \mathbf{1}^T \tag{8}$$

where the transfer matrix $W^{(i)}(t)$ is given by:

$$W_{jk}^{(i)}(t) = \begin{cases} 1, & j \notin L(i), k = j \\ \frac{1}{d_i+1} & j, k \in N(i) \\ \frac{1}{d_i+1} & k = j = i \\ 0 & elsewhere \end{cases} \tag{9}$$

where $W_{jk}^{(i)}(t)$ is the (j, k)-th entry of transfer matrix $W^{(i)}(t)$.

Proof: Based on the LAIE algorithm, the transfer matrix $W^{(i)}(t)$ is obvious. We then prove that vector $\mathbf{1}$ is the left eigenvector. According to (9), it is easy to prove that all rows of all matrices $W^{(i)}(t), i = 1, 2, \cdots, n$ satisfy $W^{(i)}(t)\mathbf{1} = \left(\sum_{k=1}^n W_{1k}^{(i)}(t), \cdots, W_{jk}^{(i)}(t), \cdots, W_{nk}^{(i)}(t)\right)^T$, where $(\cdot)^T$ denotes the transpose of a matrix. Without loss of generality, for any subscript $j, j = 1, 2, \cdots, n$, we analyze the sum $\sum_{k=1}^n W_{jk}^{(i)}(t)$ as follows:

Case I: when $j \notin L(i)$, based on (9), if $k = j$, $W_{jj}^{(i)}(t) = 1$; otherwise $k \neq j$, $W_{jj}^{(i)}(t) = 0$. Then, $\underline{1}\{j \notin L(i)\} = 1$, and $\underline{1}\{j \in L(i)\}\frac{1}{d_i+1} = 0$:

$$\sum_{k=1}^{n} W_{jk}^{(i)}(t) = \underline{1}\{j \notin L(i)\} + \underline{1}\{j \in L(i)\}\frac{1}{d_i + 1} = 1, \ j = 1, 2, \cdots, n$$

where $\underline{1}\{\cdot\}$ is the indicator function, i.e.,

$$\underline{1}\{j \notin L(i)\} = \begin{cases} 1, & if j \notin L(i) succeed \\ 0, & elsewhere \end{cases}$$

Case II: when $j \in L(i)$, based on (9), $\underline{1}\{j \in L(i)\} \cdot \frac{1}{d_i+1} = (d_i + 1) \cdot \frac{1}{d_i+1} = 1$, and $\underline{1}\{j \notin L(i)\} = 0$, then

$$\sum_{j=1}^{n} W_{jk}^{(i)}(t) = \underline{1}\{j \notin L(i)\} + \underline{1}\{j \in L(i)\}\frac{1}{d_i + 1} = 1, \ j = 1, 2, \cdots, n$$

In summary, $W^{(i)}(t)\mathbf{1} = \mathbf{1}$. The right equation can be proved similarly, i.e., $\mathbf{1}^T W^{(i)}(t) = \mathbf{1}^T$.

■

According to Lemma 1, we analyze convergence time $T_{ave}(n, \varepsilon)$, which is defined in (2), of the LAIE algorithm under random graphs. In order to obtain the lower bound on the convergence time of the LAIE algorithm, we first derive the average weight matrix as follows:

***Lemma* 2**: Given graph $G = (V, E)$, $\forall i \in V$, at time slot $t \geq 1$, the degree of node i is d_i, $i = 1, 2, \cdots, n$. All weight matrices are $\{W^{(i)}(t) : i = 1, 2, \cdots, n\}$ of the LAIE algorithm; the expectation matrix $E(W(t))$ is given by:

$$EW_{jk}(t) = \begin{cases} \frac{1}{n}\left(n - d_i - 1 + \sum_{l \in L(i)} \frac{1}{d_l+1}\right), & if j = k \\ \frac{1}{n}\left(\frac{1}{d_j+1} + \frac{1}{d_k+1}\right), & if (j,k) \in E \\ \frac{1}{n}\left(\sum \frac{1}{d_r+1}\right), & if r \in V, (r,j) \in E, (r,k) \in E, (j,k) \neq E \\ 0 & elsewhere \end{cases} \tag{10}$$

where $E(W_{jk}(t))$ is the (j,k)-th entry of matrix $E(W(t))$, and $L(i) = \{i\} \bigcup N(i)$.

Proof: Without loss of the generality, we only consider the (j,k)-th entry $E(W_{jk}(t))$ of expectation matrix $E(W(t))$. $\forall j \in V$, not only does it start a computation itself, but it is also be chosen by other nodes in its neighborhood to partake in computation. The total computation time is equal to $d_j + 1$. Otherwise, its value remains unchanged, and the diagonal entry is given by:

$$E(W_{jj}(t)) = \frac{1}{n}\left(n - d_j - 1 + \sum_{l \in L(i)} \frac{1}{d_l + 1}\right) = 1 - \frac{d_j + 1}{n} + \frac{1}{n}\sum_{l \in L(i)} \frac{1}{d_l + 1} \qquad j = 1, 2, \cdots, n \tag{11}$$

Node j and k are a pair of neighbor nodes, i.e., $(j,k) \in E$. If either j or k is chosen as the initiator node of information exchange in the LAIE algorithm, another node must be involved in its computation. Then, the (j,k)-th entry of expectation matrix $E(W(t))$ is given by:

$$E(W_{jk}(t)) = \frac{1}{n}\left(\frac{1}{d_j + 1} + \frac{1}{d_k + 1}\right)$$

$\forall r \in V$, assume that both node j and node k are neighbors of node r, but $j \neq r, k \neq r$, and $(j,k) \notin E$; then the two nodes j and k participate in computation with respect to node r. Thus, the (j,k)-th entry of expectation matrix $E(W(t))$ is given by:

$$E(W_{jk}(t)) = \frac{1}{n}\left(\sum_{r \in V, (r,j) \in E, (r,k) \in E, (j,k) \neq E} \frac{1}{d_r + 1}\right)$$

■

Lemma 2 gives the average weight matrix $E(W(t))$ of the LAIE algorithm under random graphs. Reference [12] showed that the convergence time of the random gossip algorithm depends on the second-largest eigenvalue of the average weight matrix. Therefore, we analyze the convergence time of the LAIE algorithm by the second-largest eigenvalue of the average weight matrix associated with the random graphs as follows.

***Theorem* 1**: Given graph $G = (V, E)$, assume that the nodes are placed in a random and independent way. At any time slot $t \geq 1$, assume that to $\forall i \in V$, its degree is d_i and its weight matrices $\{W^{(i)}(t) : i = 1, 2, \cdots, n\}$. The expectation matrices $E(W(t))$ of the LAIE algorithm are defined as (9) and (10), respectively. Then, the convergence time of the LAIE algorithm is given by:

$$T_{ave} = \Omega\left(\frac{(n-1)\log n}{\Delta}\right). \tag{12}$$

Proof: Based on the definition of the trace of a matrix in [27], the following equation is given by:

$$\sum_{i=1}^{n} \lambda_i^{EW} = \sum_{i=1}^{n} E(W_{ii}(t)) = tr(E(W(t)). \tag{13}$$

where $tr(E(W(t)))$ is the trace of the average weight matrix. According to (11),

$$\begin{aligned}\sum_{i=1}^{n} E(W_{ii}(t)) &= \sum_{i=1}^{n} \frac{1}{n}\left(n - d_i - 1 + \sum_{l\in L(i)} \frac{1}{d_l+1}\right) = \sum_{i=1}^{n}\left(1 - \frac{1}{n} - \frac{d_i}{n} + \frac{1}{n}\sum_{l\in L(i)}\frac{1}{d_l+1}\right)\\ &= (n-1) - \sum_{i=1}^{n}\frac{d_i}{n} + \frac{1}{n}\sum_{i=1}^{n}\sum_{l\in L(i)}\frac{1}{d_l+1} = (n-1) - \sum_{i=1}^{n}\frac{d_i}{n} + 1 = n - \sum_{i=1}^{n}\frac{d_i}{n}.\end{aligned} \tag{14}$$

Then, the sum of diagonal entries of the average weight matrix can be bounded as follows:

$$\sum_{i=1}^{n} E(W_{ii}(t)) \geq n - \sum_{i=1}^{n}\frac{\Delta}{n} = n - \Delta \tag{15}$$

Let $\{\lambda_n, \lambda_{n-1}, \ldots, \lambda_1\}$ denote the set of all the eigenvalues of the average weight matrix $E(W(t))$ in increasing order, and the largest eigenvalue λ_1 equals 1. Then, the lower bound of the second-largest eigenvalue λ_2 is given by:

$$\lambda_2 \geq \frac{1}{n-1}\left(\sum_{i=1}^{n}\lambda_i - 1\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}E(W_{ii}(t)) - 1\right) \geq \frac{n-\Delta-1}{n-1} = 1 - \frac{\Delta}{n-1}$$

based on the results in [12] , we have

$$T_{ave} = \Theta\left(\frac{\log n}{1-\lambda_2^{EW}}\right) \geq \frac{\log n}{1-(1-\frac{\Delta}{n-1})} = \frac{\log n}{\frac{\Delta}{n-1}} = \frac{(n-1)\log n}{\Delta}.$$

which denotes $T_{ave} = \Omega(\frac{(n-1)\log n}{\Delta})$. ■

We now analyze the expected communication cost of the LAIE algorithm required under the random graph.

***Corollary* 1**: The expected communication cost, which is defined as (4), required by the LAIE algorithm under the random graph is bounded as:

$$EC(n,\varepsilon) = \Omega\left(\frac{(\delta+2)(n-1)\log n}{\Delta}\right) \tag{16}$$

Proof: Based on equation (4), we must compute $E(R(t))$ and T_{ave} respectively. The parameter $R(t)$, which is defined in Definition 2, denotes the amount of transmission required for a given node to communicate with some other node at time slot t. In the LAIE algorithm, each round of transmission includes, at least, the broadcast operation twice, and δ instances of one-by-one transmission. Then we have $E(R(t)) \geq (\delta+2)$. Based on Theorem 2, the total transmission cost is given by:

$$EC(n,\varepsilon) = E(R(t))T_{ave}(n,\varepsilon,\eta) = \Omega\left(\frac{(\delta+2)(n-1)\log n}{\Delta}\right)$$

■

C. Analysis of Convergence Time under d-regular Graph

In this section, we will study the convergence time of the LAIE algorithm under d-regular graph $G(V,E)$. Based on the special connectivity of the d-regular graph, we first develop a novel analysis approach to evaluate the convergence time of the LAIE algorithm. Let Φ and D denote the adjacency matrix and diagonal matrix of the d-regular graph, respectively. Then we analyze the average weight matrix EW for the LAIE algorithm under the d-regular graph $G(V,E)$, which is depicted as follows:

***Lemma* 3**: The average weight matrix EW of the LAIE algorithm under the d-regular graph $G(V,E)$ is given by:

$$EW = \frac{n-d-1}{n}I + \frac{(L-(d+1)I)^2}{n}. \tag{17}$$

and the average weight matrix EW satisfies the following equations:

$$EW\mathbf{1} = \mathbf{1},\ \mathbf{1}^T EW = \mathbf{1}^T,\ \rho(EW - J) < 1. \tag{18}$$

Proof: Based on Corollary 1, the average weight matrix is given by:

$$EW = \frac{n-d}{n} \cdot I + \frac{1}{n(d+1)} \left(\Phi^2 + 2\Phi - D \right)$$

As the Laplacian matrix of a graph is $L = D - \Phi$, then $\Phi = D - L$. The average weight matrix EW can be formulated as:

$$\begin{aligned} EW &= \frac{n-d}{n} I + \frac{1}{n(d+1)} (\varPhi^2 + 2\varPhi - D) \\ &= \frac{1}{n(d+1)} \left[\varPhi^2 - dI + 2\varPhi + (d+1)(n-d)I \right] \\ &= \frac{1}{n(d+1)} \left[(D-L)^2 + 2(D-L) + (nd + n - d^2 - 2d)I \right] \\ &= \frac{1}{n(d+1)} \left[L^2 - 2(d+1)L \right] + I \\ &= \frac{n-d-1}{n} I + \frac{[L - (d+1)I]^2}{n} \end{aligned}$$

As all the weight matrices $\{W^{(i)} : i = 1, 2, \ldots, n\}$ are double stochastic matrices, the vector **1** is both a left and right eigenvector of the average weight matrix. With this in mind, the spectral radius of the weight matrix is less than one. ■

According to the above analysis, we can conclude that the convergence time of the LAIE algorithm under the d-regular graph $G(V, E)$ is as follows.

***Theorem* 2** Consider a d-regular graph $G = (V, E)$, where $d > 2$. Assume that the nodes are placed in a random and independent way. Then the convergence time of the LAIE algorithm is given by:

$$T_{ave} = \varTheta \left(\frac{n(d+1) \log n}{(2 + d + 2\sqrt{d-1})(d - 2\sqrt{d-1})} \right). \tag{19}$$

Proof: The convergence time of the distributed averaging algorithm depends on the second largest eigenvalue λ_2 [12] . Thus, we hope to find the second largest eigenvalue of average weight matrix EW, and minimize it to derive a tight bound of convergence time. According to (17), we have:

$$\lambda^{EW} = \frac{n-d-1}{n} + \frac{(\lambda^L - d - 1)^2}{n(d+1)}$$

where λ^L is the eigenvalue of Laplacian matrix L. For the d-regular graph, the smallest eigenvalue and the second smallest eigenvalue of Laplacian matrix are 0 and $d - 2\sqrt{(d-1)}, (d > 2)$, respectively [25]. Let λ_{n-1}^L be the second smallest eigenvalue of Laplacian matrix, which is given by:

$$\lambda_2^{EW} = \frac{n-d-1}{n} + \frac{(2\sqrt{d-1}+1)^2}{n(d+1)} = 1 - \frac{d+1}{n} + \frac{(2\sqrt{d-1}+1)^2}{n(d+1)}.$$

based on the results of the reference [12], we conclude that

$$\begin{aligned} T_{ave} &= \Theta\left(\frac{\log n}{1-\lambda_2^{EW}}\right) = \Theta\left(\frac{\log n}{\frac{d+1}{n} - \frac{(2\sqrt{d-1}+1)^2}{n(d+1)}}\right) = \Theta\left(\frac{n(d+1)\log n}{(d+1)^2 - (2\sqrt{d-1}+1)^2}\right) \\ &= \Theta\left(\frac{n(d+1)\log n}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right) \end{aligned}$$

■

***Corollary* 2**: The expected communication cost required for the LAIE algorithm under the d-regular graph is given by:

$$EC(n,\varepsilon) = \Theta\left(\frac{n(d+1)(d+2)\log n}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right). \tag{20}$$

Proof: Based on (4), the expected communication cost required for the LAIE algorithm under the d-regular graph is given by:

$$\begin{aligned} EC(n,\varepsilon) &= E(R(t))T_{ave}(n,\varepsilon,\eta) = (d+2)\cdot T_{ave}(n,\varepsilon,\eta) \\ &= \Theta\left(\frac{n(d+1)(d+2)\log n}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right) \end{aligned}$$

■

Finally, we conclude that the averaging time and the expected communication cost are all decreasing functions as the degree of graph increases, which is proved as follows:

***Theorem* 3**: Given a regular graph $G = (V, E)$ with degree d, where $d > 2$, assume that the nodes are placed in a random and independent way. For a large n, the averaging time decreases as the degree of graph increases, as well as the expected communication cost.

Proof: According to Theorem 2, we have

$$T_{ave}(n,\varepsilon,\eta) = \Theta\left(\frac{n(d+1)\log n}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right) = \Theta\left(\frac{n\log n}{(1+\frac{1+2\sqrt{d-1}}{d+1})(d-2\sqrt{d-1})}\right)$$

As the parameter $d > 2$, we have $1 < 1+\frac{1+2\sqrt{d-1}}{d+1} < 2$, and $\frac{1}{d-2\sqrt{d-1}}$ is a decreasing function as d increasing. Consequently, $T_{ave}(n,d)$ is a decreasing function about d for large n. The proving process regarding the expected communication cost required for convergence is similar, which is omitted here. ■

Theorem 3 shows that the larger the degree of a graph is, the better the performance of the LAIE algorithm, which will be further verified by the extensive numerical simulation in the next section.

IV. Numerical Examples

In this section, we compare the performance of the LAIE algorithm with the randomized gossip algorithm [12], geographical gossip algorithm [13], and broadcast gossip algorithm [19] under d-regular graphs and random graphs through numerical simulation. The simulation parameters are listed as follows. The initial value(state) support of each node is [0,1]. The number of the nodes is from 0 to 200. The maximum number of rounds is 2000, and $\epsilon = 0.01$. In these four algorithms, we mainly compare the relationship between the required time achieve the global convergence as the number of nodes increases. We first plot the convergence time curves of different node degrees under k-regular graphs.

Fig.1 shows the convergence time versus the number of nodes under a 3-regular graph. LAIE algorithm results in the quickest convergence time compared to other algorithms. The gaps among these algorithms are enlarged as the number of nodes increases. This shows that the LAIE algorithm is more suited to a large number of nodes. The curves between the LAIE algorithm and the broadcast gossip algorithm are very similar with increasing number of nodes. It may be because a broadcast operator is used in these algorithms. The geographic gossip algorithm was worse than the broadcast algorithm and the LAIE algorithm because its long route experienced too many intermediate nodes, which led to wasted time.

Fig.2 shows the convergence time versus the number of nodes under 6-regular graph. As the number of nodes increases, the convergence time of the randomized gossip algorithm varies sharply. However, the convergence time of the LAIE algorithm increases slowly. Although the gap between the LAIE algorithm and the broadcast gossip algorithm is obviously thinning compared to the curves shown in Fig.1, they still outperform the random gossip algorithm and the geographic gossip algorithm.

Fig.3 shows the convergence time versus the number of nodes under 9-regular graph. The performance of the random gossip algorithm is still worse than that of other algorithms, especially

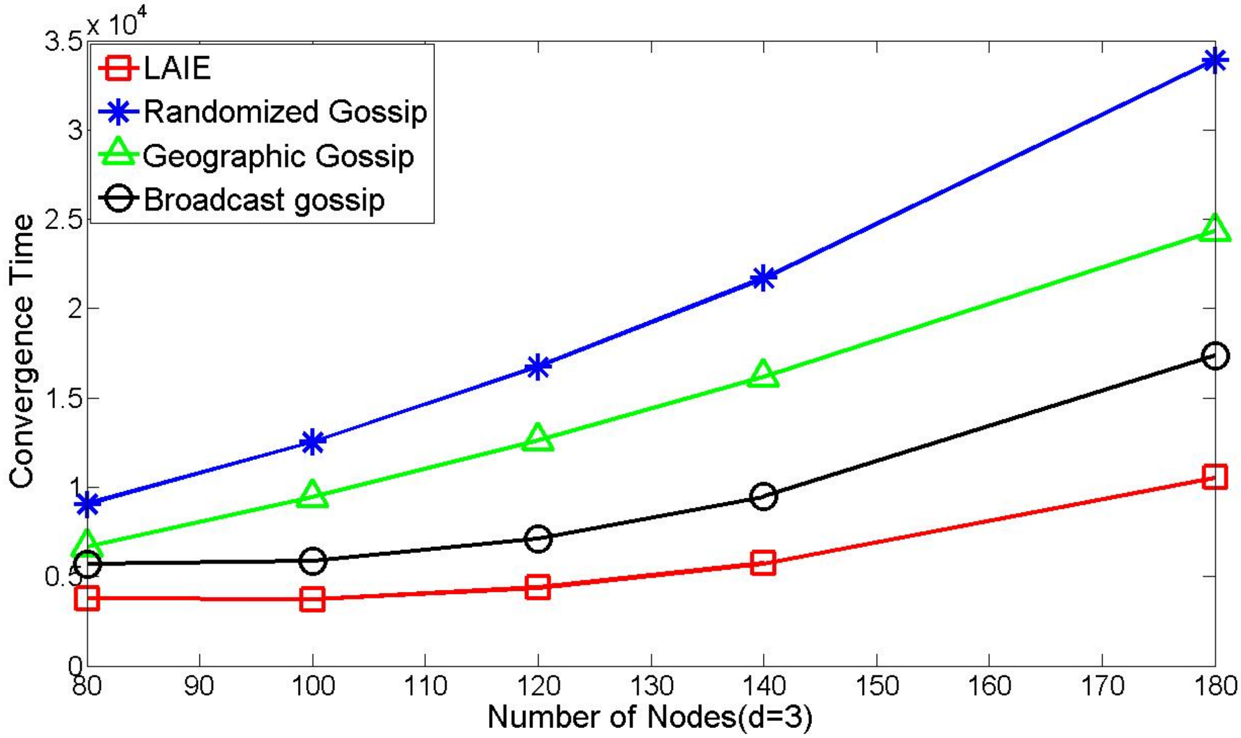

Fig. 1. The convergence time versus the number of nodes under 3-regular graph

with the number of nodes increasing. As shown from Fig.1 to Fig.3, the convergence time required for the LAIE algorithm is always smaller than that of the other three existing algorithms in the different d-regular graphs. However, as the degree of the graph increased, the gap in convergence time between the LAIE algorithm and the broadcast gossip algorithm varied little as the number of nodes increased. This implies that the broadcast operator plays an important role in these two algorithms.

Relaxing the constraint of the degree of node, the convergence time curves of the four algorithms under random graph are shown in Fig. 4. With the number of nodes increasing, the convergence time of the LAIE algorithm is quicker than other algorithms. The broadcast gossip algorithm has almost the same performance as the LAIE algorithm, which further verifies the variation trend shown from Fig.1 to Fig.3 once again. Although the random gossip algorithm is worse than other algorithms, it plays an irreplaceable role in the simplicity and scalability, especially in the dynamic scenarios.

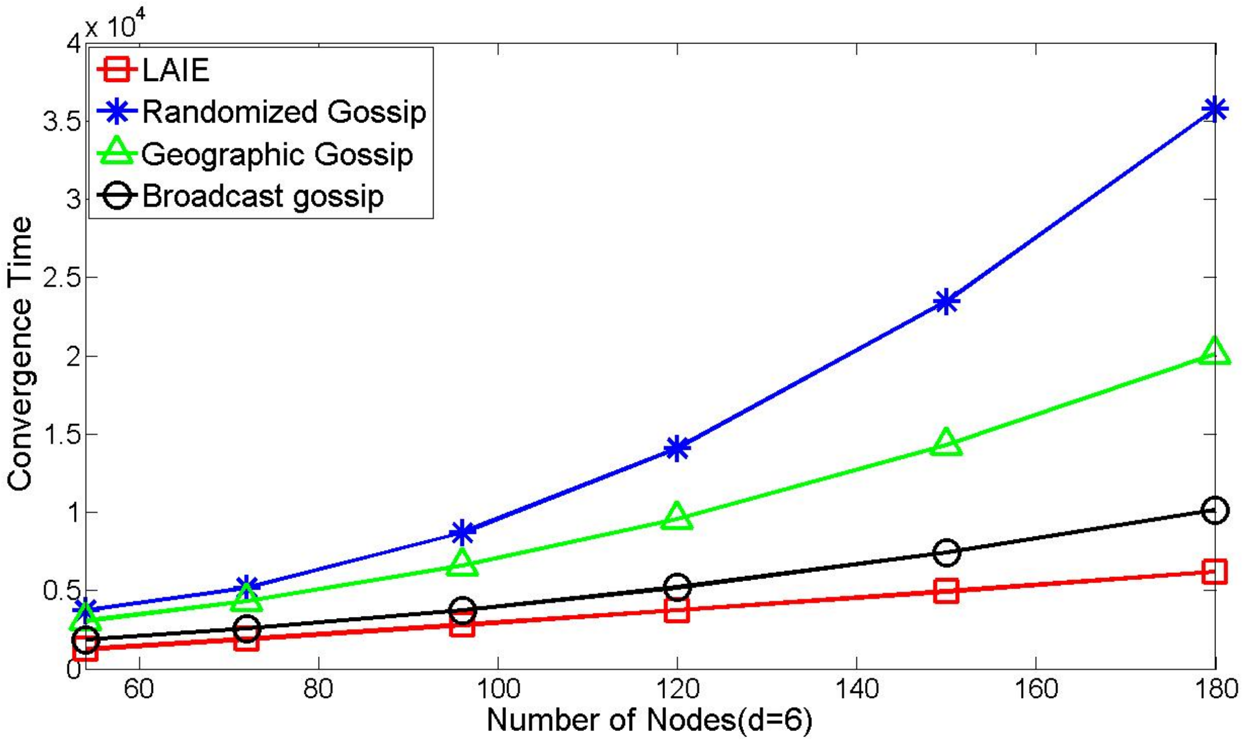

Fig. 2. The convergence time versus the number of nodes under 6-regular graph

V. Conclusion

In this paper, we proposed a novel local average broadcasting algorithm to accelerate global consensus on random and d-regular networks. In the proposed algorithm, each node interacts with all its neighboring nodes in each round of the diffusion process, such that all nodes can reach a global consensus in a distributed manner very quickly. This is in contrast to the conventional random gossip scheme, where each node only interacts with one of its neighbors, leading to a very long convergence time. By exploring the nature of broadcasts, each node acts based on only local information, without a centralized controller. It was shown that in a random network with n nodes, the convergence time of the LAIE algorithm is lower bounded by $\varOmega\left(\frac{(n-1)logn}{\Delta}\right)^1$, where the parameter Δ denotes the graph with the largest degree. For n nodes in d-regular topology graphs ($d > 2$), where each node has the same number of neighbors d, based on a novel mathematical tool—the spectrum of the Laplacian matrix—the lower bound of the convergence time of the LAIE algorithm $\varTheta\left(\frac{n(d+1)logn}{(2+d+2\sqrt{d-1})(d-2\sqrt{d-1})}\right)$ was developed. This was tighter than the bound derived by the randomized gossip algorithm $\varTheta\left(n^2 \log n\right)$ and the geographic gossip

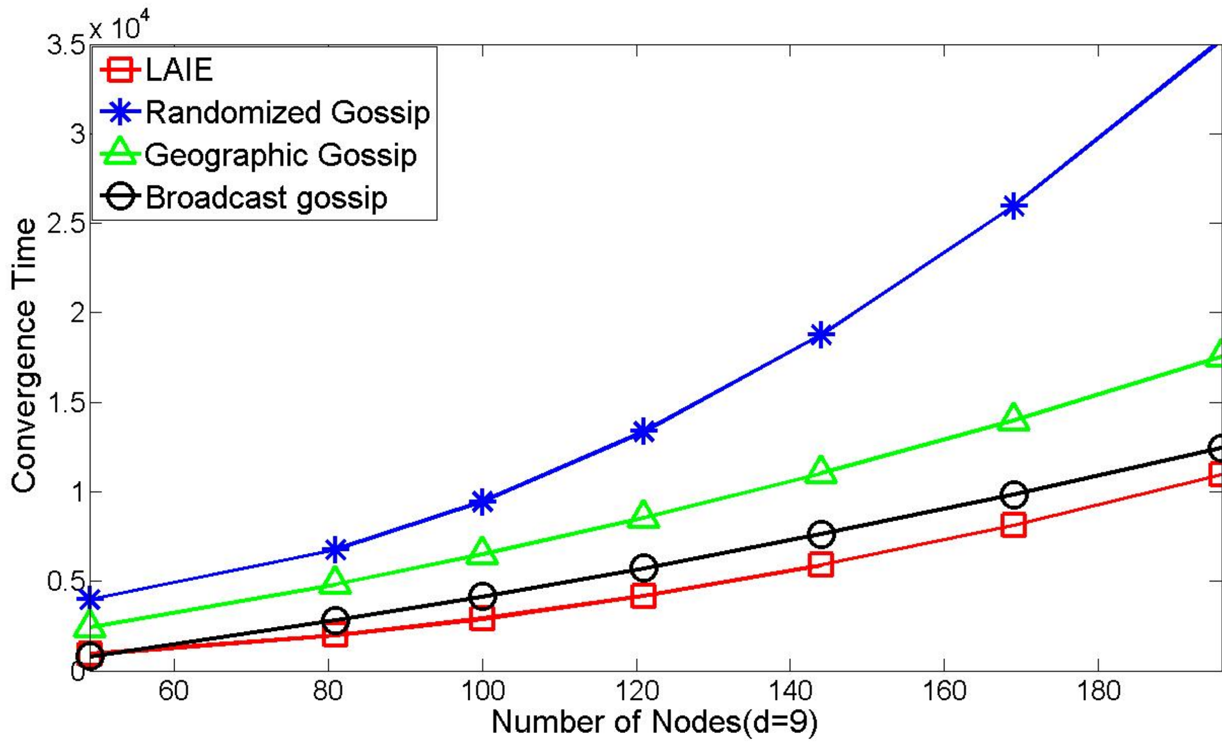

Fig. 3. The convergence time versus the number of nodes under 9-regular graph

algorithms $\Theta\left(n\log n\right)$, etc.. The results showed that this algorithm can significantly speed up the distributed average consensus process and reduce the amount of transmission in the consensus process in comparison with existing gossip algorithms. The effects on performance of the failures of nodes and links will be the focus of our future work.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61271194), by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 61521091).

REFERENCES

[1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbors Rules," IEEE Trans. Autom. Control, Vol. 48, NO. 6, pp. 988-1001, Jun. 2003.

[2] Z.Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," IEEE J. Selected Areas Communication, vol. 23, NO. 4, pp.735-744, Apr. 2005.

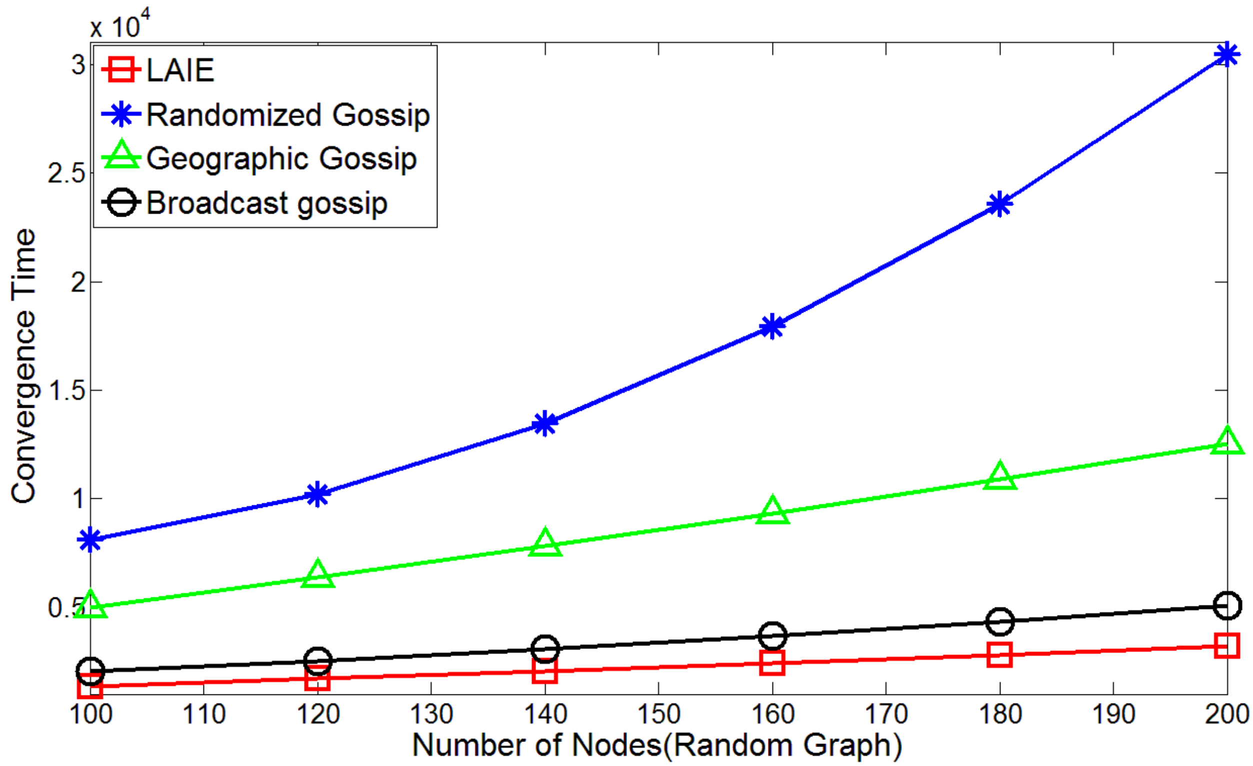

Fig. 4. The expected convergence time versus the number of nodes under random graph

[3] Z. Li, F. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios," IEEE Trans. Veh. Technol, Vol. 59, NO. 1, pp. 383-393, Jan. 2010.

[4] A. Savkin, "Coordinated collective motion of groups of autonomous mobile robots: Analysis of Vicsek's model," IEEE Trans. Autom. Control, Vol. 49, NO. 6, pp. 981-982, Jun. 2004.

[5] A. Olshevsky and J. Tsitsiklis, "Convergence rates in distributed consensus and averaging," in Proc. IEEE Conf. Decision Contr., San Diego, CA, Dec. 2006.

[6] S. Strogatz, "Exploring complex networks," Nature, Vol. 410, NO. 6825, pp. 268-276, 2001.

[7] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," in Proc. IEEE, Vol. 95, pp. 215-233, Jan. 2007.

[8] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," IEEE Trans. Autom. Control, Vol. AC-31, NO. 9, pp. 803?812, Sep. 1986.

[9] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," IEEE J. Sel. Areas Commun.,vol. 26, no. 4, pp. 634-649, May 2008.

[10] D. M. Yuan, S. Y. Xu, H. Y. Zhao, Y. M. Chu, "Distributed average consensus via gossip algorithm with real-valued and quantized data for $0 < q < 1$," Systems & Control Letters, Vol. 59, pp. 536-542, 2010.

[11] L. Xiao, S. Boyd, and S.J. Kim, "Distributed average consensus with least-mean-square deviation," J. Parallel Distrib.Comput., vol. 67, no.1, pp. 33-46, 2007.

[12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," IEEE Trans. Inf. Theory, Vol. 52, NO.

6, pp. 2508-2530, Jun. 2006.

[13] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, “Geographic gossip: Efficient aggregation for sensor networks,” IEEE Trans. Signal Processing, Vol. 56, NO. 3, pp. 1205-1216, Mar. 2008.

[14] W. Li and H. Dai, “Cluster-based distributed consensus,” IEEE Trans. Wireless Communications, Vol. 8, NO. 1, pp. 28-31, Jan. 2009.

[15] K.Censor-Hillel, B. Haeupler, P. Maymounkov, and J. Kelner, “Global computation in a poorly connected world: Fast rumor spreading with no dependence on conductance,” in Proceedings of the 44st Annual ACM Symposium on Theory of Computing (STOC), 2012.

[16] Bernhard Haeupler, “Simple, fast and deterministic gossip and rumor spreading,” in Proc. Of 24th SODA: ACM-SIAM Symposium on Discrete Algorithms, 2013.

[17] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione, “Broadcast gossip algorithms for consensus,” IEEE Trans. Signal Processing, Vol. 57, NO. 7, pp. 2748-2761, Jul. 2009.

[18] F. Fagnani and P. Frasca, “Broadcast Gossip Averaging: Interference and Unbiasedness in Large Abelian Cayley Networks,” IEEE J. Sel. Topics Signal Processing, vol. 5, no. 4, pp. 866-875, Aug. 2011.

[19] S. Wu and M. G. Rabbat, “Broadcast Gossip algorithms for consensus on strongly connected digraphs,” IEEE Trans. Signal processing, Vol. 61, NO. 16, pp. 3959-3971, August, 2013.

[20] S. C. Wu, J. Y. Zhang, Y. G. Hou, and X. Bai, “Convergence of Gossip Algorithms for Consensus in Wireless Sensor Networks with Intermittent Links and Mobile Nodes,” Mathematical Problems in Engineering, Vol 2014, 2014, Article ID 836584, http://dx.doi.org/10.1155/2014/836584

[21] H. W. Wang, X. F. Liao, Z. D. Wang, T. W. Huang, G. Chen, “Distributed parameter estimation in unreliable sensor networks via broadcast gossip algorithms,” Neural Networks, Vol. 73, pp. 1-9, 2016.

[22] Sukhavasi R T, Hassibi B, “Tree codes improve convergence rate of consensus over erasure channels,“ arXiv preprint arXiv: 1204.0301, 2012.

[23] Tran, T. M. D, Kibangou, A. Y, “Consensus-based distributed estimation of Laplacian eigenvalues of undirected graphs,” in IEEE Control Conference (ECC), European, pp. 227-232, July. 2013.

[24] Michael William Newman, “The laplacian spectrum of graphs,” PhD thesis, Department of Mathematics, University of Manitoba, July. 2000.

[25] B. Mohar, “The Laplacian spectrum of graphs,” in Graph Theory, Combin. Applicat, Y. A. G. Chartrand, O. R. Ollerman, and A. J.Schwenk, Eds, vol. 2, pp. 871-898, 1991.

[26] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. Shamma, “Belief consensus and distributed hypothesis testing in sensor networks,” Netw. Embedded Sens. Control, pp. 169-182, 2006.

[27] R.A. Horn, C. R.Johnson, Matrix Analysis (2nd Edition), Cambridge Publisher, 2014.