MDPI

Article

AIPI: Network Status Identification on Multi-Protocol Wireless Sensor Networks

Peng Jiang [1,2], Xinglin Feng [1], Renhai Feng [1,*] and Junpeng Cui [3]

[1] School of Electrical and Information Engineering, Weijin Road Campus, Tianjin University, Nankai District, Tianjin 300072, China; jiangpeng@motimo.com.cn (P.J.); fengxinglin@tju.edu.cn (X.F.)

[2] Tianjin Motimo Membrane Tech. Co., Ltd., Tianjin 300457, China

[3] Software and Communication School, TianjinSino-German University of Applied Sciences, Tianjin 300350, China; cuijunpeng@tsguas.edu.cn

[*] Correspondence: fengrenhai@tju.edu.cn; Tel.: +86-176-9554-1156

Abstract: Topology control is important for extending networks lifetime and reducing interference. The accuracy of topology identification plays a crucial role in topology control. Traditional passive interception can only identify the connectivity among cooperative sensor networks with known protocol. This paper proposes a novel method called Active Interfere and Passive Interception (AIPI) to identify the topology of non-cooperative sensor networks by using both active and passive interceptions. Active interception uses full duplex sensors to disrupt communication until frequency hopped to acquire distance information, and thus, infer their connectivity and calculate the location after modifying error in a non-cooperative sensor network. Passive interception uses Granger causality to infer the connectivity between two communication nodes after getting the time frame structure in physical layer. Passive interception is applied to conserve power consumption after obtaining physical information via active interception. Simulation results indicate that AIPI can identify the topology of non-cooperative sensor networks with a higher accuracy than traditional method.

Keywords: topology identification; active interfere; frequency hopping; passive interception; Granger causality

1. Introduction

Wireless Sensor Networks (WSN) consisting of a large number of self-organizing nodes are widely deployed in environmental monitoring, smart homes, and healthcare. In these applications, network topology plays a crucial role. Topology not only affects capacity and energy consumption but also directly impacts the data collection and reliability of the network. Some algorithms including routing algorithms [1], data aggregation algorithms, load balancing, and some data aggregation algorithms [2] need to know the topology to ensure optimal performance. Therefore, identifying the topology of WSN is meaningful for maintaining their operation. Traditional passive interception methods include the correlation-based method [3], information-theory-based method [4], Granger-causality-test-based method [5], compression-sensing-based method [6], drive-response-based method [7], and graph-based neural network method [8]. Passive interception obtains the connections between nodes by capturing and decoding data packets, and it is susceptible to channel noise. The existing literature demonstrates that passive interception can identify the topology of sensor network assuming known protocol [3–8], in which the type and structure of the data package is known. Unlike cooperative WSN, the communication protocols

of non-cooperative WSN are unknown. Such assumption limited their application scope under non-cooperative WSN.

In response to non-cooperative networks, Xu et al. [9] first introduced the concept of active interception, utilizing full-duplex on interception device. This method can interfere signal reception of the suspicious receiver, thus intercepting suspicious information, and thereby, enhancing eavesdropping. Subsequently, some scholars have further applied the strategy of Xu. Hui et al. [10] studied the problem of maximizing the average monitoring rate of a lawful monitoring system which actively intercepts on suspicious wireless links through interfering Rayleigh fading channels. Yunlong et al. [11] investigated millimeter-wave information monitoring systems, where suspicious transmitters in the network send messages to suspicious receivers under the supervision of a monitoring controller, achieving lawful surveillance of suspicious links. In order to increase the probability of successful intercepting, Yitao et al. [12] proposed cognitive interference to alter the long-term trust of the suspicious links under parallel channels, thereby encourage the links to transmit on a smaller subset of unblocked channels with a lower transmission rate. Hu et al. [13] investigated active intercepting of half-duplex lawful monitor in a wireless power supply multi-channel suspicious system, where the suspicious transmitters harvest energy from power beacons and then communicate with the suspicious destination through parallel channels. Additionally, single-hop relay networks [14], unmanned aerial vehicle (UAV)-assisted downlink multicast [15], broadcast and uplink multi-access [16], multiuser communication [17], collaborative monitoring systems [18], and intelligent reflecting surface-assisted intercepting networks [19] all applied active interception. Compare to passive interception, active interception exchange the information through interfering actively, which avoids decoding data package. In order to identify the topology of sensor network accurately, combining localization and active interception is a promising direction. The location of the communication nodes provide extra information about topology. Existing localization algorithm include range-dependent [20–22] and range-independent [23,24]. These techniques often require additional hardware and consume significant resources. Recent advancements in sensing technologies [25] and distributed systems [26,27] offer potential solutions to these challenges. To address the limitations of these methods, the level of the power of signal from interfering device can used to infer position information (PI).

Therefore, this paper combines active interception and local information extraction to identify the topology of sensor network with unknown communication protocol and use passive interception with the Granger causality test to identify the topology of sensor network while the communication protocol is decoded. The topology sensing model shown in Figure 1 is considered in this paper, where three full duplex sensors transmit the information from target WSN to fusion center. Fusion center can apply active interception or passive interception in different stage. Figure 2 illustrates the overall framework of the AIPI method. In the initial stage, the AIPI method applies active interception to identify topology and obtain physical layer information. To conserve power consumption, the AIPI method applies passive interception to track dynamic topology constantly.

The three main contribution of this paper are as follows:

(1) Active interception identifies topology through interfering node communication with an energy-effective dichotomy;

(2) Local information extraction with self-adaptive error distribution provides extra effective information for active interception;

(3) Passive interception with the Granger causality test is proposed to track topology of WSN with physical layer information.

Traditional passive interception such as the Granger causality test [28] and Hawkess process [29] have advantages regarding power consumption because they only need to

sense the signal. However, they are susceptible to noise and are only applicable to cooperative WSN. AINL [30] is applicable to non-cooperative and cooperative WSN, but it can only extract the node PI with dichotomy in a small noise environment. AIPI can not only extract node PI with energy effective dichotomy, but it is also insensitive to noise in non-cooperative WSN, which means that it has superior performance on universality, ennergy efficiency, and anti-interference. AIPI applies active and passive interception at different stages, causing higher power consumption than pure passive interception and less than pure AINL. A comparison of AIPI with existing works on topology inference is shown in Figure 3.

The rest of this paper is organized as follows. In Section 2, we introduce the use of active interference for local information extraction, and explain self-adaptive error correction mechanism in our proposed method using the maximum likelihood function. Finally, we apply passive interception to infer topology based on Granger causality. The simulation results are given in Section 3 and the conclusions for the work are provided in Section 4.
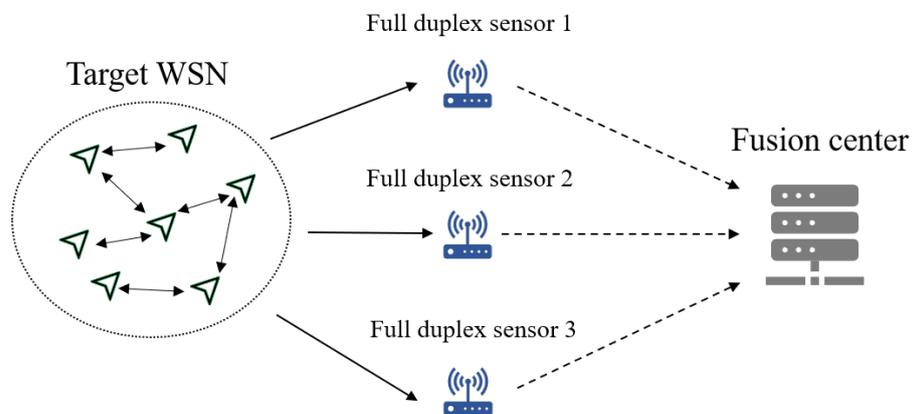
Figure 1. The system model of topology sensing.
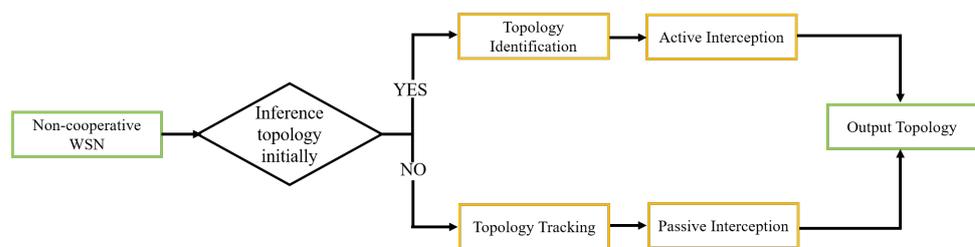
Figure 2. Overall framework of AIPI.

	Optimization variable				
	Energy efficiency	Power	Location	Universality	Anti-interference
[28]		✓			
[29]		✓			
[30]	✓		✓	✓	
AIPI	✓		✓	✓	✓

Figure 3. The comparison of AIPI with existing works on topology inference.

2. Theoretical Basis

2.1. Problem Formulation

In mathematical notation, a network graph can be represented by $g = \{v, \sigma\}$, where $v = \{1, 2, \cdots N\}$ denotes the set of vertices, N is the number of elements in v, that is,

the number of nodes in the network, and σ denotes the set of edge. The element in the adjacency matrix A of the network can be represented as (1),

$$a_{i,j} = \begin{cases} 1, & (i,j) \in \sigma \\ 0, & (i,j) \notin \sigma \end{cases} \tag{1}$$

where $i, j \in N$. We formulate the topology inference problem as (2),

$$\min \|X^- - X\|_2 + \|A^- - A\|_1 \tag{2}$$

where X^- represents estimated node coordinate, X represents real node coordinate, and A^- represents the inference topology.

2.2. Active Interception

We use active interception to identify the topology of WSN with unknown protocol. Assuming WSN has N communication nodes, while all the nodes can transmit and 20% nodes can receive. In the target network, different node pairs communicate using different frequency bands, preventing collisions. Three full duplex sensors locate the target node by colliding on the corresponding frequency bands. As shown in Figure 4, T denotes the transmitter, R denotes the receiver, and EVE nodes represent three interfering nodes. T and R are communicating at X band, and then three full duplex sensors EVE1, EVE2, and EVE3 send interference signals with different power levels until the frequency hopped in communication. PI and connectivity are deduced through trilateration after transforming power level into distance. We can distinguish different communication node when the error between the real coordinate and the estimated coordinate is less than a threshold.
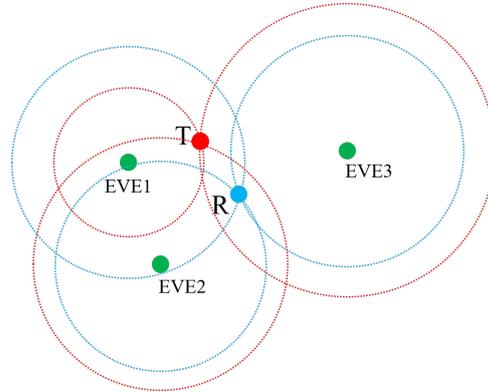
Figure 4. Full duplex sensors extract transmitter and receiver position information.

Algorithm 1 explains the strategy of power adapting for EVE. Define the upper and lower bounds on EVE transmission pseudo-noise power as $P_{\max}$ and $P_{\min}$, the transmission distance is bounded by $R_{\max}$ and $R_{\min}$. Define the R_{real} to be the real distance between R and EVE. Change power of EVE R_1 through dichotomy coefficient until the error between $R_1^{\max}$ and $R_1^{\min}$ is less than error δ. Apply same strategy to T can also obtain the coordinate.

The error of distance between R and EVE ϵ will inevitably occur when using the dichotomy to calculate it. The ϵ lead to the three circles cannot intersect at a single point based on trilateration as shown in Figure 5. The error is actually subtle, and thus, for interpreting, we enlarge it. In this case, we consider the midpoint (green) of the line connecting the two closest intersection points (red) as the result of local information extraction.

Algorithm 1 Interference Strategy of Active Interception.

1: **Input:** dichotomy coefficient α, error δ, R_{real}
2: **Initialization:** $R_1^{\min} = R_{\min}, R_1^{\max} = R_{\max}$
3: Deploy transmitting node, receiving node, and EVE
4: EVE transmits interference signal
5: **while** $R_1^{\max} - R_1^{\min} > \delta$ **do**
6:     $R_1 = R_1^{\min} + \alpha(R_1^{\max} - R_1^{\min})$
7:     **if** $R_1 < R_{\text{real}}$ **then**
8:         set $R_1^{\min} = R_1$
9:     **else**
10:         set $R_1^{\max} = R_1$
11:     **end if**
12: **end while**
13: Calculate node coordinate with Formula (4)
14: **Output:** (x, y)

It is impossible to extract the local information with a very small error δ because of the environment noise. To extract the local information with a bigger error δ accurately, we deploy data from train sample with same strategy to extract local information of unknown WSN with self-adaptive error. Assuming a WSN whose PI is known, we use (x, y) represent the result of local information extraction through trilateration and (x', y') represent the real PI, so that the error between them can be denoted as $(\Delta x, \Delta y) = (x' - x, y' - y)$. Calculate the errors of coordinates for all communication nodes. Since the coordinates of the nodes are represented in two dimensions, we use two Gaussian functions to fit the obtained errors through Algorithm 2.

The Gaussian Mixture Models can be represented as a linear combination of Gaussian distributions as shown in (3):

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid u_k, \Sigma_k), \tag{3}$$

where $\mathcal{N}$ follows a Gaussian distribution, u_k and Σ_k are the mean and covariance matrix corresponding of the Gaussian distribution, and π_k is the probability of each Gaussian distribution with π_k satisfying $\sum_{k=1}^{K} \pi_k = 1$. Solve the variable of Gaussian distribution $n_1 \sim \mathcal{N}(u_1, \Sigma_1)$ fitting Δx and $n_2 \sim \mathcal{N}(u_2, \Sigma_2)$ fitting Δy through Expectation-Maximization (EM) algorithm.
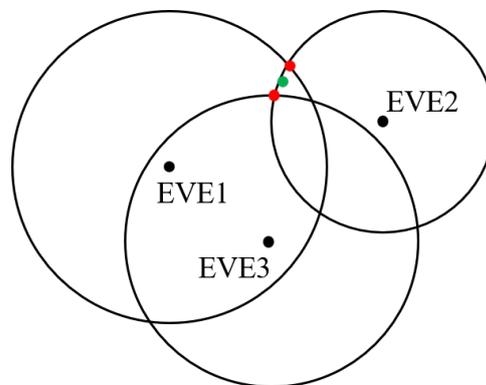
Figure 5. The error from trilateration.

Algorithm 2 EM Algorithm for Gaussian Mixture Model Parameter Estimation

---

1: **Input:** Observed dataset X, number of components k, convergence threshold η
2: **Initialization**: Randomly initialize each Gaussian model parameter θ_k (the mean μ_k, covariance matrix Σ_k, and weights π_k for each Gaussian component), ensuring that $\sum_k \pi_k = 1$
3: **while** $\theta^t - \theta^{t-1} > \eta$ **do**
4:     **E-Step (Expectation)**:
5:     **for** each observation x_n **do**
6:         Compute the posterior probability that x_n belongs to each Gaussian component:

$$p(z_{nk} \mid x_n) = \frac{\pi_k \cdot N(x_n \mid \mu_k, \Sigma_k)}{\sum_j \pi_j \cdot N(x_n \mid \mu_j, \Sigma_j)}$$

7:     **end for**
8:     **M-Step (Maximization)**:
9:     **for** each Gaussian component k **do**
10:        Update the weight:

$$\pi_k = \frac{N_k}{N}, \quad N_k \text{ is the number of data points belonging to the } k\text{-th component}$$

11:        Update the mean:

$$\mu_k = \frac{\sum_n p(z_{nk} \mid x_n) \cdot x_n}{\sum_n p(z_{nk} \mid x_n)}$$

12:        Update the covariance matrix:

$$\Sigma_k = \frac{\sum_n p(z_{nk} \mid x_n) \cdot (x_n - \mu_k)(x_n - \mu_k)^T}{\sum_n p(z_{nk} \mid x_n)}$$

13:     **end for**
14: **end while**
15: **Output:** the mean μ_k, covariance matrix Σ_k, and weights π_k for each Gaussian component

---

For acquiring PI of an unknown WSN, apply the parameter Σ and u to extract position information. The coordinates of R is (x_n, y_n). The coordinates of EVE are (x_i, y_i), and the distance between R and EVE are r_i, where $i = 1, 2, 3$. Construct a set of trilateration equations as shown in (4):

$$\begin{cases} r_1^2 = (x_1 - x_n)^2 + (y_1 - y_n)^2 \\ r_2^2 = (x_2 - x_n)^2 + (y_2 - y_n)^2 \\ r_3^2 = (x_3 - x_n)^2 + (y_3 - y_n)^2 \end{cases} \tag{4}$$

Derive (5) from (4):

$$\begin{cases} r_2^2 - r_1^2 = x_2^2 - x_1^2 - 2(x_2 - x_1)x_n + y_2^2 - y_1^2 - 2(y_2 - y_1)y_n \\ r_3^2 - r_1^2 = x_3^2 - x_1^2 - 2(x_3 - x_1)x_n + y_3^2 - y_1^2 - 2(y_3 - y_1)y_n \end{cases} \tag{5}$$

Transform (5) into matrix to obtain (6),

$$\begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix} \cdot \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}k_2^2 + \frac{1}{2}r_2^2 - \frac{1}{2}r_1^2 \\ -\frac{1}{2}k_3^2 + \frac{1}{2}r_3^2 - \frac{1}{2}r_1^2 \end{bmatrix} \tag{6}$$

where $k_i = x_i^2 - x_1^2 + y_i^2 - y_1^2$, and the matrix equation can be simplified to represent as $HX = Y$. Using (6) and the two Gaussian functions obtained from fitting the errors, construct (7):

$$Y = HX + n_1 + n_2 \tag{7}$$

Derive (8) from (7):

$$Y - HX - n_2 = n_1 \tag{8}$$

Since n_1 follows a Gaussian distribution, $Y - HX - n_2$ also follows a Gaussian distribution with the same parameters. Let M represent $Y - HX - n_2$, then the distribution of M is expressed as (9),

$$\mathcal{N}(M|u_1, \Sigma_1) = \frac{1}{2\pi} \frac{1}{|\Sigma_1|^{1/2}} \exp\left\{ -\frac{1}{2}(M - u_1)^T \Sigma_1^{-1}(M - u_1) \right\} \tag{9}$$

of which log-likelihood function, after taking the logarithm, can be expressed as (10):

$$\ln p(M|u_1, \Sigma_1) = $$
$$- \ln(2\pi) - \frac{1}{2}\ln|\Sigma_1| - \frac{1}{2}(M - u_1)^T \Sigma_1^{-1}(M - u_1) \tag{10}$$

Opening the $(M - u_1)^T \Sigma_1^{-1}(M - u_1)$, we can obtain (11):

$$
\begin{aligned}
&(Y - HX - n_2 - u_1)^T \Sigma_1^{-1}(Y - HX - n_2 - u_1) \\
&= Y^T \Sigma_1^{-1} Y - Y^T \Sigma_1^{-1} HX - Y^T \Sigma_1^{-1} n_2 - Y^T \Sigma_1^{-1} u_1 \\
&\quad - X^T H^T \Sigma_1^{-1} Y + X^T H^T \Sigma_1^{-1} HX \\
&\quad + X^T H^T \Sigma_1^{-1} n_2 + X^T H^T \Sigma_1^{-1} u_1 \\
&\quad - n_2^T \Sigma_1^{-1} Y + n_2^T \Sigma_1^{-1} HX + n_2^T \Sigma_1^{-1} n_2 + n_2^T \Sigma_1^{-1} u_1 \\
&\quad - u_1^T \Sigma_1^{-1} Y + u_1^T \Sigma_1^{-1} HX + u_1^T \Sigma_1^{-1} n_2 + u_1^T \Sigma_1^{-1} u_1
\end{aligned}
\tag{11}
$$

Since n_2 is independent, derive (12) from (10) and (11):

$$
\begin{aligned}
&\frac{\partial \ln p(M|u_1, \Sigma_1)}{\partial X} = \\
&H^T \Sigma_1^{-1} HX + (H^T \Sigma_1^{-1} H)^T X + H^T \Sigma_1^{-1} u_1 \\
&+ H^T \Sigma_1^{-1^T} u_1 - H^T \Sigma_1^{-1^T} Y - H^T \Sigma_1^{-1} Y
\end{aligned}
\tag{12}
$$

Let Equation (12) equal zero, solve for X, and thus, obtain the PI. Distinguishing the different communication nodes in WSN, we can identify the topology of WSN.

2.3. Passive Interception

After active interception distinguish communication mode in physical layer, we apply steadily passive interception to monitor the topology of WSN in a long time, which can conserve energy effectively. Assume the WSN applies Transmission Control Protocol/Internet Protocol (TCP/IP) and Automatic Repeat-reQuest (ARQ) protocol, which means that communication node will receive an Acknowledge character package (ACK) from the other communication node after transmitting data package successfully. The EVE captures the package from WSN at a period T and then acquire its transmission time and transmitter information. Divide period T time into L slots to acquire the time series. In the IEEE 802.15.4 standard, the baud rate is 250 kbps. The time slot length t should be less than the data transmission time $T_i = L_{data}/Br$, where L_{data} is the data package length and the Br is the baud rate. The optimal number of slots L can be obtain through $L = T/L_s$.

Considering the WSN is sparse and undirected, apply the Granger causality to infer two time series from two communication nodes whether exist causality relationship and then identify the connectivity. We can identify the topology of WSN after deducing all pairs of communication if there exists connectivity.

For each node $i \in v$, a time series is generated when the node transmits a data package and an ACK package. More precisely, for the set of packet transmitting times T_i, the binary time series $x_{i,I}$ is represented as (13),

$$x_{i,I} = \begin{cases} 1, IT_s < T_i < (L-I)T_s \\ 0, \text{others} \end{cases} \tag{13}$$

where i represents the node ID, I indicates the time slot index, L is the total number of time slots, and T_s represents the length of the time slots. The schematic diagram of the node's time series can be represented as shown in Figure 6.
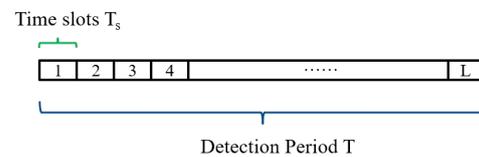
Figure 6. Schematic diagram of the time series data.

Granger causality is used to compare prediction error to determine whether two variables are related. For the time series $\mathbf{X} = \{x_1, x_2, x_3, \ldots\}$ and $\mathbf{Y} = \{y_1, y_2, y_3, \ldots\}$, if the knowledge of past values of Y contributes some additional information about x_n more than the knowledge of just past of X, we consider that Y Granger causes X. Granger causality is always modeled as a hypothesis test problem, where the null hypothesis H_0 means that Y does not Granger cause X and the alternative hypothesis H_1 means that Y Granger X. The hypothesis test is represented as (14) and (15),

$$H_0: X[n] = \sum_{i=1}^{k} a_i X[n-i] + \varepsilon[n] \tag{14}$$

$$H_1: X[n] = \sum_{i=1}^{k} b_i X[n-i] + \sum_{i=1}^{k} c_i Y[n-i] + \eta[n], \tag{15}$$

where $X[n]$ is the value of X in the nth slot, a_i and b_i are the parameters of the regression, which quantify the relative importance of the past values, k is the order of the model, and $\epsilon[n]$ and $\eta[n]$ are the errors of H_0 and H_1, respectively. If the knowledge of the past of Y cannot contribute additional information about x_n, H_0 will have less error variance $\epsilon[n]$ than H_1.

We denote the squared-sum residuals of H_0 as $RSS_0 = \sum_{i=1}^{T}(\varepsilon[i])^2$, and similarly, the squared-sum residuals of H_1 as $RSS_1 = \sum_{i=1}^{T}(\eta[i])^2$, where $T = L - k$ and L is the number of data points. Therefore, the F-statistic is given by (16),

$$G_{Y \to X} = \frac{(RSS_0 - RSS_1)/k}{RSS_1/(T-2k-1)} \sim F(d_1, d_2), \tag{16}$$

where $d_1 = k$ and $d_2 = T - 2k - 1$. H_0 will be rejected if $G_{Y \to X}$ is greater than the critical value of F-distribution. The connectivity of two nodes will be existed if the time series of them display Granger causality. We acquired all the connectivity of WSN after deducing Granger causality among all nodes.

3. Experiment and Analysis

The console object specifies the number of communication nodes, randomly generates the positions of communication nodes within a 100*100 rectangular area, and stipulates that the distance between two nodes cannot be less than a certain threshold. All communication nodes have the capability to transmit information within the specified range. Additionally, it defines the connectivity, creating channels with ten frequency bands in each one. The console is also responsible for allocating channels between communication nodes and reallocating frequency bands after frequency hopping.

3.1. Experimental Setup

Considering the small error δ of dichotomy in a known WSN, the result of fitting the coordinate error with a Gaussian mixture model is shown in Figure 7, where the vertical axis represents the error on the y-axis, and the horizontal axis represents the error on the x-axis. The two Gaussian distributions are denoted by two sets of contour lines.

We deployed the parameter to extract local information with self-adaptive error distribution. The comparison of the local information extraction error obtained through the maximum likelihood function and the local information extraction error obtained through trilateration is shown in Figure 8. The horizontal axis represents the node number. The values of the red bars indicate the distance between the estimated coordinates obtained through trilateration and the true coordinates. The values of the blue bars represent the distance between the estimated coordinates applied maximum likelihood function and the true coordinates. It is evident that the overall blue bars are lower than the red bars, which means that extract PI with the maximum likelihood function significantly reduces the error.

The console generates the coordinates of 90 communication nodes and the connectivity, as shown in Figure 9a.The visualization of the topology matrix is shown in Figure 9b, whose transverse and longitudinal axis represent the node numbers and a white block is displayed when there is a connectivity between two nodes and a black block is displayed when there is no connectivity.

The console generates channels with each channel having 10 frequency bands, and each frequency band corresponds to a different Signal-to-Noise Ratio (SNR). The corresponding relationships are shown in Table 1. The console also allocates channels and frequency bands to each pair of communication nodes as shown in Table 2.

Table 1. Channel-band-SNR allocation chart.

channel	1	1	1	1	1	1	1	1	1	1
band	1	2	3	4	5	6	7	8	9	10
SNR	114	99	90	79	123	129	178	30	172	10

Table 2. Channel and band of the communication node.

Transmitting Node	1	2	3	3	3	4
Receiving Node	39	25	6	24	42	29
Channel Number	7279	5483	1456	3665	3027	7904
Band Number	5	5	7	6	3	3

EVE draws interfering circles around the transmitting and receiving nodes. After interfering all communication nodes, the identified topology matrix visualization is shown in Figure 9c.
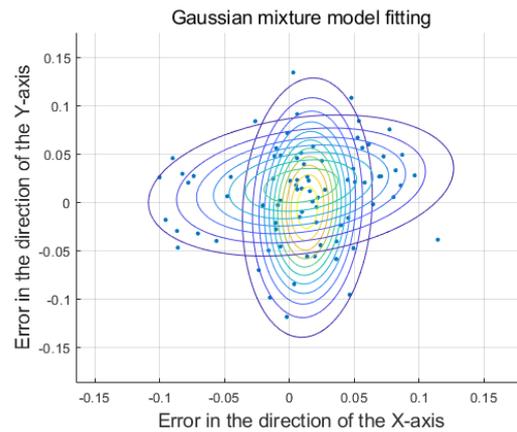
Figure 7. Gaussian mixture model fitting error.
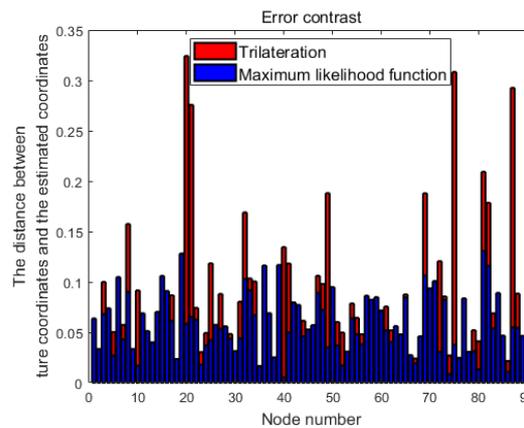
Figure 8. Comparison of PI errors.

In order to reduce the error, we extract PI with the maximum likelihood function. Deploying the extra effective information, we can identify the more precise topology, as shown in Figure 9d.

As time goes on, the topology of WSN changes, whose matrix visualization is shown in Figure 9e. We utilized the BIC model [31] to determine the maximum lag order of the time series and deployed the Granger causality test to infer whether there is a connectivity between two nodes. The identified topology matrix is shown in Figure 9f.

(**a**)

(**b**)

Figure 9. *Cont.*

(**c**)

(**d**)

(**e**)

(**f**)
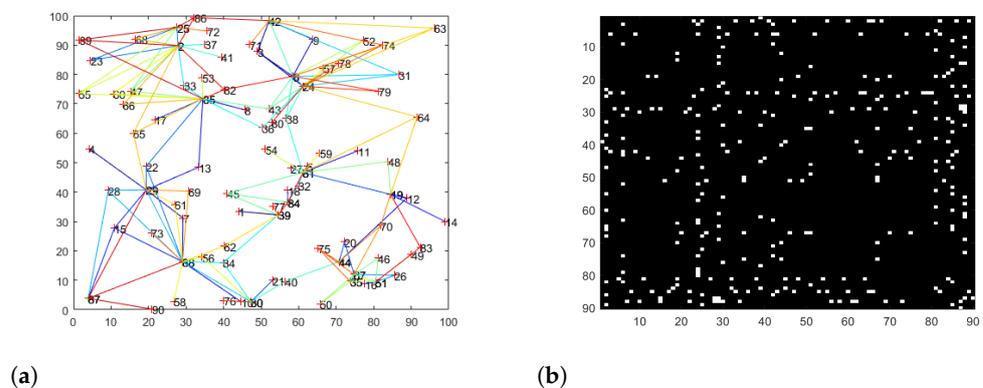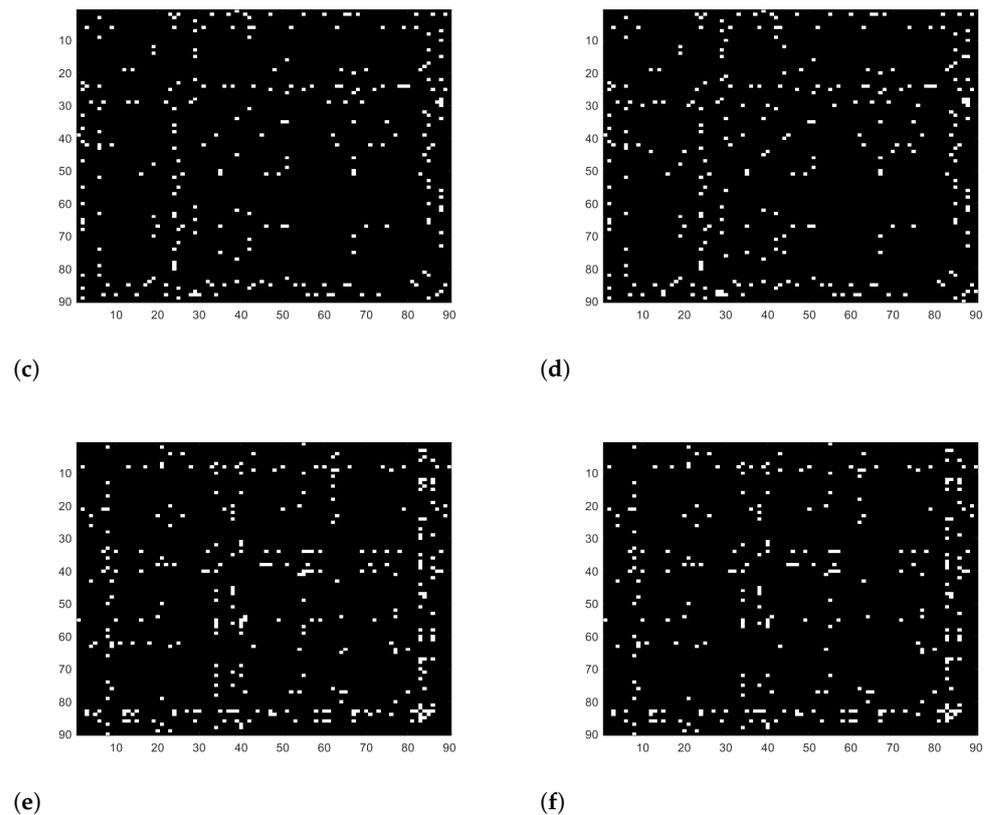
Figure 9. Topology of WSN. (**a**) Network topology diagram. (**b**) Initial topology matrix. (**c**) Topology matrix without modifying. (**d**) Topology matrix after modifying. (**e**) Dynamic topology matrix. (**f**) Dynamic topology deduced by passive interception.

3.2. Comparison Result

Based on similar framework AINL also proposed interference strategy. However, the author's simulation experiments only considered the low noise of interference strategy, and the number of simulations was relatively small. When reproducing the experiment, we increased the noise and continuously changed the positions of the communication nodes to obtain more convincing simulation results.

The number of communication nodes can be regarded as an important indicator of wireless sensor networks scale. We compared the accuracy rates of AIPI, AINL, and passive interception for network topology identification under different network scales and physical layer communication modes. We changed the network scale and the communication mode of the physical layer with Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Orthogonal Frequency Division Multiplexing (OFDM), and Code Division Multiple Access (CDMA). The accuracy rates of the identified topology are shown in Figure 10a, Figure 10b, Figure 10c, and Figure 10d, respectively. These four simulation diagrams are obtained by taking the average five repetitions, in order to avoid contingency. It can be seen from Figure 10 and Table 3 that when facing different network scales and physical layer mode, the accuracy rate of our method for network topology identification is above 99%. It has an obvious advantage over the AINL method especially in small-scale networks. The WSN network is sparse, with most nodes not being connected. As the number of nodes increases, the number of non-connected relationships far exceeds the connected ones. AINL method has low positioning accuracy, making it difficult to distinguish between different nodes, which leads to a significant drop in accuracy when dealing with small-scale networks. The passive interception method is generally inferior to

the other two methods. This is because passive interception depends on the accuracy of packet decoding. When the channel conditions are poor, the source or type of data packets cannot be accurately decoded, resulting in a low accuracy rate of topology identification. The impact of noise on the four communication modes is similar. Taking OFDM as an example, we define the network size as 40 and conduct topology identification simulations under different SNR levels. As shown in Figure 11, with the increasing of SNR, the accuracy of passive interception, including the Granger causality test and Hawkess process, shows an upward trend. AINL shows an steady performance in different SNR, while AIPI shows an better performance. We considered using four different methods to perform topology identification on the same network five times. As shown in Figure 12, with the increase of the network scale, the power consumption of four methods show the same trend. Given the premise that topology inference requires successfully deciphering non-cooperative network packets, passive interception only requires signal sensing to infer the topology, resulting in the lowest energy consumption. AINL, on the other hand, conducts active interception in every topology identification process, leading to the highest energy consumption. In contrast, AIPI performs active interception only during the initial identification process and utilizes the acquired information for passive interception in subsequent topology identification processes, resulting in moderate energy consumption.

(**a**)　　　　　　　　　　　　　　　　(**b**)

(**c**)　　　　　　　　　　　　　　　　(**d**)
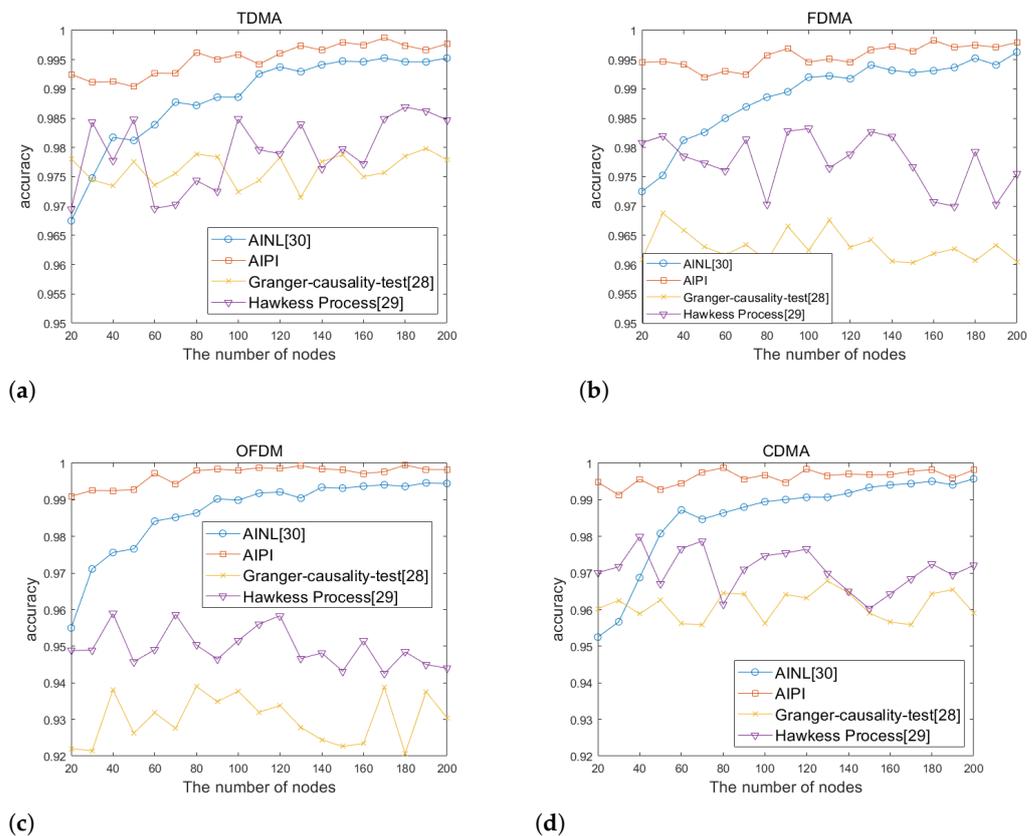
Figure 10. The accuracy rates of the identified topology with different communication modes: (**a**) TDMA, (**b**) FDMA, (**c**) OFDM, and (**d**) CDMA.

Table 3. Comparison of the average accuracy of different methods with various network scales.

Methods	Accuracy				
	15 Nodes	50 Nodes	100 Nodes	150 Nodes	200 Nodes
AIPI	99.32%	99.20%	99.63%	99.73%	99.80%
AINL	96.18%	98.03%	99.00%	99.35%	99.54%
Hawkess Process	96.45%	97.06%	97.40%	96.57%	97.13%
Granger causality test	95.53%	95.73%	95.72%	95.51%	95.69%
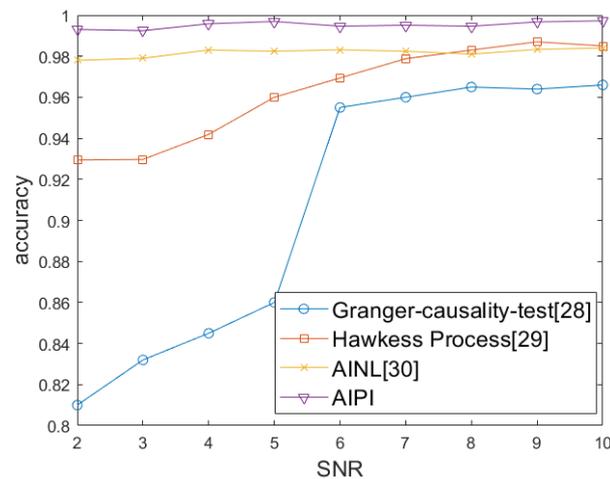
Figure 11. The accuracy rates of the identified topology with FDMA based on different SNRs.
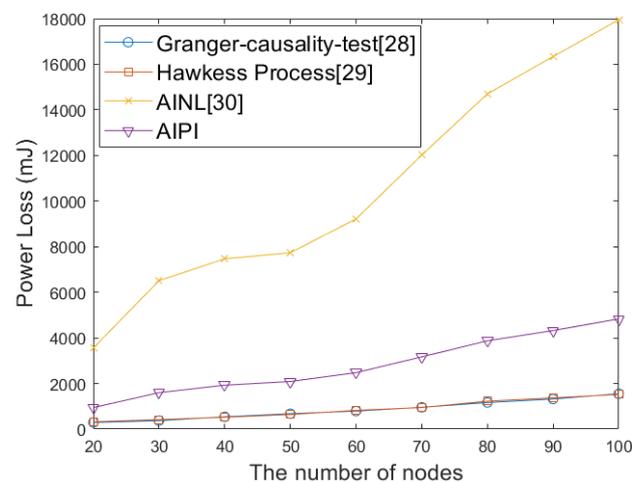
Figure 12. Power loss of four methods with different network scales for five-time identification.

4. Conclusions

Even if the network nodes are in a disconnected state, these four methods can still perform real-time identification of the network topology, and AIPI has a higher accuracy rate. The simulation experiments indicate that AIPI combine active and passive interception can inference topology of WSN with high accuracy rates. Active interception is suitable for any protocol but consumes large amounts of energy, and passive interception consumes less energy but is only suitable for the specified protocol. AIPI balances consumption and applicability properly.

The system generates communication nodes within a certain rectangular area, but when the number of communication nodes is excessive, the system's local information extraction error is relatively large. However, establishing a mixed Gaussian model requires

a larger dataset to avoid randomness. When the number of nodes is small, the Gaussian mixture model cannot fit the errors well, which also means that the correction of errors does not achieve a satisfactory effect, leaving considerable room for improvement in error correction.

References

1. Zagrouba, R.; Kardi, A. Comparative Study of Energy Efficient Routing Techniques in Wireless Sensor Networks. *Information* **2021**, *12*, 42.
2. Kenyeres, M.; Kenyeres, J. Distributed Mechanism for Detecting Average Consensus with Maximum-Degree Weights in Bipartite Regular Graphs. *Mathematics* **2021**, *9*, 3020.
3. Kullmann, L.; Kertész, J.; Kaski, K. Time-dependent cross-correlations between different stock returns: A directed network of influence. *Phys. Rev. E* **2002**, *66*, 026125.
4. Stetter, O.; Battaglia, D.; Soriano, J.; Geisel, T. Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput. Biol.* **2012**, *8*, e1002653.
5. Runge, J.; Bathiany, S.; Bollt, E.; Camps-Valls, G.; Coumou, D.; Deyle, E.; Glymour, C.; Kretschmer, M.; Mahecha, M.D.; Muñoz-Marí, J.; et al. Inferring causation from time series in Earth system sciences. *Nat. Commun.* **2019**, *10*, 2553.
6. Wang, W.X.; Lai, Y.C.; Grebogi, C. Data based identification and prediction of nonlinear and complex dynamical systems. *Phys. Rep.* **2016**, *644*, 1–76.
7. Nitzan, M.; Casadiego, J.; Timme, M. Revealing physical interaction networks from statistics of collective dynamics. *Sci. Adv.* **2017**, *3*, e1600396.
8. Battaglia, P.W.; Hamrick, J.B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; et al. Relational inductive biases, deep learning, and graph networks. *arXiv* **2018**, arXiv:1806.01261.
9. Xu, J.; Duan, L.; Zhang, R. Proactive Eavesdropping Via Jamming for Rate Maximization Over Rayleigh Fading Channels. *IEEE Wirel. Commun. Lett.* **2016**, *5*, 80–83.
10. Cai, H.; Zhang, Q.; Li, Q.; Qin, J. Proactive Monitoring Via Jamming for Rate Maximization Over MIMO Rayleigh Fading Channels. *IEEE Commun. Lett.* **2017**, *21*, 2021–2024.
11. Cai, Y.; Zhao, C.; Shi, Q.; Li, G.Y.; Champagne, B. Joint Beamforming and Jamming Design for mmWave Information Surveillance Systems. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1410–1425.
12. Han, Y.; Duan, L.; Zhang, R. Jamming-Assisted Eavesdropping Over Parallel Fading Channels. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 2486–2499.
13. Hu, G.; Cai, Y. Proactive Eavesdropping Via Jamming for Ergodic Rate Maximization Over Wireless-Powered Multichannel Suspicious System. *IEEE Commun. Lett.* **2020**, *24*, 1830–1834.
14. Hu, D.; Zhang, Q.; Yang, P.; Qin, J. Proactive Monitoring via Jamming in Amplify-and-Forward Relay Networks. *IEEE Signal Process. Lett.* **2017**, *24*, 1714–1718.
15. Lu, H.; Zhang, H.; Dai, H.; Wu, W.; Wang, B. Proactive Eavesdropping in UAV-Aided Suspicious Communication Systems. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1993–1997.

16. Hu, G.; Si, J.; Cai, Y.; Zhu, F. Proactive Eavesdropping via Jamming in UAV-Enabled Suspicious Multiuser Communications. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 3–7.
17. Li, B.; Yao, Y.; Chen, H.; Li, Y.; Huang, S. Wireless Information Surveillance and Intervention Over Multiple Suspicious Links. *IEEE Signal Process. Lett.* **2018**, *25*, 1131–1135.
18. Moon, J.; Lee, H.; Song, C.; Kang, S.; Lee, I. Relay-Assisted Proactive Eavesdropping With Cooperative Jamming and Spoofing. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6958–6971.
19. Yao, J.; Wu, T.; Zhang, Q.; Qin, J. Proactive Monitoring via Passive Reflection Using Intelligent Reflecting Surface. *IEEE Commun. Lett.* **2020**, *24*, 1909–1913.
20. Luo, Q.; Peng, Y.; Li, J.; Peng, X. RSSI-based localization through uncertain data mapping for wireless sensor networks. *IEEE Sens. J.* **2016**, *16*, 3155–3162.
21. Wang, T.; Ding, H.; Xiong, H.; Zheng, L. A compensated multi-anchors TOF-based localization algorithm for asynchronous wireless sensor networks. *IEEE Access* **2019**, *7*, 64162–64176.
22. Wang, T.; Xiong, H.; Ding, H.; Zheng, L. TDOA-based joint synchronization and localization algorithm for asynchronous wireless sensor networks. *IEEE Trans. Commun.* **2020**, *68*, 3107–3124.
23. Gui, L.; Xiao, F.; Zhou, Y.; Shu, F.; Val, T. Connectivity based DV-hop localization for Internet of Things. *IEEE Trans. Veh. Technol.* **2020**, *69*, 8949–8958.
24. Hartung, S.; Bochem, A.; Zdziarstek, A.; Hogrefe, D. Applied Sensor-Assisted Monte Carlo Localization for Mobile Wireless Sensor Networks. In Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks, Graz, Austria, 15–17 February 2016; pp. 181–192.
25. Yousaf, M.Z.; Tahir, M.F.; Raza, A.; Khan, M.A.; Badshah, F. Intelligent Sensors for dc Fault Location Scheme Based on Optimized Intelligent Architecture for HVdc Systems. *Sensors* **2022**, *22*, 9936.
26. Feng, R.; Yuan, W.; Ge, L.; Ji, S. Nonintrusive Load Disaggregation for Residential Users Based on Alternating Optimization and Downsampling. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12.
27. Feng, R.; Dai, M.; Wang, H. Distributed Beamforming in MISO SWIPT System. *IEEE Trans. Veh. Technol.* **2017**, *66*, 5440–5445.
28. Liu, Z.; Wang, W.; Ding, G.; Wu, Q.; Wang, X. Topology Sensing of Non-Collaborative Wireless Networks With Conditional Granger Causality. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 1501–1515.
29. Liu, Z.; Ding, G.; Wang, Z.; Zheng, S.; Sun, J.; Wu, Q. Cooperative Topology Sensing of Wireless Networks With Distributed Sensors. *IEEE Trans. Cogn. Commun. Netw.* **2021**, *7*, 524–540.
30. Feng, R.; Cui, J. AINL: Network Topology Identification of Multiple Communication Modes via Active Intercepting and Node Locating. *IEEE Access* **2023**, *11*, 121399–121409.
31. Akaike, H. Canonical correlation analysis of time series and the use of an information criterion. In *Mathematics in Science and Engineering*; Elsevier: Amsterdam, The Netherlands, 1976; Volume 126, pp. 27–96.