WLAN Indoor Passive Intrusion Detection Method Based on SVDD

Yong Wang, Xiaoya Zhang(✉), Luoying Gao, Mu Zhou, and Lingxia Li

School of Communication and Information Engineering,
Chongqing University of Posts and Telecommunications, Chongqing 400065, China
xiaoyazhang0@163.com

Abstract. The existing passive intrusion detection technology has poor adaptability under different monitoring environments and low detection performance, this paper proposes a wireless local area network (WLAN) indoor passive intrusion detection method based on Support Vector Domain Description (SVDD). A-distance is adopted to evaluate multiple features to correctly distinguish the average contribution of the two states of silence and intrusion, screening the extreme difference and variance as the characteristic quantity of the signal change. Then, the paper introduces the single classification method SVDD to train the hypersphere anomaly detection boundary in the high dimensional feature space. We can achieve accurate anomaly detection by determining whether the current sample point is within the hypersphere. In a typical indoor environment, compared with the existing detection algorithms, the proposed method achieves better detection performance under low overhead conditions. F1-measure which is the system evaluation index increased by nearly 4%.

Keywords: WLAN · Passive intrusion detection · SVDD

1 Introduction

Passive intrusion detection means that the detected target does not carry any signal transmitting and receiving equipments, and detects the intrusion target via the radio wave [1]. Its potential application range from smart home and elderly guardianship to police security [2].

Our work proposes a WLAN indoor passive intrusion detection algorithm based on Support Vector Domain Description (SVDD). This paper uses A-distance to evaluate multiple features to correctly distinguish the average contribution of the two states of silence and intrusion. Silence stage selects the two features with the highest contribution to construct the feature matrix that trains the SVDD [3–5] model. This paper verifies the effectiveness of the proposed algorithm in the home environment. The results show that the F1-measure of the proposed algorithm can reach 96%, which is much better than the existing WLAN-based passive intrusion detection technology.

M. Jia et al. (Eds.): WiSATS 2019, LNICST 280, pp. 235–241, 2019.
https://doi.org/10.1007/978-3-030-19153-5_24

2 System Model

The proposed indoor intrusion detection system of this paper consists of two phases: offline training phase and online detection phase. The framework is in Fig. 1.

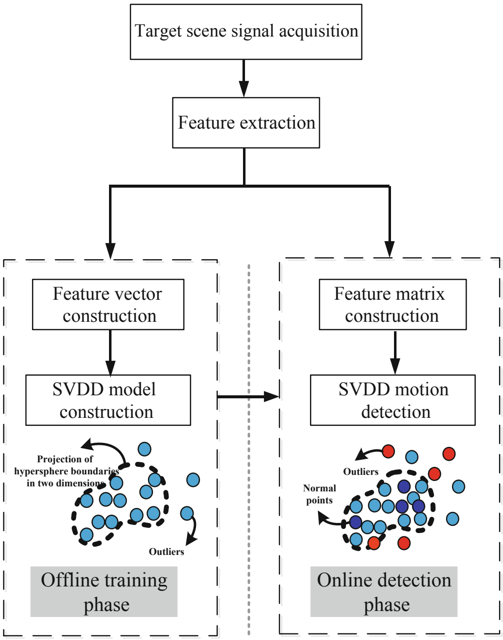


Fig. 1. System model.

2.1 Offline Training Phase

Collecting silent data to build an intrusion detection model for detecting abnormalities in the environment. At this stage, no activity in the sensing area, the receiving devices collect the parameters of received signal strength of each link. Then, we extract the corresponding feature values of each link and constructs a feature matrix based on the collected multi-link signal features. Finally, the SVDD intrusion detection model is trained by the feature matrix.

(a) Characteristic distribution analysis
In order to select features effectively, we use A-distance to evaluate the different distribution of each feature in the state of silence and walking, and select the feature with clear distribution difference to construct the detection model. Then, we use a decision tree as a linear classifier, the A-distance can

be described as $d^{\mathcal{A}}(D_s, D_t) = 2(1 - 2E(h))$, where D_s and D_t are two data sets for constructing the second classifier, $E(h)$ represents the loss of the classifier.

This paper proposes t o design a passive intrusion detection system based on the feature and select the extreme difference and variance as two detection features.

(b) Feature matrix construction

Let k be the number of wireless links in the system. k is equal to the number of signal transmitting devices, multiplied by the number of signal receiving devices. The receiving frequency of receivers is 1 time/second. We use sliding window function (sliding window length is L) to extract signal characteristics $\mathbf{s}_{t,j}$ corresponding to each link j at time t. $\mathbf{s}_{t,j} = [s^1_{t,j}, s^2_{t,j}]$, $s^1_{t,j}$ is the extreme difference and $s^2_{t,j}$ is the variance. Feature vector $\mathbf{S}_t = [\mathbf{s}_{t,1} \cdots \mathbf{s}_{t,j} \cdots \mathbf{s}_{t,k}]$, j represents the wireless links in the system. After acquiring the signals of T moments, $T - L + 1$ feature vectors are obtained, and the feature matrix $\mathbf{S} = [\mathbf{S}_1 \cdots \mathbf{S}_t \cdots \mathbf{S}_N]$ is constructed by the feature vectors, $N = T - L + 1$.

(c) Construction of SVDD passive intrusion detection model

In the SVDD intrusion detection model is trained by using the feature matrix $\mathbf{S}$. The goal of SVDD is to find the smallest sphere in a high-dimensional space that contains all or most of the training samples. In order to avoid the influence outliers effect, a slack variable $\xi_i \geq 0$ is introduced for each training sample point, allowing some points to be outside the hypersphere. Multiplying the penalty parameter C for each slack variable ξ_i, making the system has certain anti-noise performance. The objective function is:

$$\begin{aligned} &\min_{R,\mathbf{a},\xi_i} R^2 + C\sum_i \xi_i, \\ &s.t. \ \ \|\mathbf{S}_i - \mathbf{a}\|^2 \leq R^2 + \xi_i, \xi_i \geq 0, i = 1, \ldots, N. \end{aligned} \tag{1}$$

Lagrangian function can be constructed according to Eq. (2).

$$L = R^2 + C\sum_i \xi_i - \sum_i \alpha_i \left\{R^2 + \xi_i - \|\mathbf{S}_i - \mathbf{a}\|^2\right\} - \sum_i \gamma_i \xi_i. \tag{2}$$

The Lagrange multiplier $\alpha_i \geq 0$, $\gamma_i \geq 0$.

$$\begin{cases} \partial L/\partial \mathbf{a} = 2R(1 - \sum_i \alpha_i) = 0 \Rightarrow \sum_i \alpha_i = 1, \\ \partial L/\partial R = 2\sum_i \alpha_i(\mathbf{S}_i - \mathbf{a}) = 0 \Rightarrow \mathbf{a} = \sum_i \alpha_i \mathbf{S}_i, \\ \partial L/\partial \xi_i = C - \alpha_i - \gamma_i = 0 \Rightarrow C = \alpha_i + \gamma_i, \end{cases} \tag{3}$$

where $\alpha_i \geq 0$ and $\gamma_i \geq 0$, we can get $0 \leq \alpha_i \leq C$.

According to the Lagrangian duality, the minimum value problem of Eq. (2) equals to a maximum value problem of its dual problem. We introduce Gaussian kernel functions $\mathrm{K}(\mathbf{S}_i, \mathbf{S}_j)$ to implement nonlinear mapping from low-dimensional input space to high-dimensional feature space. The following optimization problem is given by:

$$\begin{aligned} &\max_{\alpha} L = \sum_i \alpha_i \mathrm{K}(\mathbf{S}_i, \mathbf{S}_i) - \sum_{i,j} \alpha_i \alpha_j \mathrm{K}(\mathbf{S}_i, \mathbf{S}_j), \\ &s.t. \sum_i \alpha_i = 1, 0 \le \alpha_i \le Ci = 1, 2, \ldots, N. \end{aligned} \tag{4}$$

The optimal solution α_i can be obtained by using Sequential Minimal Optimization (SMO) [6]. When $\alpha_i = 0$, it indicates that the corresponding sample point is located inside the hypersphere; when $\alpha_i = C$, it indicates that the corresponding sample point is outside the hypersphere, which is the limited support vector; when $0 < \alpha_i < C$, it indicates the corresponding sample. The point is on the hypersphere, which is the support vector, and the set of support vectors is represented as $\{V_q, q = 1, 2, \ldots, n\}$. The radius R of the sphere can be found by the distance from any support vector V_q on the hypersphere to the center of the sphere $\mathbf{a}$.

$$R^2 = \left\|V_q - \mathbf{a}\right\|^2 = K(V_q, V_q) - 2\sum_{i=1} \alpha_i K(V_q, \mathbf{S}_i) + \sum_{i,j=1} \alpha_i \alpha_j K(\mathbf{S}_i, \mathbf{S}_j). \tag{5}$$

In most cases, the hypersphere boundary trained by the Gaussian kernel function is better for describing the sample, so the Gaussian kernel function $K_G(\mathbf{S}_i, \mathbf{S}_j) = \exp(-\|\mathbf{S}_i - \mathbf{S}_j\|^2 / \sigma^2)$ is chosen in this paper, and the Eqs. (4) and (5) become:

$$L = 1 - \sum_i \alpha_i^2 - \sum_{i \ne j} \alpha_i \alpha_j K_G(\mathbf{S}_i, \mathbf{S}_j). \tag{6}$$

$$R^2 = 1 - 2\sum_{i=1} \alpha_i K_G(V_q, \mathbf{S}_i) + \sum_{i,j=1} \alpha_i \alpha_j K_G(\mathbf{S}_i, \mathbf{S}_j). \tag{7}$$

2.2 Online Monitoring Phase

At this stage, we extract the characteristic matrix of signal strength to construct a feature vector $\mathbf{S}_{t_{on}} = [\,\mathbf{s}_{t_{on},1} \cdots \mathbf{s}_{t_{on},j} \cdots \mathbf{s}_{t_{on},k}\,]$ at time t_{on}. Then we use the formula $m_{t_{on}} = \|S_{t_{on}} - \mathbf{a}\|^2 / R^2$ to detect environmental anomaly degree, $m_{t_{on}} > 1$ indicates someone invaded. Otherwise, there is no activity in the current perceived environment.

3 Experiment Results

We verify performance in a typical indoor environment which area is $59.48\,\mathrm{m}^2$. We uses Huawei glory router (WS851) as the access point AP. Samsung mobile phones (MP1: GT-I9308 and MP2: GT-S7568) as monitoring points MP. Figure 2 shows the indoor scenario, which is arranged with 4 APs and 1 MP. Firstly, we perform silent data collection for 25 min and train the SVDD intrusion detection model. Then, the detection model is tested in a test environment.

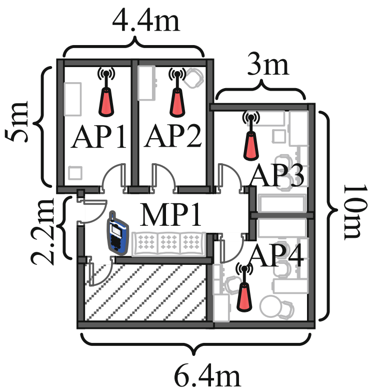
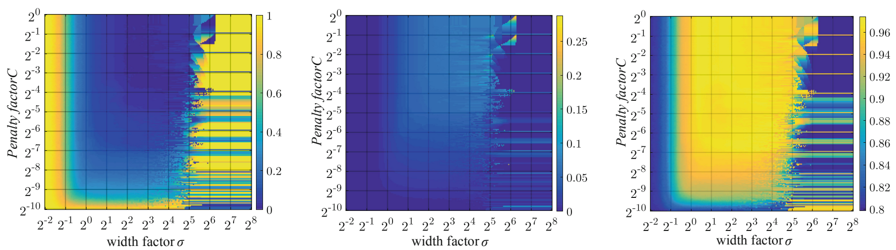


Fig. 2. Indoor scene.

In order to explore influence of penalty factor C and Gaussian kernel width factor σ on the detection performance of the system, we use FP (False Positive), FN (False Positive) and $F1 - measure$ in the experimental scene. The results are shown in Fig. 3.

As can be seen from Fig. 3, a smaller penalty factor $C(< 2^{-9})$ leads to a higher false alarm rates FP. However, as C increases, the false alarm rate FP and $F1 - measure$ gradually increase. When $C > 2^{-7}$, impact on the system is negligible. But, during actual deployment, excessive abnormal points can make the radius too large, and the FN too high in the training data. C should not be too large, therefore, we set $C = 2^{-4}$. The Gaussian kernel width factor σ is too small($\sigma < 1$) or too large ($\sigma > 2^6$), the false alarm rate FP will be higher. $F1 - measure$ shows a trend of rising first and then falling. In the scene, the most optimal $\sigma_{opt} = (2^{0.3} + 2^{5.05})/2 \approx 17.18$.



(a) FP. (b) FN. (c) F1-measure.

Fig. 3. Effect of C and σ on detection performance in indoor scene.

In this section, the performance of the proposed method is compared with other passive RSSI-based passive intrusion detection techniques. The results are

Table 1. Comparison of detection performance with existing detection technologies.

Performance		MA	MV	Ichnaea	SVDD
Indoor scene	FN	0.110	0.119	0.093	0.054
	FP	0.103	0.097	0.095	0.040
	F1-measure	0.017	0.891	0.906	0.960

given in Table 1. It can be seen from Table 1 that the proposed method has a certain degree of reduction in false alarm rate FP and missed detection rate FN, which is more than 4% higher than the classical Ichnaea system, which proves the effectiveness of the proposed algorithm.

4 Conclusion

In order to achieve human activity detection in a real wireless environment. This paper proposes a WLAN indoor passive intrusion detection algorithm based on SVDD. The algorithm extracts the salient features of the signal. In this paper, the extreme difference and variance are used to train the SVDD passive intrusion detection model to identify the normal state and the abnormal state. Accuracy is significantly improved. FN and FP are lower than 5% and 9%, respectively. $F1-measure$ is higher than 96%, compared with algorithms based on statistical properties, it has improved by nearly 4%, which proves our system has better performance.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (61771083,61704015), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental and Frontier Research Project of Chongqing (cstc2017jcyjAX0380, cstc2015jcyjBX0065), University Outstanding Achievement Transformation Project of Chongqing (KJZH17117), and Postgraduate Scientific Research and Innovation Project of Chongqing (CYS17221).

References

1. Moustafa, Y.: Challenges: device-free passive localization for wireless environments. In: IEEE International Conference on Pervasive Computing Communication Workshops, Sydney, pp. 1–2 (2016)
2. Youssef, M., Mah, M., Agrawala, A.: Challenges: device-free passive localization for wireless environments. In: ACM International Conference on Mobile Computing and NETWORKING, Canada, pp. 222–229 (2007)
3. Tax, D.M.J., Duin, R.P.W., Agrawala, A.: Support vector domain description. Pattern Recognit. Lett. **20**(11–13), 1191–1199 (1999)
4. Jinhong, Y., Tingquan, D.: A one-cluster Kernel PCM based SVDD method for outlier detection. Acta Electron. Sin. **45**(4), 813–819 (2017)

5. Long, L., Zheng, L.: Identifier for Radar ground target based on distribution of space of training features. J. Electron. Inf. Technol. **38**(4), 950–957 (2016)
6. Keerthi, S., Gilbert, E.: Convergence of a generalized SMO algorithm for SVM classifier design. Mach. Learn. **46**(1–3), 351–360 (2002)