Performance Analysis of Consensus-Based Distributed System Under False Data Injection Attacks

Xiaoyan Zheng[1], Lei Xie[1,2], Huifang Chen[1,3(✉)], and Chao Song[1]

[1] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
{21631151,xiel,chenhf,songchao31097}@zju.edu.cn
[2] Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Hangzhou 310027, China
[3] Zhoushan Ocean Research Center, Zhoushan 316021, China

Abstract. This paper investigates the security problem of consensus-based distributed system under false data injection attacks (FDIAs). Since the injected false data will spread to the whole network through data exchange between neighbor nodes, and result in continuing effect on the system performance, it is significant to study the impact of the attack. In this paper, we consider two attack models according to the property of the injection data, the deterministic attack and the stochastic attack. Then, the necessary and sufficient condition for the convergence of distributed system under the attack are derived, and the attack feature making the system unable to converge is provided. Moreover, the convergence result under resource-limited attack is deviated. On the other hand, the statistical properties of the convergence performance under zero-mean and non-zero-mean stochastic attacks are analyzed, respectively. Simulation results illustrate the effects caused by FDIAs on the convergence performance of distributed system.

Keywords: Consensus-based distributed system · False data injection attack (FDIA) · Performance analysis · Convergence

1 Introduction

Recently, the consensus-based distributed system has been received wide attention with the widespread use of wireless networks [1]. For a distributed system, each node is treated equally, and the load of each communication link is almost balanced. The distributed structure can reduce communication load, computation burden and energy consumption compared to the traditional centralized structure [2, 3]. Furthermore, the distributed structure is flexible for dynamical network topologies [4, 5].

This work was partly supported by National Natural Science Foundation of China (No. 61671410, No. 61471318) and Zhejiang Provincial Natural Science Foundation of China (No. LGG18F010005, No. 2018R52046).

H. Gao et al. (Eds.): ChinaCom 2019, LNICST 312, pp. 483–497, 2020.
https://doi.org/10.1007/978-3-030-41114-5_36

However, due to the open characteristic of the distributed system, it is vulnerable to malicious attacks [6]. The false data injection attack (FDIA) is one of typical attacks for the distributed system, which injects false data into the unprotected procedure during information exchange process, intending to degrade the performance and threat the security of the distributed system.

In recent years, researchers have been beginning to pay attention to the impact of random or artificial injected false data on the system performance. In [7], the necessary and sufficient condition to guarantee the convergence of distributed system under bounded noise is proved, and the closed expression of the relationship between the noise bound and the consensus accuracy is derived. In [8], the expression of weighted least-squared error of nodes when the distributed system reaches the steady state under noise is deduced. In [9], the author further explored the influence of noise on performance of distributed system under different topologies on the basis of [8], and indicated that the error caused by noise is related to the depth of the graph. In [10] and [11], authors analyzed the effects of zero-mean random noise and non-zero-mean random noise on the convergence performance of the broadcast-based consensus algorithm, and derived the upper and lower performance bounds under noise interference. In [12], the stability of the distributed detection system under the interference of two kinds of energy limited signals is discussed for the resilient consensus problem.

In this paper, we study the security problem of the consensus-based distributed system and analyze the impact of FDIA on the system performance. Two types of FDIAs, the deterministic attack and the stochastic attack, are considered in the fusion phase of distributed system. The effect of the deterministic attack and the stochastic attack on the convergence performance of distributed system is analyzed in detail. For the distributed system under different FDIAs, some interesting theoretical results are derived. Finally, the theoretical results are verified by simulations.

The rest of the paper is organized as follows. Section 2 describes the system model. Section 3 analyzes the impact of the FDIA on the performance of distributed system. Section 4 gives simulation results and discussions. Finally, the paper is concluded in Sect. 5.

Notations: Boldfaced capital and lower-case letters represent matrices and vectors, respectively. $\mathbf{I}_n$ denotes the identity matrix of dimension $n \times n$. $\mathbf{1}_n$ denotes the n-column vector with all elements to be 1. $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. Besides, $\det(\cdot)$ denotes the determinant operation. $\mathrm{diag}(\cdot)$ denotes the creation operation of the diagonal matrix with supplied elements.

2 System Model

2.1 Network Model

Considering a distributed network consisting of N sensor nodes, where the topology structure is connected, and communication links are steady. The network is described by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of nodes in the network, and $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$, $\mathcal{E}$ represents the set of edges in the network, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. If $(v_i, v_j) \in \mathcal{E}$, node i and node j are adjacent and they can exchange information with each other. Hence, the neighboring set of node i is defined as $\mathcal{N}_i = \{v_j \in \mathcal{V} \cap (v_i, v_j) \in \mathcal{E}\}$.

The characteristics of a network can be expressed using a set of matrices. Adjacent matrix $\mathbf{A}$ denotes the neighborhood relationship. If node i and node j are adjacent, the corresponding entry in $\mathbf{A}$ is $a_{ij} = 1$; otherwise, $a_{ij} = 0$. Degree matrix $\mathbf{D}$ is a diagonal matrix whose diagonal element d_i is the sum of neighbors of node i, $d_i = |\mathcal{N}_i|$ and $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_N)$. Laplacian matrix $\mathbf{L}$ is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. That is, the element is $l_{ij} = d_{ij}$ if $i = j$; otherwise, $l_{ij} = -a_{ij}$.

2.2 Attack Model

It is assumed that there exists malicious attacker(s) in the distributed network. In order to degrade the performance of the distributed system, an attacker captures part of unprotected sensor node(s) and makes them be malicious node(s). Malicious sensor nodes will launch the FDIA to degrade the system performance.

The operation of the distributed system over the network includes three phases: local measurement, distributed fusion and distributed inference (detection, classification and estimation). In this work, we focus on the FDIA in the distributed fusion phase.

As the FDIA is launched by malicious node(s), the false data are added to the state update step in the distributed system. Let $\mathcal{V}_\text{a}$ be the set of malicious sensor nodes, and $\mathcal{V}_a \subseteq \mathcal{V}$. Let $\mathcal{V}_\text{s}$ be the set of normal sensor nodes, and $\mathcal{V}_s = \mathcal{V} \backslash \mathcal{V}_a$.

In the fusion phase, each sensor node exchanges information with its one-hop neighboring nodes. The local state of the target will be updated until the whole network reaches a consensus if there is no attack. When malicious nodes launch the attack, the state update process can be denoted as

$$x_i(k+1) = x_i(k) + \frac{\varepsilon}{w_i} \sum_{j \in \mathcal{N}_i} \left[x_j(k) - x_i(k) \right] + u_i(k), \tag{1}$$

where $x_i(k)$ is the state of the target at sensor node i at the kth iteration, ε is the iteration step, and w_i is the weight coefficient at sensor node i. In order to ensure the convergence of the algorithm, the iteration step needs to satisfy $0 < \varepsilon < \min(w_i/d_i)$. The injection data $u_i(k) = 0$ if node i is normal; otherwise, $u_i(k) \neq 0$ if node i is malicious.

For convenience, the update process in (1) can also be represented in the form of matrix as:

$$\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k) + \mathbf{u}(k), \tag{2}$$

where $\mathbf{x}(k)$ is the state vector, and $\mathbf{x}(k) = [\text{x}_1(k), \text{x}_2(k), \ldots, x_N(k)]^\text{T}$, $\mathbf{W}$ is the weight matrix, and $\mathbf{W} = \mathbf{I}_N - \varepsilon \,\text{diag}(1/w_1, 1/w_2, \ldots, 1/w_N)\mathbf{L}$.

Let the weight coefficient vector be $\mathbf{w}^\text{T} = [w_1, w_2, \ldots, w_N]^\text{T}$. Thus, $\mathbf{W} = \mathbf{I}_N - \varepsilon \,\text{diag}(\mathbf{1}_N/\mathbf{w}^\text{T})\mathbf{L}$. We assume that the weight matrix $\mathbf{W}$ is invariant.

FDIA can be divided into two categories based on the injected data of the attacker, the deterministic attack and the stochastic attack. For the deterministic attack, the injected data of the attacker is a deterministic variable. And in the stochastic attack, the injected data of the attacker is a random variable.

3 Performance Analysis of Distributed System Under FDIAs

In this section, we analyze the effect of different FDIAs on the performance of the distributed system.

3.1 Performance Metrics

We introduce the definition of the convergence and consensus of the distributed system.

Definition 1: For the distributed system $\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k)$ with any initial state $\mathbf{x}(0)$, $\exists \mathbf{x}^* \in \mathbb{R}^N$, if $\lim_{k\to\infty} \mathbf{x}(k) = \mathbf{x}^*$, the system is convergent.

Definition 2: For the distributed system $\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k)$ with any initial state $\mathbf{x}(0)$, $\exists x^* \in \mathbb{R}$, if $\lim_{k\to\infty} \mathbf{x}(k) = x^*\mathbf{1}_N$, i.e., $\lim_{k\to\infty} x_i(k) = x^*$, the system is consensus.

Lemma 1: For the consensus-based distributed system $\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k)$, the convergence is equivalent to the consensus.

Proof: 1. If the distributed system is consensus, i.e., when $k \to \infty$, $\mathbf{x}(k) \to x^*\mathbf{1}_N$, where $x^*\mathbf{1}_N$ is a constant vector. Hence, the system is convergent according to the Definition 1.

2. If the distributed system is convergent, i.e., when $k \to \infty$, $\mathbf{x}(k) \to \mathbf{x}^*$, where $\mathbf{x}^*$ is a constant vector. Since $\lim_{k\to\infty} \mathbf{x}(k+1) = \lim_{k\to\infty} \mathbf{W}\mathbf{x}(k) = \mathbf{W}\mathbf{x}^* = \mathbf{x}^*$, and according to the definition of weight matrix $\mathbf{W}$, $\mathbf{W}\mathbf{x}^* - \mathbf{x}^* = \varepsilon\,\mathrm{diag}(\mathbf{1}_N/\mathbf{w})\mathbf{L}\mathbf{x}^* = \mathbf{0}$, $\mathbf{L}\mathbf{x}^* = \mathbf{0}$, $(\mathbf{x}^*)^T\mathbf{L}\mathbf{x}^* = \sum_{i=1}^{N}\sum_{j=1}^{N}(x_i^* - x_j^*)^2 = 0$ for any $\mathbf{x}^*$. That is, $x_i^* = x_j^*, \forall i, j \in \mathcal{V}$, i.e., $\mathbf{x}^* = x^*\mathbf{1}_N$. Hence, the system is consensus. ■

3.2 Under Deterministic Attack

The attack may make the distributed system unable to reach consensus or converge to a wrong result.

First, we derivate the necessary and sufficient condition for the convergence of the distributed system.

Theorem 1: The necessary condition for the convergence of the distributed system is that $\lim_{k\to\infty} \mathbf{u}(k) = \mathbf{0}$.

Proof: Assume that the convergent state of nodes is constant x^*, the state vector is $\mathbf{x}^* = x^*\mathbf{1}_N$. When $k \to \infty$, $\lim_{k\to\infty} \mathbf{x}(k+1) = \lim_{k\to\infty} (\mathbf{W}\mathbf{x}(k) + \mathbf{u}(k))$.

Since $\lim_{k\to\infty} \mathbf{x}(k+1) = \lim_{k\to\infty} \mathbf{x}(k) = x^*\mathbf{1}_N$ and the weight matrix is a stochastic matrix, $\mathbf{W}\mathbf{1}_N = \mathbf{1}_N$, $x^*\mathbf{1}_N = \mathbf{W}x^*\mathbf{1}_N + \lim_{k\to\infty} \mathbf{u}(k) = x^*\mathbf{1}_N + \lim_{k\to\infty} \mathbf{u}(k)$, $\lim_{k\to\infty} \mathbf{u}(k) = \mathbf{0}$. ■

Theorem 2: The sufficient condition for the convergence of the distributed consensus system is that $\lim_{k\to\infty}\sum_{\tau=0}^{k}||\mathbf{u}(\tau)|| = C$, where $C \in \mathbb{R}$.

Proof: According to Perron-Frobenius Theorem, $\lim_{k\to\infty}\mathbf{W}^k = \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}$, where $\boldsymbol{\pi}^{\mathrm{T}} = (\mathbf{w}^{\mathrm{T}}\mathbf{1}_N)^{-1}\mathbf{w}^{\mathrm{T}}$. When $k \to \infty$,

$$\begin{aligned}\lim_{k\to\infty}\mathbf{x}(k+1) &= \lim_{k\to\infty}\mathbf{W}^{k+1}\mathbf{x}(0) + \lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau)\\ &= \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\mathbf{x}(0) + \lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau). \end{aligned} \tag{3}$$

Using the properties of norms, we have

$$\begin{aligned}\lim_{k\to\infty}\|\mathbf{x}(k+1)\| &= \left\|\mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\mathbf{x}(0) + \lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau)\right\|\\ &\le \left\|\mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\mathbf{x}(0)\right\| + \lim_{k\to\infty}\sum_{\tau=0}^{k}\left\|\mathbf{W}^{k-\tau}\mathbf{u}(\tau)\right\|. \end{aligned} \tag{4}$$

Since $\lim_{k\to\infty}\sum_{\tau=0}^{k}||\mathbf{u}(\tau)|| = C$, $\lim_{k\to\infty}\sum_{\tau=0}^{k}\boldsymbol{u}(\boldsymbol{\tau}) = \boldsymbol{c}$, where $\boldsymbol{c} \in \mathbb{R}^N$, and $\lim_{k\to\infty}\mathbf{u}(k) = \mathbf{0}$. According to Cauchy criterion, $\forall\varepsilon > 0, \exists M > 0$, $||\mathbf{u}(\tau)|| < \varepsilon$ when $\tau > M$; $\forall\upsilon > 0, \exists N > 0$, $||\mathbf{W}^{\tau} - \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}|| < \upsilon$, when $\tau > N$. Therefore, $||\mathbf{W}^{\tau}|| < \tau + ||\mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}||$. Since $k - N >> M$ when $k \to \infty$, $||\mathbf{u}(k-N)||, \ldots, ||\mathbf{u}(k)|| < \varepsilon$. Hence,

$$\begin{aligned}\lim_{k\to\infty}\left\|\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau)\right\| &\le \lim_{k\to\infty}\Big(\left\|\mathbf{W}^{k}\right\|\|\mathbf{u}(0)\| + \ldots + \left\|\mathbf{W}^{N+1}\right\|\|\mathbf{u}(k-N-1)\|\\ &\quad + \left\|\mathbf{W}^{N}\right\|\|\mathbf{u}(k-N)\| + \ldots + \|\mathbf{u}(k)\|\Big)\\ &< \left(\upsilon + \left\|\mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\right\|\right)C + \varepsilon\sum_{\tau=0}^{N}\|\mathbf{W}\|^{\tau}\end{aligned}$$

Moreover, norm and spectral radius of matrix satisfy that $||\mathbf{W}|| \ge \rho(\mathbf{W}) = 1$. When $||\mathbf{W}|| = 1, \varepsilon\sum_{\tau=0}^{N}||\mathbf{W}^{\tau}|| < \varepsilon N$. When $||\mathbf{W}|| > 1, \varepsilon\sum_{\tau=0}^{N}||\mathbf{W}^{\tau}|| = \varepsilon(1-||\mathbf{W}||^{N+1})(1-||\mathbf{W}||)^{-1}$. Hence, Eq. (4) is convergent.

Therefore, when the attack vector satisfies $\lim_{k\to\infty}\sum_{\tau=0}^{k}||\mathbf{u}(k)|| = C$, the norm of state vector is convergent, i.e., the states of all nodes converge. ■

In Theorems 1 and 2, the necessary condition and sufficient condition to convergent for the distributed system under FDIA are proved, respectively. Therefore, the attack strategy making the network unable converge should have

$$\lim_{k\to\infty}\sum\nolimits_{\tau=0}^{k}\|\mathbf{u}(\tau)\| = \infty. \tag{5}$$

That is, when the series of vectors are not converge, the system cannot reach consensus although the iteration step increases.

From (2), we find that the impact of FDIA on the performance of distributed system is continuous. That is, once the attacker launches the attack (although the attack subsequently stops), the injected false data will spread to the surrounding nodes during nodes exchange information. The false data injected by the attacker will remain in the convergence result. Hence, it is necessary to analyze the deviation between the wrong convergence result and the normal convergence result under the FDIA.

Theorem 3: When the attack vector series converge, $\lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{u}(\tau) = \boldsymbol{c}$ and $\boldsymbol{c} \in \mathbb{R}^N$, the convergence result of nodes is

$$\lim_{k\to\infty}\mathbf{x}(k) = \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}(\mathbf{x}(0)+\mathbf{c}). \tag{6}$$

Proof: According to (2), the convergence result of state vector under FDIA is

$$\lim_{k\to\infty}\mathbf{x}(k+1) = \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\mathbf{x}(0) + \lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau).$$

As $\boldsymbol{\pi}^{\mathrm{T}}\mathbf{W} = \boldsymbol{\pi}^{\mathrm{T}}$, $\boldsymbol{\pi}^{\mathrm{T}}\mathbf{W}^{\mathrm{k}} = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{W}^{\mathrm{k-1}} = \ldots = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{W} = \boldsymbol{\pi}^{\mathrm{T}}$. And then,

$$\begin{aligned}\boldsymbol{\pi}^{\mathrm{T}}\lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau) &= \lim_{k\to\infty}\boldsymbol{\pi}^{\mathrm{T}}\mathbf{W}^{k}\mathbf{u}(0)+\ldots+\boldsymbol{\pi}^{\mathrm{T}}\mathbf{W}\mathbf{u}(k-1)+\boldsymbol{\pi}^{\mathrm{T}}\mathbf{u}(k)\\ &= \lim_{k\to\infty}\boldsymbol{\pi}^{\mathrm{T}}[\mathbf{u}(0)+\ldots+\mathbf{u}(k-1)+\mathbf{u}(k)] = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{c}\end{aligned}$$

As $\mathbf{b} = \lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau)$, $\boldsymbol{\pi}^{\mathrm{T}}\mathbf{b} = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{c}$. When $\lim_{k\to\infty}\sum_{\tau=0}^{k}\mathbf{u}(\tau) = \mathbf{c}$, the states of nodes are convergent. Thus, $b_i = b_j, \forall i, j \in \mathcal{V}$. If $\mathbf{b} = b\mathbf{1}_N$, $b\boldsymbol{\pi}^{\mathrm{T}}\mathbf{1}_N = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{c}$, *i.e.*, $b = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{c}$. Therefore, we have

$$\lim_{k\to\infty}\sum\nolimits_{\tau=0}^{k}\mathbf{W}^{k-\tau}\mathbf{u}(\tau) = \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}\mathbf{c}. \tag{7}$$

From (7), one finds that the series of injected attack vector will directly affect the convergence result of the distributed system. ■

Lemma 2: If and only if the weighted average of attack vector series converges to zero, $\lim_{k\to\infty}\boldsymbol{\pi}^{\mathrm{T}}\sum_{\tau=0}^{k}\mathbf{u}(\tau) = 0$, the convergence result will be the weighted average of initial states of nodes.

3.3 Under Stochastic Attack

When the injection attack vector is random, we consider the case of Gaussian random variable. That is, the malicious node i injects false data $u_i(k)$ at step k, $u_i(k) \sim \mathcal{N}(\mu_i, \sigma 2i)$, and the attack vector satisfies $\mathbf{u}(k) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

To facilitate the subsequent derivation, we introduce the nature of normalized weighted average matrix, $\mathbf{Q} = \mathbf{1}_N\boldsymbol{\pi}^{\mathrm{T}}$, and its relationship with the weight matrix $\mathbf{W}$.

Property 1: The normalized weighted average matrix $\mathbf{Q}$ has the following properties:

(1) $\mathbf{Q} = \lim_{k\to\infty} \mathbf{W}^k$;
(2) $\mathbf{QW} = \mathbf{Q}$;
(3) $\mathbf{WQ} = \mathbf{Q}$;
(4) $\mathbf{Q}^2 = \mathbf{Q}$;
(5) $(\mathbf{W} - \mathbf{Q})^k = \mathbf{W}^k - \mathbf{Q}$;
(6) $(\mathbf{W} - \mathbf{Q})^k(\mathbf{I}_N - \mathbf{Q}) = \mathbf{W}^k - \mathbf{Q}$;
(7) $\rho(\mathbf{W} - \mathbf{Q}) < 1$.

Since the injected attack vector is random, the state vector under its influence is also random. In the view of the statistical characteristics, we analyze the impact of the stochastic attack on the performance of distributed system.

The divergence vector of current nodes' state is defined as

$$\boldsymbol{\delta}(k) = \mathbf{x}(k) - \mathbf{Q}\mathbf{x}(k), \tag{8}$$

which can be used to measure the convergence performance of distributed system. The system state converges as the divergence is zero. The larger the divergence vector, the more obvious the divergence of distributed system.

Lemma 3: The divergence vector $\boldsymbol{\delta}(k)$ satisfies the recursive relationship as

$$\boldsymbol{\delta}(k+1) = (\mathbf{W} - \mathbf{Q})\boldsymbol{\delta}(k) + (\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k). \tag{9}$$

Proof: According to the definition of $\boldsymbol{\delta}(k)$ and the properties of matrix $\mathbf{Q}$, we have

$$\begin{aligned}\boldsymbol{\delta}(k+1) &= \mathbf{x}(k+1) - \mathbf{Q}\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k) + \mathbf{u}(k) - \mathbf{QW}\mathbf{x}(k) - \mathbf{Q}\mathbf{u}(k)\\ &= (\mathbf{W} - \mathbf{Q})\mathbf{x}(k) + (\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k).\end{aligned}$$

While $(\mathbf{W} - \mathbf{Q})\mathbf{Q}\mathbf{x}(k) = \left(\mathbf{WQ} - \mathbf{Q}^2\right)\mathbf{x}(k) = 0$,

$$\begin{aligned}\boldsymbol{\delta}(k+1) &= (\mathbf{W} - \mathbf{Q})\mathbf{x}(k) - (\mathbf{W} - \mathbf{Q})\mathbf{Q}\mathbf{x}(k) + (\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k)\\ &= (\mathbf{W} - \mathbf{Q})(\mathbf{x}(k) - \mathbf{Q}\mathbf{x}(k)) + (\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k)\\ &= (\mathbf{W} - \mathbf{Q})\boldsymbol{\delta}(k) + (\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k).\end{aligned}$$

■

First, we analyze the mean of divergence vector, $\mathrm{E}[\boldsymbol{\delta}(k)]$ or $\bar{\boldsymbol{\delta}}(k)$.
Using the recursive relationship in (9), we have

$$\boldsymbol{\delta}(k+1) = (\mathbf{W} - \mathbf{Q})^{k+1}\boldsymbol{\delta}(0) + \sum_{\tau=0}^{k} (\mathbf{W} - \mathbf{Q})^{\tau}(\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k-\tau). \tag{10}$$

Since $\boldsymbol{\delta}(0) = \mathbf{x}(0) - \mathbf{Q}\mathbf{x}(0) = (\mathbf{I}_N - \mathbf{Q})\mathbf{x}(0)$,

$$\boldsymbol{\delta}(k+1) = (\mathbf{W} - \mathbf{Q})^{k+1}(\mathbf{I}_N - \mathbf{Q})\mathbf{x}(0) + \sum_{\tau=0}^{k} (\mathbf{W} - \mathbf{Q})^{\tau}(\mathbf{I}_N - \mathbf{Q})\mathbf{u}(k-\tau). \tag{11}$$

Since the mean of attack vector is $\mathrm{E}[\mathbf{u}(k)] = \bar{\mathbf{u}}$, $k \in \mathrm{N}$, taking expectation and limitation on both sides of (11), we get

$$\lim_{k\to\infty} \mathrm{E}[\boldsymbol{\delta}(k+1)] = \lim_{k\to\infty} \sum_{\tau=0}^{k} (\mathbf{W} - \mathbf{Q})^{\tau} (\mathbf{I}_N - \mathbf{Q})\bar{\mathbf{u}}. \tag{12}$$

As $\rho(\mathbf{W}-\mathbf{Q}) < 1$, the series is convergent and $\lim_{k\to\infty} \sum_{\tau=0}^{k} (\mathbf{W}-\mathbf{Q})^{\tau} = [\mathbf{I}_N - (\mathbf{W}-\mathbf{Q})]^{-1}$. Hence, the mean of divergence vector is convergent, and the convergence result is

$$\lim_{k\to\infty} \mathrm{E}[\boldsymbol{\delta}(k)] = [\mathbf{I}_N - (\mathbf{W} - \mathbf{Q})]^{-1} (\mathbf{I}_N - \mathbf{Q})\bar{\mathbf{u}}. \tag{13}$$

From (2), we find that, as the number of iterations increases, the mean of the divergence will converge, and the convergence result is related to the mean of attack vector. If the mean of attack vector is non-zero, the mean of divergence is also non-zero, which means that the state of nodes cannot converge.

Second, we analyze the covariance of divergence, $\boldsymbol{\Phi}(k)$.

Suppose that $\boldsymbol{\delta}(k)$ is independent of $\mathbf{u}(k)$ and the covariance of $\mathbf{u}(k)$ is $\boldsymbol{\Sigma}$, the covariance can be further simplified as

$$\boldsymbol{\Phi}(k+1) = (\mathbf{W} - \mathbf{Q})\boldsymbol{\Phi}(k)(\mathbf{W} - \mathbf{Q})^{\mathrm{T}} + (\mathbf{I}_N - \mathbf{Q})\boldsymbol{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}. \tag{14}$$

Then,

$$\begin{aligned}\boldsymbol{\Phi}(k+1) = {} & (\mathbf{W} - \mathbf{Q})^{k+1}\boldsymbol{\Phi}(0)[(\mathbf{W} - \mathbf{Q})^{\mathrm{T}}]^{k+1} \\ & + \sum_{\tau=0}^{k} (\mathbf{W} - \mathbf{Q})^{\tau} (\mathbf{I}_N - \mathbf{Q})\boldsymbol{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}[(\mathbf{W} - \mathbf{Q})^{\mathrm{T}}]^{\tau}.\end{aligned} \tag{15}$$

Since $\boldsymbol{\delta}(0) = \mathbf{x}(0) - \mathbf{Q}\mathbf{x}(0) = \bar{\boldsymbol{\delta}}(0)$ and $\boldsymbol{\Phi}(0) = 0$,

$$\lim_{k\to\infty} \boldsymbol{\Phi}(k+1) = \lim_{k\to\infty} \sum_{\tau=0}^{k} (\mathbf{W} - \mathbf{Q})^{\tau} (\mathbf{I}_N - \mathbf{Q})\boldsymbol{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}[(\mathbf{W} - \mathbf{Q})^{\mathrm{T}}]^{\tau}. \tag{16}$$

As $\rho(\mathbf{W}-\mathbf{Q}) < 1$, the series is convergent, $\lim_{k\to\infty} \sum_{\tau=0}^{k} (\mathbf{W}-\mathbf{Q})^{\tau}$ and $\lim_{k\to\infty} \sum_{\tau=0}^{k} \left[(\mathbf{W}-\mathbf{Q})^{\mathrm{T}}\right]^{\tau}$ are convergent. Hence, the covariance of divergence vector is also convergent. Moreover,

$$\lim_{k\to\infty} (\mathbf{W} - \mathbf{Q})\boldsymbol{\Phi}(k+1)(\mathbf{W} - \mathbf{Q})^{\mathrm{T}} = \lim_{k\to\infty} \boldsymbol{\Phi}(k+1) - (\mathbf{I}_N - \mathbf{Q})\boldsymbol{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}.$$

Suppose that the convergence result is $\mathbf{X} = \lim_{k\to\infty} \boldsymbol{\Phi}(k)$, $\mathbf{F} = \mathbf{W} - \mathbf{Q}$, and $\mathbf{G} = (\mathbf{I}_N - \mathbf{Q})\boldsymbol{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}$. That is,

$$\mathbf{X} - \mathbf{F}\mathbf{X}\mathbf{F}^{\mathrm{T}} = \mathbf{G}. \tag{17}$$

The formula of (17) is a discrete Lyapunov equation. To resolve the equation, the Kronecker product, vectorization function and Matricization function will be used.

Theorem 4: For $m \times n$ dimension matrix $\mathbf{A}$, $n \times p$ dimension matrix $\mathbf{B}$, $p \times q$ dimension matrix $\mathbf{B}$,

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^{\mathrm{T}} \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}).$$

For convenience, we vectorize both sides of matrix equation $\mathbf{X} - \mathbf{FXF}^{\mathrm{T}} = \mathbf{G}$. According to Theorem 4, we have $(\mathbf{I}_{N^2} - \mathbf{F} \otimes \mathbf{F})\mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\mathbf{G})$. Since $\det(\mathbf{I}_{N^2} - \mathbf{F} \otimes \mathbf{F}) \neq 0$, $\mathrm{vec}(\mathbf{X}) = (\mathbf{I}_{N^2} - \mathbf{F} \otimes \mathbf{F})^{-1}\mathrm{vec}(\mathbf{G})$. Using the matricization function, we get the convergence result of covariance matrix as

$$\lim_{k \to \infty} \mathbf{\Phi}(k) = \mathrm{unvec}_{N,N}\{[\mathbf{I}_{N^2} - (\mathbf{W} - \mathbf{Q}) \otimes (\mathbf{W} - \mathbf{Q})]^{-1} * \mathrm{vec}[(\mathbf{I}_N - \mathbf{Q})\mathbf{\Sigma}(\mathbf{I}_N - \mathbf{Q})^{\mathrm{T}}]\}. \tag{18}$$

According to the steady-state expression of the covariance matrix of divergence vector, we find that when the covariance matrix of attack vector is zero, the covariance of divergence is also zero. When the covariance matrix of attack vector is non-zero, the covariance of divergence is also non-zero. It means that the deviation will fluctuate around its mean. The larger the covariance of attack vector, the more dramatic the fluctuation.

Furthermore, to analyze the impact of stochastic attack on the system convergence results, the deviation vector is defined as $\boldsymbol{\varphi}(k) = \mathbf{x}(k) - \mathbf{Q}\mathbf{x}(0)$, which represents the deviation from the state of the kth step and the weighted average result of the initial state.

Lemma 4: The deviation vector $\boldsymbol{\varphi}(k)$ satisfies the recursive relation as

$$\boldsymbol{\varphi}(k+1) = \boldsymbol{W}\boldsymbol{\varphi}(k) + \mathbf{u}(k). \tag{19}$$

Proof: According to the definition of $\boldsymbol{\varphi}(k)$ and the property of matrix $\mathbf{Q}$, we have

$$\begin{aligned} \boldsymbol{\varphi}(k+1) &= \mathbf{x}(k+1) - \mathbf{Q}\mathbf{x}(0) = \mathbf{W}\mathbf{x}(k) + \mathbf{u}(k) - \mathbf{W}\mathbf{Q}\mathbf{x}(0) \\ &= \mathbf{W}[\mathbf{x}(k) - \mathbf{Q}\mathbf{x}(0)] + \mathbf{u}(k) = \boldsymbol{W}\boldsymbol{\varphi}(k) + \mathbf{u}(k). \end{aligned}$$

■

First, we calculate the mean of the deviation vector, $\mathrm{E}[\boldsymbol{\varphi}(k)]$ or $\overline{\boldsymbol{\varphi}}(k)$.
Using the recursive relation,

$$\boldsymbol{\varphi}(k+1) = \mathbf{W}^{k+1}\boldsymbol{\varphi}(0) + \sum_{\tau=0}^{k} \mathbf{W}^{\tau}\mathbf{u}(k-\tau). \tag{20}$$

Since $\boldsymbol{\varphi}(0) = \mathbf{x}(0) - \mathbf{Q}\mathbf{x}(0)$,

$$\boldsymbol{\varphi}(k+1) = (\mathbf{W}^{k+1} - \mathbf{Q})\mathbf{x}(0) + \sum_{\tau=0}^{k} \mathbf{W}^{\tau}\mathbf{u}(k-\tau). \tag{21}$$

Calculating the expectation both sides of (21), we have

$$\lim_{k \to \infty} \mathrm{E}[\boldsymbol{\varphi}(k+1)] = \lim_{k \to \infty} \sum_{\tau=0}^{k} \mathbf{W}^{\tau}\bar{\mathbf{u}}. \tag{22}$$

As $\rho(\mathbf{W}) = 1$, the matrix series $\sum_{\tau=0}^{k} \mathbf{W}^{\tau}$ is not convergent. Thus, only when the mean of injected attack vector is zero, the mean of deviation is convergent to zero. Moreover, the state of nodes converges to the convergence result without attacks; otherwise, the state of nodes cannot converge.

Second, we calculate the covariance of deviation vector, $\mathbf{\Psi}(k)$.

Suppose that $\boldsymbol{\varphi}(k)$ is independent to $\mathbf{u}(k)$, and the covariance of $\mathbf{u}(k)$ is $\mathbf{\Sigma}$,

$$\mathbf{\Psi}(k+1) = \mathbf{W}\mathbf{\Psi}(k)\mathbf{W}^{\mathrm{T}} + \mathbf{\Sigma}. \tag{23}$$

According to recursive relation, we have

$$\mathbf{\Psi}(k+1) = \mathbf{W}^{k+1}\mathbf{\Psi}(0)(\mathbf{W}^{\mathrm{T}})^{k+1} + \sum_{\tau=0}^{k} \mathbf{W}^{\tau}\mathbf{\Sigma}(\mathbf{W}^{\mathrm{T}})^{\tau}. \tag{24}$$

Since $\boldsymbol{\varphi}(0) = \mathbf{x}(0) - \mathbf{Q}\mathbf{x}(0) = \overline{\boldsymbol{\varphi}}(0)$ and $\mathbf{\Psi}(0) = 0$,

$$\lim_{k\to\infty} \mathbf{\Psi}(k+1) = \lim_{k\to\infty} \sum_{\tau=0}^{k} \mathbf{W}^{\tau}\mathbf{\Sigma}(\mathbf{W}^{\mathrm{T}})^{\tau}. \tag{25}$$

As $\rho(\mathbf{W}) = 1$, the matrix series $\sum_{\tau=0}^{k} \mathbf{W}^{\tau}\mathbf{\Sigma}\left(\mathbf{W}^{\mathrm{T}}\right)^{\tau}$ is not convergent When $\mathbf{\Sigma} \neq \mathbf{0}$. Thus, the range of deviation between the node state and the true convergence result is always changing, and the range of deviation is increasing. The series converges to zero when $\mathbf{\Sigma} = \mathbf{0}$, which means that the deviation between the node state and the true convergence result is fixed.

4 Simulation Results and Discussions

The impact of FDIA on the system convergence are simulated. Suppose that there are 10 nodes in the network, among which there are several malicious nodes.

Considering a network with 10 nodes, as shown in Fig. 1, among which several nodes are malicious. It is assumed that the communication channel is error-free. The initial state of nodes is randomly set as $\mathbf{x}(0) = [18.13, 19.90, 24.34, 18.50, 23.72, 12.65, 7.33, 21.04, 21.32, 13.96]^{\mathrm{T}}$. The weight coefficient of all nodes is 0.1. Thus, the theoretical convergent value is the weighted average of initial state is 18.09.

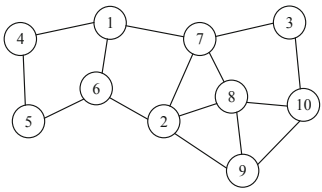
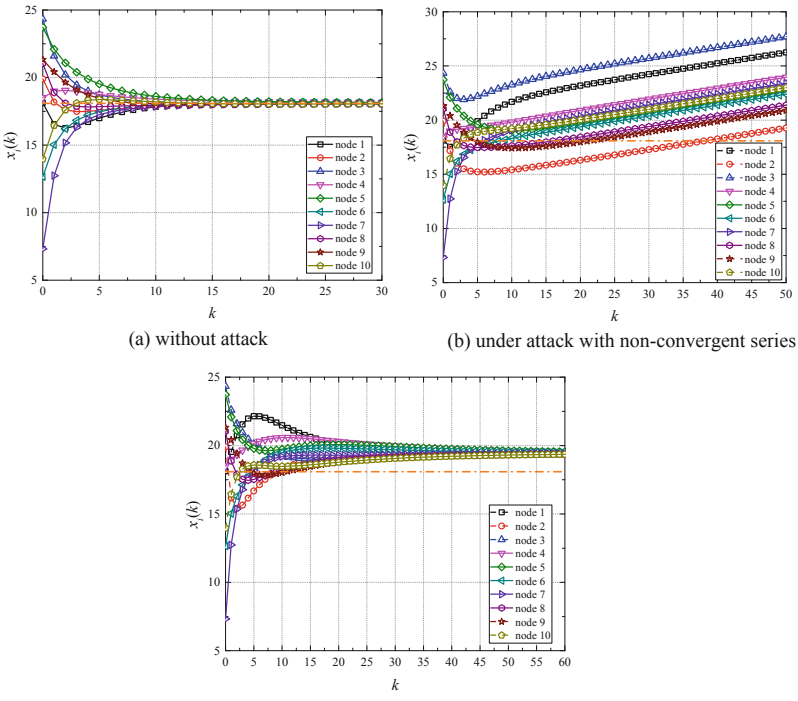


Fig. 1. Network model.

4.1 Under Deterministic Attack

Figure 2 shows the convergence performance of distributed system under deterministic attack. In the figures, the solid line denotes the state of each node, the dash line indicates the state of malicious node, and the orange horizontal dotted line indicates the theoretical convergence result.



(a) without attack

(b) under attack with non-convergent series

(c) under attack with convergent series

Fig. 2. The convergence performance of distributed system under deterministic attack.

Figure 2(a) shows the convergence performance of distributed system without attack. From Fig. 2(a), we observe that the state of nodes gradually converges as the number of iterations increases. After iterating 18 times, the state of each node converges to the desired weighted average of 18.09.

Figure 2(b) shows the convergence performance of distributed system under attack, where the attack series is not convergent, nodes 1, 2 and 3 are malicious, and the injected false vector is $\mathbf{m}(k) = \left[3 \times 0.8^k, -2 \times 0.5^k, 0.6^k\right]^{\mathrm{T}}$, $k \in \mathrm{N}$. From Fig. 2(b), we observe that when the distributed system encounters the attack with non-convergent series, the system state cannot converge although the number of iterations increases. Hence, Fig. 2(b) verifies that the attack with non-convergent series will make the system state unable to converge.

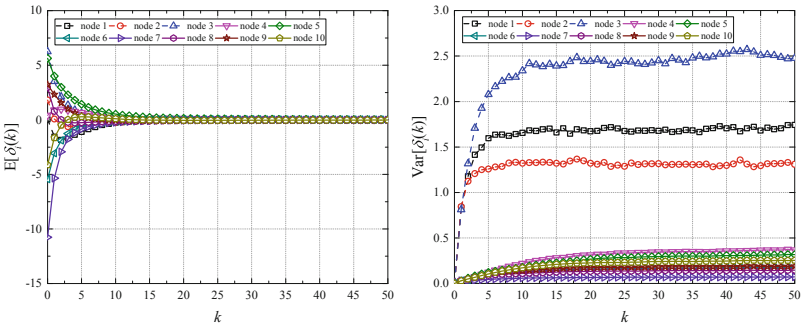
Figure 2(c) shows the convergence performance of distributed system under attack, where the attack series is convergent, nodes 1, 2 and 3 are malicious, and the injected

false vector is $\mathbf{m}(k) = [1, -1, 1]^{\mathrm{T}}, k \in N$. From Fig. 2(c), we observe that when the distributed system encounters the attack with convergent series, the system state will reach convergence as the number of iterations increases. The convergence result is about 19.44, which is different from the expected convergence result without attack, 18.09. That is, the convergent attack will cause the system state to converge to a wrong result, and the convergence result deviation equals to the weighted average of attack vector series.

4.2 Under Stochastic Attack

In the following simulations, we assume that the covariance of injected attack vector is an identity matrix.

Divergence can be used to measure whether the system converges or not. When the divergence is zero, the system state converges.



(a) mean (b) variance

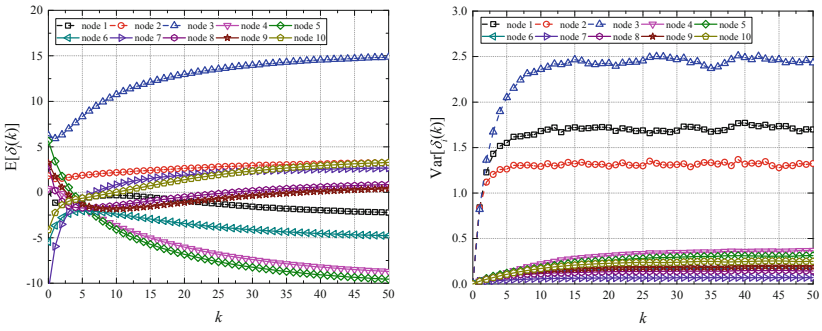
Fig. 3. The divergence performance of distributed system under zero-mean stochastic attack.

Figure 3 shows the divergence performance of distributed system under zero-mean stochastic attack, where Fig. 3(a) and (b) gives the mean and variance of divergence, respectively. From Fig. 3(a), we observe that the divergence tends to zero as the number of iterations increases, which means that the node state converges. Moreover, from Fig. 3(b), we observe that the variance of divergence first increases, and then stabilizes as the number of iterations increases. This indicates that the fluctuation of divergence at each node gradually stabilizes as the number of iterations increases.

Figure 4 shows the divergence performance of distributed system under non-zero-mean stochastic attack, where Fig. 4(a) and (b) gives the mean and variance of divergence, respectively. From Fig. 4(a) and (b), we observe that as the number of iterations increases, the mean of divergence at each node converges to a non-zero value, and the variance of divergence also converges. Moreover, the mean of divergence at each node converges different value, which means the node state cannot converge. According to the theoretical results, the divergence convergent value is $[-2.57, 3.43, 15.32, -9.43, -10.28, -5.14, 2.85, 1.22, 0.81, 3.78]^{\mathrm{T}}$, which is the same as simulation result.

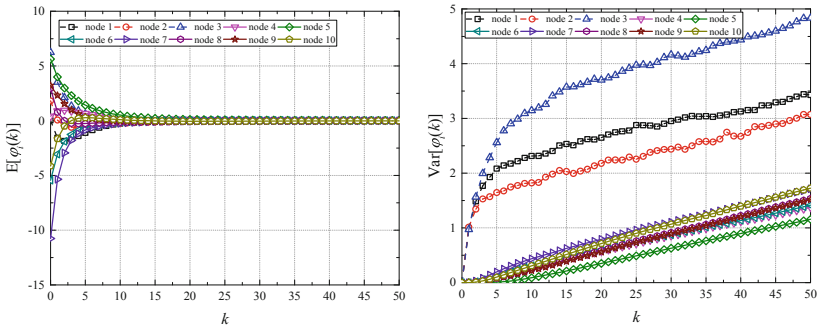
When the deviation equals to zero, the system state converges to the true result. The deviation can be used to measure how convergence result deviates from the true result.

Figure 5 shows the deviation performance of distributed system under non-zero-mean stochastic attack, where Fig. 5(a) and (b) gives the mean and variance of deviation, respectively. From Fig. 5(a), we observe that the mean of deviation tends to zero as the number of iterations increases, which means that the node state converges to the true value. Thus, the theoretical analysis is verified. However, from Fig. 5(b), we observe that the variance of deviation increases as the number of iterations increases. The results in Fig. 5(b) indicate that the internal fluctuation increases gradually although the node state converges to the true result as expected.



(a) mean (b) variance

Fig. 4. The divergence performance of distributed system under non-zero-mean stochastic attack.

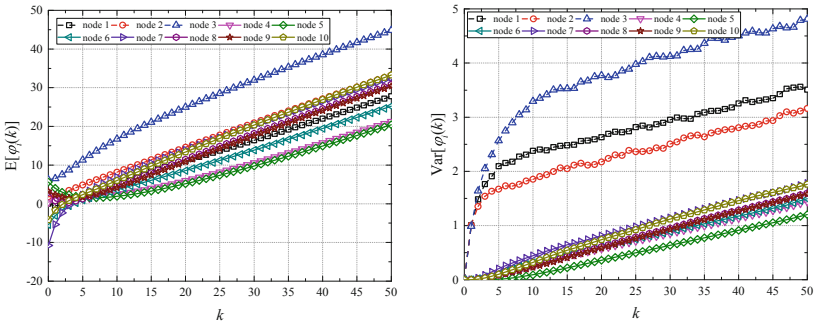


(a) mean (b) variance

Fig. 5. The deviation performance of distributed system under zero-mean stochastic attack.

Figure 6 shows the deviation performance of distributed system under non-zero-mean stochastic attack, where Fig. 6(a) and (b) gives the mean and variance of divergence, respectively. From Figs. 6(a) and (b), we observe that the mean and variance of deviation at each node do not converge as the number of iterations increases. The results in Fig. 6

indicate that the node state gradually deviates from the true result under the interference of non-zero mean stochastic attack. Hence, the theoretical analysis is verified.



(a) mean (b) variance

Fig. 6. The deviation performance of distributed system under non-zero-mean stochastic attack.

5 Conclusions

We studied the problem of FDIA in the consensus-based distributed system in this paper. According to the malicious behavior of the attacker, the attack is classified into two types, deterministic attack and stochastic attack. The impact of the FDIA on the performance of distributed system is emphatically analyzed. Under the deterministic attack, we first addressed the necessary and sufficient condition for the system convergence, and derived the attack characteristics making the network unable to converge. Moreover, the convergence result with limited attack resources is deduced. On the other hand, we analyzed the statistical properties of the convergence performance and convergence result at nodes under zero-mean and non-zero-mean stochastic attacks, respectively. Finally, the effects of deterministic attack and stochastic attack on the performance of distributed system are verified by simulation results. In the future, we will study the defense strategy against FDIA in the consensus-based distributed system.

References

1. Pasqualetti, F., Bicchi, A., Bullo, F.: Consensus computation in unreliable networks: a system theoretic approach. IEEE Trans. Autom. Control **57**(1), 90–104 (2012)
2. Kar, S., Moura, J.M.F.: Consensus + innovations distributed inference over networks: cooperation and sensing in networked systems. IEEE Signal Process. Mag. **30**(3), 99–109 (2013)
3. Zhang, W., Wang, Z., Guo, Y., Liu, H., Chen, Y., Mitola III, J.: Distributed cooperative spectrum sensing based on weighted average consensus. In: Proceedings of IEEE GLOBECOM, Houston, TX, USA, pp. 1–6 (2011)
4. Olfati-Saber, R., Murray, R.M.: Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans. Autom. Control **49**(9), 1520–1533 (2004)

5. Kailkhura, B., Brahma, S., Varshney, P.K.: Data falsification attacks on consensus-based detection systems. IEEE Trans. Signal Inf. Process. Netw. **3**(1), 145–158 (2017)
6. Yan, Q., Li, M., Jiang, T., Lou, W., Hou, Y.T.: Vulnerability and protection for distributed consensus-based spectrum sensing in cognitive radio networks. In: Proceedings of IEEE INFOCOM, Orlando, FL, USA, pp. 900–908 (2012)
7. He, J., Zhou, M., Cheng, P., Shi, L., Chen, J.: Consensus under bounded noise in discrete network systems: an algorithm with fast convergence and high accuracy. IEEE Trans Cybern. **46**(12), 2874–2884 (2016)
8. Jadbabaie, A., Olshevsky, A.: On performance of consensus protocols subject to noise: role of hitting times and network structure. In: Proceedings of 2016 IEEE CDC, Las Vegas, NV, pp. 179–184 (2016)
9. Jadbabaie, A., Olshevsky, A.: Scaling laws for consensus protocols subject to noise. IEEE Trans. Autom. Control **64**(4), 1389–1402 (2019)
10. Aysal, T.C., Barner, K.E.: Convergence of consensus models with stochastic disturbances. IEEE Trans. Inf. Theory **56**(8), 4101–4113 (2010)
11. Yang, Y., Blum, R.S.: Broadcast-based consensus with non-zero-mean stochastic perturbations. IEEE Trans. Inf. Theory **59**(6), 3971–3989 (2013)
12. Meng, D., Moore, K.L.: Studies on resilient control through multiagent consensus networks subject to disturbances. IEEE Trans Cybern. **44**(11), 2050–2064 (2014)