Intruder Identification in IEEE 802.11 Wireless Infrastructure Using Localization

Asish Kumar Dalai and Sanjay Kumar Jena

***Abstract*—The method of obtaining the physical co-ordinates of a device is known as localization. Location information of the wireless devices in a network has many applications. Generally we use GPS (Global Positioning System) to get the location but it fails to provide the service in an indoor environment. Therefore we need a local positioning system, which can provide the relative location information in a network. Many approaches has been made for wireless localization using the distance obtained through Time of Arrival (ToA), Time Difference of Arrival (TDoA), Received Signal Strength (RSS) etc.. The proposed method uses RSS with minimum overhead compared to conventional multiliteration techniques. Due to the growing importance of security in Wi-Fi, the location information has been used for intrusion detection. The proposed model compares the intruder information with preset profile of genuine devices for full proof detection of the intrusion. It has been observed that the model outperforms its counterparts.**

***Index Terms*—Intrusion detection, wireless security, localization.**

I. Introduction

A wireless local-area network (WLAN) uses radio waves to connect devices such as laptops, smart phones and other wireless devices. Unlike wired LAN (Local Area Network), WLAN provides all features of a network along with seamless mobility. Increased use of laptops, mobiles and hand-held devices within the enterprise, and the increase in user mobility have fueled the demand for wireless networks. According to reports [1]: number of Wi-Fi hotspots will become triple by 2015. In future Wi-Fi devices will use more bandwidth than wired devices, according to Cisco's Global Mobile Data Traffic Forecast Update [2]. The use of wireless devices and networks is increasing at an exponential rate. The devices are characterized by low cost, flexibility and ease to use. IEEE

802.11 [3], a Institute of Electrical and Electronics Engineers (IEEE) standard, is the widely used standard for WLAN. The scope of this standard is to develop a Medium Access Control (MAC) and physical layer specification for wireless connectivity for fixed, portable, and moving stations within a local area. In the existing IEEE standards for WLAN, there is no provision for obtaining the location information of the wireless devices. The location based information of the Wi-Fi devices has many different applications, which leads to the development of a system which can provide the approximate location of the device. Wi-Fi-based local positioning system is used where GPS is inadequate due to various causes including multi path and signal blockage indoors. Such systems include indoor positioning systems. Wi-Fi positioning takes advantage of the rapid growth in the early 21st century of wireless access points in urban areas. In the proposed technique the RSS value is used for obtaining the location using Fangs [4] 3D positioning system. Out of the several application of Wi-Fi localization, this method uses the location information for security, i.e. to identify intruders in a wireless network.

Due to the inherent characteristic of not being bounded by walls and perimeters, the wireless network has many security issues. In order to provide security equivalent to its predecessor, i.e. wired network, 802.11 specified security standards known as Wired Equivalent Privacy, Wi-Fi Protected Access and IEEE 802.11i. But all those enhancements failed to achieve the desired objectives especially network injection, man-in-the-middle attacks, identity theft (MAC spoofing), malicious association / disassociation, honeypot / evil twin attack and Denial of Service Attack. Unauthenticated management and control frames provide no protection against identity theft and are the main cause of such vulnerabilities. Therefore, there is a need of a system which can address these issues. Our proposed method targets the malicious node in the network, which is accessing the network by not being an authorized member. As the wireless radio frequency is not bounded by any walls it can spread beyond the boundary. Hence an attacker can capture the signals sitting far away by using high gain antennas. Furthermore, the attacker can misuse our network. Hence to protect such scenario, we have set some boundary conditions for the devices to be operated in our network. If any of the devices is not satisfying the boundary condition then that device is considered as malicious, then the malicious devices is sent for further verification to test its authenticity. If the device fails then it is spotted and kept out of the network.

The rest of the paper is structured as follows: In Section II, the related works for Wi-Fi localization and intruder identification are discussed. The proposed method for intruder detection using wireless localization has been discussed in Section III. Results and Discussion of the experimental work is given in Section IV. Concluding remarks are given in Section V.

II. Related Works

Localization and intrusion detection in Wi-Fi are active area of research. Here we provide a brief overview of some

Manuscript received May 5, 2014; revised July 13, 2014.

The authors are with the Department of Computer Science and Engineering, National Institute of Technology Rourkela, India (e-mail: dalai.asish@gmail.com, skjena@nitrkl.ac.in).

DOI: 10.7763/IJFCC.2014.V3.330

key research contributions to these areas. Location estimation is mainly done by GPS satellites. As GPS cannot track devices in an indoor environment, some local positioning systems has been proposed.

The initial stage for localization is based on distance angle estimation between nodes. Different techniques for localization using distance/angle estimation are classified into: time of arrival [5], time difference of arrival [6], received signal strength indicator [7]–[10], and angle of arrival [11]. Currently, the Wi-Fi RSS-based positioning is considered to have great importance for localization [9]. RSSs are more cost effective, because they are compatible with existing wireless devices without hardware modification. The basic idea in Wi-Fi location systems is to determine the relationship between the measured RSS and the user's location based on a previous set of measurements [10]. When a mobile device requests services, it compares the online RSS from nearby APs with values stored in the database to determine its location. This approach is known as "radio fingerprinting" in the literature [12], [13]. Many works have indicated that the positioning accuracy is non monotonically improved as the number of access points increases, because the rising number of access points results in the addition of more information, thus incur- ring more noise and information duplication [12], [14], [15]. After obtaining the distance value various techniques used to estimate a node's location such as trilateration, multilateration, and triangulation. Estimated distance and the position of tracking devices is used to estimate the location. Triliteration and multiliteration are conventional techniques and involves intense calculation. Triangulation is a geometric technique that uses the trigonometry laws of sine and cosines on the angles of incoming signal to estimate a unique location. Angle of arrival measurement requires bulkier and expensive hardware such as multi-sectored antennae. This makes triangulation unsuitable in Wi-Fi Localization. Hence our method uses a location estimation technique which is simpler than its counterparts. Inspired by [4] and [16], our experiments extends the projection to a more generalized form in which the obtained location information is used in intrusion detection.

Intrusion detection system for wireless networks have been researched from different perspectives. Many of them focus on network topology monitoring, some of them consider different and independent layers traffic analysis and others assume knowledge of network infrastructure. A significant amount of research has been done in the design and analysis of wireless intrusion detection system. Wireless networks are employed both in private and corporate networks. Many proprietary as well as open source solutions were developed. Some open source solutions are: Snort-Wireless and Kismet. Snort- Wireless is an IDS that checks each 802.11 frame against a rule-set. If the rule set is violated an alarm is raised. Kismet is a more complex IDS. It is a signature-based distributed IDS that checks for known attacks. Proprietary solutions are: AirMagnet, Red-M and Air Defense. All these solutions are far more complex than previous ones since they are developed to be a comprehensive IDS solution. Therefore they are not really classifiable in a specific category. They are designed to monitor specific networks, therefore they use information's about network topology, policies and network infrastructure. Although IEEE 802.11i provides effective countermeasures through strong authentication, Confidentiality and integrity algorithms but there are networks where IEEE 802.11i is not used for many reasons such as presence of legacy devices, complexity of 802.11i deployment, presence of anonymous users etc.. Our method uses two step detection process, first it obtain the location of the devices operating from outside the network and then it compares the device with stored profiles for accurate detection.

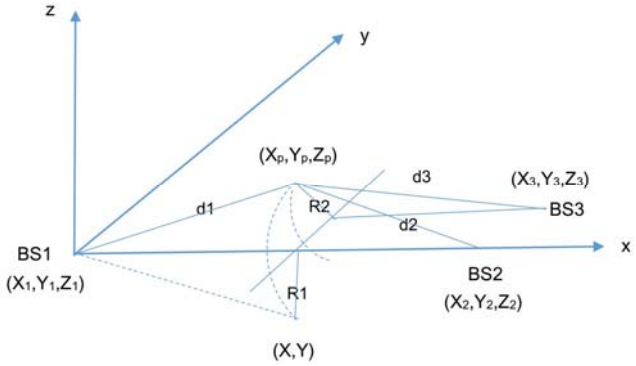

Fig. 1. Architecture of the model.

III. Proposed Method

A. Wi-Fi Localization

The localization algorithm uses three base stations to de termine the location in an IEEE 802.11 wireless infrastructure. The base stations are chosen with range, which covers the entire Wi-Fi deployment. The positioning of the base stations and the deployment of our method is given in the Fig. 1.

For our localization method we have taken Fang's simple solution for hyperbolic and related position technique [4]. Let us assume that the coordinates of BS1, BS2 and BS3 in a local right handed orthogonal coordinate system are (X_1 , Y_1, Z_1), (X_2 , Y_2 , Z_2) and (X_3 , Y_3 , Z_3), respectively. The base station BS1 is at the origin, BS2 is on the axis along the station baseline, and another station is on the XY plane. Therefore according to the RSS from each of the three base stations, the device position (Xp, Yp, Zp) is the intersection of the three spheres centered at BS1 (0, 0, 0), BS2(X_2, 0, 0) and BS3(X_3, $Y_{3,0}$) with radius d_1, d_2 and d_3 respectively.

For calculating the distance of the device from the base stations, we have used the RSS value. Radio signal attenuates when the distance between the transmitter and receiver in- creases. With the increase in distance, strength of radio signal decreases exponentially. The attenuation in signal strength is measured by the receivers received signal strength indicator circuit. Received signal strength indicator estimates the distance covered by a signal to the receiver by measuring the power of received signal. Decrease in transmitted power at the receiver can be calculated and translated into an estimated distance. The usage of RSSI in distance calculation can be interpreted as [17].

$$Pr\,(dBm) = A - 10.\,\eta.\,\mathrm{Log}\,(d) \qquad (1)$$

where Pr is the received signal power given in dBm, A is the signal power at a distance of one meter and η is the path loss factor. Using the above Equation we can easily calculate the distance. We have used Kismet, an open source Wi-Fi sniffer to calculate the received signal power and the distance value. Using formula of right angle triangle the equation for BS1 is:

$$d_1^2 = (X - X_1)^2 + (Y - Y_1)^2 \quad (2)$$

And the equation for BS2 is

$$d_2^2 = (X - X_2)^2 + (Y - Y_2)^2 \quad (3)$$

Since $Y_2 = Y_1$, Equation (3) can be rewritten as

$$d_2^2 = (X - X_2)^2 + (Y - Y_1)^2 \quad (4)$$

Combining Equation (2) and (4), we can find the X coordinates of the intersection points as

$$X = \frac{d_2^2 - d_1^2 + X_1^2 - X_2^2}{2(X_1 - X_2)} \quad (5)$$

Substituting Equation (5) in (4), we get the Y coordinates of the intersection points as

$$Y = \sqrt{d_1^2 - X_1^2 - X^2 + 2XX_1} + Y_1 \quad (6)$$

To obtain the 3D location it is decomposed into 2D rotation,

Which results: $Xp = X$ and radius of the circle generated after rotation is $R_1 = Y$. Then the projection of vector d_3 on the circle plane is calculated as:

$$R_2 = \sqrt{d_3^2 - (\mathrm{X}_3 - X_P)^2} \quad (7)$$

The next step is used to calculate the coordinates of two intersection points for two circles with radius R_1 and R_2:

$$Y_p = \frac{(R_1^2 - R_2^2 + Y_3^2)}{2Y_3} \quad (8)$$

$$Z_p = \sqrt{R_1^2 - Y_p^2} \quad (9)$$

Only the positive value of Z_p is considered, because as per our topology of base stations a node can never have negative value for the Z coordinate. The position (X_p, Y_p, and Z_p) provides the 3D location of the node.

B. Intruder Identification

The obtained coordinates are compared with the preset boundary conditions. The maximum coordinate limit has been set for the wireless devices, let's say (X_m, Y_m, Z_m). If the node exceeds the limit, i.e. (X_p, Y_p, Z_p) > (X_m, Y_m, Z_m) then the node is considered as suspicions and it needs further investigation. A profile of every registered device is maintained in the base station BS1. The profiles are generated using Aircrack-ng and its associated modules. The profiles of each device contain the following parameters: Basic Service Set Identifier (BSSID) / MAC Address, Channel No, Potential bandwidth and Round-Trip Delay time (RTD). BSSID / MAC Address and the Channel Number can be directly obtained from the management frames, but these parameters are subject to spoofing. Therefore we have chosen the above features which needs an intensive calculation and will differ, even if the attacker tries some technique to impersonate the registered device.

Time	Length	Source	Protocol
0.000000	86	erasy_6b:68	802.11
0.139215	160	erasy_6b:68	802.11
0.144215	30	erasy_6b:68	802.11
0.149090	46	erasy_6b:68	802.11

Fig. 2. Calculation of the bandwidth consumption by a node.

The theoretical bandwidth varies due to several factors like; distance, interference and shared bandwidth. Hence the actual bandwidth consumption of each device is a unique characteristic. For example IEEE 802.11b has a maximum raw data rate of 11 Mbps. Due to the CSMA/CA protocol overhead, in practice the maximum 802.11b throughput that an application can achieve is about 5.9 Mbps. To calculate the bandwidth consumption of a particular device we have filtered the traffic with the device's source address (wlan.sa ==00: 11: 88: XX: XX: XX) and mark the packets at the beginning of the file transfer. After one second the cumulative length field is calculated as given in Fig. 2, which gives the data rate. The bandwidth consumption of the device is obtained by dividing the maximum throughput of the channel with the calculated data rate. The attacker machine will definitely consume more bandwidth than other devices, which makes it easy for identification.

The round-trip delay time (RTD) or round-trip time (RTT) is the length of time it takes for a signal to be sent plus an acknowledgment of that signal to be received. This time delay therefore consists of the propagation times between the two points of a signal. To obtain the RTT value we have used the ICMP ECHO request. The average time for the response is considered as the RTT value of the device.

The features are calculated for the suspicious device and then matched with the stored profile of registered devices. If no such device is found with matching profile, then the device is spotted and kept out of the network.

IV. Results and Discussion

The comparison of the conventional multiliteration techniques and the proposed technique for 3D localization has been given in Table I. It has been observed that the computational overhead of the multiliteration technique is higher than the simplified technique used in our model, which makes the proposed model faster than its counterparts.

The method has been implemented in a testbed. The base stations are arranged as per the architecture discussed in Section III.

The packet capturing is done by using a packet sniffer tool known as Wireshark. During the training phase the profiles are generated for the registered devices as discussed in Section III-B. Table II shows the profile of the devices.

TABLE I: CO MPA RISON IN TERMS OF WEIGHT SO FOPER

Method	Addition/ Substitution / Shift	Multiplication /Division	Square/ Square root
Multiliteration	52	13	36
Proposed	22	3	26

TABLE II: PROFILE OF REGIS TERED DE VICES

BSSID / MAC Address	Channel No	Bandwidth Consumption	Signal Power (dBm)	RTT (ms)
00:21:29:E9:38:XX	6	11.8	-44	10
00:11:88:98:3A:XX	6	14.6	-87	4
00:17:C4:1C:95:XX	6	12.1	-91	4
00:20:A6:51:E3:XX	6	9.5	-77	5

The testbed contains three base stations and four authorized devices. Initially a preset boundary condition has been set, specifying the 3D perimeter for the devices to operate. To check the trustworthy of the model one impostor devices has been implanted intentionally. Now by using the pro- posed localization technique the 3D location of the devices operating in the network has been obtained. We compare the obtained coordinates with the preset boundary conditions. The maximum coordinate limit has been set for the network is 30 meters, 20 meters and 15 meters in X, Y and Z coordinates respectively. The node that exceeds the limit, i.e. $(X_p, Y_p, Z_p) >$ (30, 20, 15) is considered as an outsider. The Table III shows the result of the 3D localization algorithm.

TABLE III: 3D LOCATION OF THE DE VICES OPERATING IN THE NETWORK

Device no	BSSID / MAC Address	3D Location	Decision
1	00:21:29:E9:38:XX	< 6, 5, 7 >	Insider
2	00:11:88:98:3A:XX	< 15, 11, 9 >	Insider
3	00:17:C4:1C:95:XX	< 21, 10, 11	Insider
4	00:20:A6:51:E3:XX	< 4, 15, 10 >	Insider
5	00:11:88:98:3A:XX	< 7, 25, 4 >	Outsider

From the result given in Table III it has been found that only one device i.e. device no 5 with MAC id "00:11:88:98:3A: XX" is identified as an outsider. The result obtained by local- ization algorithm is not giving 100% precision, as there are many factors like noise, signal interference, obstacles etc. gives an error rate of 2 to 3 meter. Therefore the devices are sent for further investigation by using the intruder identification module. Here the selected features for the impostor device has been obtained by using the methods discussed in Section III-B. The features of the device no 5 is as found as < 00:11: 88: 98: 3A: X X, 6, 24.3, −114dBm, 10ms >. It has been observed that, the MAC Address is registered with our network and also the device is operating in the same channel i.e. channel no. 6. Even though the impostor has tempered the MAC address with one of the registered device and also operating in the same channel, it fails to meet the other criteria to prove its authenticity. Therefore the device is identified and disconnected from our network.

V. CONCLUSION

The proposed model is a novel technique for outlier detection in Wi-Fi infrastructure. It uses the localization information to track the devices operating from outside the preset boundary. It uses the localization technique which is faster than the conventional multiliteration approach. Considering that a genuine device might be using the network from outside, it is further verified against the stored profiles. Once the intrusion is finalized we can spot the device using the location information. As the features like MAC address, channel number can be spoofed, our method includes many device parameters which is difficult for the attacker to impersonate. The method can be further enhanced to strengthen the security over Wi-Fi network in detection of Rouge AP and Denial of Service attack.

REFERENCES

[1] I. Plc. Informa Plc. (2014). [Online]. Available: http://www.informa.com/Media-centre/Press-releases–news/Latest-News/Wifi-hotspots-set-to-more-than-triple-by-2015/.

[2] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update. (2014). [Online]. Available: www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/whitepaperc11520862.html/.

[3] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 wireless local area networks," *Communications Magazine*, vol. 35, no. 9, pp. 116–126, 1997.

[4] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748–753, 1990.

[5] C. H. Chen, K. T. Feng, C. L. Chen, and P. H. Tseng, "Wireless location estimation with the assistance of virtual base stations," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 93–106, 2009.

[6] B. C. Liu and K. H. Lin, "Sssd-based mobile positioning: on the accuracy improvement issues in distance and location estimations," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1245–1254, 2009.

[7] R. W. Ouyang, A. S. Wong, and C. T. Lea, "Received signal strength-based wireless localization via semidefinite programming: noncoopera- tive and cooperative schemes," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1307–1318, 2010.

[8] A. J. Weiss, "On the accuracy of a cellular location system based on rss measurements," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 6, pp. 1508–1518, 2003.

[9] S. Mazuelas, A. Bahillo, R. M. Lorenzo, P. Fernandez, F. A. Lago, E. Garcia, J. Blas, and E. J. Abril, "Robust indoor positioning provided by real-time rssi values in unmodified wlan networks," *Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 821–831, 2009.

[10] S. H. Fang and T. N. Lin, "A dynamic system approach for radio location fingerprinting in wireless local area networks," *IEEE Transactions on Communications*, vol. 58, no. 4, pp. 1020–1025, 2010.

[11] C. Botteron, A. H. Madsen, and M. Fattouche, "Effects of system and environment parameters on the performance of network-based mobile station position estimators," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 163–180, 2004.

[12] V. Honkavirta, T. Perala, S. A. Loytty, and R. Piche, "A comparative survey of wlan location fingerprinting methods," *Positioning, Navigation and Communication*, 2009, pp. 243–251.

[13] M. B. Kjærgaard, "A taxonomy for radio location fingerprinting," *Location-and Context-Awareness*, 2007, pp. 139–156.

[14] Y. Chen, Q. Yang, J. Yin, and X. Chai, "Power-efficient access- point selection for indoor location estimation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 7, pp. 877–888, 2006.

[15] T. King, T. Haenselmann, and W. Effelsberg, "Deployment, calibration, and measurement factors for position errors in 802.11-based indoor positioning systems," *Location-and Context-Awareness*, Springer, 2007, pp. 17–34.

[16] E. Doukhnitch, M. Salamah, and E. Ozen, "An efficient approach for trilateration in 3d positioning," *Computer Communications*, vol. 31, no. 17, pp. 4124–4129, 2008.

[17] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *Signal Processing Magazine*, vol. 19, no. 2, pp. 17–29, 2002.

Asish Kumar Dalai received his B.Tech in computer science and engineering from Gandhi Institute of Engineering and Technology, Odisha, India and M. Tech from the Department of Computer Science and Engineering, National Institute of Technology Rourkela in 2005 and 2013, Respectively. He is currently working toward his Ph.D. on application intrusion detection system. His research interests include web security, intrusion detection system.

Sanjay Kumar Jena received his M.Tech in computer science and engineering from the Indian Institute of Technology Kharagpur and Ph.D. from the Indian Institute of Technology Bombay in 1982 and 1990, Respectively. He is a fulltime professor in the Department of Computer Science and Engineering, National Institute of Technology Rourkela. He is a senior member of IEEE and ACM and a life member of IE(I), ISTE and CSI. His research interests include data engineering, information security, parallel computing and privacy preserving techniques