International Scholarly Research Network
ISRN Communications and Networking
Volume 2011, Article ID 860936, 10 pages
doi:10.5402/2011/860936

Research Article

Self-Aware Distributed Consensus Building for Sensor Networks

Lei Chen and Jeff Frolik

School of Engineering, The University of Vermont, Burlington, VT 05405, USA

Correspondence should be addressed to Lei Chen, lchen2@uvm.edu

Received 9 August 2011; Accepted 15 September 2011

Academic Editors: M. Maier and R. Montemanni

Distributed consensus building promises to improve the robustness and reliability of sensor networks and thus is an active topic of research. Whereas extensive study has been done on the theoretical analysis of the asymptotic behavior of consensus building, one important issue that is crucial to the practical implementation of sensor networks was rarely explored, namely, the criteria to determine whether consensus has been attained. In this paper, we propose an approach that allows each node in a network to make the decision by itself, based on the second derivatives of its own state. The approach does not rely on the states of other nodes, leads to substantial saving of communication resources, and is resilient to connection failure. We perform a systematic analysis of the approach and, as a consequence, derive the optimal parameters that minimize the upper bound of the number of required iterations to reach consensus.

1. Introduction

In past decades, sensor networks have been widely applied in civilian, military, and industrial systems. Despite tremendous efforts that have been devoted to sensor techniques, it remains a great challenge to build a sensor system that works reliably and responsively under varying conditions. Distributed consensus building has emerged as an approach to address such shortcomings. Each individual node updates its "*state*" (e.g., view of local or global sensed values, actions to be taken, and/or network conditions) while exchanging information with others, until consensus of the state is attained. While individual sensors could have unreliable information or be vulnerable to environmental or network dynamics, all sensors in the network, via effective cooperation, can produce decisions that are higher in reliability because of the reached consensus. Since such consensus building enables the network to act locally based on distributed interactions, we contend that this approach can benefit networks that also involve actuation or otherwise interact with the physical world.

In previous work, there has been extensive theoretical study of the distributed consensus building problem. Xiao and Boyd [1] considered the convergence of a distributed averaging problem and proposed to seek the optimal updating weights by solving a semidefinite optimization problem. We note that the derivation in this work was based on the symmetric communication link assumption, which is not always satisfied in real networks. Olshevsky and Tsitsiklis [2] performed a more comprehensive investigation of this problem and derived lower bounds on the worst-case convergence time for different types of distributed consensus methods. Based on the theoretical analysis, they also developed an algorithm whose convergence time can essentially match these bounds. Olfati-Saber et al. [3, 4] developed a theoretical framework of information consensus over a multiagent network. They inspected various theoretical aspects of the problem and established the connections between the spectral properties of complex networks and the convergence rate of the consensus algorithms.

Whereas this previous work has laid down a solid theoretical foundation for the research on distributed consensus, it mostly focused on the analysis of asymptotic behaviors. As a consequence, practical issues, which are equally important to the effectiveness of sensor networks, have received less attention. In this paper, we particularly study one of these issues, namely, how might an individual node determine whether consensus (within a tolerable range) has been attained. This is motivated by the consideration that if the consensus status can be detected as soon as it is reached, the sensor network can then be put to rest earlier, which should lead to considerable energy saving.

Specifically, our goal here is to develop a consensus determining scheme with two desirable properties. First, it is *self-aware*, namely, each node in the network decides whether this scheme has reached consensus solely based on its own state, without resorting to the communication with other nodes. Compared to the schemes that detect consensus via information exchange with other nodes, a self-aware scheme does not consume extra communication bandwidth and tends to be more responsive. Second, the scheme should be resilient against noise as well as the variation of propagation delay and link weights. Our approach to achieve this target is inspired by the continuous-time work of Barbarossa et al. [5–7]. In their work, the network consensus is defined with respect to the first derivative of the states, rather than the states themselves. They showed that this new definition can result in improved robustness against noises, delays, and even topology changes. Hence, we adopt this formalism as the basis of our work.

The work presented here extends our earlier work [8] which initiates the basic idea of our approach. In that paper, we derived a discrete-time counterpart of the continuous-time formulation given in [5], making it implementable in physical, discrete-time hardware. We also verified the method through both numerical analysis and hardware implementation. In this paper, we further extend our previous exploration. Particularly, we consider the self-aware criterion given by

$$\|\ddot{\mathbf{x}}(t)\| \leq \varepsilon, \tag{1}$$

where $\mathbf{x}(t)$ is the state of a node as a function of time and ε is the tolerance threshold. With this criterion, we obtain an upper bound of the number of iterations needed to achieve consensus and thereon derive the optimal step coefficient δ that minimizes this bound. Moreover, we examined how the optimal solution is related to different factors, including the spectrum of the graph Laplacian as well as the tolerance range, via both theoretical analysis and numerical simulation. We note that our analysis is not restricted to symmetric network, as opposed to some previous analysis, such as that in [1].

To sum up, the main contributions of this paper lies in the following aspects. First, we develop a self-aware criterion to detect consensus based on the discrete-time formulation presented in our earlier work [8]. It is simple, cost-effective, and robust to noises as well as network changes. Second, we perform theoretical analysis of the criterion for both symmetric and asymmetric networks and thereon obtain an upper bound of the number of iterations needed to reach consensus. Third, we derive an optimal updating weight and examine its relations to the spectral structure of the network as well as the tolerance threshold.

A brief outline of the rest of the paper follows. In Section 2, we formulate a discrete-time consensus model and analyzes its asymptotic behavior. In Section 3, we obtain an upper bound of the convergence time based on this model and thereon derive the optimal choice of δ. In Section 4, we give some numerical examples, illustrate the influence of ε to the consensus for both symmetric and asymmetric scenarios, and verify the analytical solution of δ. Finally, Section 5 concludes this work and identifies areas of future research.

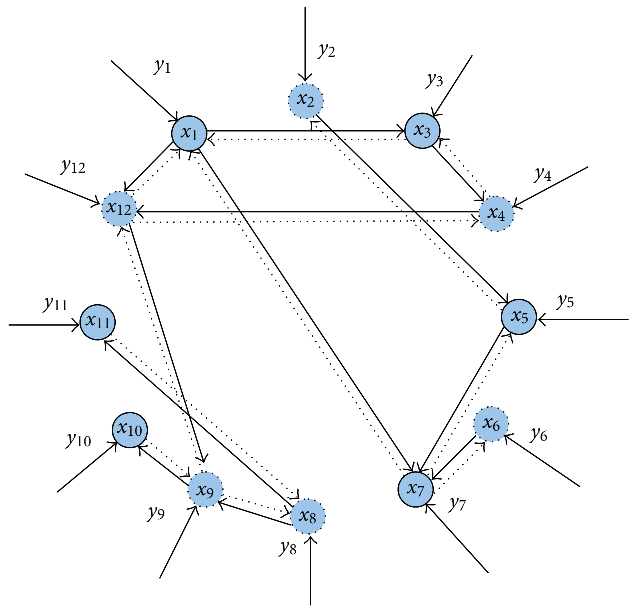

Figure 1: Schematic diagram of a sensor network with $N = 12$ nodes.

2. The Formulation of Consensus Model

In this paper, we consider a sensor network with N nodes (e.g., $N = 12$ in Figure 1). Each node receives input from its sensors, and combines it with the states of its neighboring nodes to adjust its own state. These N nodes constitute a dynamical system. Specifically, for each node i, its dynamical behavior can be described by the following equation:

$$\dot{x}_i(t) = y_i(t) + \sum_{j\in\mathcal{N}(i)} a_{ij}\left(x_j(t) - x_i(t)\right). \tag{2}$$

Here, x_i is the state of node i as a function of time, y_i is the corresponding input signal, $\mathcal{N}(i)$ is the set of neighboring nodes of i, and a_{ij} is the coupling coefficient that reflects the contribution of node j to the evolution of node i. Note here that the coupling coefficients need not to be symmetric ($a_{ij} = a_{ji}$) in general. In this model, the change of x_i is driven by a linear combination of the input signal and state differences from its neighbors. For simplicity, we assume that the input signals are static, that is, $y_i(t) = y_i$ for each $i = 1, \ldots, N$.

The model given in (2) can be rewritten into the following vectorized form:

$$\dot{\mathbf{x}}(t) = -\mathbf{L}\mathbf{x}(t) + \mathbf{y}. \tag{3}$$

Here, $\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^T$, $\mathbf{y} = [y_1, \ldots, y_N]^T$, and $\mathbf{L}$ denotes the Laplacian matrix of the communication graph whose edge weights are given by the coupling coefficients a_{ij}. Particularly, we have

$$\mathbf{L}(i,j) = \begin{cases} \sum_{j\in\mathcal{N}(i)} a_{ij} & (i = j), \\ -a_{ij} & (j \in \mathcal{N}(i)), \\ 0 & (\text{otherwise}). \end{cases} \quad (4)$$

We note that $\mathbf{L}$ is symmetric if and only if the coupling coefficients a_{ij} are symmetric. A symmetric system implies that all communication links are bidirectional. However, this is not always the case due to packet drops, and different transmit powers and receiver sensitivities of the hardware.

2.1. Discrete-Time Formulation. To implement (3) on physical, discrete-time hardware, we reformulate the problem under discrete-time setting, via finite difference approximation, as

$$\begin{aligned} \mathbf{x}(n) &= \mathbf{x}(n-1) + \delta\dot{\mathbf{x}}(n-1) \\ &= (\mathbf{I} - \delta\mathbf{L})\mathbf{x}(n-1) + \delta\mathbf{y}. \end{aligned} \quad (5)$$

We note that the updating at each step is controlled by an updating weight δ. Through recursively expanding the formula above, we find the analytic expression for $\mathbf{x}(n)$, as

$$\mathbf{x}(n) = \mathbf{\Phi}^n\mathbf{x}(0) + \delta\sum_{k=0}^{n-1}\mathbf{\Phi}^k\mathbf{y}, \quad (6)$$

where $\mathbf{\Phi} = \mathbf{I} - \delta\mathbf{L}$. From this, we further obtain a discrete version of the first derivative, as

$$\begin{aligned} \dot{\mathbf{x}}_\delta(n) &\triangleq \frac{1}{\delta}(\mathbf{x}(n) - \mathbf{x}(n-1)) \\ &= \mathbf{\Phi}^{n-1}(-\mathbf{L}\mathbf{x}(0) + \mathbf{y}). \end{aligned} \quad (7)$$

Here, $\dot{\mathbf{x}}_\delta$ refers to the discrete derivative obtained with updating weight δ.

2.2. The Conditions of Convergence. According to Barbarosa's definition, the consensus is attained when the first derivatives of the states of all nodes converge to stationary values, which we denote by a vector $\boldsymbol{\omega}^*$, as illustrated by Figure 2. However, as we will show below, convergence is guaranteed only under specific conditions, but not in general.

Lemma 1. *The first derivatives of the system given by* (7) *converge to stationary values* $\boldsymbol{\omega}^*$ *with arbitrary input* $\mathbf{y}$ *if and only if*

$$\rho(\mathbf{\Phi}) \le 1. \quad (8)$$

In this case, $\mathbf{\Phi}^\infty \triangleq \lim_{n\to\infty}\mathbf{\Phi}^n$ *is finite, and the vector of stationary values* $\boldsymbol{\omega}^*$ *is given by*

$$\boldsymbol{\omega}^* = -\mathbf{\Phi}^\infty\mathbf{L}\mathbf{x}(0) + \mathbf{\Phi}^\infty\mathbf{y}. \quad (9)$$

In particular, when $\|\mathbf{\Phi}^\infty\mathbf{L}\| = 0$*, one has* $\boldsymbol{\omega}^* = \mathbf{\Phi}^\infty\mathbf{y}$*, which is independent of the initial states* $\mathbf{x}(0)$.

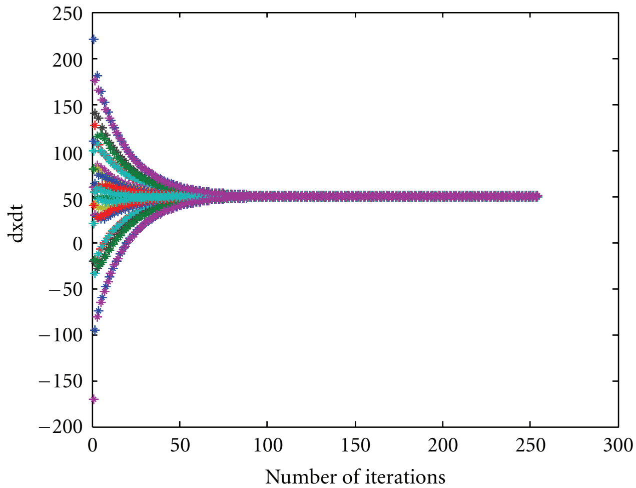

FIGURE 2: Convergence of the state of the sensor network in Figure 1.

Here, $\rho(\mathbf{\Phi})$ is the *spectral radius* of $\mathbf{\Phi}$, that is, the maximum absolute value of its eigenvalues.

Proof of Lemma 1. If $\rho(\mathbf{\Phi}) \le 1$, $\mathbf{\Phi}^n$ is a convergent sequence and thus so is $\dot{\mathbf{x}}_\delta(n)$ due to continuity of linear operation. Otherwise, if $\rho(\mathbf{\Phi}) > 1$, there exists an eigenvalue λ of $\mathbf{\Phi}$ such that $|\lambda| > 1$. Since $\mathbf{y}$ is arbitrary, we can choose $\mathbf{y}$ such that $-\mathbf{L}\mathbf{x}(0) + \mathbf{y} = \mathbf{e}$, where $\mathbf{e}$ is the eigenvector associated with λ. As a result, $\dot{\mathbf{x}}_\delta(n) = \lambda^{n-1}\mathbf{e}$, which diverges. Therefore, the system converges for every $\mathbf{y}$ if and only if $\rho(\mathbf{\Phi}) \le 1$.

The fact that $\mathbf{\Phi}^n$ converges implies that $\mathbf{\Phi}^\infty$ is finite. Again, by continuity of linear operation, we obtain

$$\begin{aligned} &\lim_{n\to\infty}\mathbf{\Phi}^{n-1}(-\mathbf{L}\mathbf{x}(0) + \mathbf{y}) \\ &\quad = \left(\lim_{n\to\infty}\mathbf{\Phi}^{n-1}\right)(-\mathbf{L}\mathbf{x}(0) + \mathbf{y}) \\ &\quad = \mathbf{\Phi}^\infty(-\mathbf{L}\mathbf{x}(0) + \mathbf{y}). \end{aligned} \quad (10)$$

This completes the proof of the second statement. □

Based on this lemma, we derive the conditions of convergence, which as we will see are closely related to the eigenvalues of $\mathbf{L}$, denoted by $\lambda_1, \ldots, \lambda_N$. Without losing generality, we assume that $|\lambda_1| \le \cdots \le |\lambda_N|$. The properties of these eigenvalues are summarized below.

(1) $\text{rank}(\mathbf{L}) = N - K$, where K is the number of connected components of the communication graph. The number of null eigenvalues indicates the number of connected components.

(2) $\mathbf{L}$ has a null space which contains $\mathbf{1}$, that is, $\mathbf{L1} = \mathbf{0}$. In particular, when $K = 1$, the null space is a one-dimensional space spanned by $\mathbf{1}$.

(3) $|\lambda_N| = \rho(\mathbf{L})$, the spectral radius of $\mathbf{L}$.

(4) When $\mathbf{L}$ is symmetric, $\lambda_1, \ldots, \lambda_N$ are nonnegative real values.

The eigenvalues of $\mathbf{\Phi} = \mathbf{I} - \delta\mathbf{L}$ are then given by $1 - \delta\lambda_1, \ldots, 1 - \delta\lambda_N$. Hence, the spectral radius of $\mathbf{\Phi}$ is given by

$$\rho(\mathbf{\Phi}) = \max\{|1 - \delta\lambda_i| : i = 1, \ldots, N\}. \tag{11}$$

This immediately leads to an important condition of convergence, as given by the following corollary.

Corollary 2. *The system given by* (7) *converges if and only if*

$$|1 - \delta\lambda_i| \leq 1, \quad \forall i = 1, \ldots, N. \tag{12}$$

In particular, when $\mathbf{L}$ *is symmetric, all eigenvalues are real, and thus this condition is equivalent to*

$$\delta\lambda_N \leq 2. \tag{13}$$

Proof of Corollary 2. By Lemma 1, the system converges if and only if $\rho(\mathbf{\Phi}) \leq 1$. By (11), this is equivalent to $\max_{i=1,\ldots,N}|1-\delta\lambda_i| \leq 1$, that is, $|1-\delta\lambda_i| \leq 1$ for each i. When $\mathbf{L}$ is symmetric, all eigenvalues are real; hence, the maximum here can be either $|1 - \delta\lambda_1|$ or $|1 - \delta\lambda_N|$. Recall that $\lambda_1 = 0$, thus $|1-\delta\lambda_1| = 1$. Therefore, the condition holds if and only if $|1 - \delta\lambda_N| \leq 1$. Considering that $\lambda_N > 0$, this is equivalent to $\delta\lambda_N \leq 2$. □

Suppose that the *condition of convergence* given above is satisfied; we have the following discussions in regard to the matrices $\mathbf{\Phi}$ and $\mathbf{\Phi}^\infty$.

(1) Since $\mathbf{L1} = \mathbf{0}$, we have $\mathbf{\Phi 1} = (\mathbf{I} - \delta\mathbf{L})\mathbf{1} = \mathbf{1}$, implying that $\mathbf{\Phi}$ has an eigenvalue 1, which is associated with an eigenvector $\mathbf{1}$.

(2) Suppose that the number of connected components of the communication graph is K, then the geometric multiplicity of the eigenvalue 1 of $\mathbf{\Phi}$ is K, meaning that the dimension of the eigenspace corresponding to the eigenvalue 1 is K.

(3) $\mathbf{\Phi}^n\mathbf{1} = \mathbf{1}$ for each $n = 1, 2, \ldots$. Taking this sequence to the limit leads to an important result, namely, $\mathbf{\Phi}^\infty\mathbf{1} = \mathbf{1}$.

These results provide necessary basis for our later derivation.

2.3. Consensus over Connected Networks. In this paper, we focus on the case where the communication graph is connected and $\mathbf{L}$ is diagonalizable. For this case, we derive the following results that characterize the asymptotic behavior of the dynamical system.

Lemma 3. *Suppose that* $\mathbf{L}$ *is the Laplacian matrix of a connected graph and is diagonalizable,* $\mathbf{\Phi} = \mathbf{I} - \delta\mathbf{L}$ *such that* $\rho(\mathbf{\Phi}) = 1$, *then*

$$\mathbf{\Phi}^\infty = \mathbf{1}\boldsymbol{\gamma}^T, \tag{14}$$

where $\boldsymbol{\gamma}$ *is a left eigenvector of* $\mathbf{\Phi}$ *associated with the eigenvalue* 1 *that satisfies* $\boldsymbol{\gamma}^T\mathbf{1} = 1$.

Proof of Lemma 3. Since $\mathbf{\Phi}$ is diagonalizable, it can be written as

$$\mathbf{\Phi} = \mathbf{PDP}^{-1}, \tag{15}$$

where $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}_{ii} = 1 - \delta\lambda_i$, and $\mathbf{P}$ is a matrix whose columns are the corresponding eigenvectors. Hence,

$$\mathbf{\Phi}^n = \mathbf{PD}^n\mathbf{P}^{-1}. \tag{16}$$

Note that $\mathbf{D}^n$ is also a diagonal matrix with $\mathbf{D}^n_{ii} = (1 - \delta\lambda_i)^n$. Under the condition of convergence, as $n \to \infty$, $\mathbf{D}^n_{11}$ remains 1, while all other entries converge to 0. Consequently,

$$\mathbf{\Phi}^\infty = (\mathbf{P})_1 (\mathbf{P}^{-1})_1^T. \tag{17}$$

Here, we use the notation $(\cdot)_1$ to indicate the first column of a matrix. The first column of $\mathbf{P}$ is the eigenvector associated with the eigenvalue $1 - \delta\lambda_1 = 1$, which is $\mathbf{1}$, while the first column of $\mathbf{P}^{-1}$ is a left eigenvector associated with the same eigenvalue, which we denote by $\boldsymbol{\gamma}$. Consequently, we can write $\mathbf{\Phi}^\infty = \mathbf{1}\boldsymbol{\gamma}^T$. Moreover, as we have discussed above, $\mathbf{\Phi}^\infty\mathbf{1} = \mathbf{1}$, thus $\mathbf{1}\boldsymbol{\gamma}^T\mathbf{1} = \mathbf{1}$, which implies that $\boldsymbol{\gamma}^T\mathbf{1} = 1$. The proof is completed. □

Following this lemma, we derive the following results that characterize the asymptotic behaviors of the dynamical system on a connected communication network.

Theorem 4. *Suppose that the dynamical system given by* (7) *is on a connected network and the Laplacian matrix* $\mathbf{L}$ *is diagonalizable. If* δ *is chosen such that* $\rho(\mathbf{\Phi}) \leq 1$, *then the first derivative of each node converges to a common stationary value* ω^* *given by*

$$\omega^* = \boldsymbol{\gamma}^T\mathbf{y} = \sum_{i=1}^{N}\gamma_i y_i. \tag{18}$$

Here, $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_N]^T$ *is a left eigenvector of* $\mathbf{\Phi}$ *associated with the eigenvalue* 1 *that satisfies* $\sum_{i=1}^N \gamma_i = 1$.

Here, convergence to a common stationary value means that all entries of $\boldsymbol{\omega}^*$ are the same, that is, $\boldsymbol{\omega}^* = [\omega^*, \ldots, \omega^*]^T$. In addition, we can see that the value ω^* is a weighted average of the input signals, which reflects the global consensus formed over the sensor network.

Proof of Theorem 4. From Lemma 3, we can see that, under the condition of convergence, $\mathbf{\Phi}^n$ converges to $\mathbf{\Phi}^\infty = \mathbf{1}\boldsymbol{\gamma}^T$ as $n \to \infty$. And since $\mathbf{L1} = \mathbf{0}$, $\mathbf{L\Phi}^\infty = \mathbf{L1}\boldsymbol{\gamma}^T = \mathbf{0}$. By Lemma 1, we know that the states converge to stationary values given by

$$\boldsymbol{\omega}^* = \mathbf{\Phi}^\infty\mathbf{y} = \mathbf{1}\boldsymbol{\gamma}^T\mathbf{y}. \tag{19}$$

Let $\omega^* = \boldsymbol{\gamma}^T\mathbf{y}$; this can be further rewritten as $\boldsymbol{\omega}^* = \omega^*\mathbf{1}$. The proof is completed. □

2.4. Determining Consensus. For any practical implementation, the stationary value ω^* is not available in advance; hence, one has to resort to other means to determine whether a sensor network has reached consensus. A natural idea is to compare the first derivatives of all nodes and see whether they are close to each other. However, implementation of this strategy might require a centralized process to collect information from every node, which could be expensive or even infeasible in practice.

In this paper, we take a different approach. Instead of relying on the first derivatives to detect consensus, we consider the second derivatives. Based on our discrete-time formulation, we define the vector of *second derivatives* to be

$$\begin{aligned}\ddot{\mathbf{x}}_\delta(n) &= \frac{1}{\delta}\left(\dot{\mathbf{x}}_\delta(n) - \dot{\mathbf{x}}_\delta(n-1)\right)\\ &= \mathbf{L}\mathbf{\Phi}^{n-2}\left(\mathbf{L}\mathbf{x}(0) - \mathbf{y}\right).\end{aligned} \tag{20}$$

Suppose that the dynamical system is on a connected communication network with diagonalizable Laplacian matrix $\mathbf{L}$; we have $\mathbf{L}\mathbf{\Phi}^n \to \mathbf{0}$ as $n \to \infty$. It immediately follows that

$$\lim_{n\to\infty} \ddot{\mathbf{x}}_\delta(n) = \mathbf{0}. \tag{21}$$

This observation suggests that we may decide whether consensus has been attained by examining whether the second derivatives vanish. Specifically, we devise the following rule. Given a tolerance threshold $\varepsilon > 0$, a sensor network is regarded to reach consensus if

$$\|\ddot{\mathbf{x}}_\delta(n)\|_\infty < \varepsilon. \tag{22}$$

Here, $\|\cdot\|_\infty$ denotes the infinity norm. Intuitively, the network is considered to attain consensus when the magnitude of the second derivative at each node is below the given threshold ε.

It is worth noting that this rule is *self-aware*, meaning that each node can decide whether it meets the condition locally, without talking to other nodes. Compared to other rules, its main advantages are three-fold: (1) simplicity of implementation, (2) no consumption of extra communication energy, and (3) resilience to the changes of network conditions.

3. Convergence Time and Optimal δ

In the previous section, we have derived the conditions under which a sensor network is guaranteed to converge to the consensus status. The next important question is *how long does it take to converge?* In this section, we are going to seek an answer to this question.

For discrete-time implementations, the convergence time is measured by the number of iterations n needed to reach consensus. Generally, it is impossible to obtain the exact value of this number in design stage, as it depends on both the initial states and the input signals. It is nonetheless possible to derive upper bounds, which would provide guidance for us to choose the optimal design.

Recall that we have obtained the analytic expression of the second derivatives, as given by

$$\ddot{\mathbf{x}}_\delta(n) = \mathbf{L}(\mathbf{I} - \delta\mathbf{L})^{n-2}(\mathbf{L}\mathbf{x}(0) - \mathbf{y}). \tag{23}$$

Given a tolerance threshold ε, our goal here is to derive an upper bound $M(\varepsilon)$ such that

$$\|\ddot{\mathbf{x}}_\delta(n)\|_\infty \le \varepsilon, \quad \forall n > M(\varepsilon). \tag{24}$$

However, directly working with $\ddot{\mathbf{x}}_\delta$ is difficult in general. Our basic idea is to first seek an upper bound $U(n;\delta)$ of $\|\ddot{\mathbf{x}}_\delta(n)\|$ and then find $M(\varepsilon)$ such that

$$U(n;\delta) \le \varepsilon, \quad \forall n > M(\varepsilon). \tag{25}$$

Note here that we introduce δ as an argument of U for the purpose of emphasizing the fact that such an upper bound is often closely related to the updating weight δ. Since $\|\ddot{\mathbf{x}}_\delta(n)\| \le U(n;\delta)$, the condition given by (25) implies the one given by (24).

3.1. Upper Bound Based on Spectral Analysis. Our approach to obtain an upper bound is based on the concept of induced norm. Given a matrix $\mathbf{A}$, its *induced norm*, denoted by $\|\mathbf{A}\|$, is defined as follows:

$$\|\mathbf{A}\| \triangleq \sup_{\mathbf{x}:\|\mathbf{x}\|\le 1} \|\mathbf{A}\mathbf{x}\|. \tag{26}$$

From this definition, we can immediately see that

$$\|\mathbf{A}\mathbf{x}\| \le \|\mathbf{A}\| \cdot \|\mathbf{x}\|. \tag{27}$$

The induced norm is related to the spectral radius. In general, we have $\|\mathbf{A}\| \ge \rho(\mathbf{A})$. Particularly, when $\mathbf{A}$ is symmetric, the equality holds, that is, $\|\mathbf{A}\| = \rho(\mathbf{A})$.

For induced norm, we also have the following result, which is important for our derivation.

Lemma 5. *Given a real square matrix $\mathbf{A}$ and $\alpha_0, \alpha_1, \dots, \alpha_K \in \mathbb{R}$, one has*

$$\left\|\alpha_0\mathbf{I} + \sum_{k=1}^{K}\alpha_k\mathbf{A}^k\right\| \le \alpha_0 + \sum_{k=1}^{K}|\alpha_k|\,\|\mathbf{A}\|^k. \tag{28}$$

This is a well-known result of linear algebra, which states that the induced norm of the matrix obtained by applying a polynomial to a matrix $\mathbf{A}$ is upper bounded by the corresponding polynomial of $\|\mathbf{A}\|$ with the coefficients replaces by their absolute values.

Coming back to our problem and considering the norm $\|\ddot{\mathbf{x}}_\delta(n)\|$, we have the following result.

Theorem 6. *Let $\mathbf{x}$ be the states of a dynamical system on a connected network whose Laplacian matrix $\mathbf{L}$ is diagonalizable. Then, one has*

$$\|\ddot{\mathbf{x}}_\delta(n)\| \le \left\|\mathbf{I} - \delta\mathbf{L} - \mathbf{1}\boldsymbol{\gamma}^T\right\|^{n-2} \|\mathbf{L}\| \cdot \|\mathbf{h}\|. \tag{29}$$

Here, $\boldsymbol{\gamma}$ is a left eigenvector of $\mathbf{L}$ associated with the eigenvalue 1 that satisfies $\boldsymbol{\gamma}^T\mathbf{1} = 1$, and $\mathbf{h} = \mathbf{L}\mathbf{x}(0) - \mathbf{y}$.

Proof of Theorem 6. We first rewrite $\ddot{\mathbf{x}}_\delta(n)$ as follows:

$$\begin{aligned}\ddot{\mathbf{x}}_\delta(n) &= \mathbf{L}(\mathbf{I}-\delta\mathbf{L})^{n-2}\mathbf{h}\\ &= \mathbf{L}\left(\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T+\mathbf{1}\boldsymbol{\gamma}^T\right)^{n-2}\mathbf{h}\\ &= \mathbf{L}\left(\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right)^{n-2}\mathbf{h}.\end{aligned}\tag{30}$$

Here, we make use of the fact that $\mathbf{L}\mathbf{1}\boldsymbol{\gamma}^T = \mathbf{0}\boldsymbol{\gamma}^T = \mathbf{1}$. With this expression, (29) in this theorem immediately follows from Lemma 5 and (27). The proof is completed. □

From this theorem, it is not difficult to show that when

$$n \geq \frac{\log\|\mathbf{L}\| + \log\|\mathbf{h}\| + \log(1/\varepsilon)}{\log(1/\|\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\|)} + 2. \tag{31}$$

The right-hand side of this inequality gives an upper bound of the number of iterations needed to attain consensus.

3.2. Optimal Choice of δ. To minimize the upper bound given by (31), one can choose the optimal updating weight $\hat{\delta}$ that minimizes $\|\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\|$, as

$$\hat{\delta} = \operatorname*{argmin}_{\delta\geq 0}\left\|\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right\|. \tag{32}$$

3.2.1. Symmetric Cases. We first consider the cases where the underlying communication network is symmetric. In these cases, the Laplacian matrix $\mathbf{L}$ is symmetric. Let $0 = \lambda_1 \leq \cdots \leq \lambda_N$ be the eigenvalues of $\mathbf{L}$, which are all nonnegative real values. Since the component $\mathbf{1}\boldsymbol{\gamma}^T$ that corresponds to the eigenvalue 1 of $\boldsymbol{\Phi}$ is subtracted, the eigenvalues of $\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T$ are $0, 1-\delta\lambda_2, \ldots, 1-\delta\lambda_N$. Since $\mathbf{L}$ is symmetric, we have

$$\begin{aligned}\left\|\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right\| &= \rho\left(\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right)\\ &= \max\{|1-\delta\lambda_2|, |1-\delta\lambda_N|\}.\end{aligned}\tag{33}$$

Hence, the optimal choice of δ is given by

$$\begin{aligned}\hat{\delta} &= \operatorname*{argmin}_{\delta>0}\max\{|1-\delta\lambda_2|, |1-\delta\lambda_N|\}\\ &= \frac{2}{\lambda_2+\lambda_N}.\end{aligned}\tag{34}$$

Note that $\hat{\delta} \leq 2/\lambda_N$. According to Corollary 2, it follows, that with such an optimal choice, convergence is guaranteed. We also note that Xiao and Boyd [1] obtained similar results in different context.

3.2.2. Generic Diagonalizable Cases. In real applications the communication topology is not always symmetric due to difference of devices and deployment environment. Hence, the spectral radius does not necessarily equal the induced norm, we use the spectral radius in place of the induced norm; leading to the following objective in choosing optimal δ:

$$\hat{\delta} = \operatorname*{argmin}_{\delta\geq 0}\rho\left(\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right). \tag{35}$$

In generic cases, the spectral radius is given by

$$\rho\left(\mathbf{I}-\delta\mathbf{L}-\mathbf{1}\boldsymbol{\gamma}^T\right) = \max\{|1-\delta\lambda_i| : i = 2,\ldots,N\}. \tag{36}$$

The eigenvalues here can be complex numbers. When solving this problem, it is important to contain the solution within the domain such that convergence is guaranteed. The domain is defined by

$$|1-\delta\lambda_i| \leq 1, \quad \forall i = 2,\ldots,N. \tag{37}$$

Suppose that $\lambda_i = \alpha_i + i\beta_i$; we have that $|1-\delta\lambda_i| < 1$ if and only if

$$0 \leq \delta \leq \frac{2\alpha_i}{|\lambda_i|^2}. \tag{38}$$

Therefore, the valid range of δ is

$$0 \leq \delta \leq \min_{i=2,\ldots,N}\frac{2\alpha_i}{|\lambda_i|^2}. \tag{39}$$

We note that there exists no analytic solution to this problem. However, the objective function is a piece-wise quadratic function, which can be readily solved by first delimiting the pieces and then comparing their respective minimum.

The derivation above provides an upper bound of the convergence time and a guideline for choosing the optimal δ. We can see that they are closely related to the spectral characteristics of the Laplacian matrix $\mathbf{L}$. In general, the network converges rapidly when the eigenvalues of $\mathbf{I}-\delta\mathbf{L}$ (except the one that equals 1) are small.

3.3. More Accurate Analysis of the Convergence Time. As mentioned above, the actual convergence time also depends on the initial states $\mathbf{x}(0)$, and the input signals $\mathbf{y}$, as well as the tolerance threshold ε. Here, we take a close examination of such dependency.

Suppose that $\mathbf{L}$ is diagonalizable with eigenvalues $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N$. Let $\mathbf{e}_j$ be the eigenvector associated with the eigenvalue λ_i for each $i = 1,\ldots,N$. Since $\mathbf{L}$ is diagonalizable, these eigenvectors span the entire state space, and as a consequence, the vector $\mathbf{h} = \mathbf{L}\mathbf{x}(0) - \mathbf{y}$ can be expressed as a linear combination of them, as

$$\mathbf{h} = \sum_{i=1}^{N} c_i\mathbf{e}_i. \tag{40}$$

Let $\mathcal{R}(\mathbf{L}) = \{\mathbf{L}\mathbf{v} : \mathbf{v} \in \mathbb{R}^N\}$ be the range space of $\mathbf{L}$. Since the underlying network is connected, the null space of $\mathbf{L}$ is one dimensional. Hence, $\mathcal{R}(\mathbf{L})$ has dimension $N-1$ and is spanned by $\mathbf{e}_2,\ldots,\mathbf{e}_N$. Clearly, the second derivatives $\ddot{\mathbf{x}}_\delta(n) = \mathbf{L}(\mathbf{I}-\delta\mathbf{L})^{n-2}\mathbf{h} \in \mathcal{R}(\mathbf{L})$ and thus can be expressed to be a linear combination of $\mathbf{e}_2,\ldots,\mathbf{e}_N$, as

$$\ddot{\mathbf{x}}_\delta(n) = \sum_{i=2}^{N} c_i\lambda_i(1-\delta\lambda_i)^{n-2}\mathbf{e}_i. \tag{41}$$

We can see that the ith term decreases geometrically with shrinking ratio $|1-\delta\lambda_i|$. Asymptotically, the term with greatest shrinking ratio, namely, the one that attenuates most slowly, will dominate in a long run.

Let $C : \mathcal{R}(\mathbf{L}) \to \mathcal{R}(\mathbf{L})$ be a coordinate transform that maps each vector in $\mathcal{R}(\mathbf{L})$ to the coefficients with respect to the basis $\{\mathbf{e}_2, \ldots, \mathbf{e}_N\}$. This is a bijective linear map. For each vector $\mathbf{v} \in \mathcal{R}(\mathbf{L})$, we define

$$\|\mathbf{v}\|_{\infty/C} = \|C(\mathbf{v})\|_\infty. \tag{42}$$

It can be easily shown that $\|\cdot\|_{\infty/C}$ is a norm on $\mathcal{R}(\mathbf{L})$. To characterize the relations between $\|\cdot\|_{\infty/C}$ and $\|\cdot\|_\infty$, we define

$$\kappa_C \triangleq \sup_{\mathbf{v}\in\mathcal{R}(\mathbf{L})} \frac{\|\mathbf{v}\|_\infty}{\|\mathbf{v}\|_{\infty/C}}. \tag{43}$$

In other words, κ_C is the greatest real value that satisfies

$$\|\mathbf{v}\|_\infty \le \kappa_C \|\mathbf{v}\|_{\infty/C}, \quad \forall \mathbf{v} \in \mathcal{R}(\mathbf{L}). \tag{44}$$

By the equivalence of norms on a finite dimensional vector space, we have $0 < \kappa_C < +\infty$; that is, κ_C is a finite positive value. Consider the second derivatives $\ddot{\mathbf{x}}_\delta(n)$. To have $\|\ddot{\mathbf{x}}_\delta(n)\|_\infty < \varepsilon$, it suffices to have

$$\|\ddot{\mathbf{x}}_\delta(n)\|_{\infty/C} < \frac{\varepsilon}{\kappa_C}. \tag{45}$$

From the definition of $\|\cdot\|_{\infty/C}$ and (41), we get

$$\|\ddot{\mathbf{x}}_\delta(n)\|_{\infty/C} = \max_{i=2,\ldots,N} \left| c_i\lambda_i (1-\delta\lambda_i)^{n-2} \right|. \tag{46}$$

As a result, when

$$n > \max_{i=2,\ldots,N} \frac{\log(\kappa_C |c_i\lambda_i|/\varepsilon)}{\log(1/|1-\delta\lambda_i|)} + 2, \tag{47}$$

$\|\ddot{\mathbf{x}}_\delta(n)\| < \varepsilon$, implying that the sensor network attains the consensus within the tolerable range. Based on this bound, the optimal choice δ is given by

$$\hat{\delta} = \operatorname*{argmin}_{\delta>0} \max_{i=2,\ldots,N} \frac{\log(\kappa_C |c_i\lambda_i|/\varepsilon)}{\log(1/|1-\delta\lambda_i|)}. \tag{48}$$

This result indicates that with the knowledge of input signals, one can derive an improved upper bound, and a proper choice for δ, which relates not only to the underlying graph, but also to the initial states, input signals, and the tolerance threshold ε.

However, directly applying this result in practice is infeasible, as the input signals are unavailable in advance; otherwise, we do not even have to deploy a sensor network to measure them. The true significance of this result lies is that it suggests that we can incorporate domain-specific knowledge to make a better design of the sensor network.

Specifically, while the exact input is unknown before we really perform the measurement, we generally have some rough knowledge about what the input might be, which could be formulated as a prior distribution of the input signals and then be exploited to guide the choice of δ. From such a prior model, one can derive the prior distribution of the coefficients $\mathbf{c} = [c_2, \ldots, c_N]^T$, denoted by $p(\mathbf{c})$. Then, we obtain an upper bound of the expected convergence time, as

$$U_T(\delta; p) = \int \max_{i=2,\ldots,N} \frac{\log(\kappa_C |c_i\lambda_i|/\varepsilon)}{\log(1/|1-\delta\lambda_i|)} p(\mathbf{c}) d\mathbf{c}. \tag{49}$$

With this model, one can choose an optimal δ that minimizes $U_T(\delta; p)$. Though it might be difficult to directly solve this problem, one can resort to various techniques to make further approximation.

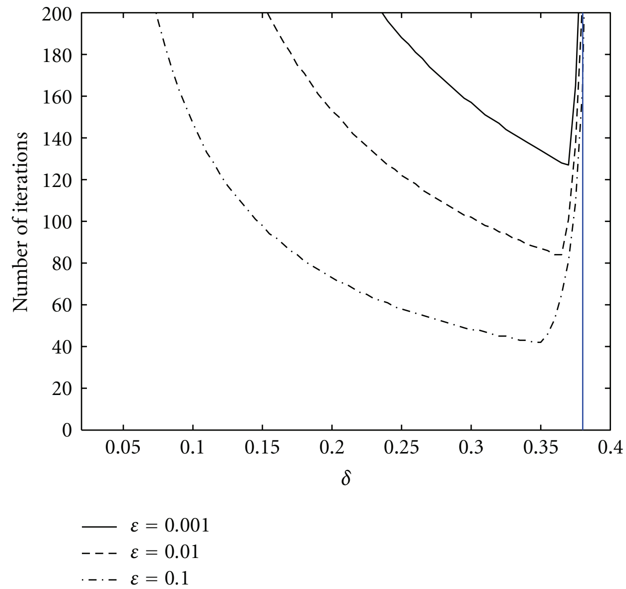

Figure 3: Minimal number of iterations for a $N = 12$ symmetric network with different consensus criteria ε's.

4. Simulation

In this section, we will illustrate the following key results. First, the consensus rate and the choice of optimal iteration weight δ are dependent on the self-aware criterion ε for a sensor network where communication links are symmetric and asymmetric. Second, the analysis of determining the optimal δ provided in previous section is consistent with the simulation result with various ε's, which greatly helps designing efficient and energy-saving sensor network for different topologies.

4.1. Symmetric Topology. We first consider the symmetric network shown in Figure 1, which has 12 nodes and 24 links. This network was randomly generated with arbitrarily chosen initial state value $\mathbf{x}(\mathbf{0})$ and input $\mathbf{y}$ as follows:

$$\begin{gathered} \mathbf{x}(\mathbf{0}) = [0, 10, 20, \ldots, 110], \\ \mathbf{y} = [30, 30, 30, 30, 70, 70, 70, 70, 50, 50, 50, 50]. \end{gathered} \tag{50}$$

The coupling coefficient a_{ij} is defined to equal 1 if there is signal transmitted from node j to node i; otherwise, $a_{ij} = 0$. For the network shown in Figure 1, Figure 3 shows through

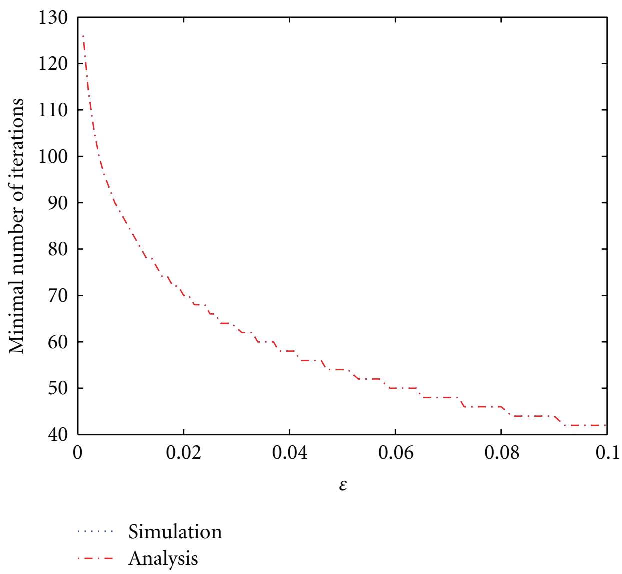

Figure 4: Minimal number of iterations versus ε for a symmetric topology.

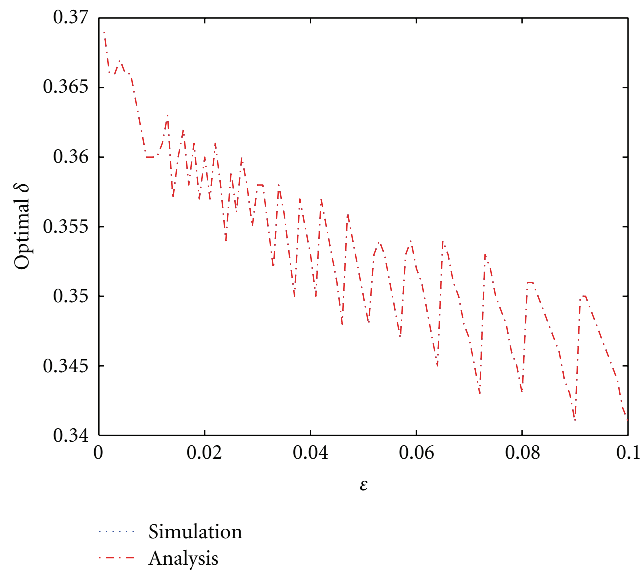

Figure 5: Optimal iteration weight δ versus ε for a symmetric topology.

simulation the number of iterations required to achieve consensus as a function of the weight parameter (δ—x-axis) and the self-aware criteria (ε—various curves). For this network, the best iteration weight found from earlier work by (34) is $\hat{\delta} = 0.3802$ (illustrated in Figure 3 by the vertical line). However, we can see from Figure 3 that the three curves with $\varepsilon = 0.001$, $\varepsilon = 0.01$, and $\varepsilon = 0.1$ reach the lowest point at about 0.37, 0.36, and 0.35, which means setting the iteration weight δ to each of those values would get convergence by fewest iterations based on each network operator ε. Consensus with these values is achieved in significantly fewer iterations, and, most importantly, convergence is determined by local criterion. As ε increases, the whole curve moves to the left slightly. Though the change in the selection of δ is comparatively small with different threshold ε, we can find that the minimal iterations for reaching consensus drop largely, from 124 to 85, then fall off to 40. If one uses $\hat{\delta} = 0.3802$ as determined by (34) and the criteria presented in [1], the system would require much more than 200 iterations to achieve consensus.

By solving (41) as a function of criterion ε, we found the relation between ε and minimal number of iterations and the relation between ε and optimal δ. As shown in Figures 4 and 5, the analytical results are very consistent with the simulation. we can see that the minimal number of iterations decreases as the threshold ε increases. When ε goes from 0.001 to 0.1, the minimal number of iterations keeps dropping from 125 to 40, which shows that achieving consensus with slightly different ε would make a difference in the number of communication iterations. There is a tradeoff between the consensus precision and the number of communications, or energy consuming. Realizing that, it is desirable to choose a larger criterion ε to make sensor nodes live longer, as long as it meets the specific application requirement of the sensor network.

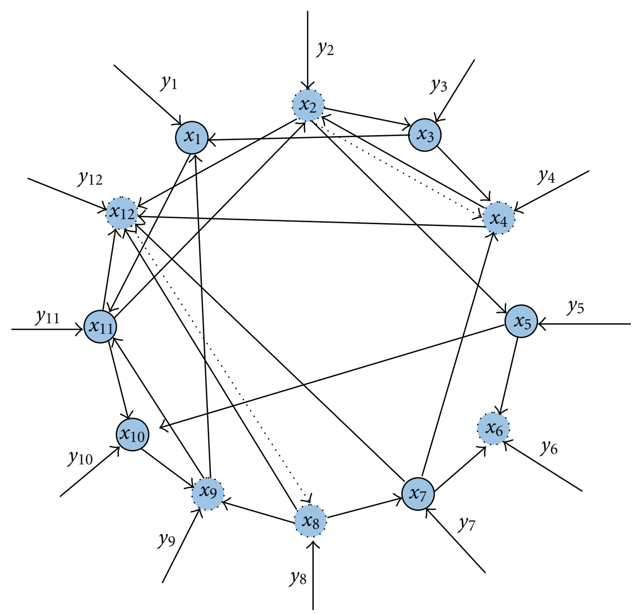

Figure 6: Schematic diagram of a asymmetric network, $N = 12$; degree = 2.

4.2. Asymmetric Topology. Next, we consider another 12-node network shown in Figure 6 with an asymmetric communication topology. Using the same initial value and input as the above example, we show that the optimal iteration weight δ as well as the minimal number of iterations varies with different network operators ε's in Figure 7. Similar to the symmetric example, it can be seen that as ε increases, the corresponding δ would be reduced slightly, while the minimal number of iterations would decrease in a big way. For this asymmetric network, the theoretical best weight found from (35) is 0.3466, which is also denoted in Figure 7 by the vertical line. Again, this theoretical "best weight" is obviously

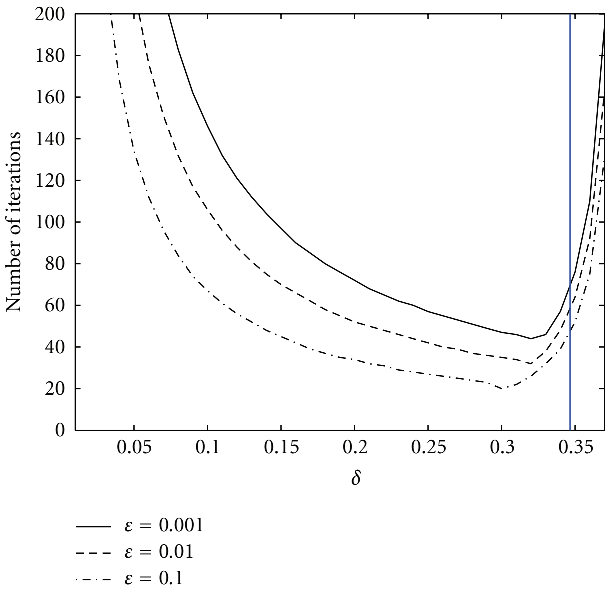

Figure 7: Minimal number of iterations for a $N = 12$ asymmetric network with different consensus criteria ε's.

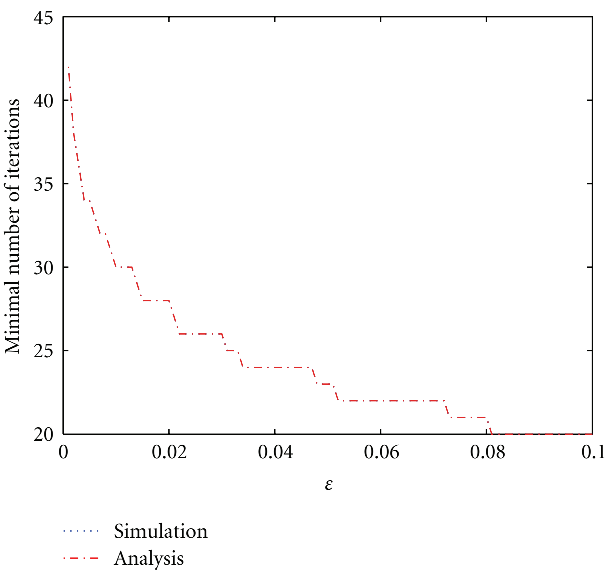

Figure 8: Minimal number of iterations versus ε for an asymmetric topology.

different with the real practical best choice of δ; applying the "best weight" to the practical systems would yield many more iterations for network to reach consensus. As illustrated by Figures 8 and 9, the performances of minimal number of iterations and optimal δ with ε for asymmetric topology present similar behaviors as those of symmetric topology.

5. Conclusion

In this work, the problem of achieving consensus among distributed sensor nodes has been considered. In particular, a potential answer to the question on how do individual

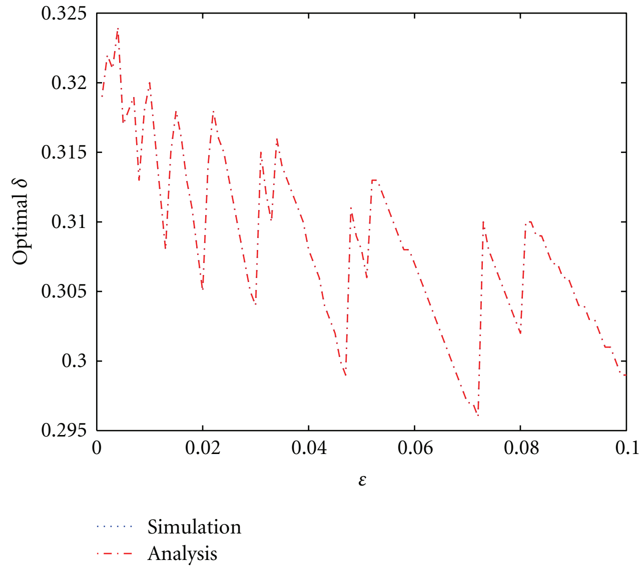

Figure 9: Optimal iteration weight δ versus ε for an asymmetric topology.

nodes "know" when consensus is achieved is presented. The work introduces a local "self-aware" criterion, ε, which in practice is consensus rate of the parameter under consideration. The work illustrates (1) that ε, as expected, will influence the number of iterations required but more importantly (2) that existing criteria for the weighting functions is not appropriate for that will lead to an increased number of energy-consuming iterations. The work also considers the impact of nonsymmetric links on achieving consensus. These results address practical considerations regarding implementing distributed consensus building in "real" sensor networks. As such, it is hoped that this work will motivate others to consider implementation issues in this domain.

References

[1] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, vol. 5, pp. 4997–5002, December 2003.

[2] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.

[3] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[4] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[5] S. Barbarossa and G. Scutari, "Bio-inspired sensor network design," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 26–35, 2007.

[6] S. Barbarossa and G. Scutari, "Decentralized maximum-likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3456–3470, 2007.

[7] G. Scutari, S. Barbarossa, and L. Pescosolido, “Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels,” *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1667–1684, 2008.

[8] L. Chen, G. Carpenter, S. Greenberg, J. Frolik, and X. S. Wang, “An implementation of decentralized consensus building in sensor networks,” *IEEE Sensors Journal*, vol. 11, no. 3, pp. 667–675, 2011.