IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 22, 2022, accepted April 20, 2022, date of publication April 26, 2022, date of current version May 3, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3170478

An Improvement of Security Scheme for Radio Environment Map Under Massive Attacking

YING GAO AND TAKEO FUJII, (Member, IEEE)
Advanced Wireless and Communication Research Center (AWCC), The University of Electro-Communications, Tokyo 182-8585, Japan
Corresponding author: Ying Gao (gaoying@awcc.uec.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 18H01439, Grant 18KK0109, and Grant 21H01322; and in part by the National Institute of Information and Communications Technology (NICT), Japan.

ABSTRACT Radio environment maps (REM) are widely used to enhance communication efficiency in spatial spectrum sharing. This can be generated using reports from the terminals. However, an open characteristic environment always leads to a security problem; for example, when malicious terminals exist in the environment and data falsification attacks occur, the accuracy of the REM is affected by the malicious action. In this study, we improve the double-layer monitor algorithm by optimizing the reward penalty function using a similarity comparison and sustainable monitor. The proposed algorithm can remove malicious terminal reports from the database to improve accuracy. Additionally, we propose a new algorithm involving interpolation based on spatial information to solve the problem of unequal information obtained from the meshes. By interpolation, a database can obtain sufficient datasets from each mesh. Furthermore, we defined an optimal attack strategy using the proposed security algorithm. The influence of malicious terminals can be the strongest among data falsification attacks; thus, we can check the performance more comprehensively. The simulation results indicate that the proposed method can eliminate the influence of malicious terminals and that a highly accurate REM can be obtained under massive malicious terminal attacks.

INDEX TERMS Radio environment map, radio propagation model, interpolation, wireless network, data falsification attack.

I. INTRODUCTION

The radio environment map (REM) is used to manage the inter-transmitter interference. It can store a large amount of information, including the average received signal power of the communication area. In addition, the average received signal power can be estimated using the measurement datasets. For example, in TV white space (TVWS) systems, secondary users usually recognize the white spaces and the allowable interference power according to the REM, which is stored in the spectrum database. As reported in [1] and [2], the REM can be used to obtain an efficient spectrum-sharing system.

The generation method of the REM requires the use of data from a database comprising of environmental information. The terminals can be located randomly in the communication area, and they report the received signal power, terminal ID, and the terminal location, etc. REM can be generated using a database with sufficient datasets of the a certain communication area. Therefore, the accuracy of the information that terminals report to the database is important for the REM, and the estimation accuracy of REM construction is an essential index that can directly affect the spectral efficiency.

In [3], REM was constructed without a shadowing impact. In [4], spatial spectrum sharing was modeled over log-normal channels, and the results indicated that the model can be simplified by using the Kriging-aided method. In [5], both a fixed transmitter system and a distributed transmitter system were considered, and a neural network was used to increase accuracy. Experimental measurement datasets can be used to generate the REM. In [5]-[7], a model of a distributed transmitter system was built, the frequency correlation of shadowing was examined, and the V2V communication environment was modeled via measurements.

One of the methods used to generate REMs is using information from reports of terminals in the communication

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

area. The accuracy of the REM is significantly influenced by the received datasets. If malicious terminals exist in the communication area, they rewrite the information and upload the wrong datasets to the database, and they attempt to blind the database so that the error of the REM satisfies their selfish requirements. As reported in [8]–[10], data falsification attacks are well known in the spectrum sensing field. Examples of always present, always absent, and always opposite attacks are presented in [11]. In [12], an example of malicious terminals performing independent and collaborative attacks is presented. In [13]–[15], the optimal likelihood ratio test was used to address attacks. In [16]–[20], a penalty-based Dempster–Shafer theory of evidence was proposed to address uncertainty representation. However, when the number of datasets is insufficient, particularly after malicious terminals are removed, fewer datasets may not be sufficient to generate REM with a high accuracy. In [21]–[25], it was shown that shadowing has a spatial correlation. The interpolation method can be used to estimate the information of unknown points in the communication area to enhance the accuracy of the REM.

In this study, we propose an algorithm based on the spatial information for anti-malicious terminals in the process of generating the REM. The contributions of this study are as follows:

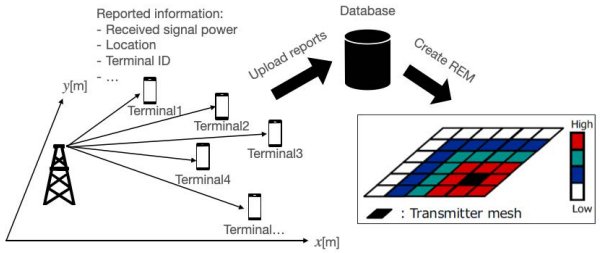
- The double-layer monitor (DLM) algorithm was improved by optimizing the reward-penalty function.
- A DLM based on spatial information algorithms, including inverse distance weighting (IDW) and spatial correlation, was proposed to enhance the performance of the network.
- The optimal attack strategy under security algorithms was defined to comprehensively measure the performance of our algorithms, and the maximum error under the optimal attack was determined.

The remainder of this paper is organized as follows: Section II presents the system model, including the REM and data falsification attack models. Section III introduces the proposed DLM-based algorithm to remove the malicious terminal datasets. Section IV presents multiple attack strategies, including the optimal attack. Section V presents the results of the numerical simulations. Finally, the conclusions are presented in Section VI.

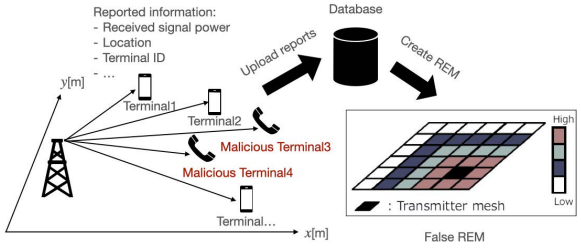
II. SYSTEM DESCRIPTION

A. RADIO ENVIRONMENT MAP

To draw an REM for the communication area, we assume that several mobile terminals are used to collect spectrum information in the communication area. The terminal location is randomly generated according to their random movement. As reported in [3], terminals can collect information and send it to the database. The database can be installed in the cloud or base station, and can store large amounts of data. After the database obtains sufficient data from the terminals, it can generate an REM based on this information. In our research, we only considered the use of information



(a) Without malicious terminals (True Map)



(b) With malicious terminals (Fake Map)

FIGURE 1. A concept of the conventional REM.

such as the terminal ID, location, and received signal power. Here, we use received signal power to generate the REM because the mobile terminals can easily obtain the received signal power, for example, the smartphone can obtain the received signal power of the cellular system and wireless LAN by using the API of the Android OS. Therefore, the measurement cost of the received power is very low. The concept of the conventional REM is shown in Fig.1(a).

To remove the influence of small-scale fading, we split the communication area into two-dimensional meshes. Datasets from the same mesh were used to generate the average power of the mesh. When the mesh is sufficiently small, the influence of shadowing can be ignored, and the accuracy of the REM can be improved. However, if malicious terminals exist in the communication area, the malicious information from them can reduce the accuracy of the REM, the algorithm which can distinguish the malicious information which is reported by the malicious terminals needs to be considered [26].

B. DATA FALSIFICATION ATTACK MODEL

We consider that malicious terminals exist in the communication environment, and they can attack the database to damage the accuracy of REM to satisfy their selfish requirements: affect the primary terminal frequency bands or take the free frequency bands, etc. REM construction need to collect received signal power from spatially distributed mobile terminals and calculate the average power of each mesh. Data falsification attack is an efficiency attack to damage the REM.

Data falsification attacks are also called Byzantine attacks, which refer to malicious terminals that change the data of their sensing power of the spectrum to blind the database. Examples of this type of attack models were presented in [8]–[12]. Herein, we present an attack strategy for a data falsification attack as follows: malicious terminals change their data by comparing them with the power threshold η. In every sensing slot, if the sensing power of the malicious terminal exceeds the threshold, they rewrite the sensing power by multiplying a certain index, called the attacking index, by the probability P_a and report the wrong dataset to the database. Otherwise, the malicious terminals send the correct dataset to the database. The reported power of a malicious terminal is as follows,

$$P'(i) = \begin{cases} P(i) \cdot \delta, & \text{if } P(i) > \eta \text{ with } P_a \\ P(i), & \text{otherwise} \end{cases}, \tag{1}$$

where δ represents the attacking index, and P_a represents the attack probability. If $\delta > 1$, the attacking strength increases with the attacking index, if the attacking index is $0 < \delta < 1$, the attacking strength decreases as the attacking index increases. The concept of REM damage from a data falsification attack is shown in Fig.1(b).

III. SENSING SCHEME AGAINST DATA FALSIFICATION ATTACK

A. DOUBLE-LAYER MONITOR (DLM)

1) SIMILARITY COMPARISON

To identify malicious datasets in the database, we calculate the similarity degree in the first layer. We calculate the similarity degree for each pair of datasets in the same mesh. If the mesh size is sufficiently small, the shadowing index can be considered uniform in any given mesh, thus, different terminals only suffer different path loss and fading in the same mesh. Hence, calculating the similarity degree as an index to judge malicious datasets is reasonable.

The malicious terminals conduct the data falsification attacks during their movement, that is, they rewrite the received information and send incorrect information to the database. In this condition, the malicious terminals' datasets differ from the honest terminals' datasets even if they are at the same location, although the datasets from the same location should be different because of the different fading. After calculating the similarity degree of each pair of data, we perform a comparison. If the similarity degrees of most of the terminals are high, we consider this terminal to have a high probability as an honest terminal, otherwise, we consider it to be malicious.

The similarity degree is calculated as follows,

$$sim(i, j) = \frac{max(P'(i), P'(j))}{\frac{1}{2}(P'(i) + P'(j))}, \tag{2}$$

therefore, for the received signal power in the entire mesh, the following similarity matrix can be constructed,

$$Sim = \begin{bmatrix} 1 & \cdots & sim(1,j) & \cdots & sim(1,n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ sim(i,1) & \cdots & 1 & \cdots & sim(i,n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ sim(n,1) & \cdots & sim(n,j) & \cdots & 1 \end{bmatrix}, \tag{3}$$

then, the support level of each data can be calculated as follows,

$$Sup(i) = \sum_{j=1}^{n} sim(i,j) \quad j \neq i \quad i,j = 1 \cdots n, \tag{4}$$

consequently, the reliability of the terminal i can be determined by normalizing the support as follows,

$$Rel(i) = \frac{Sup(i)}{max(Sup(1), Sup(2), \cdots, Sup(n))}. \tag{5}$$

The similarity comparison method is like Algorithm 1 presented below, where $m = 1, 2, 3, \ldots, D$ represents the mesh number, and we assume that there are D meshes split from the communication area in total. $i = 1, 2, 3, \ldots, N$ represents the number of reports, and we assume that there are N reports in each mesh. $\mathcal{HT}$ and $\mathcal{MT}$ represent the set names of honest terminals and malicious terminals, respectively. α represents the similarity threshold, and when $Rel(i) > \alpha$, the terminal is judged as honest, otherwise, it is malicious.

However, when the malicious datasets occupy a large part of the datasets in one mesh, only considering about using similarity to set the reliability is not sufficient to remove them, moreover, if the malicious terminals perform a dynamic attack (changing the attacking index), it is not reasonable for us to reset the threshold continuously during their movement. To address this situation, a sustainable monitor method is considered.

2) SUSTAINABLE MONITOR

As aforementioned, using a similarity comparison is not sufficient to remove malicious information from the power fusion. To solve this problem, we considered using a sustainable monitor to enhance the performance. A sustainable monitor uses terminal ID information. We also monitor the performance of the terminal when it moves to another mesh, thus, the reliability can be judged continuously. If the terminal's performance is judged as malicious continuously, we can remove all the data from this terminal from the power fusion.

a: REAL-STEP CONFIDENCE

Real-step confidence refers to the difference between the reported power and the average power after power fusion in each mesh. If the difference is small, we consider the terminal to be confident, otherwise, the confidence should be

Algorithm 1 Similarity Comparison

Require:
The parameters related to the Database, such as $P'(i)$, $\mathcal{HT}$, $\mathcal{MT}$, α, etc.

Ensure:
for Each mesh $m \in D$ **do**
  Initialize $\mathcal{MT} = \phi$,$\mathcal{HT} = N$,$P_{all} = 0$
  for Each reported power in each mesh $i \in N$ **do**
    Calculate the similarity degree
    Generate the similarity matrix
    Calculate the $Sup(i)$ for each dataset
    Normalization get $Rel(i)$
    if $Rel(i) \geq \alpha$ **then**
      $P_{all} \leftarrow P_{all} + P'(i)$
    else
      $\mathcal{MT} \leftarrow \mathcal{MT} + \{i\}$
      $\mathcal{HT} \leftarrow \mathcal{HT} - \{i\}$
    end if
  end for
  Calculate the average power $\overline{P}_m = \frac{P_{all}}{|\mathcal{HT}|}$
end for
Generate the REM

decreased. The confidence index *Bias* can be calculated as follows,

$$Bias(i) = |P'(i) - \overline{P}|, \tag{6}$$

where $Bias(i)$ represents the absolute value of the difference between the reported power and the average power in the mesh for terminal i.

b: HISTORICAL RELIABILITY

The historical reliability is updated step-by-step. When the historical reliability is lower than the threshold, the terminal is judged to be malicious, and the data from the same ID are removed from the power fusion.

$$HisRe_i^H = l * HisRe_i^{H-1} + (-1)^{para} * \Re(\cdot), \tag{7}$$

where $HisRe_i^H$ represents the historical reliability for the *ith* terminal at the step $H(H = 2, 3, 4, \ldots)$, l is the impact factor of the historical reliability, and a larger value of l, the impact of the historical reliability is larger. $para = 0, 1$ is the reward-penalty parameter, when $Bias < \zeta$, $para = 0$, the historical reliability increases, otherwise, when $para = 1$, the historical reliability should decreases. Here, ζ represents the threshold for the real-step confidence. Comparing with our previous study [26], we improved the historical calculation by resetting the reward–penalty function $\Re(\cdot)$. $\Re(\cdot)$ is the fitting function that is related to the *Bias*. When $Bias < \zeta$, $\Re(\cdot)$ should decrease slowly as *Bias* increases, whereas when $Bias \geq \zeta$, $\Re(\cdot)$ should increase rapidly as *Bias* increases.

c: WEIGHT ALLOCATION

After calculating the historical reliability, we update the weight for each terminal. A terminal with a higher reliability has a higher weight. Otherwise, the weight decreases. When the weight decreases to zero, the reliability is set to zero.

$$w\_HisRe_i^H = \frac{HisRe_i^H}{max(HisRe)}, \tag{8}$$

where w_HisRe represents the weight, and $max(HisRe)$ represents the maximum $HisRe$ for all the terminals in the same mesh.

The DLM is like Algorithm 2 represented below, where β represents the historical reliability threshold and H represents the step number of the terminal. By using the terminal ID information, we monitor the performance of the terminals during their movements, thus, the reliability can be judged continuously. In this case, even under a dynamic attack (malicious terminals change their attacking indexes during their movement), or a small-scale attack, their historical reliability can be judged step-by-step, and once they are judged as malicious terminals, they can be removed from the power fusion.

Algorithm 2 Double Layer Monitor

Require:
The parameters related to the Database, such as $P'(i)$, $\mathcal{HT}$, $\mathcal{MT}$, α, β, terminal ID, etc.

Ensure:
for Each reported power in each mesh $i \in N$ **do**
  Initialize $\mathcal{MT} = \phi$,$\mathcal{HT} = N$,$P_{all} = 0$,$HisRe_i^H = 0.1$
  for Each terminal's step H **do**
    Calculate the Similarity comparison get the $Rel(i)$
    if $Rel(i) > \alpha$ && $w\_HisRe_i^H \geq \beta$ **then**
      $P_{all} \leftarrow P_{all} + P'(i)$
    else
      $\mathcal{MT} \leftarrow \mathcal{MT} + \{i\}$
      $\mathcal{HT} \leftarrow \mathcal{HT} - \{i\}$
    end if
    Calculate the average power $\overline{P}$
    Calculate the Real-step confidence $Bias(i)$ and $para$
    Calculate the Historical reliability $HisRe_i^H$
    Do the weight allocation
  end for
  Update the $HisRe_i^H$ and w_HisRe_iH
  Move to next step $H \leftarrow H + 1$
end for
Generate the REM

B. IMPROVEMENT OF DLM

As mentioned previously, a DLM can identify reports that are transferred by malicious terminals and remove these datasets from power fusion. In this case, the number of reports for each mesh is different because of the malicious terminals and the random movements of the terminals. The database generates the REM according to the reports from the terminals. If the number of datasets from one mesh is insufficient, the REM can deviate significantly from reality. As a collaborative sensing network with more reliable datasets in the database,

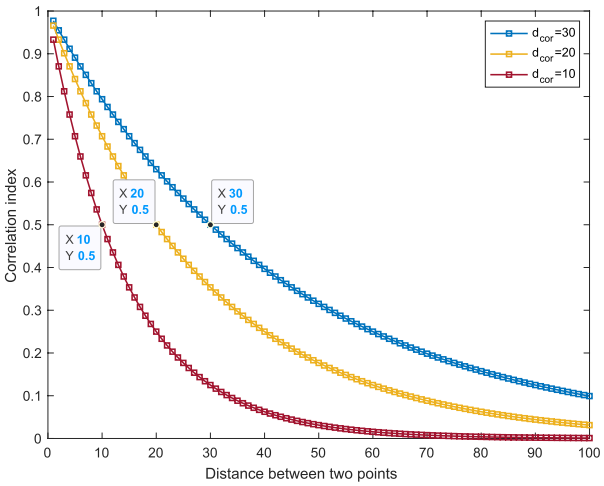


FIGURE 2. Relationship between distance and the correlation index.

the accuracy of the REM can be increased. To solve the problem of unequal amounts of information in meshes, interpolation can be considered.

1) DLM BASED ON SPATIAL CORRELATION

It is well known that the spatial correlation of shadowing is represented by an exponential decay model, and the correlation index is defined as follows [21],

$$\rho_{ij} = \frac{E[W(x_i)W(x_j)]}{\sigma^2} = \exp\left(-\frac{\Delta d_{ij}}{d_{\text{cor}}}\ln 2\right), \tag{9}$$

where Δd_{ij}[m] represents the distance between two different terminals i and j, and d_{cor}[m] represents the correlation distance, which is defined as the point where $\rho_{ij} = 0.5$. In an urban area, the correlation distance is approximately 20[m] according to an experiment[21].

As shown in Fig.2, when the correlation distance is fixed, a shorter distance between two points corresponds to a stronger correlation. Fig.3 shows the cumulative distribution function(CDF) curves under different thresholds. Here θ represents the threshold for the correlation index ρ. Only when $\rho > \theta$ can the observation reports be used to estimate the test points' information. Table 1 presents the values plotted in Fig.3. As shown, higher accuracy can be obtained with a higher threshold θ.

To deal with the unequal information in the meshes and increase the accuracy of the REM, we use a DLM based on a spatial correlation algorithm, as shown in Algorithm 3. Sufficient data can be obtained for environmental estimation through reasonable interpolation. This addresses the shortage of unequal information caused by the terminals' random movements and solves the problem of information loss caused by malicious terminal attacks.

2) DLM BASED ON INVERSE DISTANCE WEIGHTING (IDW)

The IDW is a deterministic method for interpolating with a large number of known observation points. The assigned values are based on the weighted average of the known observation points, and the weight is assigned according to the inverse of the distance between the observation point and the unknown point.

We assume that there are N observation points located randomly in the area of interest with the location information (x_i, y_i), where $i = 1, 2, 3, \ldots, N$. Here, we only consider the situation of two dimensions, therefore, the x_i, y_i represent the horizontal and vertical distances from the observation point, respectively. $P(x_i, y_i)$ represents the power information corresponding to the coordinates. The ordinary IDW interpolation

Algorithm 3 DLM Based on Spatial Correlation

Require:
The parameters related to the Database, such as $P'(i)$, $\mathcal{HT}$, $\mathcal{MT}$, α, β, terminal ID, etc.

Ensure:
1: **for** Each reported power in each mesh $i \in N$ **do**
2: Do the Double Layer Monitor
3: Get $\mathcal{MT}$ and $\mathcal{HT}$
4: Calculate the amount of data which need to be interpolation
5: Generate the random location in the mesh
6: Interpolation by Spatial Correlation
7: **end for**
8: Generate the REM

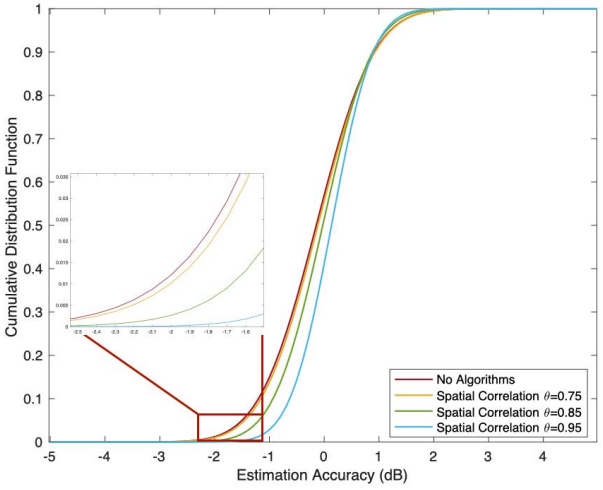


FIGURE 3. Cumulative distribution function of spatial correlation.

TABLE 1. Detail value of spatial correlation.

	error	μ	σ
No Algorithms	3.4799	-0.1443	0.8231
$\theta = 0.75$	3.4090	-0.1147	0.8122
$\theta = 0.85$	2.8055	-0.0260	0.7078
$\theta = 0.95$	2.3808	0.1338	0.5947

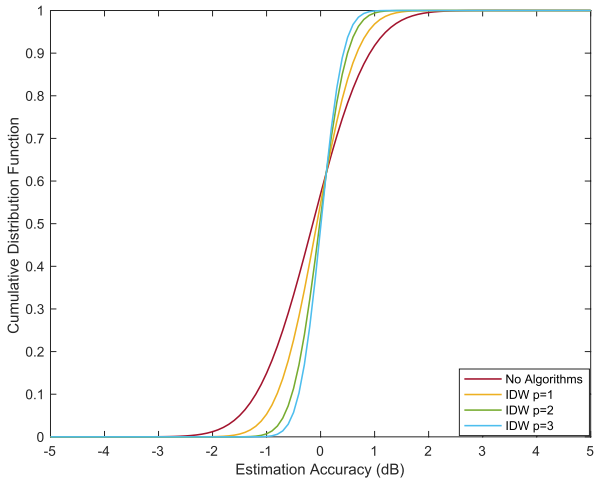


FIGURE 4. Cumulative distribution function of IDW.

function is expressed as follows,

$$\begin{cases} \widetilde{P}(x_0, y_0) = \sum_{i=1}^{N} w_i P'(x_i, y_i), & \text{if } d_i \neq 0 \text{ for all } i \\ \widetilde{P}(x_0, y_0) = P'(x_i, y_i), & \text{if } d_i = 0 \text{ for some } i \end{cases}, \tag{10}$$

where, x_0, y_0 represents the coordinate location of the interpolated point, $\widetilde{P}(x_0, y_0)$ represents the interpolated power at the coordinate location, and w_i represents the assigned weight from each observation point, which can be expressed as follows,

$$w_i = \frac{d_i^{-p}}{\sum_{i=1}^{N} d_i^{-p}}, \tag{11}$$

where, d_i represents the distance between the ith observation point and interpolated point, which can be calculated as $d_i = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}$. p is a positive real number, which is called the IDW power parameter. For a larger value of p, the closest points have a stronger influence on the interpolated point. As references [27]-[28] shown, the range of p is generally [0.5,3].

As shown in Fig.4, the interpolation performance is better for a larger value of p, because the dependence on the neighboring points is greater. The sensing power is influenced by path loss, shadowing, and fading. Moreover the path loss and shadowing are affected by the locations of the terminals. Thus, with a shorter distance between the two terminals, they have more similar sensing power values, and smaller interpolation error. Table 2 presents the detailed values in Fig.4. A larger IDW power parameter corresponds to a smaller error in the interpolated point.

As aforementioned, interpolation can be used to address the problem of unequal information in meshes. Similar to Algorithm 3 using the knowledge of spatial information, IDW interpolation can solve this problem to obtain sufficient data to generate the REM. Additionally, by adding the weight according to the distance, the estimated data generated by the IDW are more related to the nearer points, which should lead to better performance compared with the spatial correlation. The DLM based on IDW is described in Algorithm 4.

TABLE 2. Detail value of IDW.

	error	μ	σ
No Algorithms	3.4799	-0.1443	0.8231
$p = 1$	2.3127	-0.0694	0.5760
$p = 2$	1.4599	-0.0155	0.4013
$p = 3$	1.1435	0.0077	0.3219

Algorithm 4 DLM Based on IDW

Require:
The parameters related to the Database, such as $P'(i)$, $\mathcal{HT}$, $\mathcal{MT}$, α, β, terminal ID, etc.

Ensure:
1: **for** Each reported power in each mesh $i \in N$ **do**
2: Do the Double Layer Monitor
3: Get $\mathcal{MT}$ and $\mathcal{HT}$
4: Calculate the amount of data which need to be interpolation
5: Generate the random location in the mesh
6: Interpolation with Inverse Distance Weighting
7: **end for**
8: Generate the REM

IV. DIFFERENT ATTACK STRATEGIES

Ordinary algorithms for constructing REMs always face the threat of malicious terminal attacks. Malicious terminals can influence REM using different attack strategies. In this section, we present different types of attack strategies under both black-box and white-box conditions to evaluate the performance comprehensively.

A. BLACK-BOX ATTACK

A black-box attack means that malicious terminals do not have the knowledge of the database, they may have the head of the attacker to make the decision of the attack strategy but they do not know the algorithms and parameter setup information in the database.

1) STATIC ATTACK AND DYNAMIC ATTACK

Without loss of generality, we consider both static and dynamic attacks in our study. In static attacks, all the malicious terminals use the same attack strategy. Malicious terminals use the same attacking index all the time, and they cannot stop attacking or change the attacking index midway. In contrast, in a dynamic attack, the malicious terminals can change their attack strategy during the sensing process, they can change their attacking index if they wish, even if they choose not to attack several sensing slots to protect themselves from being detected by the database.

2) INDEPENDENT ATTACK AND COLLABORATIVE ATTACK

In an independent attack, malicious terminals can independently make decisions. They can select the attacking index and attacking slot individually. In this case, malicious terminals have stronger subjective initiative than do collaborative attacks, and this type of attack can increase the complexity of the attack system. Because attackers make decisions independently, they can significantly reduce the risk of being detected. By contrast, in a collaborative attack, all malicious terminals select the same attack strategy. Compared with independent attacks, this type of attack may have a stronger influence on the sensing system, however, it always has a higher risk of being detected.

B. WHITE-BOX ATTACK

In contrast to a black-box attack, in a white-box attack, malicious terminals have full knowledge of the security system of the entire network, which can be used to function as a database, however, their goal is to increase the error of the REM and obtain benefits from the wrong map.

Using the information of the security system, malicious terminals can launch an optimal attack through calculations. As indicated by Equation (1), when the attacking index $\delta(\delta \geq 1)$ increases, the attacking strength increases, the reliability $Rel(i)$ of malicious terminals decreases, and the detection probability of malicious terminals increases. Conversely, when the attacking index $\delta(0 < \delta < 1)$ decreases, the attacking strength increases, the reliability $Rel(i)$ of the malicious terminals decreases, and the detection probability of malicious terminals increases. The reason to consider $Rel(i)$ is that when $Rel(i) \leq \alpha$, although $Bias(i)$ is small, the data can be judged as malicious terminal's at the first layer similarity comparison step, hence there is a trade-off problem that is δ needs to be strong and cannot be distinguished at the same time. To determine the optimal attacking index, δ must be defined as follows,

$$\arg\max_{\delta} \sum_{i=1}^{M} [Rel(i) > \alpha] * Bias(i). \tag{12}$$

We assume that malicious terminals have an attack center, similar to a database, and we consider that the attack center may hack the database and obtain information. Therefore, the attack center has full knowledge of the DLM operation method and knows their parameters, therefore, the malicious terminals can generate the optimal attack strategy under our security network.

The most effective method is to ensure that malicious datasets can participate in the power fusion step and generate errors to the greatest extent possible. In this case, *HisRe* is monitored, as long as it is lower than the threshold, the attacker should set the attack model to silence and send the correct data to increase the *HisRe*, when the *HisRe* is safe, the attacker can continue attacking, leading to an error in the REM.

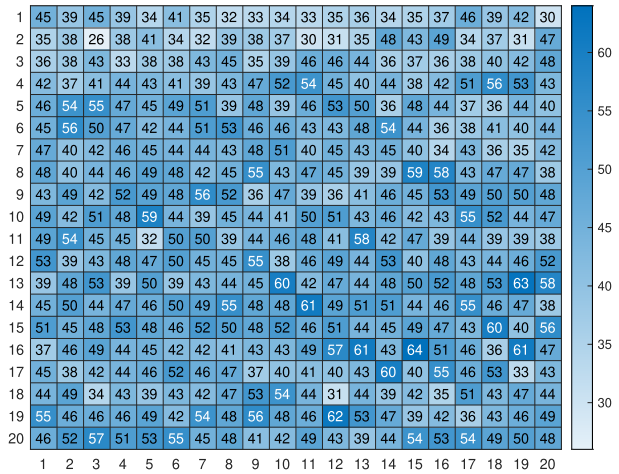


FIGURE 5. The amount of the reports in every meshes.

V. RESULTS AND DISCUSSION

In this section, we present simulation results to verify the performance of the aforementioned algorithms. The "histo" here means the algorithms consider the historical reliability, so it is the shortened form for DLM. Moreover, "sim" is the shortened name for similarity comparison, "corr" is the shortened name for spatial correlation. The methods in the first column of Table 4 are as follows: DLM based on IDW, DLM based on spatial correlation, DLM alone, similarity comparison (first layer) based on IDW, similarity comparison based on spatial correlation, similarity comparison alone, bi_weight from reference [29], average combination from reference [30], IDW alone, spatial correlation alone, and no algorithms.

A. SIMULATION SETUP

We divided the communication area into a 10×10 m^2. For simplicity, we assume random routes (random walk from a mesh to the next mesh) for each terminal, and they all go through from one side of the communication area to the other side in 20 steps. We also assume that they only upload one dataset at every step. The number of reports is different for each mesh because the movement of each terminal is random, and a mesh by which more terminals pass can obtain more reports. We randomly generated 900 routes in total and selected 100 routes as malicious terminals'. The simulation parameters are listed in Table 3. The number of reports for each mesh is shown in Fig.5, and the number of malicious terminals' reports in each mesh is shown in Fig.6.

B. RADIO PROPAGATION MODEL

Let x_{T_x} denote the primary user's transmitter location. Therefore, we assume that the received signal power of the terminal which is located at x is given as follows [6],

$$P(x) = P_{T_x} - L(d_0) - 10\gamma \log_{10}\left(\frac{d_m}{d_0}\right) + W + F, \tag{13}$$

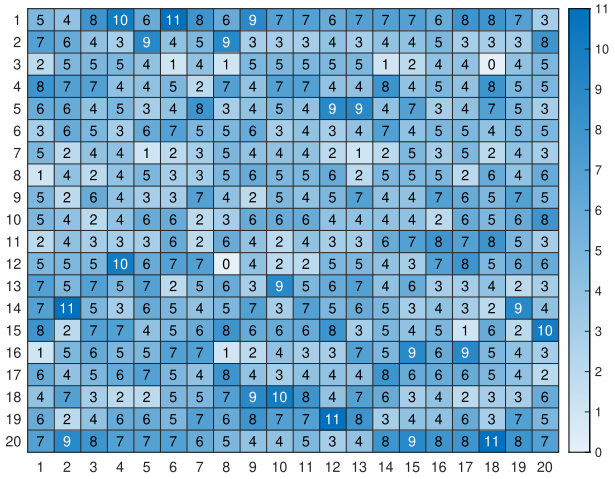


FIGURE 6. The amount of the malicious reports in every meshes.

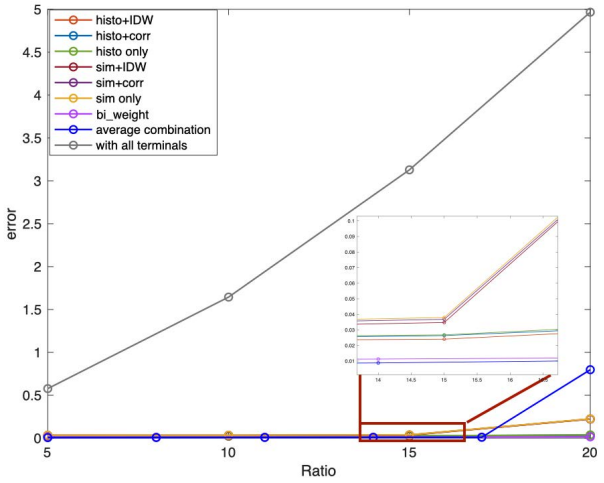


FIGURE 7. Error versus of fixed reports.

TABLE 3. Simulation parameters.

Parameter	Value
Mesh size [m^2]	10×10
Mesh amount	20×20
Center frequency [GHz]	3.5
Transmission power [dBm]	29
Reference distance [m]	10
path loss index γ	3.5
Standard deviation of W	6
Similarity threshold α	0.95
Bias threshold ζ	0.8
Historical reliability threshold β	0.1
Correlation threshold	0.75
IDW power parameter	3
The number of routes	900
Percentage of malicious routes	11.11
Steps for each route	20
Desired SIR Γ_d [dB]	20
Desired outage probability p_{out}	0.1
attacking index	0.5-1.5

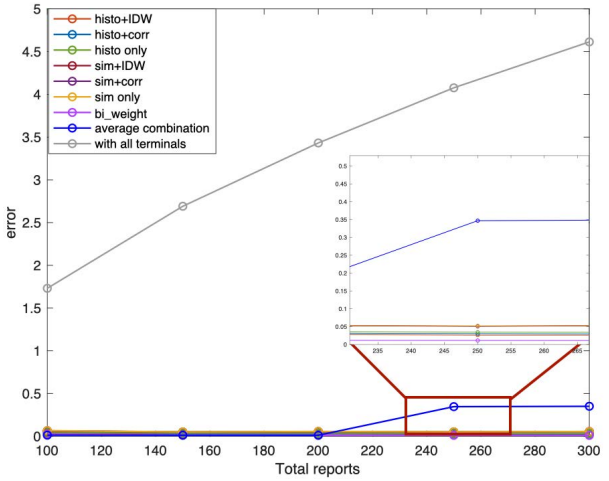


FIGURE 8. Error versus of fixed ratio.

where P_{T_x} represents the primary transmission power in the dBm domain, $d_m = ||x_{T_x} - x||$ represents the distance [m] between the transmitter and the sensing terminal which located at x, d_0 is the reference distance [m], γ represents the path loss index, W represents the shadowing loss [dB] at location x and W follows a log-normal distribution with a standard deviation of σ [dB]. F represents small-scale fading [dB]. $L(d_0)$ denotes the free-space path loss [dB], which is calculated as follows,

$$L(d_0) = 10\log_{10}\left(\frac{4\pi d_0}{\lambda}\right)^2, \tag{14}$$

where λ represents the wavelength [m].

C. FIXED-TERMINAL CONDITION

In this subsection, we check a simple condition, in which we do not consider the random routes of the terminals. The environment is simple and ideal. We set the terminals on the map at fixed locations. First, we check that the received reports in every mesh are identical, when the number of malicious terminals reports increases, the performance of the error of the REM changes. We set 200 reports in each mesh, and the ratio of malicious reports ranged from 5% to 20%. The performance of the proposed method is illustrated in Fig.7. The error increases when the ratio of malicious terminals' reports increases. Additionally, we check the performance when the ratio of malicious reports is fixed and the number of received reports in each mesh is changing. We set the ratio of malicious terminals' reports to 15%, and the number of reports in each mesh ranged from 100 to 300. As shown in Fig.8, the error increases with the total number of reports.

However, in reality, the condition is more complex, because the movements of the terminals are random, and the number of received reports and the malicious reports in each mesh are different, as shown in Fig.5 and 6. In the following subsections, we evaluate the performance under random terminal movements and different attacking strategies.

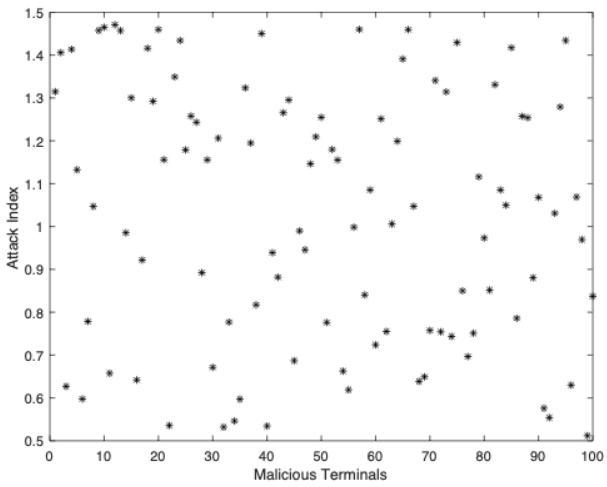


FIGURE 9. Attacking index of malicious terminals.

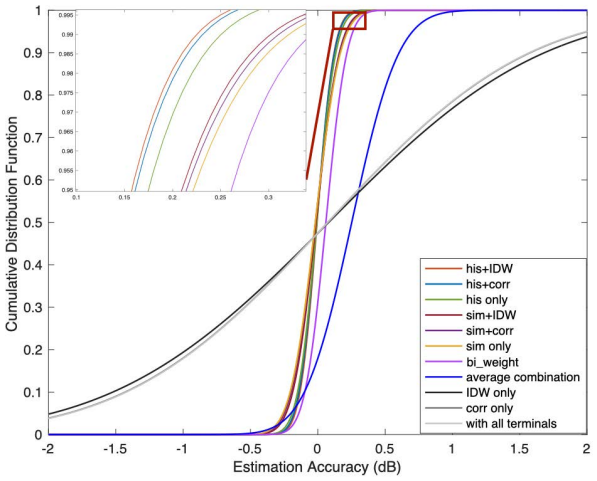


FIGURE 10. CDF of static attack under independent.

D. UNDER STATIC ATTACK

As mentioned previously, in a static attack, malicious terminals perform the same attack all the time, they do not change their attack strategy (including the attacking index and attacking slot) during their movement. After they begin their attacking behavior, they can no longer change it. Static attacks are classified as either independent or collaborative attacks. In this subsection, we consider the random routes of all the terminals, as shown in Figs.5 and 6.

1) UNDER INDEPENDENT STRATEGY

In the static attack under an independent strategy, malicious terminals make the attacking decision by themselves, and they select their attacking indexes individually and attack throughout their journeys. After selecting their attacking indexes, they do not change them during their movement. Here, we assume that malicious terminals randomly select their attacking indexes in the range of 0.5–1.5. The attacking index information for each malicious route is shown in Fig.9, and the CDF curves after the application of the security algorithm are shown in Fig.10.

As shown in Fig.10, when the malicious terminals perform a static attack under the independent strategy, DLM based on IDW has the best performance, followed by DLM based on spatial correlation. Algorithms that consider historical reliability outperform those that only consider real-time similarity. In addition, the algorithms that consider spatial information outperform those that do not.

The estimation accuracy is measured, as shown in Table 4, which also presents the details of the accuracy index, such as the error and ΔI. The error is defined as follows,

$$e = \frac{1}{D}\sum_{i=1}^{D}(\overline{P}_{true}(i) - \overline{P'}(i)), \tag{15}$$

where $\overline{P}_{true}(i)$ [dBm] represents the average received power in the ith mesh, $\overline{P}(i)$ [dBm] represents the estimated average

TABLE 4. Accuracy index under independent static attack.

	error	ΔI
histo+IDW	0.0426	17.0318
histo+corr	0.0527	21.0957
histo only	0.0579	23.1706
sim+IDW	0.0678	27.1133
sim+corr	0.0756	30.2412
sim only	0.0800	32.0080
bi_weight	0.0599	23.9570
average combination	0.8600	100.8269
IDW only	1.9370	774.7825
corr only	1.9514	780.5703
with all terminals	1.9417	776.6994

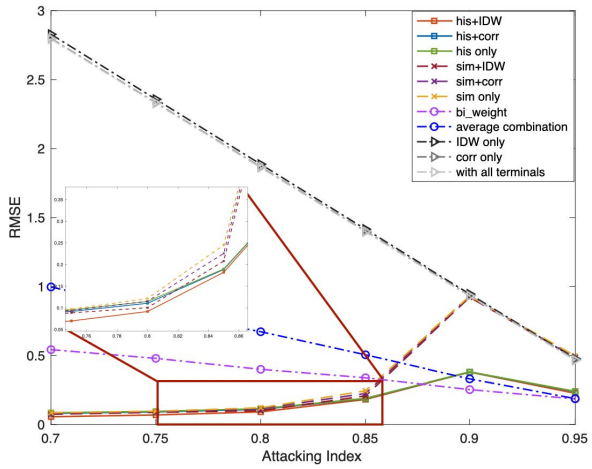
received power in the ith mesh, and D represents the number of meshes in the REM. ΔI represents the maximum interference difference between reality and estimation. The calculation method is as follows,

$$\Delta I = \sum |I_{\max\_real} - I_{\max\_est.}|, \tag{16}$$

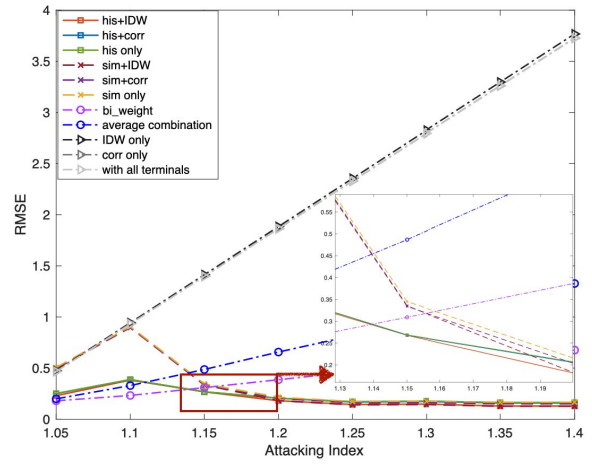
where $I_{\max\_real}$ and $I_{\max\_est.}$ represent the real and estimated values of maximum interference, respectively. $I_{\max}$ represents the maximum interference that primary users can withstand in a certain communication area, as reference [3] shown, it can be calculated as follows,

$$I_{\max} = P'(i) - \Gamma_d - \sqrt{2}\sigma_s \mathrm{erf}^{-1}(1 - 2p_{out}), \tag{17}$$

where Γ_d represents the desired signal-to-interference power ratio (SIR), p_{out} represents the desired outage probability, and σ_s represents the standard deviation of shadowing. As indicated by Table 4, the error and ΔI of the DLM based on IDW are the smallest among all the methods, followed by the DLM based on spatial correlation, which is better than using only the DLM. Adding historical information is better than using only the similarity comparison, and the use of spatial information improves the performance. When the security algorithm is not used, using only spatial information cannot improve the accuracy of REM, even if it can lead to a larger error. This is because the interpolated points may



(a) $\delta < 1$



(b) $\delta > 1$

FIGURE 11. RMSE under different attacking index.

use malicious information for estimation, which can lead to a larger error than that without interpolation. However, after the security algorithm is used even in the first layer, the database can remove most of the malicious information, and interpolation improves the accuracy.

2) UNDER COLLABORATIVE STRATEGY

Compared with the independent strategy, when malicious terminals collaborate, they cannot individually make an attack decision. Malicious terminals can select the same attacking index and report incorrect information to the database by obeying the same rule. Because this subsection focuses on the static attack strategy, all malicious terminals select the same attacking index and always report incorrect information to the database during their movements. We estimated the accuracy using the root-mean-square error (RMSE), which can be calculated as follows,

$$\mathrm{RMSE} = \sqrt{\frac{1}{D}\sum_{i=1}^{D}(\overline{P}_{true}(i) - \overline{P'}(i))^2}. \tag{18}$$

Fig.11(a) shows the RMSE under different attacking indexes. Here, we only show the attacking index in the range of 0.7–0.95. When $\delta \leq 0.8$, the RMSEs of the algorithms based on the DLM and similarity comparison are small, because when $\delta(\delta < 1)$ decreases, $Bias(i)$ increases. With a larger $Bias(i)$, $Rel(i)$ can be lower, thus, malicious terminals are easily distinguished by the database even using only one layer. When $\delta > 0.85$, the detection capability of similarity comparison becomes significantly lower than that of the DLM. When $\delta > 0.9$, the RMSE for using the similarity comparison algorithm is almost identical to that without the security algorithm, because when the attacking strength is sufficiently small, the malicious terminals cannot be distinguished by only using similarity comparison, and the use of historical information can enhance the detection performance. Additionally, the RMSE decreases after $\delta > 0.9$ because the attacking strength is too small, even they cannot be detected perfectly by the database, the error caused by them is small. For the curves obtained without using the security algorithm, when the attacking index δ increases, the attacking strength decreases, thus, the RMSE decreases. Deserve to be mentioned, as Fig.11(b) shows, when $\delta > 1$, the RMSE curves should be opposite, because when $\delta(\delta > 1)$ increases, the attacking strength increases.

E. UNDER DYNAMIC ATTACK

As mentioned previously, a dynamic attack is a smart attack strategy in which malicious terminals can change their attacking parameters during their movements. They can change their attacking indexes and even choose their attacking slots, thus, they can select a certain step to perform an attack and certain steps to not attack to protect themselves. In this case, the attacking behavior of malicious terminals is not fixed, and they can change it as often as needed. We classified dynamic attacks as independent or collaborative.

1) UNDER INDEPENDENT STRATEGY

In a dynamic attack under an independent strategy, malicious terminals can make the attacking decision by themselves and can set the attacking parameters by themselves, including selecting and changing their attacking indexes individually. They can also decide whether or not to attack at each step. In this paper, the steps where malicious terminals do not attack are denoted as the "Silent Mode," and the steps where they do attack are denoted as the "Active Mode." The silent mode refers to the slots in which malicious terminals act as honest terminals. In this mode, they should send the correct information to the database to protect themselves and to avoid detection. Thus, in silent mode, $\delta = 1$.We also assume that the attacking index ranges from 0.5 to 1.5, because from the foregoing analysis, when the attacking strength is too high, malicious terminals are easier to find. In addition, each malicious terminal randomly selects its changing points from its entire journey. The attacking-parameter conditions are shown in Fig.12, here, malicious route indicates the different

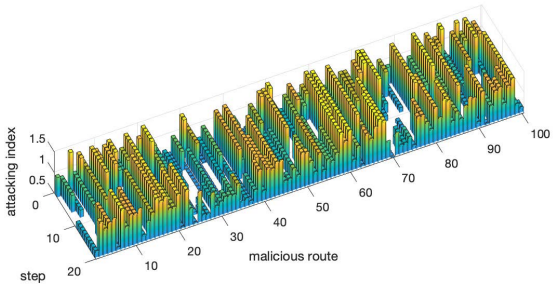


FIGURE 12. Attacking parameter condition.

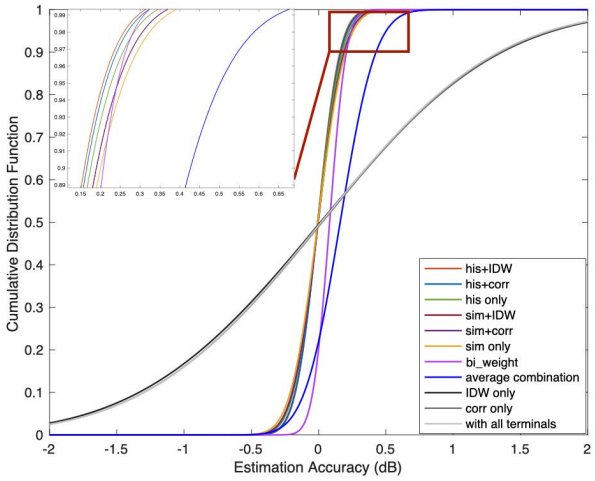


FIGURE 13. CDF of dynamic attack under independent.

TABLE 5. Accuracy index under independent dynamic attack.

	error	ΔI
histo+IDW	0.0572	22.8755
histo+corr	0.0658	26.3096
histo only	0.0706	28.2212
sim+IDW	0.0830	33.2120
sim+corr	0.0906	36.2272
sim only	0.0942	37.6802
bi_weight	0.0817	32.6852
average combination	0.1604	64.1722
IDW only	0.7889	315.5608
corr only	0.7789	311.5664
with all terminals	0.7804	312.1486

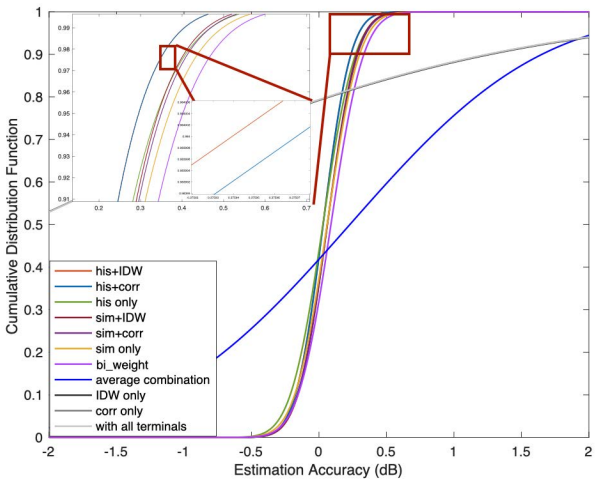


FIGURE 14. CDF of dynamic attack under collaborative.

numbering of malicious terminals. Additionally, the CDF curves after the application of the security algorithm are shown in Fig.13.

Fig.12 shows the attacking parameter. The bar height indicates the attacking index. In this figure, to show the attacking slot clearly, the attacking index in the silent mode period is set as 0, but in fact it should be 1. Thus, in Fig.12, for the steps that do not have the data, the malicious terminals act as honest terminals, sending the correct information and not performing an attack. Each terminal selects the attacking index randomly and sets the silent mode and active mode randomly. Under this attack strategy, the performance of each algorithm is shown in Fig.13. The estimation accuracy is presented in Table 5. The algorithms using the historical information outperform those using only similarity information, and they both outperform those not using the security information. The security algorithms based on the spatial information outperform those that do not consider the spatial information. Among the algorithms, DLM based on IDW achieves the best performance.

2) UNDER COLLABORATIVE STRATEGY

Compared with the independent attack strategy, the collaborative strategy is a group attack in which malicious terminals attack synchronously. We can consider them as having the same brain, and they share the same attacking parameters. After the head of the malicious terminal selects the attacking index and slot, all the other terminals follow. In this section, we let the head of the malicious terminals be the first malicious terminal and implement an independent attack strategy. We set the attacking-parameter condition as the 9^{th} malicious route shown in Fig.12, the silent mode lasts from step 9 to step 11, and the attacking indexes in the active mode are 0.979523385210219 and 1.49949162009770, respectively. All other terminals performed the same attack as the head. The CDF curves are shown in Fig.14, and the estimation accuracy is presented in Table 6. According to the simulation results, DLM based on IDW exhibits the best performance, and adding historical information can improve the estimation accuracy.

F. UNDER OPTIMAL ATTACK

We generated the condition under the white-box optimal attack strategy to comprehensively evaluate the performance. As mentioned in Section IV-B, if the malicious terminals have an attack center that can obtain the full information of the security network, including the algorithm operation and

TABLE 6. Accuracy index under collaborative dynamic attack.

	error	ΔI
histo+IDW	0.0679	27.1583
histo+corr	0.0709	28.3772
histo only	0.0788	31.5001
sim+IDW	0.0884	35.3407
sim+corr	0.0934	37.3633
sim only	0.0956	38.2493
bi_weight	0.0887	35.4679
average combination	0.2285	91.3849
IDW only	2.3621	944.8350
corr only	2.3636	945.4427
with all terminals	2.3590	943.6062

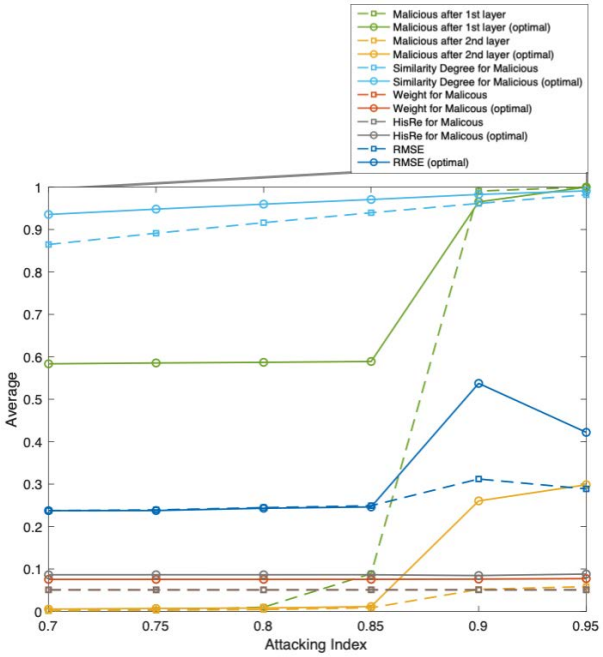


FIGURE 15. Attacking performance under different attacking index.

parameter setting, they can act as a database and simulate the reliability of each malicious terminal. In this case, the types of attacks become complex and diverse, and the damage to the security network can be severe. The attacking index δ has a trade-off problem between the detection probability and attacking strength, its value should satisfy Equation (12). Additionally, the attack strategy must consider *HisRe* to ensure that malicious terminals can join the power fusion.

Fig.15 shows the attacking performance under different attacking indexes. The dashed lines of different colors correspond to normal attacking, where the malicious terminals attack all the time. The solid line of different colors correspond to optimal attacking, where the historical reliability is monitored. As long as the *HisRe* is below the threshold, the malicious terminals should change their mode to the silent mode to increase the reliability, to ensure that the wrong information can join the power fusion, leading to an error in the REM. When the *HisRe* is sufficient to join the power fusion, they should set the mode back to the active mode. This pattern should be repeated.

In Fig.15, the green color indicates the percentage of malicious terminals that are not detected after the first layer (similarity comparison). A larger percentage indicates a better attacking performance. The simulation results show that the optimal attack strategy is better than a normal strategy. Similarly, the yellow color indicates the percentage of malicious terminals that are not detected after the second layer (DLM). Again, the optimal attack exhibits better performance. Additionally, when the attacking index is less than 0.8, the malicious terminals are easily found even when using similarity comparison, because the malicious information differs significantly from the original information. When the attacking index is greater than 0.9, the malicious terminals are not easily distinguishable even using DLM, because the difference between the malicious information and the original information is small. Considering that different terminals suffer from different channel conditions, such as path loss and fading, the algorithm cannot recognize that the difference is caused by the malicious action or channel difference. The light-blue color indicates the average similarity degree for malicious terminals, a higher value suggests that it is more difficult for the terminal to be identified as malicious. The red and grey colors represent the weight allocation and historical reliability results, respectively, for malicious terminals. The historical reliability is similar to the similarity degree, a higher value indicates a lower probability to be detected. The dark-blue color indicates the RMSE under different attacking indexes. The attacking performance is optimal when the attacking index is approximately 0.9. As the attacking index increases, the RMSE decreases because the attacking strength reduced. For example, when the attacking index is infinitely close to one, the error of the REM is very small even without the use of an algorithm.

Figs.16 and 17 show the historical reliability under the optimal attack with an attacking index $\delta = 0.9$. The former shows the total map for all the malicious routes, and the latter shows the map for the first six malicious routes. The grey bar indicates that the malicious terminal is performing an attack at that step. The absence of a grey bar indicates that the malicious terminal acts as an honest terminal at that step. The orange bar indicates the historical reliability for each step. Because the historical-reliability calculation results for each terminal are obtained at the end of the step and affect the next step, we can determine when the malicious terminal changes to the silent mode. The historical reliability increases slowly after this point. After the malicious terminal performs an attack, the historical reliability decreases rapidly.

Fig.18 shows the CDF curves under different attacking indexes obtained using the DLM, and Table 7 presents the errors under different attacking indexes for different algorithms. The "DLM(optimal)" column presents the error when attackers perform the optimal attack under the DLM algorithm, the "DLM" column presents the error when attackers perform the normal attack under the DLM algorithm, and the rightmost column presents the error when no algorithms are used. As indicated by the figure and table,

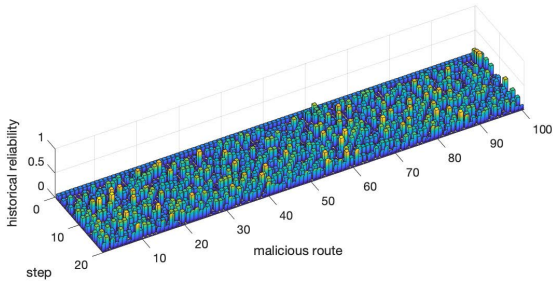


FIGURE 16. Historical reliability under optimal attack.

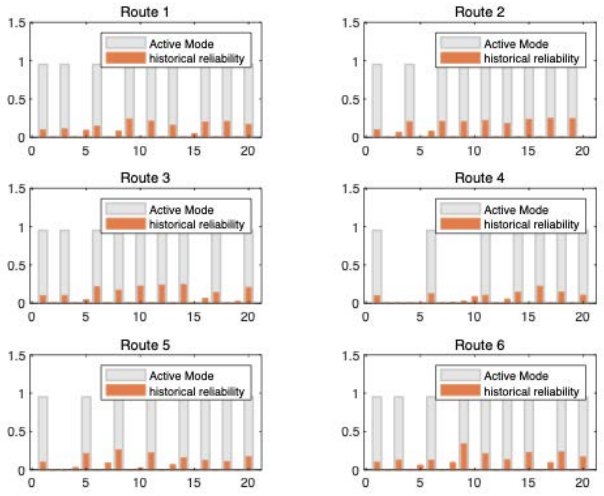


FIGURE 17. Attacking condition of first 6 malicious routes.

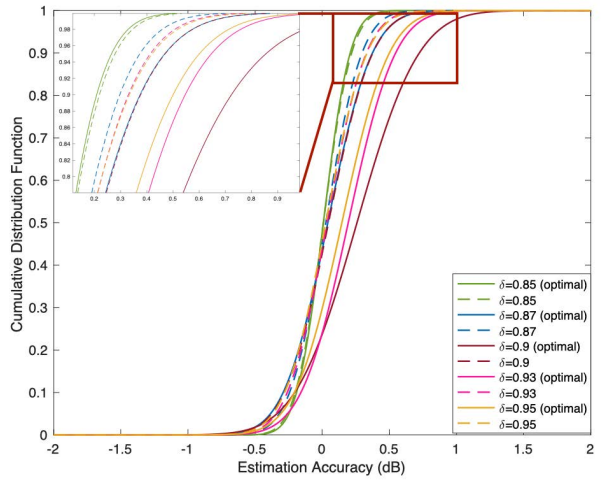


FIGURE 18. CDF under different attacking index.

the optimal attack strategy results in the largest error for the REM, but even under this type of attack, the error of DLM still less than 0.3dB for each mesh.

G. DISCUSSION

From the multiple simulation results, we found that the DLM has significant advantages for data falsification attacks. After historical information is added, the database can continually monitor the behavior of malicious terminals, and the performance improves significantly after the historical reliability is calculated. In addition, we considered using spatial information to improve the algorithm. The simulation results indicated that the DLM based on IDW is better than the DLM based on spatial correlation, and both have better performance than the DLM alone. Finally, we used the full knowledge of the network and launched the optimal attack, and the results indicated that the optimal attack has significant advantages (the error of the REM increases significantly), nonetheless, our security algorithm can ensure the accuracy of the network.

TABLE 7. Accuracy index under optimal attack.

δ	error		
	DLM(optimal)	DLM	with all terminals
0.85	0.0606	0.0621	**1.2877**
0.87	0.0867	0.0745	1.1160
0.90	**0.2888**	**0.0974**	0.8585
0.93	0.2239	0.0856	0.6009
0.95	0.1781	0.0834	0.4292

VI. CONCLUSION

We proposed a DLM based on spatial-information algorithms, including IDW and spatial correlation, to deal with various data falsification attacks in the network. Additionally, we improved the reward–penalty function for the DLM algorithm to ensure that historical reliability decreases rapidly when terminals are malicious and increases slowly when terminals are honest. We evaluated our algorithm under different attack scenarios, including static, dynamic, independent, and collaborative attacks, and the simulation results indicated that our algorithm performed well in removing malicious information and increasing the accuracy of REM. Additionally, according to our algorithm, we defined the optimal attack model and attempted to increase the error of the network. The simulation results indicated that the proposed attack strategy outperformed the normal strategy. More importantly, our security algorithm was effective under the optimal attack conditions.

REFERENCES

[1] *TV White Spaces: Pilot Database Provider Contract*, Office Communication, London, U.K., Feb. 2014.
[2] J. Perez-Romero, A. Zalonis, L. Boukhatem, A. Kliks, K. Koutlia, N. Dimitriou, and R. Kurda, "On the use of radio environment maps for interference management in heterogeneous networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 184–191, Aug. 2015.
[3] K. Sato and T. Fujii, "Kriging-based interference power constraint: Integrated design of the radio environment map and transmission power," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 1, pp. 13–25, Mar. 2017.
[4] K. Sato, K. Inage, and T. Fujii, "Modeling the Kriging-aided spatial spectrum sharing over log-normal channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 749–752, Jun. 2019.
[5] K. Sato, K. Inage, and T. Fujii, "On the performance of neural network residual Kriging in radio environment mapping," *IEEE Access*, vol. 7, pp. 94557–94568, 2019.
[6] K. Sato, K. Inage, and T. Fujii, "Frequency correlation of shadowing over TV bands in suburban area," *Electron. Lett.*, vol. 54, no. 1, pp. 6–8, Jan. 2018.

[7] K. Katagiri, K. Sato, and T. Fujii, "Crowdsourcing-assisted radio environment database for V2V communication," *Sensors*, vol. 18, no. 4, p. 1183, Apr. 2018.
[8] S. Men, P. Charge, and S. Pillement, "A robust cooperative spectrum sensing method against faulty nodes in CWSNs," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 334–339.
[9] J. Ren, Y. Zhang, Q. Ye, K. Yang, K. Zhang, and X. S. Shen, "Exploiting secure and energy-efficient collaborative spectrum sensing for cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6813–6827, Oct. 2016.
[10] Y. Gao, M. Diao, and T. Fujii, "Sensor selection based on dempster-shafer evidence theory under collaborative spectrum sensing in cognitive radio sensor networks," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–7.
[11] F. Ye, X. Zhang, and Y. Li, "Comprehensive reputation-based security mechanism against dynamic SSDF attack in cognitive radio networks," *Symmetry*, vol. 8, no. 12, p. 147, Dec. 2016.
[12] J. Ren, Y. Zhang, Q. Ye, K. Yang, K. Zhang, and X. S. Shen, "Exploiting secure and energy-efficient collaborative spectrum sensing for cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6813–6827, Oct. 2016.
[13] P. K. Varshney, *Distributed Detection and Data Fusion*. New York, NY, USA: Springer, 1997.
[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.
[15] B. Chen and P. K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 693–699, Feb. 2005.
[16] Q. H. Peng, K. Zeng, J. Wang, and S. Q. Li, "A distributed spectrum sensing scheme based on credibility and evidence theory in cognitive radio context," in *Proc. 17th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (IEEE PIMRC)*, Piscataway, NJ, USA, Sep. 2006, pp. 2511–2515.
[17] N. Nguyen-Thanh and I. Koo, "An enhanced cooperative spectrum sensing scheme based on evidence theory and reliability source evaluation in cognitive radio context," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 492–494, Jul. 2009.
[18] Y. Han, Q. Chen, and J. X. Wang, "An enhanced D-S theory cooperative spectrum sensing algorithm against SSDF attack," in *Proc. 75th IEEE Veh. Technol. Conf. (IEEE VTC Spring)*, Piscataway, NJ, USA, May 2012, pp. 1–5.
[19] N. Nguyen-Thanh and I. Koo, "Evidence-theory-based cooperative spectrum sensing with efficient quantization method in cognitive radio," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 185–195, Jan. 2011.
[20] S. Jana, K. Zeng, W. Cheng, and P. Mohapatra, "Trusted collaborative spectrum sensing for mobile cognitive radio networks," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 9, pp. 1497–1507, Sep. 2013.
[21] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.
[22] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 34–36, Jan. 2007.
[23] R. Zhang, J. Wei, D. G. Michelson, and V. C. M. Leung, "Outage probability of MRC diversity over correlated shadowed fading channels," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 516–519, Oct. 2012.
[24] R. Wan, M. Wu, L. Hu, and H. Wang, "Energy-efficient cooperative spectrum sensing scheme based on spatial correlation for cognitive Internet of Things," *IEEE Access*, vol. 8, pp. 139501–139511, 2020.
[25] D. Giancristofaro, "Correlation model for shadow fading in mobile radio channels," *Electron. Lett.*, vol. 32, no. 11, pp. 958–959, 1996.
[26] Y. Gao and T. Fujii, "Improvement of radio environment map under data falsification attack," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2021, pp. 1–6.
[27] A. T. S. Dhevi, "Imputing missing values using inverse distance weighted interpolation for time series data," in *Proc. 6th Int. Conf. Adv. Comput. (ICoAC)*, Dec. 2014, pp. 255–259.
[28] Y. Cao and X. Guan, "A class of constrained inverse bottleneck optimization problems under weighted Hamming distance," in *Proc. Int. Joint Conf. Comput. Sci. Optim.*, Apr. 2009, pp. 859–863.
[29] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Malicious user detection in a cognitive radio cooperative sensing system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2488–2497, Aug. 2010.
[30] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2008, pp. 3406–3410.

YING GAO was born in Harbin, China, in 1994. She received the B.E. and M.E. degrees from the College of Information and Communication Engineering, Harbin Engineering University, China, in 2016 and 2020, respectively. She is currently pursuing the Ph.D. degree with the Advanced Wireless and Communication Research Center, The University of Electro-Communications, Japan. She is supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Scholarship. Her current research interests include radio environment map, spectrum sensing, and spectrum sharing.

TAKEO FUJII (Member, IEEE) was born in Tokyo, Japan, in 1974. He received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1997, 1999, and 2002, respectively. He was a Research Associate with the Department of Information and Computer Science, Keio University, from 2000 to 2002. From 2002 to 2006, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. From 2006 to 2014, he was an Associate Professor with the Advanced Wireless and Communication Research Center, The University of Electro-Communications, where he is currently a Professor. His current research interests include cognitive radio and *ad hoc* wireless networks. He was a recipient of the Best Paper Award for the IEEE VTC 1999-Fall; the 2001 Active Research Award in Radio Communication Systems from the IEICE Technical Committee of RCS; the 2001 Ericsson Young Scientist Award; the Young Researcher's Award from IEICE, in 2004; the Young Researcher Study Encouragement Award from the IEICE Technical Committee of AN, in 2009; the Best Paper Award for IEEE CCNC 2013; and the IEICE Communication Society Best Paper Award, in 2016.