A single supervised learning model to detect fake access points, frequency sweeping jamming and deauthentication attacks in IEEE 802.11 networks

Andy Amoordon, Virginie Deniau, Anthony Fleury, Christophe Gransart

Machine Learning with Applications 10 (2022) 100389

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

A single supervised learning model to detect fake access points, frequency sweeping jamming and deauthentication attacks in IEEE 802.11 networks

Andy Amoordon [a,*], Virginie Deniau [a], Anthony Fleury [b,c], Christophe Gransart [a]

[a] University Gustave Eiffel, Campus Villeneuve D'Ascq, France
[b] IMT Nord Europe, CERI SN, Institut Mines Telecom, France
[c] University of Lille, Douai, France

ARTICLE INFO

Keywords:
IEEE 802.11 networks
Attacks
Jamming
Fake access point
Deauthentication
Supervised learning algorithms
Network Intrusion Detection System

ABSTRACT

Wireless networks are nowadays indispensable components of telecommunication infrastructures. They offer flexibility, mobility and rapid expansion of telecommunication infrastructures. In wireless networks, transmissions are unisolated and most commonly emitted using omnidirectional antennas. This makes wireless networks more vulnerable to some specific attacks as compared to wired networks. For instance, attacks such as fake access points, intentional jamming and deauthentication can be easily perpetrated against IEEE 802.11 networks using freely accessible software and cheap hardware. Intentional jamming and deauthentication attacks are standalone attacks, but they can be combined with the fake access point attack to increase the latter's effectiveness. In our research, we work on methods to detect the three different attacks when they are perpetrated independently (one at a time) or concurrently (several at the same time). In this contribution, we present a model that can detect the three attacks, when perpetrated independently, by analysing a set of features (frame interval, Received Signal Strength Indicator, sequence number gap and management frame subtype) extracted from IEEE 802.11 management frame and radiotap headers. We have implemented the model using several supervised learning algorithms. The model with Random Forest and the K-Nearest Neighbour predictors have best detection precision (over 96 %) for fake access point and deauthentication attacks and perfectible detection precision for the intentional jamming attack (over 81%).

Contents

* Corresponding author.
E-mail addresses: amoordon.andy@yahoo.com (A. Amoordon), virginie.deniau@univ-eiffel.fr (V. Deniau), anthony.fleury@imt-nord-europe.fr (A. Fleury), christophe.gransart@univ-eiffel.fr (C. Gransart).

https://doi.org/10.1016/j.mlwa.2022.100389
Received 3 June 2022; Received in revised form 11 July 2022; Accepted 22 July 2022
Available online 1 September 2022

1. Introduction

Wireless networks are nowadays widely spread. They are preferred to wired networks as they offer flexibility, mobility, and rapid expansion of telecommunication infrastructures. However, as compared to wired networks, they are vulnerable to some specific attacks. In wireless networks, as radio frequency (RF) signals are emitted in all directions, over a spatial coverage determined by the emission power, all devices within the spatial coverage receive this RF signal. For this reason, attackers can more easily perform eavesdropping attacks on wireless networks. In an eavesdropping attack, an attacker listens and keeps data that are not addressed to them. By doing so, they can steal sensitive information which can be dangerous and intrusive to users' privacy. To prevent eavesdropping attacks, source and destination devices can encrypt their communication using encryption key. When the communication is encrypted, the eavesdropper can still receive the communication but cannot (depending on the encryption) fully or partially understand the transmission. Encryption, however, is not the panacea and can be broken using brute-force or dictionary attacks. In practice, attackers (with traditional computers) rarely use brute-force and dictionary attacks as the former is time-consuming and the latter is not always successful. Attackers tend to favour more elaborated schemes such as the Man-in-the-Middle attack. Man-in-the-Middle (MITM) is an attack in which the attacker tries to place himself stealthily between a sender and a receiver. If the attacker succeeds, he has control of the data, and can read, modify or fabricate data (Amoordon, Gransart, & Deniau, 2020). Man-in-the-Middle attacks also render encryption ineffective, the devices are tricked and they unknowingly perform the encryption key exchange procedure with the attacker's device. Concerning the Man-in-the-Middle attack, in our work, we study the detection of fake access points in IEEE 802.11 networks (Wi-Fi)[1] in infrastructure mode. In the Wi-Fi infrastructure mode, each transmission passes through an access point (logical connection). Therefore, to perform a Man-in-the-middle attack on IEEE 802.11 networks, the attacker creates a fake access point and waits or forces devices to connect to this access point. Attackers can force users to connect to the fake access point by combining the fake access point attack with other types of attack such as the deauthentication or jamming attack.

Jamming attacks are the intentional emissions of radio interference in order to disrupt radio communications by decreasing the Signal-to-Interference-plus-Noise ratio (SINR) (Berg, 2008). The efficiency of jamming attacks depends on the power level of the jammer and the distance between the jammer and the communication receivers. When the jamming power level is high or when the jammer is close to the communication receivers, the effect of the jammer will most probably be the annihilation of radio communications or disconnection of clients. In other cases, the jamming signals may cause disruptions like a decrease in data rate or throughput (Pirayesh & Zeng, 2021). Concerning intentional jamming, in our work, we study the impact of frequency sweeping jammers because they are the most commonly found on the Internet for purchase (Deniau, Gransart, Romero, Simon, & Farah, 2017). A frequency sweeping jammer continuously emits radio interference signals in a frequency range, regardless of the presence of ongoing communications. An attacker can therefore use frequency sweeping jamming signals to disconnect devices from the licit access point and entice them to associate with his fake access point (Amoordon et al., 2020). A deauthentication attack is the excessive emission of deauthentication frames. Deauthentication frames can be sent by clients or access points to inform the receiver of the deauthentication frame that their communication must end. Since, these frames are unencrypted and unauthenticated, attackers can easily forge them and transmit them in the place of a licit access point. Similarly to jamming signals, they can use these frames to disconnect devices from the licit access point and attract devices to their fake access point. The IEEE 802.11w amendment[2] proposes to protect some management frames such as deauthentication or deassociation frames. However, in practice, many devices do not implement this protection as they are either legacy devices (manufactured before 2009) or simply because the user has not enabled this feature. Moreover, both the access point and the devices should be IEEE 802.11-w compatible to benefit from this protection.

The three aforementioned attacks can be easily perpetrated using readily available software and cheap hardware. Deauthentication frames and a fake access point can both be forged/created using commercial Wi-Fi network cards and freely available software while jamming attacks can be perpetuated using off-the-shelf, ready-to-use and portable jammers for a few dozen dollars or a software-defined radio device like the RTL-SDR[3] that costs under $20. This means that nowadays, depending on the encryption used by the users, attackers can easily attack Wi-Fi networks to disturb the network and steal at worst private information (credit card numbers, social network login...) and at best connection details (visited websites, connection date, time spent on each website...). On certain networks, if these attacks are performed successfully, it can help the attackers to inject false information or commands. For instance, in security systems using wireless networks for video surveillance or remote access, such injection attacks constitute security threats. Depending on the encryption used in the network, these attacks can affect all users in the network or target a specific user(Amoordon et al., 2020). Users who use public networks (guest networks in companies, hot-spots in airports and train stations) are the most vulnerable ones as they, nowadays, only have incomplete tools to verify if the network they want to use is safe or not. For these reasons, it is important to be able to detect and mitigate these attacks. This should help to reduce attacks, increase security and trust in wireless networks. The objective of our research is to create a single tool to detect frequency sweeping jamming signals, the excessive emission of deauthentication frames, and the presence of fake access points in IEEE 802.11 networks when they are perpetrated independently (one at a time) or concurrently (several at a time). To detect these attacks, we have adopted an anomaly-based approach. It consists in comparing attack situations with a normal situation to identify anomalies. We have reproduced four situations in laboratory experiments: a normal situation with normal Wi-Fi communications and three attack situations (fake access point, jamming, and deauthentication attacks). We have then analysed the captured frames to identify relevant features and associated thresholds above which we can assert that there is the presence of one of the aforementioned attack. Using these thresholds, we have implemented a first version of a Network Intrusion Detection System (NIDS) (Amoordon, Gransart, Deniau, & Fleury, 2021) that can detect the three attacks independently using two different indicators. However, threshold definitions can be challenging to set when there are several features, dimensions or the existence of overlapping data between two situations, to take into consideration. For this reason, in a second step, we have adopted a machine learning-based approach (Bierbrauer, Chang, Kritzer, & Bastian, 2021; Ferrag, Maglaras, Moschoyiannis, & Janicke, 2020). This machine learning approach is relatively easy to implement, which makes the NIDS relatively cheap to manufacture and use. In this paper:

- We present a machine-learning based NIDS designed to detect the presence of the three aforementioned attacks when perpetrated independently by analysing a set of features (frame interval, Received Signal Strength, Sequence number gap, and subtype) extracted from management frames. The NIDS analyses only management frames and does not analyse data or control frames.

[1] IEEE 802.11 standards. https://www.ieee802.org/11/.

[2] IEEE 802.11w-2009. https://en.wikipedia.org/wiki/IEEE_802.11w-2009.

[3] "The RTL-SDR is an ultra cheap software defined radio based on DVB-T TV tuners with RTL2832U chips. https://www.rtl-sdr.com/.

- We present, compare and comment the detection results of different supervised learning algorithms that we have used to create the model of the NIDS. With the Random Forest algorithm (number of estimators = 100, min_samples_leaf= 3 and criterion = entropy), the model can detect the deauthentication, fake access point and jamming attacks with a precision of 99.80%, 96.41% and 82.12% respectively. With the KNN (K=4) predictor, the model can detect the deauthentication, fake access point and jamming attacks with a precision of 99.67%, 95.83% and 78.30% respectively.

The paper is organized as follows: In Section 2, we provide background and highlight related works on the detection of fake access points, deauthentication and jamming attacks in Wi-Fi networks. In Section 3, we detail the experimental setup and the data collection phase. In Section 4, we explain our set of features and why we chose it. In Section 5, we detail and analyse our results. Finally, in Section 6, we conclude by giving perspectives for future works.

2. Background and related works

If Deauthentication attacks are specific to IEEE 802.11 networks, Man-in-the-Middle attacks are more generic attacks that can be perpetrated against wired and wireless networks. Jamming attacks, latterly are attacks which are specific to wireless networks and can be perpetrated against different types of wireless communication protocols. The aim of this subsection is to give background about the functioning of IEEE networks and present academic works around Network Intrusion Detection Systems concerning these attacks. Since our NIDS uses information found on frames (OSI layer 2), we will focus on research works using frames or information available at OSI layer 2 (See OSI Model[4] Other research works about Network Intrusion Detection Systems based on layer 1 and upper layers will be cited but not detailed. Concerning non-academic works or industrial solutions, some Antivirus and internet security companies (such as Norton) provide tools to detect if a Wi-Fi network is safe from attacks such as Man-in-the-Middle attacks. However, the functioning and limitations of these software are not detailed. It seems that they are limited to a MAC address based detection and that they are not able to detect all of the three attacks.

To detect these attacks in IEEE 802.11 networks on OSI layer 2, authors use information found on transmitted frames. In IEEE 802.11 networks, only three types of frames can be transmitted: data frames, management frames, and control frames. Data frames contain actual user data and are used to send data from one device to another. Control frames are frames that are, for example, used to control the emission and reception of data frames. Management frames are frames sent to manage the clients and perform supervisory functions. They are, for example, used to connect, leave the wireless network or to move association (connected devices) from one access point to another during roaming. These three types of frames have subtypes frames. For instance, deauthentication frames are a subtype of management frames. In this subsection, we present passive methods described in the literature that analyse management frame information to detect fake access points, deauthentication attack, and frequency sweeping jamming signals against IEEE 802.11 Wi-Fi networks.

[4] The OSI model, "The Open Systems Interconnection model (OSI model) is a conceptual model that describes the universal standard of communication functions of a telecommunication system or computing system, without any regard to the system's underlying internal technology and specific protocol suites. [...] In the OSI reference model, the communications between a computing system are split into seven different abstraction layers: Physical, Data Link, Network, Transport, Session, Presentation, and Application"., https://en.wikipedia.org/wiki/OSI_model.

2.1. Methods to detect Wi-Fi fake access points

In the literature, methods to detect the presence of fake access points are essentially based on the analysis of information found on beacon frames (a sub-type of management frames). As per IEEE 802.11 specifications, an access point needs to send a beacon frame every 102.4 ms to inform surrounding devices of its presence. When an attacker creates a fake access point, it will coexist with the licit access point (same RF channel) and will try to emit the same beacon information as the licit access point to trick devices. Beacon frames contain information and some information is not easy to spoof or copy. Some scholars have therefore proposed methods which consist in analysing differences between two consecutive beacon frames to detect a fake access point attack. Other authors propose to perform a quantitative analysis on the number of received beacons over a time period to identify incoherences. In Kim, Park, Jung, and Lee (2012), Wang and Wyglinski (2016), some authors propose to detect fake access point using OSI layer 1 information.

A beacon contains static information like the name of the network (Service set identifier), capacity information and dynamic information like sequence numbers or physical layer constraints information (received signal strength, data rate, modulation...). Static information can be easily copied by the attacker using networking analysing tools while dynamic information is more complex to copy. Scholars have proposed methods to detect the presence of fake access point by analysing both static and dynamic information found in beacon frames. Concerning detection based on static information, in Bambang Setiadji, Ibrahim, and Amiruddin (2019), the authors describe a method to detect the presence of fake access points by analysing MAC and BSSID addresses found in beacons. The advantage of this detection method, as highlighted by the authors, is that it is lightweight and can be easily implemented (even on mobile phones). However, the drawback is that this method is not capable of detecting fake access points made by advanced attackers who take care to copy all the static information. In Han et al. (2012), the authors compare beacon information such as SSID, authentication type, and cipher type to verify if there is a fake access point attack. Authors claim that cipher and authentication type are specified by the vendors and added in beacon frames by the Wi-Fi card's firmware. For this reason, the authors believe that these information cannot be easily copied. However, in several non-academic and academic articles, authors have given tutorials to modify the firmware of various Wi-Fi cards.[5] Other methods combine the static beacon information comparison with IP addresses or environment identifiers. These methods consider extra added information or information found on other layers than layer 2 of the OSI model.

Concerning detection based on dynamic information (information that varies and whose variations are not easily foreseeable and therefore cannot be easily copied hlby an attacker), in Chumchu, Saelim, and Sriklauy (2011), the authors demonstrate they can detect the presence of Man-in-the-Middle attacks by analysing the data rate and modulation type information indicated in beacon frames, and defined by the transmission rate adaptation algorithm. The authors underline that this algorithm is designed by Wi-Fi card manufacturers and that the data rate and modulation type vary depending on the state of the channel. Therefore, it cannot be easily forged by attackers. However, as the authors in Alotaibi and Elleithy (2016) underline, according to the IEEE 802.11 protocol, there are a limited number of modulation types and possible data rates. Consequently, there is a high probability the attacker's fake access point's transmission rate adaptation algorithm will determine the same data rate or modulation scheme than the licit access point — especially if they are close and are operating on the same frequency channel. In Guo and Chiueh (2005), the authors demonstrate that they can detect fake access points by analysing sequence number

[5] Modifying Consumer Off the Shelf Wireless LAN devices for specialized amateur use. https://www.qsl.net/kb9mwr/projects/wireless/modify.html.

gaps. Authors in Guo and Chiueh (2005) show that based on their setup, under normal circumstances, the sequence gap between two beacons is never greater than 8. Based on their observations and configuration, the authors conclude that if the sequence number gap is greater than 8, there is a fake access point. This method is solid. However, in theory, the attacker can try to study the sequence number gap over a period of time, predicts it and modify his fake access point's beacon interval to send beacon frames at selective times so as not to trigger high variations in the sequence number gap between two beacons. He can also use jamming signals or probe requests to influence the licit access point's counter or the detection mechanism. In Arackaparambil, Bratus, Shubina, and Kotz (2010), the authors demonstrate how they leverage clock difference (or clock skew) to detect fake access points. The authors show they can distinguish between the beacons emitted by the licit and the fake access point by analysing the clock skew in timestamp between beacons. Although this method is a solid approach, it is relatively hard to implement and other scholars have shown that attackers can analyse the licit access point's timestamp to reduce their clock skew before creating their fake access point.[6] Some authors proposed methods based on a quantitative analysis of received beacons. Concerning quantitative analysis of beacons, in Kao, Chen, Chang, and Chu (2014), the authors demonstrate that beacon interval incoherence can be used to detect fake access points. In their experiment, the authors presume that the attacker has been able to eliminate his clock skew, can control his sequence number gap, and have copied all static information. The authors propose to analyse a large number of beacon frames and states that at some point in time either the fake access point's beacon interval or the licit access point's beacon interval will deviate. The authors, unfortunately, do not explain how the attacker can synchronize the sequence number and eliminate its fake access point's clock sew. Moreover, if the attacker can eliminate his clock skew, this means that the attacker can control his fake access point with high time accuracy. With such an accurate control of time, he should be able to detect that any deviation and correct his fake access point's beacon interval rapidly and accordingly.

2.2. Methods to detect frequency sweeping jamming signals

In literature, scholars have proposed methods to detect jamming attacks using data frames and control frames. These detection mechanisms can be classified into two categories: detection mechanisms with and without machine learning algorithms (Threshold based). Other authors have proposed detection mechanisms based on the monitoring of the OSI layer 1 (Deniau et al., 2017; Villain et al., 2019). Concerning threshold based detection, in Cheng, Ling, and Wu (2017), the authors have developed and used a time series model to detect the presence of jamming attacks. They propose to analyse the duration of the received packets, the number of bytes in the received packets, end-to-end delay of the packets, and signal-to-interference-plus-noise ratio (SINR) to detect jamming attacks. They have also considered the throughput and inter-packet time interval (gap). The jamming detection based on these parameters was only assessed by simulation. Moreover, this method only work if data frames are present and is ineffective when devices are idle and not transmitting data frames. In Reyes and Kaabouch (2013), authors calculate BPR (Bad Packet Ratio), use CCA (Clear Channel Assessment), PDR (Packet Delivery Ratio), and RSS (Received Strength Signal) from information found on data and control frame headers to detect jamming attacks. Their analysis is based on simulations (Matlab Fuzzy logic) and real-life experiments. This method is interesting, but is ineffective in the absence of data and control frames. Concerning detection mechanisms, leveraging machine learning algorithms, in Arjoune, Salahdine, Islam, Ghribi, and Kaabouch (2020), the authors propose a novel method to detect jamming attacks by using machine learning algorithms using four features: bad packet ratio, packet delivery ratio, received signal strength, and clear channel assessment. The authors have high accuracy results when detecting jamming attacks. With neural networks, their model has an accuracy of 96.4% and with a Random Forest (with 100 estimators) predictor, their model has an accuracy of 96.6%. Even if this study is based on 5G communication networks, similar features can be found on the OSI layer 2 of IEEE 802.11 networks. However, the detection works only when data frames are transmitted. Also, the experimental setup and the data analysis procedure are not accurately described. In Sufyan, Saqib, and Zia (2013), the authors propose a multi-modal scheme that can detect different jamming attacks (reactive, constant, random, intelligent, normal) by analysing the correlation between three features: packet delivery ratio, signal strength variation, and pulse width of the received signal. Signal strength variation and pulse width of received signals are studied on the physical layer using Software Defined Radios (SDRs). According to the authors, the Packet Delivery Ratio (PDR) is the ratio between the total number of packets correctly received and the total number of packets received. The usage of PDR limits the detection to when there is data frame transmission. Moreover, it is not specified if the analysis is carried out on management or data frames. In Puñal et al. (2014), the authors propose a jamming detection approach for IEEE 802.11 networks using machine learning algorithms by analysing frame header information on management, control, and data frames. They analyse features such as inactive time, packet delivery, and maximum signal strength. They also measure noise and channel busy ratio using Wi-Fi cards when there is no frame transmission. Their approach has high detection rates in indoor and mobile outdoor scenarios even under different and challenging link conditions. This method can detect jamming attacks when there are data transmission but also when the clients are idle. This is made possible by measuring noise and channel busy ratio using a specific Wi-Fi card and are not extracted from data frames. However, the authors state that they use a cross-layer platform software to make this detection possible. They also underline that their packet delivery is calculated on the application layer and that each node of the network must be aware of the number of network members in its hearing range and of a predefined rate for generating probing packets. This detection method is therefore knowledge-based, which can be challenging to generalize. This method is also partially based on the application layer and processes user data which can make it subject to legal constraints.

2.3. Detection of deauthentication attacks

Like fake access points, deauthentication attacks can be detected by comparing sequence number gap (Guo & Chiueh, 2005; Sheng et al., 2008). Some authors propose to detect this attack on the physical layer (OSI layer 1) (Villain et al., 2019). Deauthentication frames are a subtype of management frames and share the same counter as beacon frames. When an attacker sends a deauthentication frame in the name and place of an access point, he cannot copy the same sequence number (as it is dynamic information) as the licit access point, this can lead to high variations. In Sheng et al. (2008), the authors also propose to study the Received Signal Strength Indicator (RSSI) on management frames to detect deauthentication frames. Using RSSI to detect deauthentication attack is interesting but it only works if the attacker is at a significant distance from the licit access point. In Agarwal, Biswas, and Nandi (2013), the authors propose a method to detect the deauthentication attack on IEEE 802.11 by comparing the number of deauthentication frames and the throughput calculated with data frames, against a previously determined threshold. As the authors underline, this method is easy to implement and lightweight. The authors do not precise whether their approach will have the same efficiency in the absence of data frames. In Agarwal, Pasumarthi, Biswas, and Nandi (2016), the authors propose to detect the deauthentication attacks by analysing a list of 18 features including features from layer 3 and layer 4. Their method has a

[6] Detection of Rogue APs Using Clock Skews:Does it Really Work? https://www.cs.dartmouth.edu/~sergey/skew/toorcon11-slides.pdf.

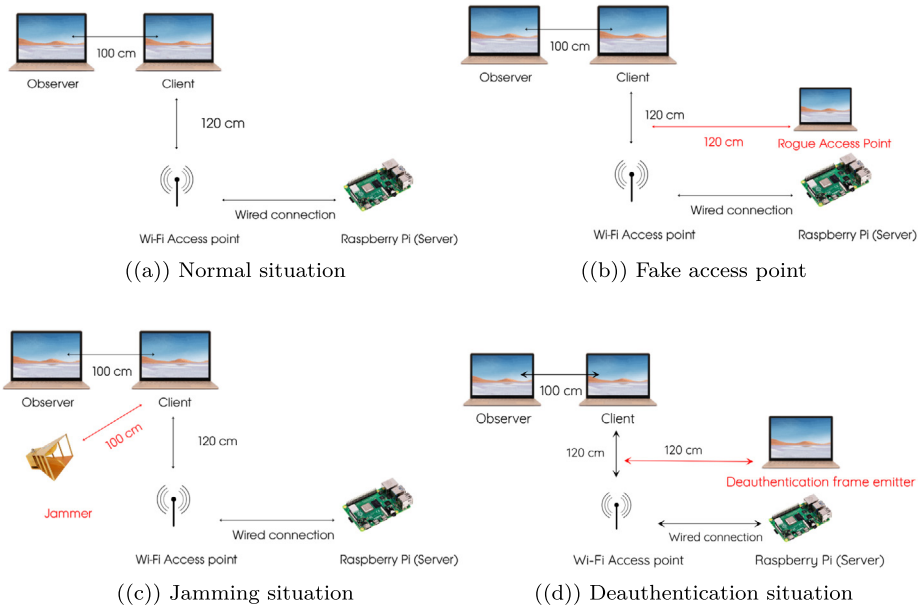

Fig. 1. The four different situations.

high accuracy rate of 0.987 with Support Vector Machine (SVM) and an accuracy of 0.954 with Naives Bayes algorithm. This method analyses information on layer 3 and layer 4 headers which imply that the NIDS will have to handle encryption keys if the Wi-Fi network is encrypted. This can particularly challenging if the Wi-Fi encryption is WPA-802.1X (WPA-EAP).[7]

We concluded that in the literature, there is no single tool to detect the three attacks. The three attacks can be combined and our work is highly motivated by the lack of a holistic approach to detect fake access points, frequency sweeping jamming signals, and deauthentication attacks on IEEE 802.11 networks. Furthermore, our detection for the three aforementioned attacks is based on management frames which are (at least for beacon frames) sent regularly and can therefore be effective even in the absence of data frame emission.

3. Experimental setup and frame collection

To implement the NIDS, we have adopted an anomaly based approach which consists in comparing different situations to identify indicators to detect the attacks. Consequently, we have reproduced a normal situation and three attack situations in laboratory experiments. We then compare the different situations against the normal situation.

In the normal situation, there are a client, a server, an IEEE 802.11 access point, and an observer. The client, a portable computer, is communicating with the server, a Raspberry PI (Fig. 1). The server is continuously sending random data to the client, via the access point, at a rate of 100 Mb/s using iPerf3.[8] The client and the server communicate in infrastructure mode. That is, there is an access point between the client and the server. The Raspberry PI is connected to the access point via an Ethernet cable while the client is connected wirelessly to the access point via a Wi-Fi (IEEE 802.11 b/n) connection. The client and the access point operate at a frequency of 2.472 GHz (channel 13). The observer is a computer whose Wi-Fi card is in monitor mode. When the NIDS is intended to replace the observer. Monitor mode allows a device to receive and store all frames transmitted within a frequency channel. The observer operates at 2.472 GHz and is capturing all frames transmitted on channel 13 during the experiment. Before choosing channel 13, we have verified that no other device was, apart from the client, and the access point were operating on channel 13. The distances between the different devices are indicated in Fig. 1: the client is at a distance of 120 cm from the access point and the observer is at a distance of 100 cm from the client. The capture lasts two minutes.

In the fake access point attack situation (Fig. 1), a fake access point emitting the same **static** beacon information at the same interval as the licit access point, is added to the configuration of the normal situation. The fake access point copies all static information of the licit access point, including but not limited to, the MAC address, the network name (BSSID), supported rates, capability information... This situation represents the case where the attacker creates a fake access point and passively waits until devices connect to his access point. The attacker does not use jamming signals or forged deauthentication frames in this situation. The access point is created with hostapd[9] using a commercial Wi-Fi which operates at a central frequency of 2.472 GHz (channel 13). The observer is still capturing frames emitted on the channel 13. In this situation, three devices are operating on the WiFi channel 13: the client, the access point, and the attacker's fake access point. The capture lasts two minutes and starts a few moments after the creation of the fake access point.

In the jamming attack situation (Fig. 1), a frequency sweeping jamming signal is used to disturb the Wi-Fi communication between the client and the access point. A frequency sweeping jamming signal is emitted using a directional antenna oriented towards the client. The jamming signal is implemented using Matlab[10] and generated using an arbitrary signal generator to affect the whole 2.4 GHz frequency band. The jamming signal sweeps rapidly and continuously the frequency band from 2.4 to 2.5 GHz and affects all communications from channel 1 to channel 14. The sweep time is 10 us. The directional antenna is at a distance of 100 cm from the client. The observer operates at a

[7] Wi-Fi Extensible Authentication Protocol. https://fr.wikipedia.org/wiki/Extensible_Authentication_Protocol.

[8] iPerf3, is a tool for active measurements of the maximum achievable bandwidth on IP networks. It is released under a three-clause BSD license, https://iperf.fr/.

[9] hostapd (host access point daemon) "is a user-space daemon software enabling a network interface card to act as an access point and authentication server. There are three implementations: Jouni Malinen's hostapd, OpenBSD's hostapd, and Devicescape's hostapd", https://en.wikipedia.org/wiki/Hostapd.

[10] MATLAB "is a programming and numeric computing platform used by millions of engineers and scientists to analyse data, develop algorithms, and create models". https://www.mathworks.com/products/matlab.html.

central frequency of 2.472 GHz and captures all transmitted frames. The capture lasts two minutes and starts a few moments after the jamming attack has been activated. In the deauthentication attack situation, forged Wi-Fi deauthentication frames are continuously emitted to disconnect the client from the access point. The forged deauthentication frames are emitted using aireplay-ng.[11] As shown in Fig. 1, the deauthentication frames emitter is at a distance of 120 cm from the centre of the communication between the client and the access point. The attacker forges the deauthentication frames of the licit access point by indicating the licit access point's MAC address in the source MAC address field and the client's MAC address in the destination MAC address field. The observer still operates at a central frequency of 2.472 GHz and captures all transmitted frames. The capture lasts two minutes and starts a few moments after the deauthentication attack has been activated. At the end of the experiments, we have collected a certain number of frames. In this paper, we consider only management frames. We have collected, after management frames filtering, for the normal, deauthentication, fake access point and frequency sweeping jamming situations: **1366**, **30152**, **2681** and **959** management frames respectively.

4. Data analysis and feature selection

We observe that the majority of frames sent by the access point in normal, jamming, and fake access point attack situations, is beacon frames and deauthentication frames in the deauthentication attack situation. From the management frame and radiotap headers, we have extracted the time, Received Signal Strength Indicator (RSSI), sequence number, and subtype. With time and sequence number, we have calculated the frame time interval by subtracting frame+1 and frame time field and the sequence number gap by subtracting frame+1 and frame sequence number field. After feature extraction, we have create a structured data with five columns: Frame interval, Received Signal Strength Indicator, Sequence number gap, subtype with the corresponding values for each situation (class). Using different plots, we have compared each feature in the four situations to detect anomalies and set indicators.

4.1. Frame interval

In Fig. 2, we observe that the distribution of the frame interval is not the same in the four situations. We know that the majority of management frames received, in normal, jamming, and fake access point attack situations, is beacon frames. According to the IEEE 802.11 specification, the beacon interval of an access point should be around 102.4 ms (Target Beacon Transmission Time). In the normal situation, the mean frame interval (orange bar) is a sightly above **102.4 ms** with several beacons under 102 ms. Beacons received with a frame interval under 102.4 ms can be explained by the fact that, in our experiment configuration, there is a significant data traffic of 100 Mb/s between the client and the server. The IEEE 802.11 protocol specifies that access points can favour data frames over beacon frames. In such cases, beacons are buffered (sometimes delayed) and sent at once when possible. In the deauthentication situation, the frame interval is close to zero. This is because the forged deauthentication frames (majority of frames) sent by the attacker are mixed up with the access point's beacons (deauthentication and beacon frames are both management frames). Deauthentication frames are also sent massively at an interval smaller than 102.4 ms which decreases the mean frame interval in this

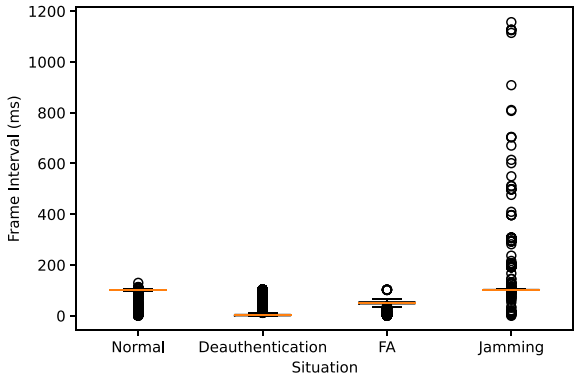

Fig. 2. Management frame interval comparison. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

situation. In the fake access point situation, the attacker has created a fake access point that sends beacon frames with the same static information (source MAC address, etc.) as the licit access point. The two access points operate simultaneously in the same Wi-Fi channel and send the same beacons. The observer cannot differentiate between the two beacons and therefore, after capture and filtering, there are twice as much beacons and the mean frame interval is divided by around 2.

Finally, in the jamming situation, we would expect significant loss of beacon frames (jamming attack should annihilate all communication), but we observe that the mean frame interval is close to 102.4 ms. This is because beacon frames are usually sent at the lowest data rate to ensure that every possible client in the range of an access point can receive the frame. Also, in the 2.4 GHz band, beacons are sent using the IEEE 802.11b physical layer modulation. Frames sent via the interface 802.11b physical interface are resilient to interference because they are modulated as Direct Sequence Spread Spectrum (DSSS) signals (Martínez et al., 2008). Other frames such as data frames are sent by the access point at a higher rate and usually using the 802.11 n physical interface (which is not as resilient to interference as the 802.11 b). Therefore we observe in our unfiltered dataset, that during the jamming attack, the transmission of data frames is degraded and ultimately stops but that management frames sent using the 802.11b interface are still present. We also observe in the jamming situation that there are some beacon intervals under and over the mean value. Beacon interval under the mean value can be explained by the fact that the access point senses that the channel is busy and buffers some beacon frames from time to time. Beacon interval over the mean value can be explained by the fact that some beacon frames are sent but not received because of collisions. DSSS signals are more resilient to interference but there can still be some collisions at the reception. As a conclusion, we can say that the mean management frame interval can be used as an indicator to easily detect the deauthentication and fake access point attacks — in literature, as mentioned earlier, many scholars have used frame intervals as an indicator to detect these attacks.

4.2. Received Signal Strength Indicator

In Fig. 3, we observe the Received Signal Strength Indicator (RSSI) values. The RSSI represents the Wi-Fi Signal power in dBm calculated by the Wi-Fi receiver of the observer when it receives a frame. As per the same figure, the range for RSSI is around −56 dBm to −34 dBm in the normal situation. We can observe that the mean RSSI value is not the same for the four situations. In the jamming attack situation, the mean RSSI of beacon frames is degraded. In the deauthentication situation, we can observe that the mean RSSI is zero. This is because in this situation, deauthentication frames are predominant and the Wi-Fi card of the observer cannot seem to be able to calculate the RSSI value. The RSSI field is therefore empty and we have replaced empty fields

[11] Aireplay-ng, "The primary function is to generate traffic for the later use in aircrack-ng for cracking the WEP and WPA-PSK keys. There are different attacks which can cause deauthentications for the purpose of capturing WPA handshake data, fake authentications, Interactive packet replay, hand-crafted ARP request injection and ARP-request reinjection". https://www.aircrack-ng.org/doku.php?id=aireplay-ng.

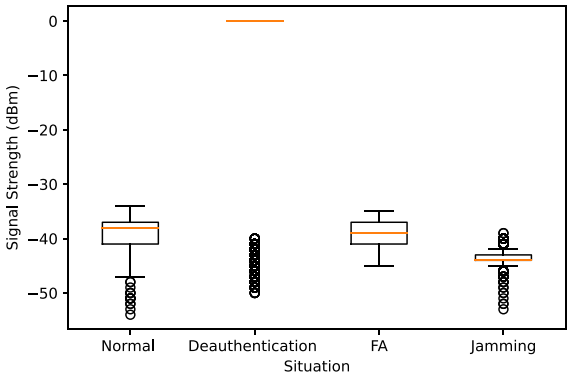

Fig. 3. Management Signal Strength box plot.

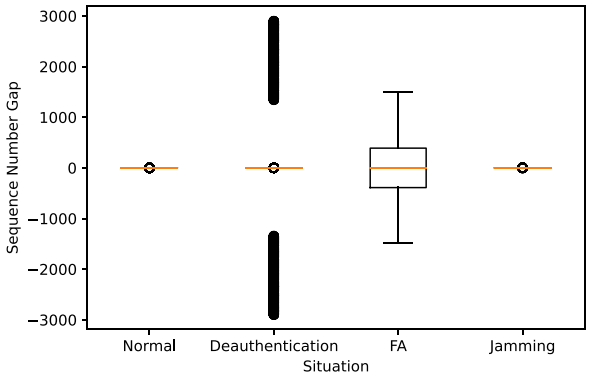

Fig. 4. Sequence Number Gap.

with zeros during pre-processing. The rest of the values in this situation corresponds to the RSSI of beacon frames. In the fake access point situation, the mean value is lower than in the normal situation but we cannot set clear boundaries to distinguish between the two situations. We can conclude that the RSSI could be used as an indicator to detect the three attacks but there will be numerous false positives especially in the jamming and fake access point situations as the RSSI range in those situations overlaps and the mean values are close. RSSI values can however be combined with other features to increase detection efficiency. Another interesting feature to study is the sequence number gap.

4.3. Sequence number gap

In Fig. 4, we observe that concerning the sequence number gap, the deauthentication and fake access point situations differ from the normal situation. The sequence number gap is the gap between two management frames. The sequence number gap is calculated after capture. In the normal situation, the mean gap is around 1 which is the normal behaviour (see section Section 2) . In the deauthentication situation, we observe that the mean gap is significantly high which does not conform with the normal situation. This is because the deauthentication frames are not sent by the licit access point. Even if the attacker usurps the MAC address of the access point, as indicated in Section 2, it is very difficult for the attacker to synchronize his sequence number counter with the access point's counter. The same observation applies to the Fake Access (FA) situation with the slight difference that there is a range in the sequence number gap. This can be explained by the fact in the Fake Access point situation, there are relatively the same amount of beacons from the licit access point and the fake access point. The gap oscillates within a range. In the deauthentication situation, deauthentication frames are predominant, the gap is usually 1 except when there are beacon frames from the access point (each 102.4 ms) which leads to greater gap. Concerning the jamming, the sequence number gap is around 1 as the jamming signal does not affect IEEE 802.11 b frames, so the behaviour is similar to normal situation. After data analysis, we concluded that using our set of features, we cannot always set precise thresholds to detect the three attacks. For this reason, we have opted for a supervised learning approach with algorithms that can consider several indicators simultaneously and more easily distinguish between overlapping situations. We have therefore used our dataset to train seven supervised learning algorithms using Scikit-Learn on Python. Each algorithm create a model capable of detecting the three attacks. We compare each model and keep the model having the highest detection precision results to detect the attacks. To prepare our data for learning phase, we have divided the dataset into two smaller train and test datasets. We have then, split the train datasets into two smaller train and validation datasets. The train and the validation datasets of different situations are concatenated and used to train and tune the model. We have used seven supervised learning algorithms namely K-Nearest Neighbour (KNN), Random Forest, Classification and Regression Trees(CART), Logistic Regression, Navies Bayes, Support Vector Machine (SVM) and Linear Discriminant Analysis.

5. Results

Each algorithm creates a model and we use a set of unused data in the learning phase to evaluate each model. Unknown values for Frame interval, RSSI, Sequence number gap, and subtype are presented to the model, and the model has to predict whether these values correspond to the normal, deauthentication, fake access point, or jamming situation. The number of true positive, true negative, false positive, and false negative is noted[12] and the accuracy score and precision are calculated. The accuracy score represents the fraction of correct predictions or true positives over the total number of positives and negatives. The precision is the number of true positives over the total number of positives (true and false positives). The training and testing phase are repeated with different hyperparameters for each supervised learning algorithm to increase accuracy and precision (Burkov, 2019).

The table in Fig. 5 summarizes the prediction results of the different obtained models. Concerning the hyperparameters used, for KNN, we have set the parameter k to 4 as we have four dimensions (four features). For Random Forest, we set the number of estimators to 100 to increase the number of trees in the Forest and therefore increase the accuracy and precision. We have set entropy as splitter as it is a better splitter for non-continuous data and we have set the min_samples_leafint to 3 to prune off the trees and eliminate overfitting of the model. For CART, we have set the splitter to entropy and min_samples_leafint to 3 for the same reasons. For Logistic Regression, we have set solver parameter to linear, multi_class to One-Versus-Rest as we have several classes (or situations) and we observed that one binary classification problem per class provides better prediction results. For Logistic Regression and Gaussian Naive Bayes, we have set solver to linear and multi_class to One-versus-Rest for the same reasons. Finally, for Linear Discriminant Analysis, we have set solver to Single Value Decomposition as it does not compute covariance matrix and can handle a significant number of features (Bonaccorso, 2017).

We have best scores for the model with Random Forest and KNN. The precision to detect the deauthentication attack (TP_{Deauth}) and the fake access point is high while the precision to detect the jamming attack ($TP_{Jamming}$) is satisfying. The two best precision scores for jamming attack are achieved with Random Forest and Linear Discriminant Analysis models. In both cases, however, precision is still under 90%. When analysing the confusion matrix, we observe that for the jamming

[12] Sensitivity and specificity. https://en.wikipedia.org/wiki/Sensitivity_and_specificity.

Algorithms	TP_{Deauth}	TP_{FA}	$TP_{Jamming}$	TP_{Normal}
Random forest	99.80%	96.41%	81.54%	82.12%
KNN (k=4)	99.67%	95.83%	78.30%	82.98%
CART	99.21%	96.53%	78.35%	75.50 %
Logistic Regression	100.0%	88.21%	68.28%	95%
Naives Bayes	100%	96.09%	77.46%	66.80%
Linear Discriminant Analysis	100%	79.88%	81.03%	85.19%
SVM	98.81%	96.95%	69.40%	83.25%

Fig. 5. Precision Results for each situation.

detection, the machine learning algorithms, in some cases, confuse the jamming situation with the normal situation. This is confirmed by satisfying detect precision for the normal situation. For both Random Forest and KNN, detection precision for normal situation is not higher than 83. It means than the NIDS cannot easily differentiate the normal situation from other situation and will give a certain amount of false positive or false alarm. Nonetheless, we need to make a trade off and decide what is important ? To be able to detect and differentiate between the attacks or to reduce false alarm. For instance, if we want to reduce false alarm, we can opt for the model created by logistic regression which has a 95% detection precision for normal situation but has less detection precision for the jamming attack. In cybersecurity, the viable solution is to first opt for a model which has the highest detection precision for the attacks and secondly opt for model having the best lowest false alarm rate. Therefore in conclusion, we should opt for either for the Random Forest or the KNN created model.

6. Conclusion

In this paper, we have provided a detailed description of OSI layer 2 based detection methods to detect intentional jamming signals, the presence of fake access points, and the emission of deauthentication frames on IEEE 802.11 networks. We have underlined that there is no holistic approach (single tool) to detect the aforementioned attacks. These three attacks which, are related and can be combined, are detected separately using one or two features and mostly without using machine learning algorithms. We have proposed a single model approach to detect the three attacks using a set of features (frame interval, RSSI, Sequence number gap, and management subtype) extracted on management frame headers — using machine learning algorithms. The model created by Random Forest and KNN algorithm has high detection precision when detecting deauthentication and fake access point attacks. It also has perfectible detection precision for the jamming attack. We are presently working on the improvement of jamming attacks detection by considering moderate and weak power jamming. Detecting Moderate power jamming can be sufficient and more easily detected on OSI layer 2 than high power jamming (presented in this paper). We are also considering different variations of traffic between the server and the client, the detection of combined attacks, legal conformity and limitations of the NIDS, and ways to counter them.

CRediT authorship contribution statement

Andy Amoordon: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Virginie Deniau:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Anthony Fleury:** Writing – review & editing, Supervision, Data curation. **Christophe Gransart:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is a part of a PhD thesis financed by the Hauts-de-France Region and University Gustave Eiffel. The authors would like to express their gratitude to the two public entities for their support.

References

Agarwal, M., Biswas, S., & Nandi, S. (2013). Detection of de-authentication denial of service attack in 802.11 networks. In *2013 Annual IEEE India conference* (pp. 1–6). IEEE.

Agarwal, M., Pasumarthi, D., Biswas, S., & Nandi, S. (2016). Machine learning approach for detection of flooding DoS attacks in 802.11 networks and attacker localization. *International Journal of Machine Learning and Cybernetics*, *7*(6), 1035–1051.

Alotaibi, B., & Elleithy, K. (2016). Rogue access point detection: Taxonomy, challenges, and future directions. *Wireless Personal Communications*, *90*(3), 1261–1290.

Amoordon, A., Gransart, C., & Deniau, V. (2020). Characterizing wi-fi man-in-the-middle attacks. In *2020 XXXIIIrd General assembly and scientific symposium of the international union of radio science* (pp. 1–4). IEEE.

Amoordon, A., Gransart, C., Deniau, V., & Fleury, A. (2021). La detection par seuil des attaques man-in-the-middle et jamming sur les reseaux wi-fi. *Neuvieme Conference Pleniere Du GDR ONDES*.

Arackaparambil, C., Bratus, S., Shubina, A., & Kotz, D. (2010). On the reliability of wireless fingerprinting using clock skews. In *Proceedings of the third ACM Conference on wireless network security* (pp. 169–174).

Arjoune, Y., Salahdine, F., Islam, M., Ghribi, E., & Kaabouch, N. (2020). A novel jamming attacks detection approach based on machine learning for wireless communication. In *2020 International conference on information networking* (pp. 459–464). Los Alamitos, CA, USA: IEEE Computer Society, [ISSN: 1976-7684] http://dx.doi.org/10.1109/ICOIN48656.2020.9016462.

Bambang Setiadji, M. Y., Ibrahim, R., & Amiruddin, A. (2019). Lightweight method for detecting fake authentication attack on Wi-Fi. In *2019 6th International conference on electrical engineering, computer science and informatics* (pp. 280–285). http://dx.doi.org/10.23919/EECSI48112.2019.8976975.

Berg, J. S. (2008). *Broadcasting on the short waves, 1945 to today*. McFarland.

Bierbrauer, D. A., Chang, A., Kritzer, W., & Bastian, N. D. (2021). Anomaly detection in cybersecurity: Unsupervised, graph-based and supervised learning methods in adversarial environments. arXiv preprint arXiv:2105.06742.

Bonaccorso, G. (2017). *Machine learning algorithms*. Packt Publishing Ltd.

Burkov, A. (2019). *The hundred-page machine learning book, vol. 1*. Andriy Burkov Quebec City, QC, Canada.

Cheng, M., Ling, Y., & Wu, W. B. (2017). Time series analysis for jamming attack detection in wireless networks. In *GLOBECOM 2017 - 2017 IEEE Global communications conference* (pp. 1–7). http://dx.doi.org/10.1109/GLOCOM.2017.8254000.

Chumchu, P., Saelim, T., & Sriklauy, C. (2011). A new MAC address spoofing detection algorithm using PLCP header. In *The International conference on information networking 2011* (pp. 48–53). IEEE.

Deniau, V., Gransart, C., Romero, G. L., Simon, E. P., & Farah, J. (2017). IEEE 802.11 n communications in the presence of frequency-sweeping interference signals. *IEEE Transactions on Electromagnetic Compatibility*, *59*(5), 1625–1633.

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, *50*, Article 102419.

Guo, F., & Chiueh, T.-c. (2005). Sequence number-based MAC address spoof detection. In *International workshop on recent advances in intrusion detection* (pp. 309–329). Springer.

Han, C., In-Jang, J., feng Shao, J., Chae, K., Seong-Soo, B., & Jung, S. (2012). A scheme of detection and prevention rogue AP using comparison security condition of AP.

Kao, K. F., Chen, W. C., Chang, J. C., & Chu, H. T. (2014). An accurate fake access point detection method based on deviation of beacon time interval. In *2014 IEEE Eighth international conference on software security and reliability-companion* (pp. 1–2). http://dx.doi.org/10.1109/SERE-C.2014.13.

Kim, T., Park, H., Jung, H., & Lee, H. (2012). Online detection of fake access points using received signal strengths. In *2012 IEEE 75th Vehicular technology conference* (pp. 1–5). IEEE.

Martínez, A., Zurutuza, U., Uribeetxeberria, R., Fernández, M., Lizarraga, J., Serna, A., et al. (2008). Beacon frame spoofing attack detection in IEEE 802.11 networks. In *2008 Third international conference on availability, reliability and security* (pp. 520–525). IEEE.

Pirayesh, H., & Zeng, H. (2021). Jamming attacks and anti-jamming strategies in wireless networks: A comprehensive survey. arXiv preprint arXiv:2101.00292.

Puñal, O., Aktas, I., Schnelke, C.-J., Abidin, G., Wehrle, K., & Gross, J. (2014). Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation. In *Proceeding of IEEE international symposium on a world of wireless, mobile and multimedia networks 2014* (pp. 1–10). http://dx.doi.org/10.1109/WoWMoM.2014.6918964.

Reyes, H. I., & Kaabouch, N. (2013). Jamming and lost link detection in wireless networks with fuzzy logic. *International Journal of Scientific & Engineering Research*, *4*(2), 1–7.

Sheng, Y., Tan, K., Chen, G., Kotz, D., & Campbell, A. (2008). Detecting 802.11 MAC layer spoofing using received signal strength. In *IEEE INFOCOM 2008-the 27th Conference on computer communications* (pp. 1768–1776). IEEE.

Sufyan, N., Saqib, N. A., & Zia, M. (2013). Detection of jamming attacks in 802.11 b wireless networks. *EURASIP Journal on Wireless Communications and Networking*, *2013*(1), 1–18.

Villain, J., Deniau, V., Fleury, A., Simon, E. P., Gransart, C., & Kousri, R. (2019). Em monitoring and classification of IEMI and protocol-based attacks on IEEE 802.11 n communication networks. *IEEE Transactions on Electromagnetic Compatibility*, *61*(6), 1771–1781.

Wang, L., & Wyglinski, A. M. (2016). Detection of man-in-the-middle attacks using physical layer wireless security techniques. *Wireless Communications and Mobile Computing*, *16*(4), 408–426.