International Journal of Innovative Computing, Information and Control
Volume **8**, Number **5(A)**, May **2012**

ICIC International ©2012 ISSN 1349-4198
pp. 3341–3359

INTELLIGENT INTRUSION DETECTION AND ROBUST NULL DEFENSE FOR WIRELESS NETWORKS

Jeich Mar, I-Fan Hsiao, Yow-Cheng Yeh, Chi-Cheng Kuo and Shin-Ru Wu

Department of Communications Engineering
Yuan-Ze University
No. 135, Yuan-Tung Road, Jungli, Taoyuan 320, Taiwan
eejmar@saturn.yzu.edu.tw

Received January 2011; revised June 2011

Abstract. *In the past few years, plenty of research effort has been devoted to using cross layer design to enhance the performance of wireless communication systems. In this study, the major research effort is devoted to consider jointly adaptive neuro-fuzzy inference system (ANFIS) intrusion detection in the MAC layer and multimodal digital beamforming (DBF) in the PHY layer, with focus on improving the average detection delay (ADD) of denial-of-service (DoS) attacks and demonstrating the defensive function of the robust null extension mode of a multimodal DBF by taking advantage of a two-state Markov chain model. A packet acquisition and analysis system (PAAS) is designed to collect and analyze the friend, intrusion and interference information generated from the wireless networks. A prototype of the ANFIS-based intrusion detection system (ANFIS-IDS) is implemented, trained and tested against a real de-authentication DoS attack to empirically demonstrate the improved performance of the proposed ANFIS-IDS, compared with the IDS using non-parametric sequential change point detection (NPSCPD) algorithm.*

Keywords: ANFIS, De-authentication DoS attack, Non-parametric sequential change point detection, Null extension, Multimodal DBF, Two-state Markov chain model

1. **Introduction.** The over-the-air nature of wireless networks opens the medium access control (MAC) sub-layer of the data link layer up to intrusions [1]. IEEE provides wired equivalent privacy (WEP) key management for the security of the frame body. Unfortunately, it does not encrypt the frame head; therefore, WLAN is vulnerable to disassociation attacks [2]. Intruders eavesdrop on a network and gather information about the client's address in the MAC layer. If the de-authentication and disassociation frames are unencrypted and unauthenticated, an intruder can easily spoof these frames and disconnect the client from the access point (AP), effectively launching a denial-of-service (DoS) attack. In order to enhance security of the IEEE 802.11 standard, the emerging IEEE 802.11i standard improves security in the WLAN with the stronger encryption and authentication functions [3]. Unfortunately, the DoS attacks persist even if IEEE 802.11i is used to protect the WLAN. It is expected that more effort is needed to enhance security of the WLAN environment. The categories of intrusion detection systems (IDS) include signature detection system (SDS) [4] and anomaly detection systems (ADS) [5,6]. An SDS can observe sudden increase in the sequence number gap (SNG) among the packets to rapidly detect an attack on the WLAN. An ADS, which detects abnormal network behaviors reflected in statistical departures from the normal pattern, can be used to detect unknown signature intrusions. However, a decrease of the false alarm rate (FAR) results in an increase of the average detection delay (ADD) due to the time-varying nature of

the intrusion data and the detection threshold. The integration of the decisions coming from different detection algorithms has emerged as a data fusion technique that could effectively improve final detection performance by collecting complementary information from multiple sources and combining them to obtain better detection performance. The Dempster-Shafer fusion of multiple constant FAR abrupt-change detectors has been proposed to improve the performance of network intrusion detection [7]. It is necessary to obey the distribution models of network traffic for the no intrusion state and intrusion state when constant FAR detectors are used. In the non-parametric case, we do not require independent and identically distributed (iid) observations, and do not assume the knowledge of the probability density functions (PDFs). It is only observed that the mean traffic increases between two different time instants. A non-parametric case can then be obtained by the threshold of either the Page's cumulative sum (CUSUM) [8] statistic or the Shiryaev-Roberts-Pollak statistic, which utilizes the log-likelihood ratio (LLR). The non-parametric sequential change point detection (NPSCPD) algorithm, based on CUSUM, was designed to detect abrupt changes in observed traffic [8]. The NPSCPD-IDS detects the DoS attacks and maintains the average false alert rate (FAR) below a prescribed low level. The multi-carrier decision making technique [9] has been adopted to demonstrate that the use of artificial neural network approach in intrusion detection systems will enhance the security of computer and network systems. The main criteria including minimum root mean square error (MSE), performance, less training overhead, memory usage and usability are built in [10] to select the most suitable approach from three different neural networks. The neural network approach learns from an example. After training, the system is able to detect intrusions. In [2], a central manager is proposed to dynamically manage APs and clients for avoiding DoS attacks. An access filtering mechanism is adopted in the IDS to defend DoS attacks in [6], where the early detection method of anomaly accesses based on the statistic analysis is explored. A letter-envelop protocol [11] is proposed to defend the de-authentication attack in 802.11 WLAN. The solution to this vulnerability is to modify the authentication framework such that APs and clients can authenticate all the management messages in 802.11 networks. A reactive defense mechanism [12] is proposed to alleviate the impact of a DoS attack on the victim by detecting the attack and responding to it. Firmware modification increases computing overhead costs in the MAC layer for both APs and clients to compensate for the prevention of DoS attacks. 802.11i chooses to enhance security in the link layer [3]. Since it does not provide a mechanism that can defend DoS attacks. It is valuable if the other PHY layer specifications will further strengthen 802.11i against DoS vulnerabilities. The aim of the present work is to reduce the average detection delay (ADD) of NPSCPD-IDS using the adaptive neuro-fuzzy inference system (ANFIS) rule [13] to combine the detection algorithms of SNG and NPSCPD. Both NPSCPD-IDS and ANFIS-IDS are tested and compared by a prototype IDS experimental platform designed in 802.11b/g WLAN environments. The proposed packet acquisition and analysis system (PAAS) of the IDS experimental platform is built on an x86 embedded system and used to collect the packet data and identity the de-authentication DoS packet, the SNG and detection time. In addition, a defensive solution in the physical layer is proposed to alleviate the computing loading of the MAC layer. The architecture of the ANFIS-IDS integrated with robust null extension mode of multimodal digital beamformer (DBF) is presented to prevent the DoS attacks. Based on the MAC layer protocol of the 802.11 standard, we employ request to send (RTS)/clear to send (CTS) packets to carry additional information of intrusion detections generated from the ANFIS-IDS and control the mode of DBF in the PHY layer. Finally, the performance improvement in the packet error rate (PER) of the WLAN with the null extension mode of multimodal DBF is also analyzed using the two-state Markov

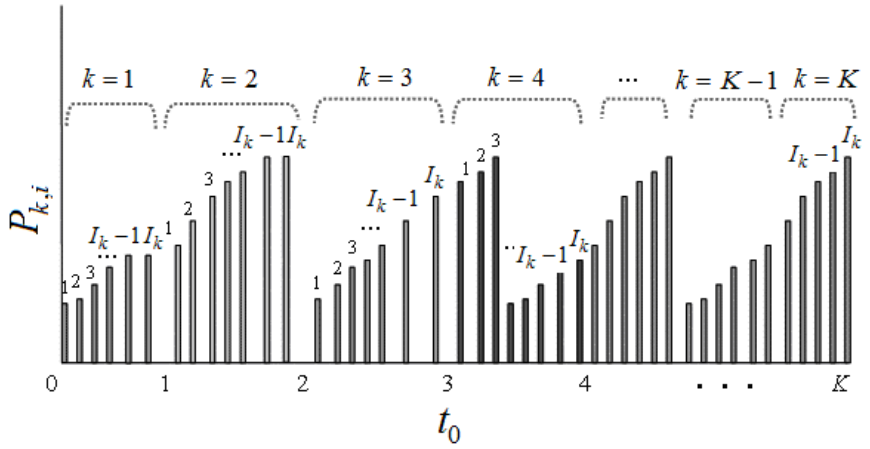

FIGURE 1. Packet sequence number diagram

chain model, where the intrusion state is determined as a bad state. The rest of this paper is organized as follows. In Section 2, we discuss the NPSCPD and SNG detection algorithms. In Section 3, the architecture and training procedure of the ANFIS-IDS are presented in detail. The IDS experimental platform for WLAN is described in Section 4, where the PAAS is able to collect the empirical information from a wireless network. In Section 5, we present our solution for defending against a de-authentication DoS attack in the PHY layer. In Section 6, the two-state Markov chain model is employed to verify that our solution can effectively defend against de-authentication DoS attacks. Finally, we conclude the paper in Section 7.

2. **Detection Algorithms.**

2.1. **NPSCPD algorithm.** When a de-authentication DoS attack occurs at an unknown time and affects the number of received de-authentication packets X_k during the k_{th} interval $X_1, X_2, \ldots$ are monitored at times $t_1, t_2, \ldots$ in the K observation time period. Each observation period is set to t_0. In the NPSCPD algorithm, the mean is assumed to increase from μ to $\theta(\mu = E_0 X_i < \theta = E_1 X_i)$. Based on Page's cumulative sum (CUSUM) statistic [8], which is a non-parametric threshold detection algorithm, the NPSCPD algorithm developed in [14] thresholds a sequential statistic S_k

$$S_k = \max\{0, S_{k-1} + X_k - \mu - \varepsilon\hat{\theta}\},\ 1 \le k \le K, \tag{1}$$

where $\hat{\theta}$ is an estimate of θ, and ε is an optimization parameter that can be generated during a training period. The design criterion of the NPSCPD algorithm is to minimize the ADD while maintaining FAR below a prescribed low level. The traffic change detection alert occurs at a detection time.

$$\tau = \min\{k \ge 1, S_k \ge h_a\}, \tag{2}$$

where h_a is defined as the detection threshold and the average false alarm rate (FAR) is defined as

$$FAR(\tau) = 1/E\{\tau\}. \tag{3}$$

The conditional ADD is defined as

$$ADD_\lambda(\tau) = E\{\tau - \lambda | \tau \ge \lambda\}. \tag{4}$$

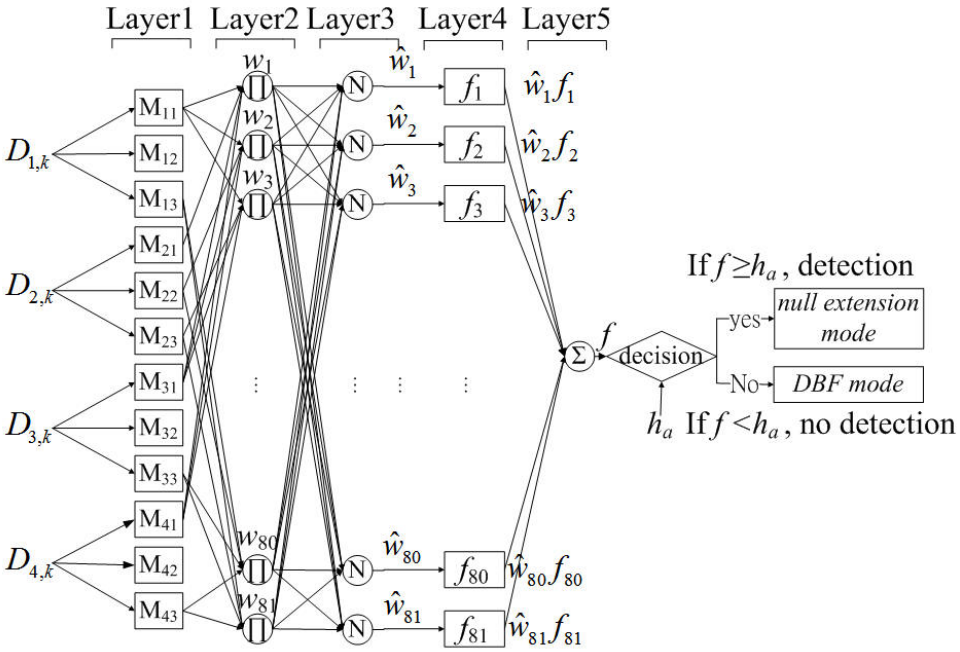

FIGURE 2. The architecture of the ANFIS-IDS and null defense system

2.2. **The SNG detection algorithm [4].** In the MAC layer of the IEEE802.11 standard, the sequence number of the received successive packets follows a monotonic relationship. If the variation of the packet sequence number does not exhibit monotonic increase, the SNG detection algorithm indicates the intrusion detection. As shown in Figure 1, the AP receives I_k packets at the k_{th} observation period of t_0. The range of the sequence number is

$$P_{k,i} \in \{0, 1, 2, ..., 4095 | \forall i = 1, 2, ..., I_k;\ k = 1, 2, ..., k\}. \tag{5}$$

The SNG at the k_{th} interval and i_{th} packet is defined as

$$\Delta P_{k,i} = (P_{k,i} - P_{k,i-1}) modulo\ 4096 = ((\Delta P'_{k,i}))_{4096},\ \forall i = 2, 3, ..., I_k. \tag{6}$$

The average SNG at the k_{th} interval is given by

$$D_{1,k} = E[\Delta P_{k,i}] = \frac{1}{I_k} \sum_{i=1}^{I_k} \Delta P_{k,i},\ \forall k = 1, 2, ..., K. \tag{7}$$

The negative exponential distribution $\Delta P'_{k,i} = ve^{-vk}$, $\forall k = 0, 1, 2, ...;\ i = 1, 2, ..., I_k$ is used to generate the SNG of the packet at the k_{th} interval and v is the mean of $\Delta P'_{k,i}$. The average SNG $D_{1,k}$ is obtained from (6) and (7). The spike intensity in $D_{1,k}$ increases with the average SNG v. The ADD of the IDS decreases with increasing value of v [15]. The feature of the average SNG spike detection is utilized to design the ANFIS-IDS in order to reduce the ADD.

3. **ANFIS-IDS.** The architecture of the ANFIS-IDS and null defense system is shown in Figure 2. The ANFIS-IDS implemented by a five-layer neural fuzzy rule [13] is designed for improving the ADD of the IDS in a wireless network. The ADD of the IDS is measured by comparing the output f of the fifth layer with the threshold h_a. Then the robust null extension mode of the multimodal DBF is initiated as soon as the intrusion is detected if $f \geq h_a$. Otherwise, the normal DBF mode of the multimodal DBF is selected at the AP of the WLAN. The threshold h_a is determined by specifying a FAR of 1.5×10^{-4}.

The ANFIS employs the adaptive network architecture to represent the fuzzy inference system, which can be applied to a wide range of areas, such nonlinear function modeling, time series prediction, and fuzzy controller design. The ANFIS-IDS has four linguistic variables including $D_{1,k}$, $D_{2,k}$, $D_{3,k}$ and $D_{4,k}$, where $D_{1,k}$ is the average SNG at the interval of observation period k, $D_{2,k}$ is the number of received de-authentication packets X_k. S_k-h_a is defined as $D_{3,k}$. The NPSCPD output S_{k-1} at the $(k-1)_{th}$ time interval is defined as $D_{4,k}$. The node functions of each layer are described as below:

Layer 1: Every node in the first layer is an adaptive node with a node function

$$O_{ij}^{1} = M_{ij}(D_{i,k}),\ \forall i = 1,2,3,4;\ j = 1,2,3;\ k = 1,2,...,K. \tag{8}$$

The membership function $M_{ij}(D_{i,k})$ is triangular-shaped [16] with the maximum equal to 1 and minimum equal to 0, such that

$$M_{ij}(D_{i,k}) = M_{ij}(D_{i,k}|m_{ij1}, m_{ij2}, m_{ij3}) = \begin{cases} 0, & D_{i,k} \le m_{ij1} \\ \frac{D_{i,k}-m_{ij1}}{m_{ij2}-m_{ij1}}, & m_{ij1} \le D_{i,k} \le m_{ij2} \\ \frac{m_{ij3}-D_{i,k}}{m_{ij3}-m_{ij2}}, & m_{ij2} \le D_{i,k} \le m_{ij3} \\ 0, & m_{ij3} \le D_{i,k} \end{cases}. \tag{9}$$

At the k_{th} time interval, $M_{ij}(D_{i,k})$ $i = 1,2,3,4$; $j = 1,2,3$ corresponds to $i \times j = 12$ membership functions. The output of each membership function is determined by (9). The premise parameter set is given by

$$\vec{w} = \{m_{ijq} | \forall i = 1,2,3,4;\ j = 1,2,3;\ q = 1,2,3\}, \tag{10}$$

where m_{ij1}, m_{ij2}, m_{ij3} pertaining to the node outputs are updated according to given training data and the steepest descent approach.

Layer 2: Every node l in the second layer is a fixed node labeled Π. The output of node l denoted by $O_{l,k}^{2}$ is the product of all the incoming signals for the l_{th} rule. It is given by

$$\begin{aligned} O_{l,k}^{2} &= w_{l,k}, \\ &= \left\{ M_{1j}^{(l)}(D_{1,k}) \times M_{2j}^{(l)}(D_{2,k}) \times M_{3j}^{(l)}(D_{3,k}) \times M_{4j}^{(l)}(D_{4,k}), \middle| \begin{array}{l} \forall j = 1,2,3 \\ l = 1,2,...,81 \end{array} \right\}, \end{aligned} \tag{11}$$

where the total number of node outputs is $L = j^{i} = 81$.

Layer 3: Every node l in the third layer is a fixed node labeled N. In the node l, the ratio of the l_{th} rule's firing strength to the sum of all the rules' firing strength is calculated. The output of node l, denoted by $O_{l,k}^{3}$, is called the normalized firing strength and is calculated as

$$O_{l,k}^{3} = \hat{w}_{l,k} = \frac{w_{l,k}}{\sum_{l=1}^{L} w_{l,k}},\ \forall l = 1,2,3,...,81. \tag{12}$$

Layer 4: Every node l in the fourth layer is an adaptive node with a node function

$$f_{l,k} = \alpha_{0l} + \alpha_{1l} D_{1,k} + \alpha_{2l} D_{2,k} + \alpha_{3l} D_{3,k} + \alpha_{4l} D_{4,k},\ \forall l = 1,2,3,...,81, \tag{13}$$

where $f_{l,k}$ is a crisp output in the consequence, and α_{0l}, α_{1l}, α_{2l}, α_{3l} and α_{4l} are the consequent parameter sets of node l. The consequent parameter vector is given by

$$\vec{\varphi} = \vec{w}_2 = [\alpha_{il} | i = 0,1,2,3,4; l = 1,2,...,81]^{T}. \tag{14}$$

The 81 fuzzy inference rules of $f_{l,k}$ corresponding to Figure 2 are constructed as follows:

$$\begin{aligned}
R_1 &: \text{if } (D_{1,k} \text{ is } M_{11}) \text{ and } (D_{2,k} \text{ is } M_{21}) \text{ and } (D_{3,k} \text{ is } M_{31}) \text{ and } (D_{4,k} \text{ is } M_{41}) \\
&\qquad \text{then (output is } f_{1,k}) \\
R_2 &: \text{if } (D_{1,k} \text{ is } M_{11}) \text{ and } (D_{2,k} \text{ is } M_{22}) \text{ and } (D_{3,k} \text{ is } M_{31}) \text{ and } (D_{4,k} \text{ is } M_{41}) \\
&\qquad \text{then (output is } f_{2,k}) \\
R_3 &: \text{if } (D_{1,k} \text{ is } M_{11}) \text{ and } (D_{2,k} \text{ is } M_{23}) \text{ and } (D_{3,k} \text{ is } M_{31}) \text{ and } (D_{4,k} \text{ is } M_{41}) \\
&\qquad \text{then (output is } f_{3,k}) \\
&\qquad \vdots \\
R_{80} &: \text{if } (D_{1,k} \text{ is } M_{13}) \text{ and } (D_{2,k} \text{ is } M_{22}) \text{ and } (D_{3,k} \text{ is } M_{33}) \text{ and } (D_{4,k} \text{ is } M_{43}) \\
&\qquad \text{then (output is } f_{80,k}) \\
R_{81} &: \text{if } (D_{1,k} \text{ is } M_{13}) \text{ and } (D_{2,k} \text{ is } M_{23}) \text{ and } (D_{3,k} \text{ is } M_{33}) \text{ and } (D_{4,k} \text{ is } M_{43}) \\
&\qquad \text{then (output is } f_{81,k})
\end{aligned} \tag{15}$$

The node output of the fourth layer is defined as

$$O_{l,k}^4 = \hat{w}_{l,k} f_{l,k} = \hat{w}_{l,k} \sum_{i=0}^{4} \alpha_{i,l} D_{i,k}, \tag{16}$$

where is the output of the third layer, $D_{i,k}$, is the input of the ANFIS-IDS and $D_{0,k} = 1$.

Layer 5: The single node in the fifth layer is a fixed node labeled Σ, which uses the weighted averaged method to compute the overall output $O_{5,k}$ as

$$O_k^5 = f_k = \frac{\sum_{l=1}^{L} w_{l,k} f_{l,k}}{\sum_{l=1}^{L} w_{l,k}} = \sum_{l=1}^{L} \hat{w}_{l,k} f_{l,k}. \tag{17}$$

The matrix form of ANFIS-IDS processing can be expressed as

$$\mathbf{A}\vec{\varphi} = \vec{f}, \tag{18}$$

where

$$\mathbf{A} = \begin{bmatrix}
\hat{w}_{1,1}D_{1,1} & \hat{w}_{1,1}D_{2,1} & \dots & \hat{w}_{1,1}D_{4,1} & \hat{w}_{1,1} & \hat{w}_{2,1}D_{1,1} & \dots & \hat{w}_{81,1} \\
\hat{w}_{1,2}D_{1,2} & \hat{w}_{1,2}D_{2,2} & \dots & \hat{w}_{1,2}D_{4,2} & \hat{w}_{1,2} & \hat{w}_{2,2}D_{1,2} & \dots & \hat{w}_{81,2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\hat{w}_{1,k}D_{1,k} & \hat{w}_{1,k}D_{2,k} & \dots & \hat{w}_{1,k}D_{4,k} & \hat{w}_{1,k} & \hat{w}_{2,k}D_{1,k} & \dots & \hat{w}_{81,k} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\hat{w}_{1,K}D_{1,K} & \hat{w}_{1,K}D_{2,K} & \dots & \hat{w}_{1,K}D_{4,K} & \hat{w}_{1,K} & \hat{w}_{2,K}D_{1,K} & \dots & \hat{w}_{81,K}
\end{bmatrix}. \tag{19}$$

$$\vec{f} = [f_1 \ f_2 \ f_3 \ \dots \ f_K]^T \tag{20}$$

Each input training vector corresponds to one row in matrix $\mathbf{A}$, which contains 405 data elements. Therefore, K must be greater than or equal to 405 in order to solve the consequent parameters α_{il}. The value of each input/output linguistic variable and the fuzzy control rules must be provided for the initial learning process. The training data includes $D_{i,k}$, $i = 1, 2, 3, 4$ and the vector of the desired ANFIS output $\vec{y}$, which is defined as

$$\vec{y} = [y_1 \ y_2 \ y_3 \ \dots \ y_K]^T, \tag{21}$$

where y_k is generated from the normalized linear combination method.

$$y_k = \frac{a \times \frac{D_{1,k}}{\max(D_{1,k})} + b \times \frac{S_k}{\max(S_k)}}{a + b}, \tag{22}$$

where a and b are the weightings of $D_{1,k}$ and S_k for the desired output y_k. Based on the fuzzy inference rules listed in (15), the training data of four input linguistic variables are applied for learning process of the ANFIS-IDS to achieve objective optimization.

A hybrid learning algorithm [16] combines the gradient descent method and the least squares estimate (LSE) to estimate the parameters in ANFIS. The hybrid learning flow chart of the ANFIS-IDS, as shown in Figure 3, is divided three steps. Firstly, a reinforcement learning scheme termed the LSE is used to optimally adjust f_i in layer 4 to approximate the desired output of the ANFIS. The consequent parameter vector is estimated as

$$\hat{\vec{\varphi}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{f} \tag{23}$$

The square error of output in fuzzy inference rules $\vec{f}$ is defined as

$$E_k = (y_k - f_k)^2,\ \forall k = 1, 2, 3, ..., K. \tag{24}$$

Secondly, a self-organized learning scheme called the gradient descent method is used to adjust the premise parameters in $\vec{w}_1$ to minimize the square error of the output in the fuzzy inference rules. That is, let

$$m_{ijq}(k+1) = m_{ijq}(k) + \Delta m_{ijq}(k),\ \forall i = 1, 2, 3, 4;\ j = 1, 2, 3;\ q = 1, 2, 3, \tag{25}$$

where the adjustment step is

$$\Delta m_{ijq} = -\eta\frac{\partial E}{\partial m_{ijq}} \tag{26}$$

Thirdly, the optimum weightings of the average SNG and NPSCPD output for establishing the desired ANFIS output are searched through repeated simulations using different values of a and b. Figure 4 shows the simulated ADD of the ANFIS-IDS for different weightings under the condition of FAR $= 1.5 \times 10^{-4}$. Therefore, we determine $a = 10$ and $b = 1$ to achieve the optimization objective of minimizing the ADD for the ANFIS-IDS.

TABLE 1. Training of premise parameters

(a).Initial MFs

$D_{1,k}$						$D_{2,k}$						
m_{121}	m_{113}	m_{122}	m_{131}	m_{123}	m_{132}	m_{212}	m_{221}	m_{213}	m_{222}	m_{231}	m_{223}	m_{232}
0.678	2048	2048	2048	4095	4095	0	0	1	1	1	2	2
$D_{3,k}$						$D_{4,k}$						
m_{321}	m_{313}	m_{322}	m_{331}	m_{323}	m_{332}	m_{412}	m_{421}	m_{413}	m_{422}	m_{431}	m_{423}	m_{432}
-3.009	-1.2	-1.2	-1.2	0.610	0.610	0	0	1.493	1.493	1.493	2.987	2.987

(b). Final MFs

$D_{1,k}$						$D_{2,k}$						
m_{121}	m_{113}	m_{122}	m_{131}	m_{123}	m_{132}	m_{212}	m_{221}	m_{213}	m_{222}	m_{231}	m_{223}	m_{232}
0.672	2048	2048	2048	4095	4095	1.61×10^{-6}	9.05×10^{-3}	0.999	1	1	2	2
$D_{3,k}$						$D_{4,k}$						
m_{321}	m_{313}	m_{322}	m_{331}	m_{323}	m_{332}	m_{412}	m_{421}	m_{413}	m_{422}	m_{431}	m_{423}	m_{432}
-3.013	-1.219	-1.182	-1.204	0.610	0.623	-3.82×10^{-4}	0.016	1.495	1.516	1.463	2.965	3.008

Four triangular-shaped membership functions $M_{ij}(D_{i,k})$ of layer 1 are defined in (9), where the initial premise parameters are determined by the training data. The minimum values of the training data $D_{i,k}$ are used to set m_{i12} and m_{i21} of the initial MF; the median values of the training data $D_{i,k}$ are used to define m_{i13}, m_{i22}, m_{i31} of the initial MF; the

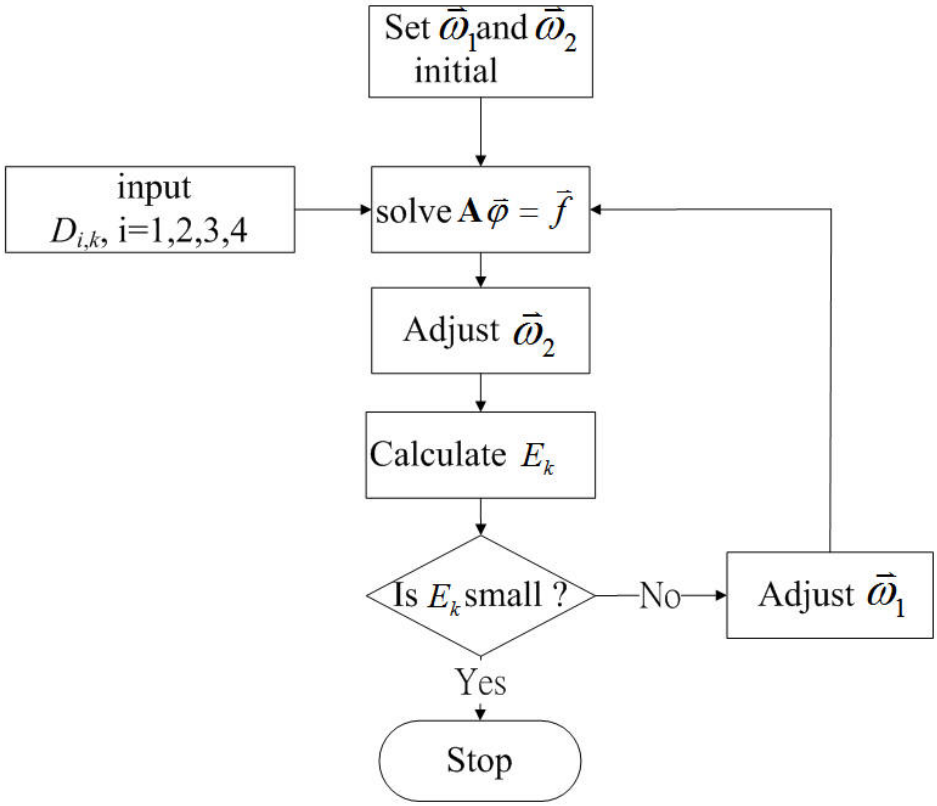

FIGURE 3. Hybrid learning flow chart of the ANFIS-IDS

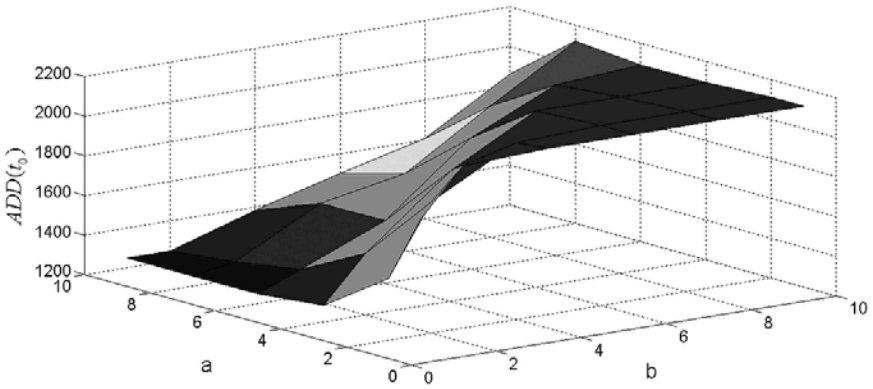

FIGURE 4. Weightings determination using the simulated ADD of the ANFIS-IDS

maximum values of the training data $D_{i,k}$ are used to define m_{i23}, m_{i32} of the initial MF. The range values of $D_{i,k}$ are (0.6779,4095), (0,2), (–3.009,0.6096) and (0,2.987) for $i = 1, 2, 3, 4$, respectively. The initial and final premise parameters of the four input MFs are listed in Tables 1(a) and (b), respectively. The root mean square error (RMSE) curve of the ANFIS-IDS, as shown in Figure 5, is calculated by (24), which demonstrates that the RMSE converges to 7.8304×10^{-4} after 200 epochs.

4. **Prototype IDS Experimental Platform.** The test scenario of the IDS experimental platform on the IEEE 802.11g WLAN is shown in Figure 6, where the algorithms of

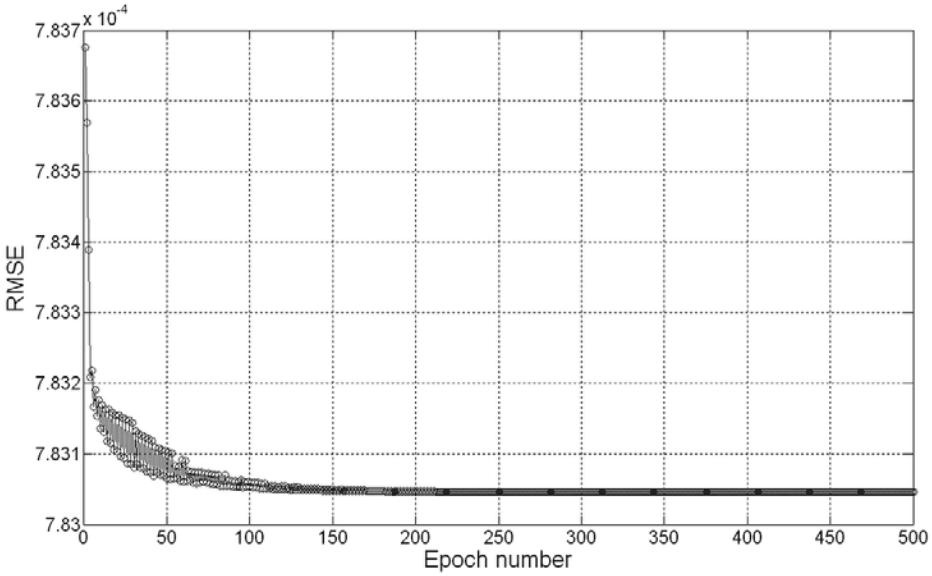

FIGURE 5. RMSE curve of the ANFIS-IDS

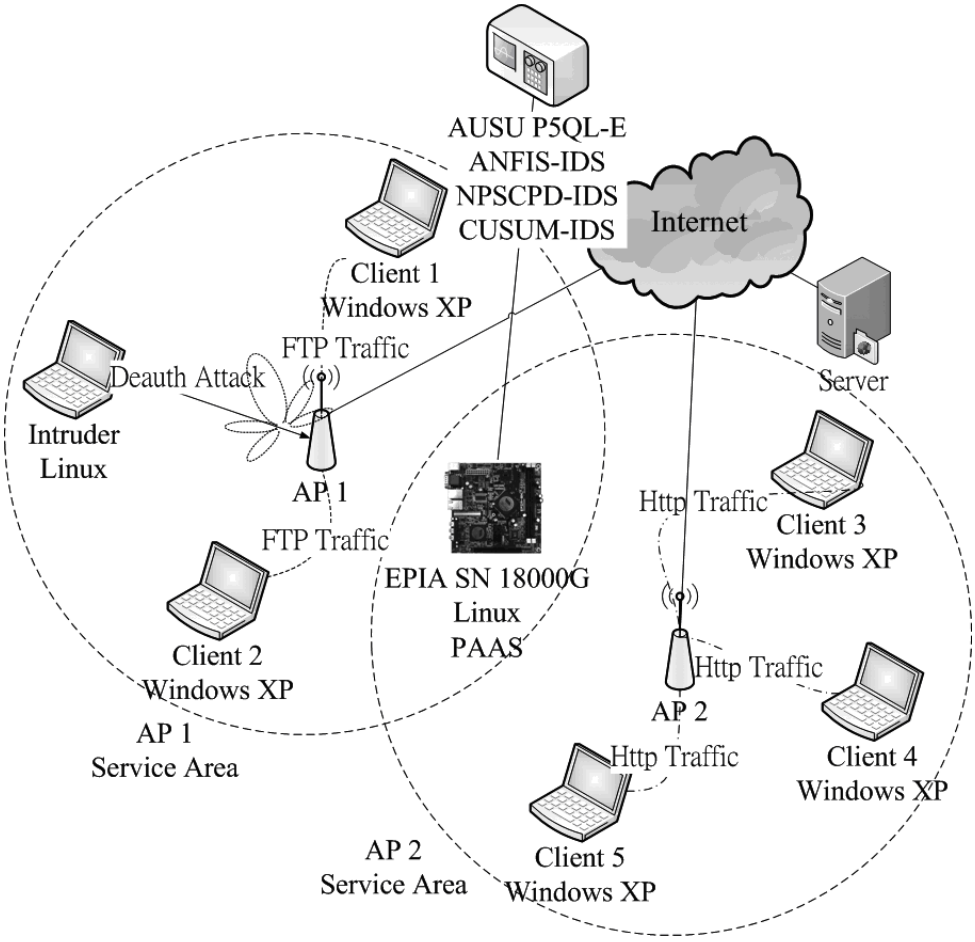

FIGURE 6. The scenario of the prototype IDS experimental platform

the ANFIS-IDS and NPSCPD-IDS are built in a personal computer (PC). Clients 1, 2 and the first AP (AP1) are used as friendly WLAN members and other WLAN members including Clients 3, 4, 5 and the second AP (AP2) are used as interference sources. The AP1, which installs a multimodal DBF, connects two WiFi clients (1, 2) and AP2 connects 3 WiFi clients (3, 4, 5) to a wired, ethernet-based network, respectively. Client 2 transmits the information to client 1 using a file transfer protocol (FTP). The 802.11g data link-layer consists of management, control, and data packets that are used to maintain the connections and carry-out the data communication. The intruder installed at a PC using a Linux operating system utilizes forged de-authentication packets to cause some or all clients to disconnect from AP1 even if they reconnect. The WLAN is then essentially disabled. The effectiveness of the de-authentication DoS attacks lies in the fact that the packets have enforcing informational nature a host must terminate the connection once it receives a de-authentication packet. The Aircrack-ng software tool [17] emulates the de-authentication DoS attacks for the Linux PC. During the test, the intruder sends one forged de-authentication packet to AP1 per 600 observation periods. Clients 3, 4 and 5 perform the hypertext transfer protocol (HTTP) data transmission by passing through AP2 and connecting to the server of the ethernet-based network. During the test, the source address (SA) filter and the destination address (DA) filter of the PAAS distinguish the uplink packets of clients 1, 2 from those of clients 3, 4, 5. The PAAS collects the packets generated in the service area of AP1 and discards the packets generated in the service area of AP2.

The PAAS is designed and implemented on the EPIA SN 18000G x86 ARM platform and using the Linux operation system to generate the signature data base. PAAS operates in the data-link layer of Open System Interconnection (OSI) to collect the packet data required to perform the practical intrusion detection performance test. With reference to IEEE 802.11g specification [18], the packet sequence number and packet type are extracted from the MAC header. The packet sequence number is used as the signature data of the SDS. When Type=01 and Subtype=1100, the received packet is identified as a de-authentication packet. The statistical value of the number of the de-authentication packets is used to analyze the variation of the traffic loading. Finally, the signature data of the SDS and ADS are applied to the ANFIS-IDS. According to the IEEE 802.11standard, the network interface card (NIC) frame is added to the ethernet-network packet head to transmit wired-data in WLAN environments [19,20]. The acquired de-authentication packet data is shown as follows:

The analysis results of a packet data collected from the PAAS includes the packet receiving time: $\lfloor 14.046059 \rfloor = 14$, the subtype of the packet: c [12 (dec)] (de-authentication packet), and the packet sequence number: 152 [338 (dec)].

5. **Defending against DoS Attack Using the Robust Null Extension Mode of Multimodal DBF.** The multimodal DBF relies on the assumption of almost-plane incident waves. First, the direction of the arrival (DOA) mode is enabled at AP to estimate the DOA of the target client signal. Then the beam is steered to the direction of the target client using the DBF mode. The DOAs of the attack signal can be estimated once the intrusion is detected. The uplink client signal is estimated and tracked in DOA mode. The tracked client DOA and the intruder DOA are used to determine the weight vectors for the null extension mode of the multimodal DBF. Since the DOA of the intruder cannot be precisely estimated [21] due to the angular spreading of the multipath fading channel generated in WLAN environments, therefore, the width of the null should widen over an extended sector to improve the robustness of the interference cancellation for the

```
No.     Time        Source                Destination
190     14.046059   D-Link_08:b9:d9       TrapezeN_15:14:44

Protocol Info
IEEE 802.11 Deauthentication, SN=338, FN=0, Flags=....R...

Frame 190 (26 bytes on wire, 26 bytes captured)
IEEE 802.11 Deauthentication, Flags: ....R...
    Type/Subtype: Deauthentication (0x0c)
    Frame Control: 0x08C0 (Normal)
        Version: 0
        Type: Management frame (0)
        Subtype: 12 (dec)
        Flags: 0x8
    Duration: 314
    Destination address: TrapezeN_15:14:44 (00:0b:0e:15:14:44)
    Source address: D-Link_08:b9:d9 (00:17:9a:08:b9:d9)
    BSS Id: TrapezeN_15:14:44 (00:0b:0e:15:14:44)
    Fragment number: 0
    Sequence number: 338 (dec)
IEEE 802.11 wireless LAN management frame
    Fixed parameters (2 bytes)
        Reason code: Class 3 frame received
        from nonassociated station (0x0007)

0000  c0 08 3a 01 00 0b 0e 15 14 44 00 17 9a 08 b9 d9   ..:......D......
0010  00 0b 0e 15 14 44 20 15 07 00                     .....D ...
```

array antenna [22]. A robust null extension algorithm that can adjust the number, width and depth of multiple nulls of the multimodal DBF in accordance with the environment requirements is proposed for the suppression of the interference caused by the intrusion of the de-authentication DoS attacks. The fast subspace decomposition (FSD) [23] is utilized to compute the inverse of the covariance matrix $\mathbf{R}_I$ in real time, the cyclic-Jacobi algorithm [24] is used to process the tri-diagonal matrix and generate the eigen-value and eigen-vector of the inverse covariance matrix $\mathbf{R}_I^{-1}$. The principal components method [25] selects q large eigen-values to robustly simplify the computation of $\hat{\mathbf{R}}_I^{-1}$ according to the eigen-decomposition procedure and to determine the optimal weighting matrix of the null extension mode.

The numerical beam pattern simulations of the null extension mode are used to demonstrate its intrusion defending capability for a de-authentication DoS attack. We consider a 16-element linear array antenna with uniform spacing $\lambda/2$, where λ is the wavelength. Therefore, the array length l is $15\lambda/2$. The flow chart of multiple null simulations is shown in Figure 7. A target client's direction at $\phi_{s\_y} = 0\,^\circ$ in conjunction with an intruder's direction at $\phi_{p\_y} = -45\,^\circ$ is assumed. The simulation result is shown in Figure 8, where the main beam is steered in the target client's direction at $\phi_{s\_y} = 0^\circ$ and four nulls at -37.3801°, -42.3533°, -47.7655° and -53.8136° are directed in the attacker's direction. The sidelobe level is -20dB and the sidelobe cancellation ratio $C = 27$dB, which is defined by the square of the maximum magnitude of beam pattern y within the null sector Δu, normalized to the maximum value of the sidelobe envelope.

$A(u)$ of the pattern after windowing, normalized to the maximum value of the sidelobe envelope $A(u)$ of the pattern after windowing, i.e.,

$$C = \frac{\max_{u\in\Delta u}\, y(u)}{\max_{u\in\Delta u}\, A(u)}. \tag{27}$$

The simulation in [26] shows that the required number of nulls is $M_p = 4$ for $\Delta u = 1.5\lambda/l = 0.2$ and $C = 27$dB. The antenna gains of the DBF at the direction of the target client signal and intrusion signal are 0dB and -20dB, respectively when the multimodal

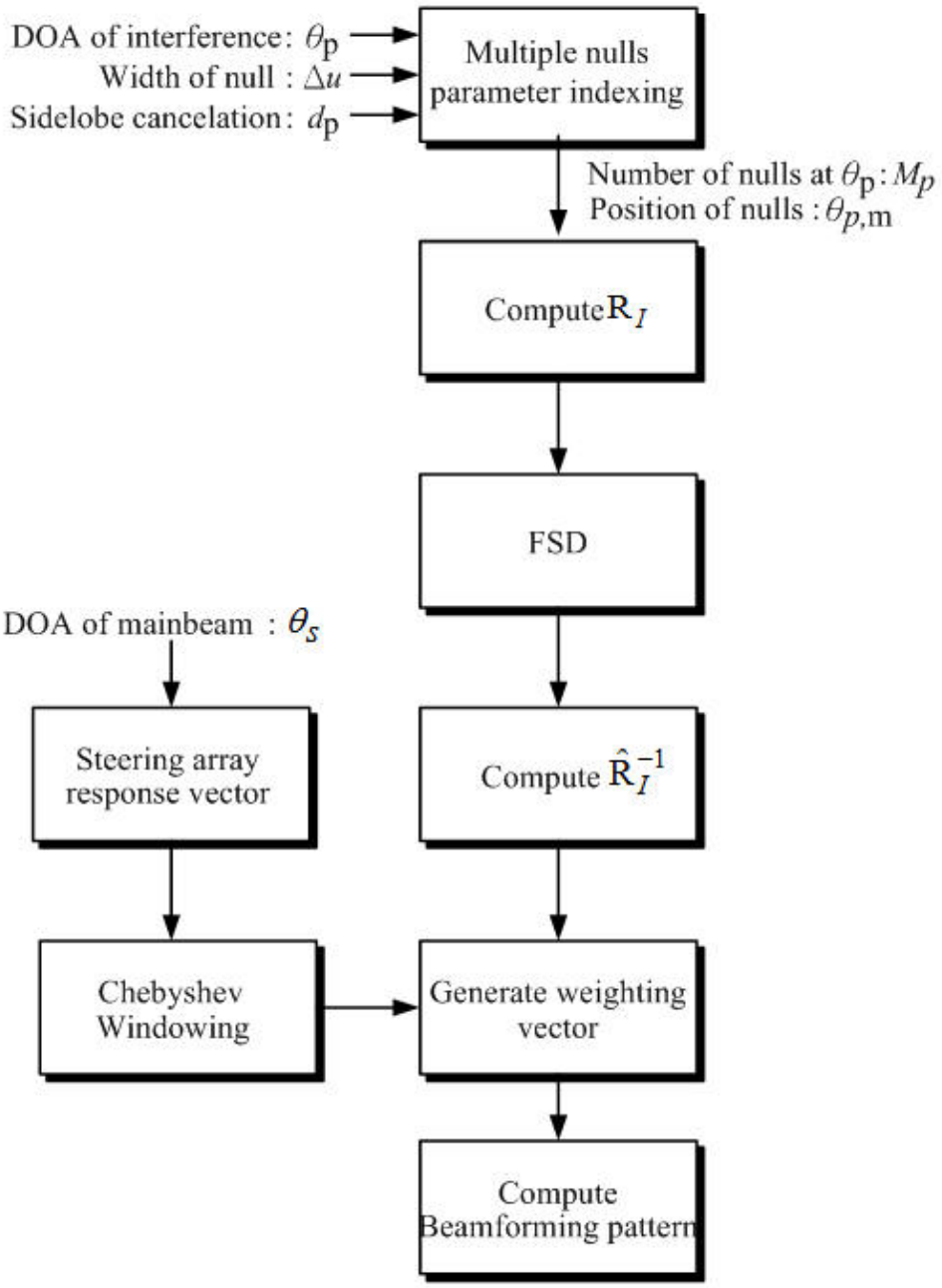

FIGURE 7. Flow chart of simulations for null extension mode

DBF is operated with the normal DBF mode. The intrusion signal is attenuated more than 27dB as soon as the null extension mode of the multimodal DBF is enabled after the intrusion is detected. Therefore, the power of the interference signal generated from the intruder is attenuated by 47dB without affecting the operation of the target client.

Finally, we show that our solution can effectively defend against de-authentication DoS attacks using the two-state Markov chain model. In a fading environment, the transmitted signals encounter a burst error [27]. Here we assume that the burst error caused by the intrusion lasts for a number of bits. Figure 9 shows a state diagram of a two-state Markov chain model for a wireless link called the Gilbert model [27]. Assume that a packet is transmitted over a binary symmetric channel under different states. The state transition period is equal to one packet time. For the target client uplink case, the state G (good)

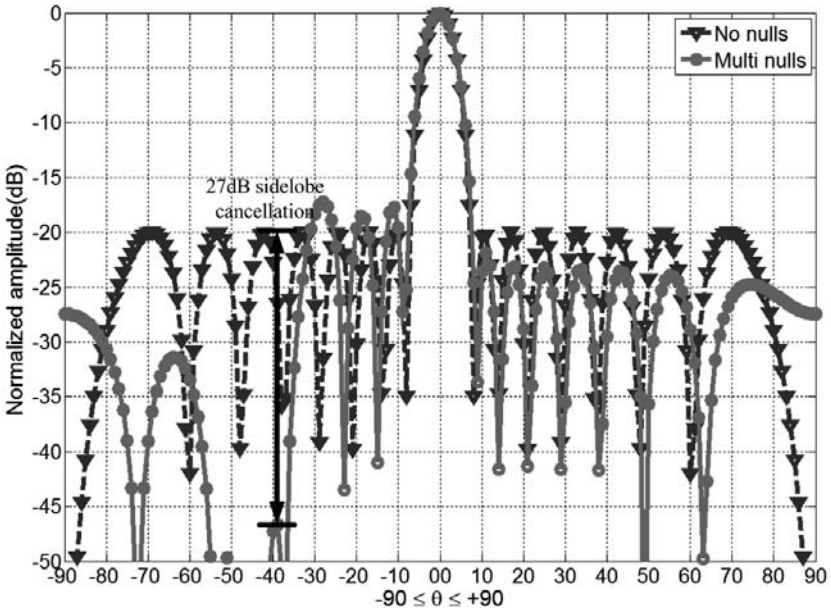

FIGURE 8. The antenna pattern over the extended null sector for $\Delta u = 0.2$, $M_p = 4$

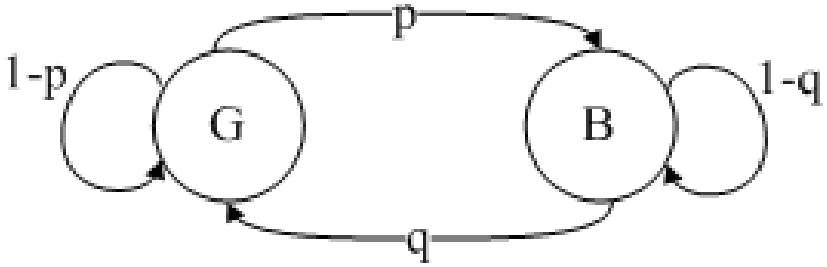

FIGURE 9. Gilbert model [27]

is the no intrusion state where the bit error rate ($BERc$) is ε_0. State B (bad) is the intrusion state where the $BERc$ is ε_1. For the intruder uplink case, state G (good) is the no client transmission state where the bit error rate ($BERa$) is also defined as ε_0. State B (bad) is the client transmission state where the $BERa$ is defined as ε_1. p is the transition probability from state G to state B and $1-p$ is the transition probability from state G to itself. q is the transition probability from state B to state G and $1-q$ is the transition probability from state B to itself. The PER given state B for both the target client uplink case and for the intruder uplink case is

$$PER = 1 - (1 - \overline{\varepsilon})^{(D+H)}, \tag{28}$$

where the length of the data field in the packet is D bits and the length of the control header in the packet is H bits. The average BER, $\overline{\varepsilon}$, in bad state is derived using the Gilbert model.

$$\overline{\varepsilon} = P_G \varepsilon_0 + P_B \varepsilon_1, \tag{29}$$

where the steady state probabilities in states G and B are $P_G = q/(p+q)$ and $P_B = p/(p+q)$, respectively.

Usually the attack signal is much stronger than the local noise under intrusion interference environment. Therefore, local noise is negligible in the BER analysis. Then, the theoretical BER of the WLAN QPSK mode in the Rayleigh fading channel is given as [28,29].

$$BER = \frac{1}{2}\left[1 - \sqrt{\frac{E_b/I_0}{1 + (E_b/I_0)}}\right]. \tag{30}$$

The bit energy to interference power density ratio (E_b/I_0) is given as

$$\frac{E_b}{I_0} = \frac{S}{I}\left(\frac{B_c}{R_b}\right), \tag{31}$$

where B_c is the channel bandwidth in hertz, R_b is the data rate in bits per second and S/I is the signal-to-noise ratio.

In the following session, the BER performance of the quadric-phase-shift keying (QPSK) modulation is simulated for the target client and intruder uplinks according to the IEEE 802.11g WLAN standard. The QPSK transceivers with and without the robust null extension mode of multimodal DBF are included in the simulations. Equation (30) is used to verify the correctness of the simulation result in state B of the target client uplink case.

6. **Experimental Results.** First, the ADD of the IDS using ANFIS, NPSCPD and CUSUM detection algorithms are simulated at a fixed FAR of 1.5×10^{-4}. The holding time of the iid random data transmission for each client in WLAN is assumed to have negative exponential distribution. As shown in Figure 10, the ADD of the ANFIS detection algorithm approximates that of the NPSCPD detection algorithm when the mean of SNG is less than 50, because the small amount of intrusion spike cannot speed up the detection of the ANFIS-IDS significantly. The ADD of the ANFIS-IDS is significantly reduced in the environment of high intrusion spike intensity, that is, the value of v becomes large. Figure 11 shows that the ANFIS detection algorithm has the shortest ADD compared with the CUSUM and NPSCPD detection algorithms and the ADD of the NPSCPD detection algorithm is shorter than that of the CUSUM detection algorithm for all values of the FAR. Then we input all practical $D_{i,k}$ values measured by the PAAS to the IDS experimental platform for the comparison of the three different detection algorithms. The ADD and FAR performance of the IDS experimental platform using three detection algorithms are listed in Table 2, which shows that the ADDs of the ANFIS, NPSCPD and CUSUM detection algorithms are 766.56 observation periods, 3068.78 observation periods, and 6010.28 observation periods, respectively, at a fixed FAR of 1.5×10^{-4}. The time savings of the IDS experimental platform using the ANFIS architecture can reach 75% and 87% compared with the NPSCPD and CUSUM detection algorithms, respectively. Note that the unit of the ADD is measured in the number of the observation period t_0, which cannot exceed 50μsec for the 802.11g WLAN. Therefore, the minimum ADD is 0.0383sec when the architecture of the ANFIS-IDS is implemented in 802.11g WLAN.

Performance simulations of IDS using multimodal DBF are carried out for the QPSK OFDM system over the Rayleigh fading channel, where the received client power (S) is set as 0dBw and noise power is -96dBm, the received attack power (I) is changed from -30dBw to 30dBw, the data rate (R_b) is 12Mbps, the convolution coding rate is 1/2 with constraint length of 7 and the bandwidth is 20MHz (B). All encoded data bits are interleaved by a block interleave with a block size of 96 bits corresponding to the number of bits in a single OFDM symbol. Packet detection, timing synchronization and coarse frequency offset estimation of the WLAN receiver are performed according to the

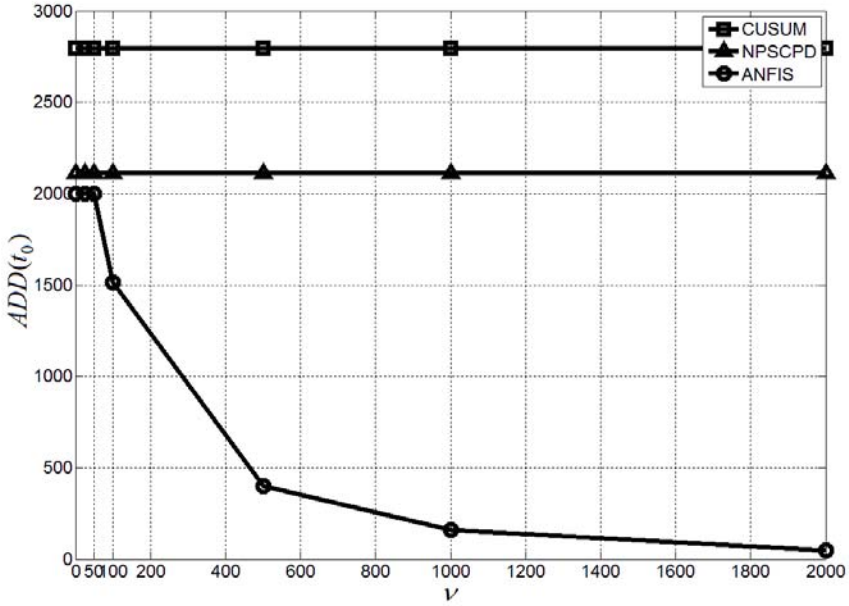

FIGURE 10. The ADD of ANFIS, NPSCPD and CUSUM

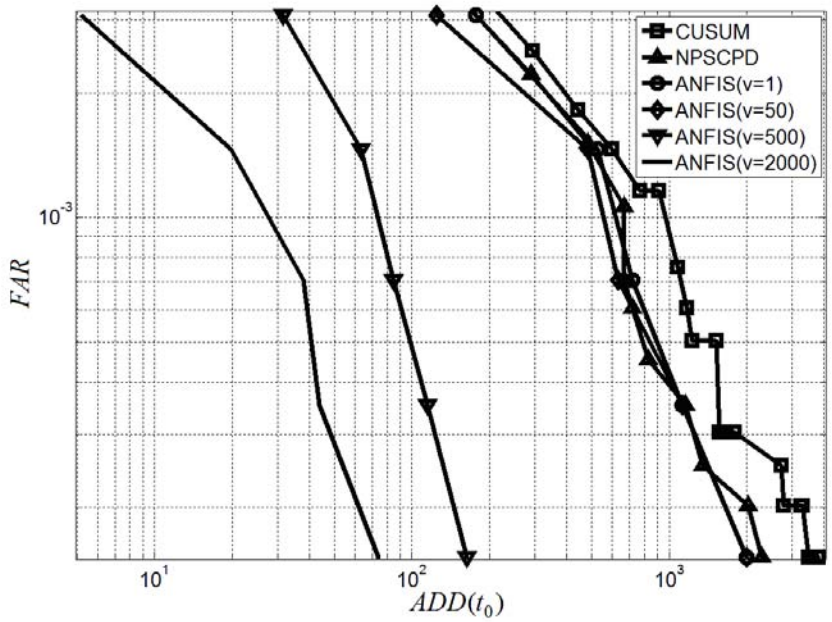

FIGURE 11. The simulated performance comparison of ANFIS, NPSCPD and CUSUM

algorithms provided in [18]. As shown in Figure 12, the theoretical BER (10^{-4}) of the non-coded QPSK system adds about 16dB of coding gain, which is quite similar to those of the simulated BER. Figure 13 shows that the BER of the target client ($BERc$) increases with the attack power. The $BERc$ is equal to about 0.5 when the attack power is equal to 10dB. Therefore, the target client is out of work when the attack power exceeds 10dB. The BER of the intruder ($BERa$) decreases with increasing attack power. After the null extension mode of the multimodal DBF is enabled at AP, the attack power is attenuated

TABLE 2. Performance comparison for different methods on IDS platform

	ANFIS	NPSCPD	CUSUM
$TFA(t_0)$	6601.67	6601.67	6601.67
FAR	0.00015	0.00015	0.00015
$ADD(t_0)$	766.56	3068.78	6010.28

by 27dB. The $BERc$ is almost reduced to zero, which is too small to be shown in the figure, and the $BERa$ is increased to 0.5, which means that the network intrusion is out of work.

Using (28), the PER of the intruder($PERa$), given the target client transmission, is illustrated in Figure 14 for five different steady state probabilities in states G and B, where state G is defined by no client transmission state and state B is defined by the client transmission state as defined earlier. The total length of 1,502 bytes is used for the FTP protocol packet of the target clients and the total length of 54 bytes consisting of 24 bytes in the MAC header and 30 bytes in the data field is used for the de-authentication packet of the intruder. As shown in Figure 14, the $PERa$ of the AP without (w/o) the robust null extension mode decreases with increasing attack power and approaches zero when the attack power exceeds 20dB. The $PERa$ of the AP decreases with an increase of P_G under a fixed attack power. The $PERa$ of the AP with (w) the robust null extension mode approaches one for all P_G. The successful attack rate (SAR) is defined as

$$SAR = 1 - PER_a \tag{32}$$

Therefore, the SAR equals one when the DBF mode of the multimodal DBF at the AP is employed and the attack power exceeds 20dB. As soon as the null extension mode is enabled, the SAR is reduced to zero because the attack power is attenuated more than 27dB by the null extension mode of the multimodal DBF. Finally, we note that as soon as the null extension mode of the multimodal DBF is enabled, the SAR is always equal to zero, that is, the designed multimodal DBF can effectively defend against de-authentication attack.

7. **Conclusions.** Intruders may attack the MAC layer of a WLAN using forged de-authentication packets that cause clients to disconnect from an AP. Early detection of DoS attacks would enable defensive actions earlier. In this study, we develop an efficient ANFIS-IDS platform for early detection of attacks. The ANFIS rule is employed to combine the NPSCPD algorithm and SNG detection algorithm to reduce the ADD of the IDS. A multimodal DBF is designed to defend against de-authentication DoS attacks in 802.11g WLAN. An ANFIS-IDS experimental platform is implemented and tested, where a PAAS is constructed on an x86 embedded system to collect packet data and analyze the packet content of the MAC layer. The experimental results demonstrate that the ANFIS-IDS can significantly reduce the ADD under heavy SNG intensity, as compared with the NPSCPD-IDS. Moreover, the developed approach is self-learning, which enables the IDS to adapt to various networks and attack patterns. Finally, the two-state Markov chain model is used to verify that the robust null extension mode of multimodal DBF is effective against de-authentication attacks as soon as the ANFIS-IDS detects a network intrusion which enables the multimodal DBF to enter a null extension mode. The multimodal DBF can be easily deployed in existing WLAN as well as in future 802.11 devices. In the meantime, because the attack signal can be suppressed in the PHY layer, the implementation overhead and management burden in the MAC

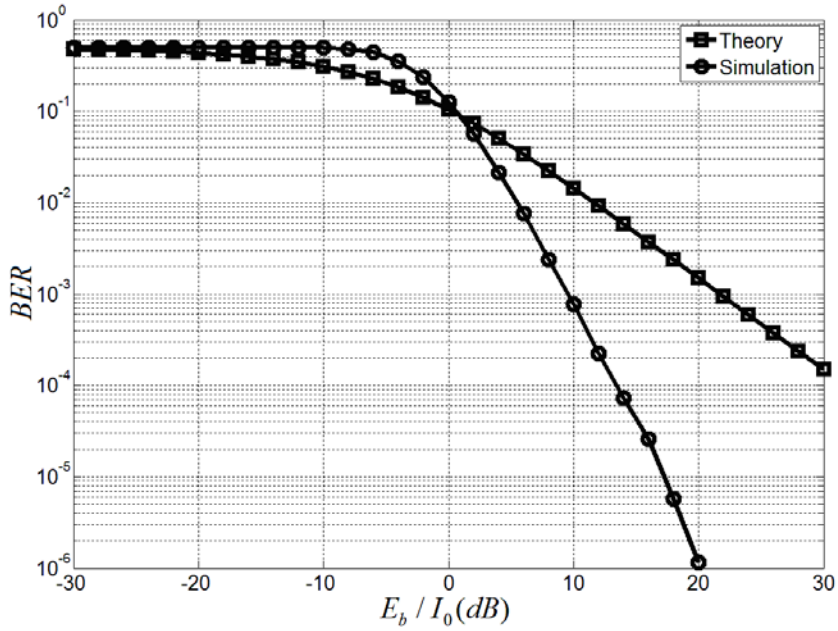

FIGURE 12. BER of QPSK over Rayleigh fading channel

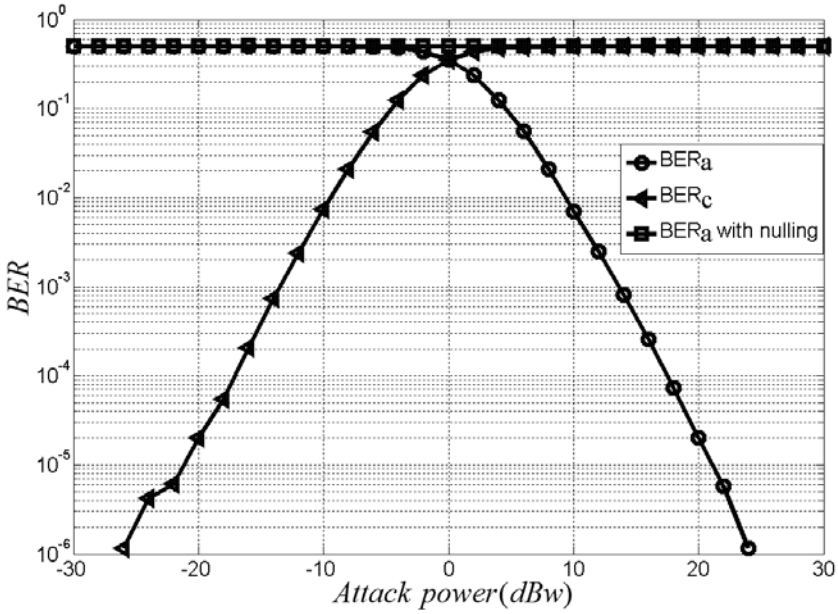

FIGURE 13. The BERs of the target client and intruder

layer can be neglected. Furthermore, the results and methods developed in this paper are helpful to further strengthen 802.11i standard against DoS vulnerabilities in the WLAN environment.

Acknowledgment. The research work was supported by the research grants from National Science Council, Taiwan (NSC 99-2221-E-155-031) and Chung-Shan Institute of Science and Technology.

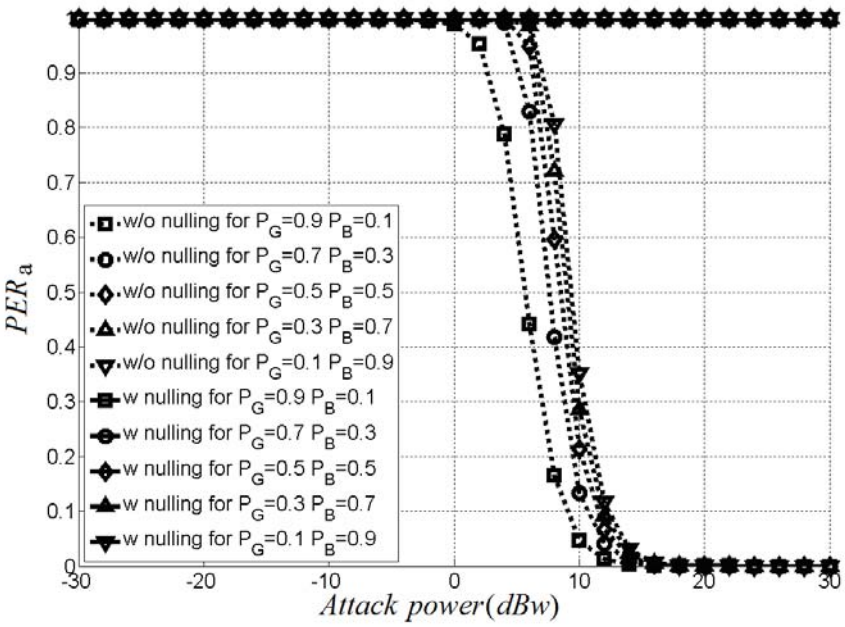

FIGURE 14. $\mathrm{PER}_{w/o\ null}$ (dotted-line) and $\mathrm{PER}_{w\ null}$ (real-line) of the intruder with different P_G and P_B

REFERENCES

[1] W. A. Arbaugh, N. Shankar, Y. Wan and K. Zhang, Your 802.11 wireless network has no clothes, *IEEE Wireless Communications*, vol.9, pp.44-51, 2002.

[2] P. Ding, J. Holliday and A. Celik, Improving the security of wireless LANs by managing 802.1x disassociation, *IEEE Consumer Communications and Networking Conference*, pp.53-58, 2004.

[3] J.-C. Chen, M.-C. Jiang and Y.-W. Liu, Wireless LAN security and IEEE 802.11i, *IEEE Wireless Communications*, vol.12, no.1, pp.27-36, 2005.

[4] F. Guo and T. Chiueh, Sequence number-based MAC address spoof detection, *Proc. of the 8th International Symposium on Recent Advances in Intrusion Detection*, vol.3858, pp.309-329, 2006.

[5] A. Makanju, P. LaRoche and A. N. Zincir-Heywood, A comparison between signature and GP-based IDSs for link layer attacks on WiFi networks, *Proc. of the 2007 IEEE Symposium on Computational Intelligence in Security and Defense Applications*, pp.213-219, 2007.

[6] S. Oshima, T. Nakashima and T. Sueyoshi, Extraction of anomaly accessed IP packets features using statistical method, *International Journal of Innovative Computing, Information and Control*, vol.6, no.8, pp.3725-3735, 2010.

[7] D. He and H. Leung, Network intrusion detection using CFAR abrupt-change detectors, *IEEE Trans. on Instrumentation and Measurement*, vol.57, no.3, pp.490-497, 2008.

[8] C.-C. Hsu, The MOSUM of squares test for monitoring variance changes, *Finance Research Letters*, vol.4, no.4, 2007.

[9] I. Ahmad, A. B. Abdullah and A. S. Alghamdi, Evaluating intrusion detection approach using multi-creteria decision making technique, *International Journal of Information Sciences and Computer Engineering*, vol.1, no.1, pp.60-67, 2010.

[10] I. Ahmad, A. Abdullah and A. Alghamdi, Investigating supervised neural networks to intrusion detection, *ICIC Express Letters*, vol.4, no.6(A), pp.2133-2138, 2010.

[11] T. D. Nguyen, D. H. M. Nguyen, B. N. Tran, H. Vu and N. Mittal, A lightweight solution for defending against de-authentication/disassociation attacks on 802.11 networks, *Proc. of the 17th International Conference on Computer Communications and Networks*, pp.1-6, 2008.

[12] M. S. Fallah, A puzzle-based defense strategy against flooding attacks using game theoty, *IEEE Trans. Dependable and Secure Computing*, vol.7, no.1, pp.5-19, 2010.

[13] J. S. R. Jang, ANFIS: Adaptive network based fuzzy inference system, *IEEE Trans. Systems, Man, and Cybernetics*, vol.23, no.3, pp.665-685, 1993.

[14] B. Rozovskii, A. Tartakovsky, R. B. Blazek and H. Kim, A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods, *IEEE Transactions on Signal Processing*, vol.54, no.9, pp.3372-3382, 2006.
[15] V. A. Siris and F. Papagalou, *Application of Anomaly Detection Algorithms for Detecting SYN Flooding Attacks*, Globecom, 2004.
[16] T. J. Ross, *Fuzzy Logic with Engineering Application*, McGraw-Hill, University of New Mexico, 1995.
[17] *http://www.aircrack-ng.org/*.
[18] IEEE 802.11g, *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2003.
[19] M. S. Gast, *802.11 Wireless Networks: The Definitive Guide*, O'Reilly Media Publisher, 2005.
[20] Y. Xu and Z. Liu, Design and implementation of wireless mesh network testbed based on layer 2 roting, *Wireless Communications, Networking and Mobile Computing*, pp.1-4, 2008.
[21] K. B. Yu and D. J. Murrow, Adaptive digital beamforming for angle estimation in jamming, *IEEE Trans. on Aerospace and Electronic Systems*, vol.37, no.2, pp.508-523, 2001.
[22] C.-H. Hsu, W.-J. Shyr, K.-H. Kuo and P.-H. Chou, Optimal radiation pattern design of adaptive linear phased array antenna using memetic algorithms, *International Journal of Innovative Computing, Information and Control*, vol.4, no.9, pp.2391-2403, 2008.
[23] G. Xu and T. Kailath, Fast subspace decomposition, *IEEE Transaction on Signal Processing*, vol.42, no.3, 1994.
[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flqnnery, *Numerical Recipes in C++*, 2nd Edition, Cambridge University Press, 2002.
[25] J. R. Guerci, *Space-Time Adaptive Processing for Radar*, Artech House, Boston, 2003.
[26] J. Mar and S.-R. Wu, Fast subspace decomposition null extension for two-dimensional array antenna, *ICIEA2010*, 2010.
[27] H. D. Robert, Hybrid ARQ schemes for point to multipoint communication over nonstationary broadcast channels, *IEEE Trans. Commun.*, vol.41, no.9, pp.1379-1387, 1993.
[28] T. S. Rappaport, *Wireless Communications Principle and Practice*, 2nd Edition, Prentice-Hall Inc., 2002.
[29] M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*, Addison-Wesley, 1987.