IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 9, 2019, accepted May 30, 2019, date of publication June 26, 2019, date of current version July 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2925009

Distributed Data-Selective DLMS Estimation Under Channel Attacks

YI HUA[1,2,3,4], FENG CHEN[1,2,3,4], (Member, IEEE), SHUKAI DUAN[1], (Member, IEEE), AND JIAGUI WU[1,2,3,4], (Member, IEEE)

[1]College of Artificial Intelligence, Southwest University, Chongqing 400715, China
[2]College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
[3]Brain-Inspired Computing and Intelligent Control Key Laboratory, Southwest University, Chongqing 400715, China
[4]Chongqing Collaborative Innovation Center for Brain Science, Chongqing 400715, China

Corresponding author: Feng Chen (fengchen.uestc@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61875168, and in part by the Chongqing Research Program of Basic Research and Frontier Technology under Grant cstc2017jcyjAX0265.

ABSTRACT With the continuous development of big data processing technology, data selection algorithms gradually attract the attention of researchers. In this paper, a distributed data-selection diffusion least mean square (DLMS) algorithm, which can improve the estimation accuracy of traditional data selection algorithms and can also censor data packets that do not bring enough innovation in wireless sensor networks, is proposed to censor the valid data in distributed iterative updates. And the adaption-then-combination strategy of the proposed algorithm is obtained. Meanwhile, in the distributed estimation system, the channel attacks are considered. When the network is under channel attacks, an adaptive credibility weight matrix is designed to improve the robustness of the distributed data-selection DLMS algorithm. We analyze the proposed algorithms in mean and mean-square performance. A series of simulations are carried out to demonstrate the effectiveness of the proposed algorithms. Moreover, it can be examined that through the comparison between the Metropolis weight strategy and the credibility weight strategy, the proposed credibility weight strategy is more robust in the face of channel attacks.

INDEX TERMS Channel attacks, credibility weight, data selection, distributed estimation, wireless sensor network.

I. INTRODUCTION

In recent years, wireless sensor networks (WSNs) have been used more and more widely in many fields, such as battlefield surveillance, smart city transportation, precision agriculture, and environmental monitoring [1]–[4]. The main reason for this is that with the development of technology, the performance of sensors has been greatly improved, such as the reduction of energy consumption and production costs, the improvement of computing and communication capabilities, and so on [5]–[9]. According to the cooperation between sensors, the WSNs can be divided into the centralized estimation system and the distributed estimation system [4], [10], [11]. In the centralized estimation system, better estimation performance can be achieved [12], but the centralized scheme is not scalable in terms of communication bandwidth and has poor robustness in link failure. However, in distributed implementations, each node communicates with a subset of nodes to estimate unknown parameters in a collaborative manner, so it exhibits good scalability and high robustness [10], [13].

In the past few years, many distributed adaptive strategies have been proposed: incremental least mean squares (LMS) [14]–[16], consensus LMS [17]–[19] and diffusion LMS [3], [20], [21]. In [14], four adaptive implementations are studied: distributed spatial LMS, distributed incremental LMS, centralized spatial LMS and centralized incremental LMS, and the performance of four adaptive implementations is compared. In [18], a fully distributed LMS type of algorithm is proposed to deal with online estimation and track (non)stationary signals by using ad hoc WSNs. In [3], the multitask distributed estimation which is different from the previous single tasks is proposed, and an appropriate mean-square error criterion with ℓ_2-regularization, which weighs the correlation between different tasks, is developed to address multitask problems.

In the age of big data, the amount of data has increased dramatically, which has increased the demand for data-selective

The associate editor coordinating the review of this manuscript and approving it for publication was Xuxun Liu.

algorithms. In the data-selective algorithms, any data can be utilized in the estimation as long as these data are not considered as outliers. In [22]–[24], many set-membership algorithms have been proposed. In [23], a set-membership constrained particle filter approach is developed to reduce communication overhead and thus enable distributed particle filter implementation. But the set-membership algorithms can only eliminate outliers. Different from those algorithms, some data-selective adaptive algorithms have been proposed to eliminate outliers and select the innovative data in [25], which do not sacrifice the estimation accuracy of distributed algorithms. However, it can be found that the proposed data-selective LMS algorithm in [25] is a non-cooperative LMS algorithm. It is well known that a distributed implementation can improve the performance of the non-cooperative LMS algorithm, which motivates us to derive the distributed diffusion LMS algorithm with data selection.

In the above articles, distributed algorithms perform parameter estimation in a secure network environment. However, wireless sensors are not always secure [26]. In step of sensing and communication, wireless sensors are easily eavesdropped by various attacks, such as replay attacks [27], false data injection attacks [28], and jamming attacks [29]. Many secure distributed algorithms have been proposed in order to obtain reliable distributed estimates from adversarial environments. In [26], a secure diffusion LMS algorithm is proposed to confront the presence of false data injection attack, and this algorithm consists of two subsystems: a noncooperative LMS and a diffusion LMS. In [30], a Gaussian mixture-model-based detection mechanism is proposed to locate the compromised sensors for obtaining an accurate state estimate. In [31], a flag raising distributed estimator (FRDE) is presented, and the FRDE algorithm proposes a consensus+innovations estimator which allows the attacked agents to obtain accurate parameter estimate and detect the antagonist agents. In these studies, security mechanisms are against specific types of attacks. However, in [32], two-channel stochastic attacks are considered in the network, and the desired distributed H∞ estimators are constructed against two-channel attacks. In [33], a novel secure data transmission scheme using chaotic compressed sensing, which is based on a T-way Bernoulli shift chaotic system, is presented to ensure efficient secure data transmission against additive noise and malicious attacks. In the appendix of [26], a reputation-based diffusion LMS (R-dLMS) algorithm is presented to obtain distributed estimation when the network is subjected to malicious attacks.

In a distributed estimation system, it is obvious that WSNs also have two sensing channels when estimating the unknown parameter: 1) sensors and unknown system receive the input signal and 2) sensors receive the output signal from the unknown system. Similar to [32], two sensing channels in WSNs are vulnerable to undetermined attacks in an adversarial environment. When the distributed estimation system is attacked by two channel attacks mentioned above, the outliers in data packets will be significantly higher. The proposed distributed data selective diffusion LMS (DS-DLMS) algorithm has a good censoring performance in the face of outliers induced by two channel attacks, but the long-term continuous channel attacks can reduce the performance of the DS-DLMS algorithm. Therefore, an adaptive credibility weight in our proposed DS-DLMS algorithm is designed to increase the robustness against two channel attacks.

The proposed algorithms not only can censor outliers, but also can censor data packets that does not bring enough innovation. Consequently, the proposed algorithms can select the innovative and effective data to update iterations. The main contributions of this paper are concluded as follows:

1. In secure environment, the distributed DS-DLMS algorithm is derived, which can improve the performance of the non-cooperative DS-LMS algorithm and can also censor data packets that does not bring enough innovation.

2. Under two channel attacks, the distributed DS-DLMS algorithm with an adaptive credibility weight matrix is designed, which can improve the robustness of the proposed algorithm when the proposed algorithm censors data packets.

3. The performance of the distributed DS-DLMS algorithms in mean and mean-square is analyzed.

The rest of this work is organized as follows. In section II, the system model of the distributed estimation and some preliminaries are briefly introduced, and the models of two channel attacks are proposed. In section III, the distributed DS-DLMS algorithm is derived without being attacked, and then an adaptive credibility weight matrix is designed to improve robustness of DS-DLMS algorithm when the WSNs are subjected to two channel attacks. In section IV, the performance of DS-DLMS algorithms in mean and mean-square is analyzed. Simulation results are presented to verify the validity of the proposed algorithms in section V and some conclusions are summarized in section VI.

II. SYSTEM MODEL AND PRELIMINARIES

In this part, the LMS model used in WSNs is briefly introduced, and then two channel attacks are described in detail.

A. THE NETWORK MODEL

A wireless sensor network is considered, which consists of a set of nodes $\mathcal{N} = \{1, 2, \ldots, N\}$ and a set of links $\mathcal{L} = \{(k, \ell)\}, k, \ell \in \mathcal{N}$, where $(k, \ell), k, \ell \in \mathcal{N}$ is a link between node k and node ℓ if and only if the two nodes can communicate with each other. The node set $\mathcal{N}$ and link set $\mathcal{L}$ form a communication graph $G = \{\mathcal{N}, \mathcal{L}\}$, which owns a preset topology with L links. The nodes, which can be connected with node k, are the neighbors of node k, and $\mathcal{N}_k$ denotes all neighbors of node k (including node k itself). The n_k is the degree of node k, and $n_k = |\mathcal{N}_k|$.

The goal of WSNs is estimating the M-dimensional parameter w° of an unknown system in a distributed estimation manner from the data packets acquired from N sensors. At each time instant i, each node k can obtain a data packet $\{d_k(i), u_{k,i}\}$, where $d_k(i)$ is an output scalar measurement, and $u_{k,i}$ is an input column regression vector. The connection

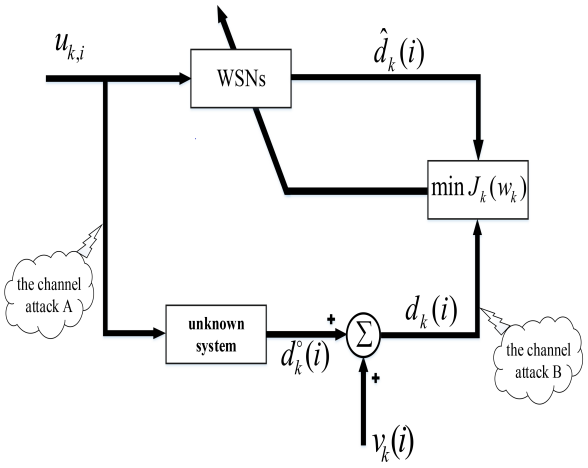


FIGURE 1. The wireless sensor system.

between $\{d_k(i), u_{k,i}\}$ can be expressed by the following linear measurement model.

$$d_k(i) = u_{k,i}^* w^\circ + v_k(i), \tag{1}$$

where the operator $*$ denotes the complex conjugate-transposition, and $v_k(i)$ is a zero mean independent and identically distributed (i.i.d) Gaussian white noise, which is independent of all other signals and owns the variance $\sigma_{v,k}^2$.

At present, many distributed algorithms have been proposed to estimate the unknown parameter w°. These distributed algorithms obtain the corresponding unknown parameter estimation by optimizing the following cost function.

$$J_k(w_k) = \sum_{\ell \in \mathcal{N}_k} \mathbb{E}\left\{\left|d_\ell(i) - u_{\ell,i}^* w_k\right|^2\right\}, \tag{2}$$

where the $\mathbb{E}(\cdot)$ denotes expectation operator. In the following algorithm derivation, the above cost function is also used as the basis for our derivation.

B. TWO CHANNEL ATTACKS

In an adaptive distributed system, there are mainly two communication channels: 1) sensors and unknown system receive the same input signal $u_{k,i}$ (channel A) and 2) sensors receive an output signal $d_k(i)$ from the unknown system (channel B). When an attacker appears on channels A and B, data packets $\{d_k(i), u_{k,i}\}$ can be tampered without being perceived by the sensors. The above two channel attacks can be expressed in the wireless sensor system, as shown in Fig. 1. Therefore two channel attacks are studied in this paper.

When channel A is attacked, the information passed will be tampered [32] as follows:

$$\tilde{u}_{k,i} = \begin{cases} \delta_{k,i}, & \textit{channel attack A} \\ u_{k,i}, & \textit{no channel attack A} \end{cases}, \tag{3}$$

where $\delta_{k,i}$ is an $M \times 1$ column vector with unknown value. For the sake of simplicity, a set of binary-valued variables $\tau_k(i) \in \{0, 1\}$ will be introduced to indicate the status of attack in channel A. In case of considering attacks in channel A, the input signal actually acquired by the unknown system can be expressed as

$$\tilde{u}_{k,i} = (1 - \tau_k(i))\, u_{k,i} + \tau_k(i)\, \delta_{k,i}. \tag{4}$$

$\tau_k(i) = 0$ means that there is no channel attack A, while $\tau_k(i) = 1$ means that channel attack A occurs. Similarly, when channel B is attacked, the received information will also be changed from $d_k(i)$ to $\theta_k(i)$ as follows:

$$\tilde{d}_k(i) = \begin{cases} \theta_k(i), & \textit{channel attack B} \\ d_k(i), & \textit{no channel attack B} \end{cases}, \tag{5}$$

where $\theta_k(i)$ is also an unknown value and is a scalar measurement. Similar to previous processing, another binary-valued variables $\eta_k(i) \in \{0, 1\}$ will be introduced to show the status of attack in channel B. When $\eta_k(i) = 0$, it denotes that no channel attack B occurs. When $\eta_k(i) = 1$, channel attack B appears. In the case of considering attacks in channel B, the output signal actually acquired by node k can be expressed as:

$$\tilde{d}_k(i) = (1 - \eta_k(i))\, d_k(i) + \eta_k(i)\, \theta_k(i). \tag{6}$$

In actual situations, it can't be known in advance whether the data packets have been attacked or not, equations (3) and (5) are used to characterize whether the packet is being attacked. Meanwhile, to facilitate the following analysis, equations (4) and (6) are introduced to denote the actually accepted information whether attacks occur or not. From the above introduction, we can find that when channel attack A occurs, it will eventually affect the scalar measurement signal received by sensors from the unknown system. Therefore, when WSNs are in an adverse environment, the data packets received by sensors are highly likely to be tampered. It indicates that it is necessary to apply a security strategy against two channel attacks in the wireless sensor system.

III. THE DISTRIBUTED DS-DLMS ALGORITHM

In this section, the distributed DS-DLMS algorithm is derived in a security environment, and then the securely distributed DS-DLMS algorithm with credibility weight is proposed under two channel attacks.

A. THE DISTRIBUTED DS-DLMS WITHOUT ATTACKS

For the sake of the following derivation, an error is defined as:

$$e_k(i) = d_k(i) - u_{k,i}^* w_k(i). \tag{7}$$

From [25], we can find that when $|e_k(i)|^2 \leq \varsigma\sigma_{v,k}^2$, where $\varsigma\sigma_{v,k}^2$ is an error squared level, the current data packet $\{d_k(i), u_{k,i}\}$ will be discarded because this data packet is considered not to generate significant new information in this iteration update at instant i. When $|e_k(i)|^2 > \varsigma_{\max}\sigma_{v,k}^2$, the current data packet $\{d_k(i), u_{k,i}\}$ is also discarded since it is considered to bring outlier.

The expression of excess mean square error (EMSE) in data selective adaptive filtering algorithm must consider the frequency of updating coefficients after the transient. In order to solve this problem, we should take into account that the update probability of the adaptive filtering algorithm is related to the frequency in which the selection factor $\gamma_k(i)$ is equal to 1. Since the coefficients mentioned above are all real and all converge to the stable value, this update probability can be defined by the following white Gaussian input signal model:

$$P_{up} = 2Q\left(\frac{\sqrt{\varsigma}}{\sqrt{1+\alpha}}\right) - 2Q\left(\frac{\sqrt{\varsigma_{\max}}}{\sqrt{1+\alpha}}\right), \tag{8}$$

where $Q(\cdot)$ is the complementary Gaussian cumulative distribution function [25] as follows:

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp\left(-t^2/2\right) dt. \tag{9}$$

By equation (8), discarding the fixed value $\varsigma_{\max}{}'s$ effect, probability of updating can be specified by appropriately selecting parameter ς as follows

$$\sqrt{\varsigma} = \sqrt{1+\alpha}\,Q^{-1}\left(\frac{P_{up}}{2}\right), \tag{10}$$

where $Q^{-1}(\cdot)$ is the inverse function of function $Q(\cdot)$.

The above parameters have a great influence on the performance of the data selective algorithm, so it is necessary to set a value range for these parameters. The parameter P_{up} should fall in the range from 0 to 1, and the misadjustment α also falls in the same range of values so that the algorithm accuracy can be accepted. In order to get better algorithm performance, we usually make $\alpha = \frac{P_{up}}{\nu - P_{up}}$ for $\nu > 4$ [25]. By determining the above parameters, the selection factor $\gamma_k(i)$ can be obtained

$$\gamma_k(i) = \begin{cases} 0, & if\ |e_k(i)| \le \sqrt{\varsigma}\sigma_{v,k} \\ 0, & if\ |e_k(i)| > \sqrt{\varsigma_{\max}}\sigma_{v,k} \\ 1, & otherwise \end{cases}. \tag{11}$$

Combining the gradient descent method to minimize the cost function $J_k(w_k)$, we can get the gradient update equation.

$$w_k(i+1) = w_k(i) - \mu_k \nabla_w J_k(w_k(i)), \tag{12}$$

where μ_k is a constant step-size, $\nabla_w J_k(w_k(i))$ denotes the complex gradient of $J_k(w_k(i))$ with respect to w. We can rewrite cost function as

$$J_k(w_k) = \left\{\left|d_k(i) - u_{k,i}^* w_k\right|^2\right\} + \sum_{\ell\in\mathcal{N}_k\setminus\{k\}} s_{\ell,k}\left\|w_k - w_\ell^{loc}\right\|^2, \tag{13}$$

where $s_{\ell,k}$ is a set of non-negative real weight coefficients, $\ell \in \mathcal{N}_k\setminus\{k\}$ denotes neighboring nodes of node k except for itself, and w_ℓ^{loc} is the optimal estimate of node ℓ [10]. We replace the optimal estimate w_ℓ^{loc} with an intermediate estimate $\varphi_\ell(i)$, which can be available at node ℓ. Therefore, the $\nabla_w J_k(w_k(i))$ can be obtained

$$\nabla_w J_k(w_k(i)) \approx u_{k,i}\left(u_{k,i}^* w_k(i) - d_k(i)\right) + 2\sum_{\ell\in\mathcal{N}_k\setminus\{k\}} s_{\ell,k}(w_k(i) - \varphi_\ell(i)). \tag{14}$$

The gradient update equation can be expressed as

$$w_k(i+1) = w_k(i) + \mu_k u_{k,i}\left(d_k(i) - u_{k,i}^* w_k(i)\right) + 2\mu_k\sum_{\ell\in\mathcal{N}_k\setminus\{k\}} s_{\ell,k}(\varphi_\ell(i) - w_k(i)). \tag{15}$$

Combining the selection factor $\gamma_k(i)$, the gradient update equation can be divided into two steps by generating an intermediate estimate $\varphi_k(i)$ as follows:

$$\varphi_k(i) = w_k(i) + \gamma_k(i)\mu_k u_{k,i}\left(d_k(i) - u_{k,i}^* w_k(i)\right), \tag{16}$$

$$w_k(i+1) = \varphi_k(i) + 2\mu_k\sum_{\ell\in\mathcal{N}_k\setminus\{k\}} s_{\ell,k}(\varphi_\ell(i) - w_k(i)). \tag{17}$$

Replacing $w_k(i)$ in (17) with $\varphi_k(i)$, we have

$$w_k(i+1) = \left(1 - 2\left(1 - s_{k,k}\right)\mu_k\right)\varphi_k(i) + \sum_{\ell\in\mathcal{N}_k\setminus\{k\}} 2\mu_k s_{\ell,k}\varphi_\ell(i). \tag{18}$$

We introduce the coefficients

$$\begin{cases} c_{k,k} = & 1 - 2\left(1 - s_{k,k}\right)\mu_k, \\ c_{\ell,k} = & 2\mu_k s_{\ell,k}, \text{for } \ell \in \mathcal{N}_k\setminus\{k\} \end{cases}. \tag{19}$$

We can obtain the adaption-then-combination (ATC) strategy of distributed DS-DLMS algorithm

$$\begin{cases} \varphi_k(i) = w_k(i) \\ +\gamma_k(i)\mu_k u_{k,i}\left(d_k(i) - u_{k,i}^* w_k(i)\right) \\ w_k(i+1) = \sum_{\ell\in\mathcal{N}_k} c_{\ell,k}\varphi_\ell(i) \end{cases}. \tag{20}$$

The coefficients $c_{\ell,k}$ need to satisfy

$$c_{\ell,k} = 0, \quad if\ \ell \notin \mathcal{N}_k\ and\ \mathbb{1}^T C = \mathbb{1}^T, \tag{21}$$

where $\mathbb{1}$ denotes the $M \times 1$ column vector with unit entries, and $(\cdot)^T$ denotes the transpose operator.

Without consideration of attacks, the Metropolis rule are generally used to obtain the matrix C as follows.

$$c_{\ell,k} = \begin{cases} 1/\max(n_k, n_\ell) & if\ \ell \in \mathcal{N}_k\setminus\{k\}, \\ 1 - \sum_{\ell\in\mathcal{N}_k\setminus\{k\}} c_{\ell,k} & if\ \ell = k, \\ 0 & otherwise \end{cases}. \tag{22}$$

Table 1 shows the implementation procedure of the proposed DS-DLMS algorithm with Metropolis weight.

TABLE 1. The distributed DS-DLMS algorithm with Metropolis weight.

Initialization: $w_k(1)$ =zero vector prescribe P_{up}, and choose $\varsigma_{\max}$ compute $\alpha = \frac{P_{up}}{\nu - P_{up}}$ for $\nu > 4$ compute $\sqrt{\varsigma}$ by (10)
At each time instant $i \geq 1$ for each node k, 1. acquire $d_k(i)$ and $u_{k,i}$, 2. use $\{d_k(i), u_{k,i}\}$ to compute $e_k(i)$ by (7), 3. compute selection factor $\gamma_k(i)$ by (11), 4. compute the intermediate estimate $\varphi_k(i)$ by the adaption step in (20), 5. exchange the intermediate estimate $\varphi_k(i)$ with neighbor nodes, 6. use Metropolis weight to compute the intermediate estimate $w_k(i+1)$ by the combination step in (20).

B. THE DISTRIBUTED DS-DLMS WITH CREDIBILITY WEIGHT UNDER TWO CHANNEL ATTACKS

Through the above derivation, we get the distributed DS-DLMS algorithm. Next, we derive the secure distributed DS-DLMS algorithm with credibility weight under two channel attacks.

When the network is subjected to two channel attacks, the number of outliers will increase significantly, which will significantly decrease the performance of the DLMS algorithm. However, we can find that the selection factor can well represent the appearance of outliers. In order to make the DS-DLMS algorithm better against two channel attacks, a credibility window $Z_{k,i}$ is introduced to achieve the weight matrix C', and the length of this credibility window is set to Z.

$$Z_{k,i} = \{\gamma_k(i), \gamma_k(i-1), \ldots, \gamma_k(i-Z+1)\}. \quad (23)$$

If node k is subjected to two channel attacks at time i, data selection factor $\gamma_k(i)$ will be equal to zero. The longer attacks last, the more elements in $Z_{k,i}$ are set to 0 in credibility window, which is a good response to the effect of attacks.

Two sum functions about the credibility window are defined as

$$sum\left(Z_{k,i}\right) \triangleq \sum_{t=0}^{Z-1} \gamma_k(\text{i-t}), \quad (24)$$

$$sum_{k,i} \triangleq \sum_{\ell \in \mathcal{N}_k} sum\left(Z_{\ell,i}\right). \quad (25)$$

The weight matrix C' with credibility window can be designed as:

$$c'_{\ell,k} = \begin{cases} \dfrac{sum\left(Z_{\ell,i}\right) + \zeta}{sum_{k,i} + n_k \times \zeta} & if\ \ell \in \mathcal{N}_k \setminus \{k\}, \\ 1 - \sum\limits_{\ell \in \mathcal{N}_k \setminus \{k\}} c'_{\ell,k} & if\ \ell = k, \\ 0 & otherwise \end{cases}, \quad (26)$$

where ζ is an extremely small positive real number.

This design may cause an extreme situation. That is, when node k and all its neighbors are attacked and the attack time exceeds the length of the credibility window $Z_{k,i}$.

TABLE 2. The distributed DS-DLMS algorithm with credibility weight under two channel attacks.

Initialization: $w_k(1)$ =zero vector prescribe P_{up}, and choose $\varsigma_{\max}$ compute $\alpha = \frac{P_{up}}{\nu - P_{up}}$ for $\nu > 4$ compute $\sqrt{\varsigma}$ by (10)
At each time instant $i \geq 1$ for each node k, 1. acquire $\tilde{d}_k(i)$ and $\tilde{u}_{k,i}$, 2. use the data packet $\{\tilde{d}_k(i), \tilde{u}_{k,i}\}$ to compute $e_k(i)$ by (7), 3. compute selection factor $\gamma_k(i)$ by (11), 4. compute the intermediate estimate $\varphi_k(i)$ by the adaption step in (20), 5. exchange the data packet $\{\varphi_k(i), Z_{k,i}\}$ with neighbor nodes, 6. compute the credibility weight $c'_{\ell,k}$ by (26), 7. use credibility weight to compute the intermediate estimate $w_k(i+1)$ by the combination step in (20).

Therefore, $\underset{\ell \in \mathcal{N}_k}{sum}\left(Z_{\ell,i}\right) = 0$ and $sum_{k,i} = 0$. When this situation happens, the weight coefficients of the corresponding node k and its neighbor nodes in the weight matrix C' become $\frac{1}{n_k}$ because of the introduction of ζ. However, this situation generally will not occur, because the WSNs may be considered to be close to paralyzing or to be already paralyzed when all sensors of a certain area are attacked.

Since we design the value of the credibility window into the fusion weight matrix, when some data packets are attacked, the fusion weights of the corresponding nodes will be decreased. This strategy allows data packets that generate new information to bring more useful information to the iteration, enabling the DS-DLMS algorithm to achieve more accurate performance. Table 2 shows the implementation procedure of the proposed DS-DLMS algorithm with credibility weight under two channel attacks.

IV. THE ANALYSIS OF PERFORMANCE

In this section, the performance of the distributed DS-DLMS algorithm is studied in detail. Regardless of whether the network is attacked or not, the difference between the above two proposed algorithms is that the fusion weighting matrices are different. However, the processes of performance analysis in mean and mean-square are same, so the analyses are performed uniformly using C. Meanwhile, for the convenience of analyses, the data packet $\{\tilde{d}_k(i), \tilde{u}_{k,i}\}$, which is similar to the introduction in section II-B, is uniformly used in the following analyses regardless of whether the network is under two channel attacks.

There are two estimate-errors defined by

$$\Delta\varphi_{k,i} \triangleq \varphi_k(i) - w^{\circ}, \quad (27)$$

$$\Delta w_{k,i} \triangleq w_k(i) - w^{\circ}. \quad (28)$$

For the convenience of analysis, we explain several notations. The operator $\otimes$ denotes the Kronecker product operator. The operator col$\{\cdot\}$ denotes a column vector. The operator $diag\{\cdot\}$ is (block) diagonal matrix. The operator Tr $(\cdot)$ denotes the trace of a matrix. I_M denotes the $M \times M$ unit matrix.

Furthermore, we define the following global variables:

$$\Delta w_i = \left[\left(\Delta w_{1,i}\right)^T, \left(\Delta w_{2,i}\right)^T, \ldots, \left(\Delta w_{N,i}\right)^T\right]^T, \quad (29)$$

$$\Delta \varphi_i = \left[\left(\Delta \varphi_{1,i}\right)^T, \left(\Delta \varphi_{2,i}\right)^T, \ldots, \left(\Delta \varphi_{N,i}\right)^T\right]^T, \quad (30)$$

$$\mathcal{M} = \mathrm{diag}\left\{\mu_1 I_M, \mu_2 I_M, \ldots, \mu_N I_M\right\}, \quad (31)$$

$$\mathcal{C} = C \otimes I_M. \quad (32)$$

Meanwhile, we also introduce the following matrices

$$\gamma_i = \mathrm{diag}\left\{\gamma_1(i) I_M, \gamma_2(i) I_M, \ldots, \gamma_N(i) I_M\right\}, \quad (33)$$

$$\mathcal{D}_i = \mathrm{diag}\left\{\tilde{u}_{1,i}\tilde{u}_{1,i}^*, \tilde{u}_{2,i}\tilde{u}_{2,i}^*, \ldots, \tilde{u}_{N,i}\tilde{u}_{N,i}^*\right\}, \quad (34)$$

$$\mathcal{G}_i = \mathrm{col}\left\{\tilde{u}_{1,i}v_1(i), \tilde{u}_{2,i}v_2(i), \ldots, \tilde{u}_{N,i}v_N(i)\right\}. \quad (35)$$

Taking expectation, we have

$$\gamma = \mathbb{E}(\gamma_i) = \mathrm{diag}\left\{p_1 I_M, p_2 I_M, \ldots, p_N I_M\right\}, \quad (36)$$

$$\begin{aligned}\mathcal{D} &= \mathbb{E}(\mathcal{D}_i)\\ &= \mathrm{diag}\left\{\mathbb{E}\left(\tilde{u}_{1,i}\tilde{u}_{1,i}^*\right), \mathbb{E}\left(\tilde{u}_{2,i}\tilde{u}_{2,i}^*\right), \ldots,\right.\\ &\quad \left.\mathbb{E}\left(\tilde{u}_{N,i}\tilde{u}_{N,i}^*\right)\right\}\\ &= \mathrm{diag}\left\{R_{u,1}, R_{u,2}, \ldots, R_{u,N}\right\}, \end{aligned} \quad (37)$$

$$\begin{aligned}\mathcal{G} &= \mathbb{E}\left(\mathcal{G}_i\mathcal{G}_i^*\right)\\ &= \mathrm{col}\left\{\sigma_{v,1}^2 R_{u,1}, \sigma_{v,2}^2 R_{u,2}, \ldots, \sigma_{v,N}^2 R_{u,N}\right\}, \end{aligned} \quad (38)$$

where p_k is the probability that $\gamma_k(i)$ is equal to 1.

A. THE MEAN PERFORMANCE

Combining the data model (1) and subtracting w° from both sides of equation (20), we can get

$$\begin{cases}\varphi_k(i) - w^\circ = w_k(i) - w^\circ\\ +\gamma_k(i)\mu_k\tilde{u}_{k,i}\left(\tilde{u}_{k,i}^* w^\circ + v_k(i) - \tilde{u}_{k,i}^* w_k(i)\right)\\ w_k(i+1) - w^\circ = \sum_{\ell\in\mathcal{N}_k} c_{\ell,k}\varphi_\ell(i) - w^\circ\end{cases}. \quad (39)$$

Using (27) and (28), the equation (39) can be written as

$$\begin{cases}\Delta\varphi_{k,i} = \left(1 - \gamma_k(i)\mu_k\tilde{u}_{k,i}\tilde{u}_{k,i}^*\right)\Delta w_{k,i}\\ +\gamma_k(i)\mu_k\tilde{u}_{k,i}v_k(i)\\ \Delta w_{k,i+1} = \sum_{\ell\in\mathcal{N}_k} c_{\ell,k}\Delta\varphi_{\ell,i}\end{cases}. \quad (40)$$

Through the global variable matrices introduced above, we have

$$\begin{cases}\Delta\varphi_i = \left(I_{NM} - \gamma_i\mathcal{M}\mathcal{D}_i\right)\Delta w_i + \gamma_i\mathcal{M}\mathcal{G}_i\\ \Delta w_{i+1} = \mathcal{C}\Delta\varphi_i\end{cases}, \quad (41)$$

or, equivalently,

$$\Delta w_{i+1} = \mathcal{C}\left(I_{NM} - \gamma_i\mathcal{M}\mathcal{D}_i\right)\Delta w_i + \mathcal{C}\gamma_i\mathcal{M}\mathcal{G}_i. \quad (42)$$

The independence assumption is defined as: all normal regression vectors $u_{k,i}$ are spatially and temporally i.i.d and $R_{u,k} = \mathbb{E}\left(\tilde{u}_{k,i}\tilde{u}_{k,i}^*\right) > 0$.

Taking the expectation of both sides of (42), we can obtain

$$\mathbb{E}\left[\Delta w_{i+1}\right] = \mathcal{C}\left(I_{NM} - \gamma\mathcal{M}\mathcal{D}\right)\mathbb{E}\left[\Delta w_i\right]. \quad (43)$$

Because $v_k(i)$ is independent of all other signals, the last term on the right side of (42) is equal to 0 when we take the expectation. A square matrix X is stable when all its eigenvalues lie inside the unit circle. That is, the spectral radius of the square matrix X satisfies the following condition.

$$\rho(X) < 1. \quad (44)$$

If the matrix $\mathcal{C}\left(I_{NM} - \gamma\mathcal{M}\mathcal{D}\right)$ is stable, the mean stability of the proposed algorithms will be guaranteed. Therefore, we have

$$\rho\left(\mathcal{C}\left(I_{NM} - \gamma\mathcal{M}\mathcal{D}\right)\right) < 1. \quad (45)$$

From [10], the equation (45) is equivalent to the following formula

$$\rho\left(I_{NM} - \gamma\mathcal{M}\mathcal{D}\right) < 1. \quad (46)$$

We can find that the different weight design methods are the same for the mean performance analysis of the proposed algorithms. From (46), we have

$$0 < \mu_k < \frac{2}{p_k\lambda_{\max}\left(R_{u,k}\right)}. \quad (47)$$

As the selection factor $\gamma_k(i)$ is binary variable, $p_k = \mathbb{E}(\gamma_k(i)) \le 1$. The following inequalities are satisfied:

$$\begin{aligned}\frac{2}{p_k\lambda_{\max}\left(R_{u,k}\right)} &\ge \frac{2}{\max\limits_{\ell=1,2,\ldots,N} p_\ell\lambda_{\max}\left(R_{u,\ell}\right)}\\ &\ge \frac{2}{\max\limits_{\ell_1=1,2,\ldots,N} p_{\ell_1}\max\limits_{\ell_2=1,2,\ldots,N}\lambda_{\max}\left(R_{u,\ell_2}\right)}\\ &\ge \frac{2}{\max\limits_{\ell=1,2,\ldots,N}\lambda_{\max}\left(R_{u,\ell}\right)}. \end{aligned} \quad (48)$$

Combining (47) and (48), a sufficient condition for step μ_k of each node k is

$$0 < \mu_k < \frac{2}{\max\limits_{\ell=1,2,\ldots,N}\lambda_{\max}\left(R_{u,\ell}\right)}. \quad (49)$$

When channel attacks occur, some extremely malicious outliers may occur, which makes the eigenvalues of $\lambda_{\max}(R)$ very large. However, due to the censorship of data selection, these outliers can be isolated from network updates. Therefore, we can conclude that as long as the step size μ_k is small enough to satisfy equation (49), the proposed algorithm is stable in mean regardless of whether the network is attacked or not.

B. THE MEAN-SQUARE PERFORMANCE

In this part, we study the mean-square performance of DS-DLMS and analyze it using energy conservation. Simultaneously, let $\|a\|_\Sigma^2 = a^*\Sigma a$. From (42), we have the networked mean-square relation

$$\|\Delta w_{i+1}\|^2 = \|\Delta w_i\|_\Gamma^2 + \|\mathcal{C}\gamma_i\mathcal{M}\mathcal{G}_i\|^2, \quad (50)$$

where $\Gamma = (I_{NM} - \gamma_i \mathcal{M}\mathcal{D}_i)^* \mathcal{C}^* \mathcal{C} (I_{NM} - \gamma_i \mathcal{M}\mathcal{D}_i)$. Evaluating the weighted norm of (50), we have

$$\mathbb{E}\|\Delta w_{i+1}\|^2 = \mathbb{E}\,\|\Delta w_i\|^2_{\Gamma_1} + \mathbb{E}[\mathcal{G}_i^* \mathcal{M}\gamma \mathcal{C}^* \mathcal{C}\gamma \mathcal{M}\mathcal{G}_i], \quad (51)$$

where $\Gamma_1 = (I_{NM} - \gamma\mathcal{M}\mathcal{D})^* \mathcal{C}^* \mathcal{C} (I_{NM} - \gamma\mathcal{M}\mathcal{D})$. Using the independence assumption, it can be obtained

$$\mathbb{E}\|\Delta w_{i+1}\|^2 = \mathbb{E}\,\|\Delta w_i\|^2_{\Gamma_1} + \mathrm{Tr}[\mathcal{C}\gamma\mathcal{M}\mathcal{G}\mathcal{M}\gamma\mathcal{C}^*]. \quad (52)$$

The equation (52) can be rewritten as

$$\mathbb{E}\,\|\Delta w_{i+1}\|^2_{I_{NM}} = \mathbb{E}\,\|\Delta w_i\|^2_{\Gamma_2} + \mathrm{Tr}[I_{NM}\mathcal{C}\gamma\mathcal{M}\mathcal{G}\mathcal{M}\gamma\mathcal{C}^*], \quad (53)$$

where $\Gamma_2 = (I_{NM} - \gamma\mathcal{M}\mathcal{D})^* \mathcal{C}^* I_{NM} \mathcal{C} (I_{NM} - \gamma\mathcal{M}\mathcal{D})$.

Let

$$\sigma = \mathrm{vec}(I_{NM}), \quad (54)$$

where the vec(·) notation is the transpose of vectorization with a matrix. For the three matrices X, Y and Q with matching dimension, we can get the following properties

$$\mathrm{Tr}\,(XY) = \left[\mathrm{vec}\left(Y^T\right)\right]^T \mathrm{vec}(X), \quad (55)$$

$$\mathrm{vec}(XYQ) = \left(Q^T \otimes X\right) \mathrm{vec}\,(Y)\,. \quad (56)$$

By taking the vectorization operation and using (53), (55) and (56), we get

$$\mathbb{E}\,\|\Delta w_{i+1}\|^2_\sigma = \mathbb{E}\,\|\Delta w_i\|^2_{\mathcal{F}_\sigma} + [\mathrm{vec}\left(\mathcal{C}\gamma\mathcal{M}\mathcal{G}\mathcal{M}\gamma\mathcal{C}^*\right)]^T \sigma, \quad (57)$$

where

$$\begin{aligned}\mathcal{F} &= \left[(I_{NM} - \gamma\mathcal{M}\mathcal{D})^T \mathcal{C}^T\right] \otimes \left[(I_{NM} - \gamma\mathcal{M}\mathcal{D})^* \mathcal{C}^*\right] \\ &= [(I_{NM} - \mathcal{D}\mathcal{M}\gamma) \otimes (I_{NM} - \mathcal{D}\mathcal{M}\gamma)] \left(\mathcal{C}^T \otimes \mathcal{C}^*\right). \end{aligned} \quad (58)$$

When $i \to \infty$, the networked mean-square deviation (MSD) can be obtained

$$\mathbb{E}\,\|\Delta w_\infty\|^2_\sigma = [\mathrm{vec}\left(\mathcal{C}\gamma\mathcal{M}\mathcal{G}\mathcal{M}\gamma\mathcal{C}^*\right)]^T \left(I_{N^2M^2} - \mathcal{F}\right)\sigma. \quad (59)$$

If $\mathcal{F}$ is stable, the iteration of (59) will be stable and convergence. All entries of $\mathcal{C}^T \otimes \mathcal{C}^*$ is non-negative and real, and all columns of it add up to one, so the stability of $\mathcal{F}$ is up to the stability of $\bar{\mathcal{F}}$.

$$\bar{\mathcal{F}} = [(I_{NM} - \mathcal{D}\mathcal{M}\gamma) \otimes (I_{NM} - \mathcal{D}\mathcal{M}\gamma)]\,, \quad (60)$$

which is stable if, and only if, $I_{NM} - \mathcal{D}\mathcal{M}\gamma$ is stable.

From the analysis of mean performance, when the step size μ_k is small enough to satisfy equation (49), $I_{NM} - \mathcal{D}\mathcal{M}\gamma$ can be stable. That is, as long as the step size μ_k satisfies equation (49), the networked MSD can be stable in the mean-square sense.

V. SIMULATIONS

In this section, we present some simulations to demonstrate the effectiveness of the proposed DS-DLMS algorithm. Meanwhile, simulations show the robustness of the proposed DS-DLMS algorithm with an adaptive credibility weight matrix under two channel attacks.

In the following simulations, a WSN is considered with $N = 20$ nodes. The unknown parameter to be estimated is set to a random vector with $M = 4$. The following examples use a small step $\mu = 0.02$ for each node so that the condition (48) can be satisfied. The following examples are carried out independently for 100 times, and the results of the following examples are averaged over 100 trails. The regression vector $u_{k,i}$ is subjected to standard normal Gaussian distribution and independent in time and space. The noise is generated from zero-mean Gaussian distribution with variances 0.2, as is shown in Fig. 2(a). And the Fig. 2(b) depicts the power of every node, which is generated from zero-mean Gaussian with variance 1.

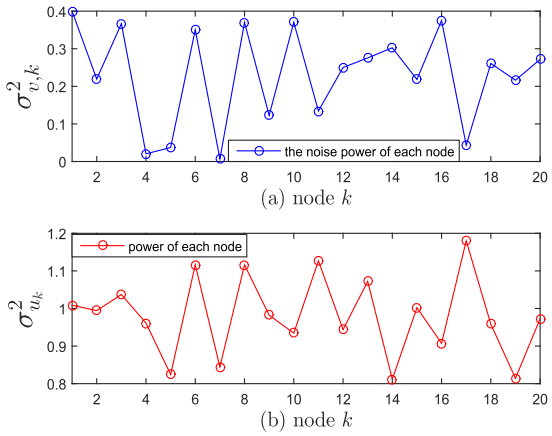


FIGURE 2. (a) The noise power of each node, (b) the power of each node without attacks.

A. THE NORMAL NETWORK

We first demonstrate the performance of the distributed DS-DLMS algorithm without being attacked. The $\varsigma_{\max}$ is set to 4. The P_{up} is set to 0.4, and $\nu = 5$. Fig. 3 shows the learning curves for different LMS algorithms in terms of MSD, where the "non-coop-LMS" denotes the non-cooperative LMS algorithm. We can find that the convergence performance of DS-LMS in [25] is same as the non-coop-LMS. However, the convergence performance of DS-DLMS is significantly better than other LMS algorithms, which indicates that the proposed algorithm not only improves the performance of the data selection LMS algorithm but also improves the performance of the DLMS algorithm.

Fig. 4 investigates the impacts of update probability P_{up} on the performance of distributed DS-DLMS algorithm by steady-state MSD. For comparison, we can see that when P_{up} is set to 0.2, the distributed DS-DLMS algorithm can achieve better performance.

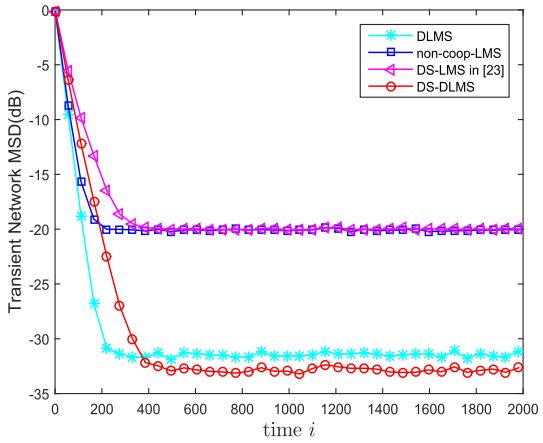


FIGURE 3. Transient network MSD of different LMS algorithms.

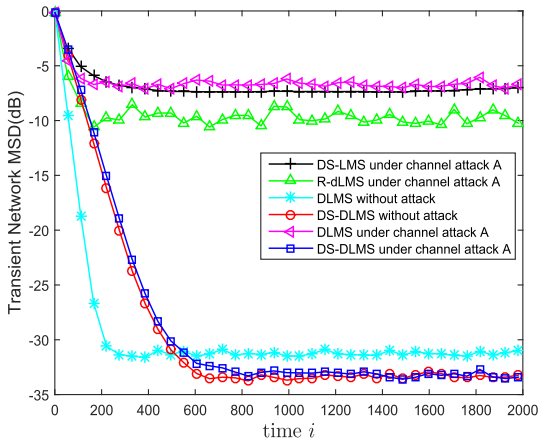


FIGURE 5. Transient network MSD of different LMS algorithms under channel attack A.

B. THE ADVERSARIAL NETWORK UNDER TWO CHANNEL ATTACKS

In this part, let the length of this credibility window Z be equal to 50. From Fig. 4, we make P_{up} equal to 0.2 in the following experiments. $\varsigma_{\max}$ is set to 4, $\nu = 5$ and $\zeta = 0.0001$. Models (4) and (6) of two channel attacks are used to create the uncertain attack. Moreover, three nodes are randomly picked to suffer channel attacks. In the channel attack A, nodes 6, 10 and 15 are chosen. In order to verify the validity of the proposed algorithm, nodes 6, 12 and 18 are selected in the channel attack B.

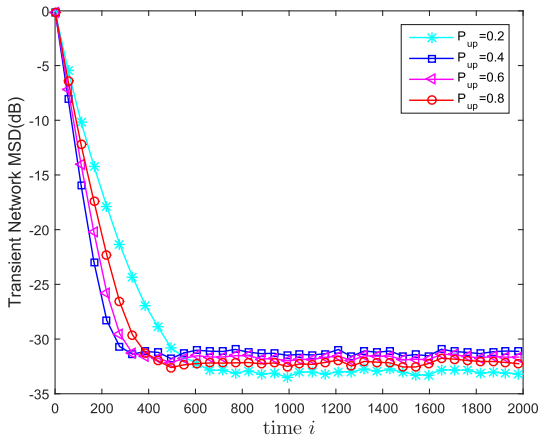


FIGURE 4. Transient network MSD of DS-DLMS algorithm under different P_{up}.

Fig. 5 shows the learning curves for different LMS algorithms, including DLMS and DS-DLMS without attacks, DLMS and DS-DLMS under channel attack A, DS-LMS in [25] and R-dLMS in [26] under channel attack A. It is obvious that when DLMS is subjected to the channel attack A, the performance of the entire network is reduced by the diffusion propagation of attack. Meanwhile, it can be found that DS-DLMS algorithm with the credibility weight is robust in the face of channel attack A.

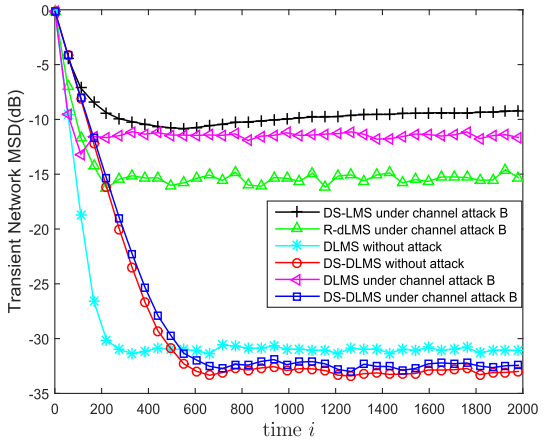


FIGURE 6. Transient network MSD of different LMS algorithms under channel attack B.

Fig. 6 shows the simulations of the proposed DS-DLMS algorithm under channel attack B comparing to other algorithms. The performance of DLMS can also be reduced because of the impact of channel attack B, which is similar to Fig. 5. From Fig. 5 and Fig. 6, we can conclude that DS-DLMS algorithm with the credibility weight can achieve the secure distributed estimation regardless of whether channel attacks occur or not.

C. THE COMPARISON OF DIFFERENT WEIGHT MATRICES

We investigate the differences in different weight matrices when the network encounters two channel attacks. Fig. 7 depicts the learning curves under channel attack A of DS-DLMS algorithm with the Metropolis weight and DS-DLMS algorithm with the credibility weight, while Fig. 8 depicts the learning curves under channel attack B. From Fig. 7 and Fig. 8, we can conclude that when two channel attacks occur in the network, the credibility weight design is obviously better than the Metropolis weight design in data selective DLMS algorithms.

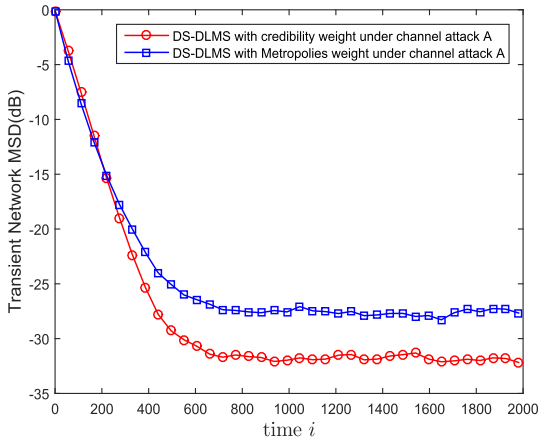


FIGURE 7. Comparison of two weights under channel attack A.

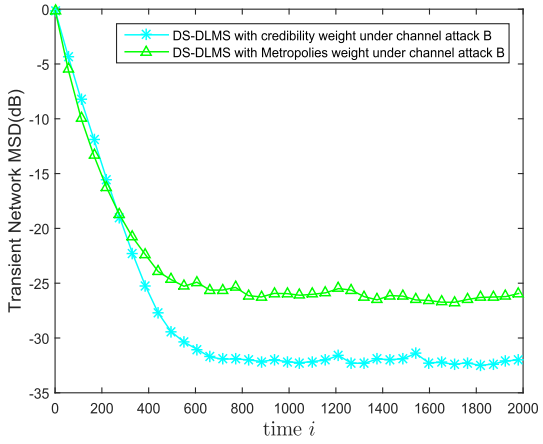


FIGURE 8. Comparison of two weights under channel attack B.

VI. CONCLUSION

In this paper, the distributed data selection diffusion LMS algorithm is proposed, and the ATC strategy of the proposed algorithm is obtained. In the distributed estimation system, two channel attacks are considered, which prompts us to design an adaptive credibility weight against channel attacks. Steady-state mean and mean-square analysis of the proposed algorithms are presented at the same time. Through simulation results, it is found that the DS-DLMS algorithm can achieve better estimation performance when outliers and less innovative data are discarded, and the DS-DLMS algorithm with credibility weight has strong robustness under two channel attacks, which is consistent with previous analysis. It is obvious that when channel attacks occur in the network, the credibility weight design is significantly better than the Metropolis weight design in the proposed DS-DLMS algorithm.

In WSNs, there are many types of attacks. To achieve the robust data selective LMS algorithms, the data selection algorithms against more types of attacks will be further studied in the next work.

REFERENCES

[1] Q. Liu, Z. Wang, X. He, and D. Zhou, "Event-based distributed filtering over Markovian switching topologies," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1595–1602, Apr. 2018.

[2] J.-T. Kong, D.-C. Ahn, S.-E. Kim, and W.-J. Song, "Robust distributed clustering algorithm over multitask networks," *IEEE Access*, vol. 6, pp. 45439–45447, 2018.

[3] J. Chen, C. Richard, and A. H. Sayed, "Multitask diffusion adaptation over networks," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4129–4144, Aug. 2014.

[4] B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Príncipe, "Generalized correntropy for robust adaptive filtering," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3376–3387, Jul. 2016.

[5] S. Ghazanfari-Rad and F. Labeau, "Formulation and analysis of LMS adaptive networks for distributed estimation in the presence of transmission errors," *IEEE Internet Things J.*, vol. 3, no. 2, pp. 146–160, Apr. 2016.

[6] R. Nassif, C. Richard, J. Chen, A. Ferrari, and A. H. Sayed, "Diffusion LMS over multitask networks with noisy links," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 4583–4587.

[7] F. Chen, T. Shi, S. Duan, L. Wang, and J. Wu, "Diffusion least logarithmic absolute difference algorithm for distributed estimation," *Signal Process.*, vol. 142, pp. 423–430, Jan. 2018.

[8] J. Chen, C. Richard, and A. H. Sayed, "Diffusion LMS over multitask networks," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2733–2748, Jun. 2015.

[9] F. Chen, X. Li, S. Duan, L. Wang, and J. Wu, "Diffusion generalized maximum correntropy criterion algorithm for distributed estimation over multitask network," *Digit. Signal Process.*, vol. 81, pp. 16–25, Oct. 2018.

[10] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035–1048, Mar. 2010.

[11] F. Chen, X. Liu, S. Duan, L. Wang, and J. Wu, "Diffusion sparse sign algorithm with variable step-size," *Circuits, Syst., Signal Process.*, vol. 38, no. 4, pp. 1736–1750, Apr. 2018.

[12] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, May 2001, pp. 2033–2036.

[13] C. Li, P. Shen, Y. Liu, and Z. Zhang, "Diffusion information theoretic learning for distributed estimation over network," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4011–4024, Aug. 2013.

[14] F. S. Cattivelli and A. H. Sayed, "Analysis of spatial and incremental LMS processing for distributed estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1465–1480, Apr. 2011.

[15] A. Rastegarnia, M. A. Tinati, and A. Khalili, "Performance analysis of quantized incremental LMS algorithm for distributed adaptive estimation," *Signal Process.*, vol. 90, no. 8, pp. 2621–2627, Aug. 2010.

[16] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4064–4077, Aug. 2007.

[17] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2365–2382, Jun. 2009.

[18] G. Mateos, I. D. Schizas, and G. B. Giannakis, "Consensus-based distributed least-mean square algorithm using wireless ad hoc networks," in *Proc. Allerton Conf.*, vol. 2, Sep. 2007, pp. 568–574.

[19] M. S. Talebi, M. Kefayati, B. H. Khalaj, and H. R. Rabiee, "Adaptive consensus averaging for information fusion over sensor networks," in *Proc. IEEE Int. Conf. Mobile Ad Hoc Sensor Syst.*, Oct. 2006, pp. 562–565.

[20] L. Hu, F. Chen, S. Duan, and L. Wang, "Diffusion logarithm-correntropy algorithm for parameter estimation in non-stationary environments over sensor networks," *Sensors*, vol. 18, no. 10, p. E3381, Oct. 2018.

[21] J. Chen and A. H. Sayed, "Diffusion adaptation strategies for distributed optimization and learning over networks," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4289–4305, Aug. 2012.

[22] D. Bertsekas and I. Rhodes, "Recursive state estimation for a set-membership description of uncertainty," *IEEE Trans. Autom. Control*, vol. AC-16, no. 2, pp. 117–128, Apr. 1971.

[23] S. Farahmand, S. I. Roumeliotis, and G. B. Giannakis, "Set-membership constrained particle filter: Distributed adaptation for sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4122–4138, Sep. 2011.

[24] X. Ge, Q.-L. Han, and Z. Wang, "A dynamic event-triggered transmission scheme for distributed set-membership estimation over wireless sensor networks," *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 171–183, Jan. 2017.

[25] P. S. R. Diniz, "Data-selective adaptive filtering," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4239–4252, Aug. 2013.
[26] Y. Liu and C. Li, "Secure distributed estimation over wireless sensor networks under attacks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1815–1831, Aug. 2018.
[27] M. Zhu and S. Martínez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 804–808, Mar. 2014.
[28] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, p. 13, 2011.
[29] Y. Guan and X. Ge, "Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 4, no. 1, pp. 48–59, Mar. 2018.
[30] Z. Guo, D. Shi, D. E. Quevedo, and L. Shi, "Secure state estimation against integrity attacks: A Gaussian mixture model approach," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 194–207, Jan. 2019.
[31] Y. Chen, S. Kar, and J. M. F. Moura, "Resilient distributed estimation through adversary detection," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2455–2469, May 2018.
[32] H. Song, P. Shi, W.-A. Zhang, C.-C. Lim, and L. Yu, "Distributed H_∞ estimation in sensor networks with two-channel stochastic attacks," *IEEE Trans. Cybern.*, to be published.
[33] H. Gan, S. Xiao, and Y. Zhao, "A novel secure data transmission scheme using chaotic compressed sensing," *IEEE Access*, vol. 6, pp. 4587–4598, 2018.

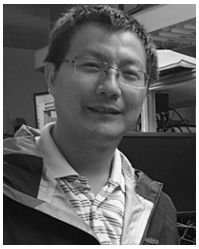
YI HUA received the B.S. degree in electronic information science and technology from the Chongqing University of Arts and Sciences, in 2017. He is currently pursuing the master's degree with the College of Electronic and Information Engineering, Southwest University, Chongqing, China. His research interests include wireless sensor networks, network security, and distributed signal processing.

FENG CHEN (M'14) received the Ph.D. degree from the University of Electronic Science and Technology of China, in 2013. In 2008, he visited the Prof. Emery Brown's research group, Massachusetts Institute of Technology, and also worked with the Massachusetts General Hospital, Harvard University. Since 2013, he has been an Associate Professor with the College of Electronic and Information Engineering, Southwest University, Chongqing, China. He has published over 20 papers in various journals and conference proceedings. His research interests include machine learning, wireless sensors networks, and signal processing.

SHUKAI DUAN (M'10) received the Ph.D. degree in computer science from Chongqing University, Chongqing, China, in 2006.

From 2005 to 2009, he was an Associate Professor and a Professor with the School of Physical Science and Technology, since 2010. From 2007 to 2009, he was a Postdoctoral Researcher with Chongqing University. He was a Visiting Professor with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, from 2010 to 2011, and the University of Windsor, Windsor, ON, Canada, in 2013. Since 1996, he has been with the College of Electronic and Information Engineering, Southwest University, Chongqing. He has taken charge of more than 10 national, provincial, or ministry level research projects, including the National Natural Science Foundation of China and the Program for New Century Excellent Talents from Ministry of Education. He has authored or coauthored four books and more than 100 papers in refereed journals and conferences. His current research interests include memristor devices and memristive systems, nonlinear circuits and systems, artificial neural networks, chaos and chaotic circuit, and intelligent signal processing.

Dr. Duan is an IEEE Computational Intelligence Society Member.



JIAGUI WU (M'14) was born in Sichuan, China, in 1981. He received the B.Sc. degree in physics and the M.Sc. degree in optics from Southwest University, Chongqing, China, in 2003 and 2006, respectively, and the Ph.D. degree in non-linear dynamics from the Sichuan University, Chengdu, China, in 2014. He is currently a Professor with the College of Electronic and Information Engineering, Southwest University, and a Visiting Scholar with the Electrical Engineering and Computer Science Department, University of California, Los Angeles, USA. He has authored or coauthored over 60 publications, including about 40 journal papers. His current research interests include the information security, the nonlinear dynamics of lasers, micro-nanophotonic devices, signal and information processing, chaos generation and synchronization, and secure communication with improved privacy.