SEVEN
publicações acadêmicas

Analysis of Performance Metrics on the Conjuncture of Intrusions in IEEE 802.11 Networks with Machine Learning at Hospital N.S.C.

Crossref doi
https://doi.org/10.56238/Connexpemultidisdevolpfut-116

Matheus Santos Andrade
Highest level of education: Post Graduate
Academic institution: National Service for Commercial Learning

Jonathas Carvalho de Freitas
Highest level of education: Graduated
Academic institution: Tiradentes University

ABSTRACT
The security present in IEEE 802.11 networks becomes more relevant every day. However, security on the IEEE 802.11 network has not kept pace with threats with as much significance. For this reason, the proposal arises to design an Intrusion Detection System-IDS based on machine learning that will be able to have self-improvement, since it will create a safe environment, capable of detecting all disguised threats, Deauthentication, Eapol -logoff (Eapol) and Beacon Flood, where they were launched on a real corporate network. With this, correlated the performance metrics, and among them, which values the quality of the classification, the Matthews Correlation Coefficient. The Deauthentication anomaly above the Naive Bayes classifier was obtained (88.71%), whereas the quality value of the Logistic Regression (Logistic) classifier was equated to (88.69%), and nevertheless, the J48 presented a lower value of (88.47%).
Despite this, the identification of the Beacon Flood attack was due to the Naive Bayes algorithm showing the highest detection rate (100.00%), followed by Logistic (99.95%) and J48 having the lowest value (98.85 %). As a result, in the detection of the Eapol anomaly, the classifications presented similarity of (100.00%) and the others, with the presentation of a detection, due to non-anomalous data (Normal), the Naive Bayes was affected by (89.92 %), followed by Logistic maintaining (89.89%), while J48 was tested with a lower rate (89.67%). With the study evidences provide the possibility that it is possible to develop an intrusion detection system based on wireless networks.

Keywords: Threats, Quality, Evidences.

1 INTRODUCTION

Computer networks arose with the need to interconnect universities or academic centers. And later embraced by companies in ways that brought benefits to industry, commerce and households (Arasaki and Della Flora, 2012). On the other hand, devices with high mobility content, whose specifications follow the methods of the IEEE 802.11 family (IEEE 802.11, 1999), such as: *laptops*, cell phones, *tablets* and among others have become common and with the diverse public being more used today (Feng, 2012). However, it is common the presence of *hackers* in electronic media, given that the number of *Internet* users and the ease of acquiring services in wireless communication networks and products, is made of the amount of financial processes arouse interest of attackers, to apply scams involving monetary gains, industrial espionage, extortion, sale of information and defamation of the image of the government.

Although the protocols aimed at the security of wireless networks: WEP *(Wired Equivalent Privacy*) (IEEE 802.11, 1999) intended to ensure a level of security similar to the wired network (Morimoto, 2008) to protect the data frames that carry data and control information through their header, but due to their diversity of vulnerabilities (Tews, 2007) and the non-guarantee of scalability to the model, however (i.e., it was outdated), emerging a new security standard called WPA (Wi-Fi Protected Access) (Wi-Fi Alliance, 2003) and *WPA2 (*Wi-Fi Protected Access Version 2) (IEEE 802.11i, 2004) assisting in the protection along with confidentiality and integrity of data communication on the network. Despite this, it does not present security to the control boards that reserve the communication channel in the confirmation of data in the network, and to the management boards in the recognition of the presence of a wireless network, to initiate the association and disassociation of stations to some AP *(Access Point*) (Linhares and Gonçalves, 2012).

However, with the emergence of the IEEE 802.11w amendment (IEEE 802.11w, 2009), which includes protection for management frameworks, which was only ratified in 2009, a decade after the emergence of the IEEE 802.11 amendment, which allowed a range of attack development, aiming at network interoperability, as well as the practice of capturing sensitive information being conducted in these frameworks.

Despite the particular nature of wireless networks, along with the amendments (IEEE 802.11i and IEEE 802.11w) resolving parts of the vulnerabilities found in this mishap in IEEE 802.11 networks, the incidents on wireless networks that are (*e.g.* caused by carrying out denial-of-service attacks, loss of information by the false transmission request that are being stored on the AP, that the linked stations would be sent and would not be ready to receive, causing the rejection of information, in addition to blocking the use of the communication channel for a stipulated period of time, and among others).

With this, there is a way to inhibit such "ills" and it is through an Intrusion *Detection System (*IDS) that provides the containment of events seeking to identify, diagnose and treat anomalies to keep a network operating (Barford et al., 2002), but with the heterogeneous scenario of wireless networks comes to make complex its fair evaluation, and therefore, the objective of this article is to present an approach in the construction of a set of data that represents a wireless network as well as the evaluation through machine learning that arises as the need to improve the performance of some activity through experience or in the discovery between similarities of homogeneous data (Mitchell, 1997), to increase security in IDS.

2 METHODOLOGY

This section describes a summary of what was adopted for the development of the IEEE 802.11-based dataset and also an overview of how to conduct the experiments that were adopted in the development of the work. The main procedure includes, the creation of the data set, pre-processing,

normalization and classification.

2.1 DATASET GENERATION

The data set generated is from a real scenario contained in the Hospital Nossa Senhora da Conceição, located in the city of Lagarto-SE, central-south region of the State and on the geographical coordinates: Latitude: -10.912929561173492, Longitude: -37.673240474073125. The way to this, the data collection happened between the days 11/01/2023 until 16/01/2023 interspersed / varying days and between the interval of $1 \pm hr$ for the generation of the *dataset*. Included from a *wireless* network to components such as HTTP/HTTPS, SMTP, POP3, IMAP, and SSH. Impersonating a corporate network with multiple authenticated users, plus WPA2 encryption contention enabled for a secure network.

The scenario pointed out (Figure 1) is limited to the point that there are several devices sending/receiving data to the network infrastructure, however, it is the responsibility of the offending station *(Attacking Laptop)* through Linux (Kali Linux) to capture and/or monitor radio data in the transmission of the MAC frame corresponding to IEEE 802.11, in addition to the generation of simultaneous and categorical attacks with denial of service techniques:

1) *Deauthentication*: This type of attack affects the management frames, with the simultaneous sending of unrealistic frames that comes to force the connected device to be deauthenticated from the network (Ahmad and Tadakamadla, 2011). Details on the use of this attack is through the *aireplay-ng* tool in the *aircrack-ng package* (AIRCRACK-NG, 2022);

2) EAPOL-Logoff: This attack disrupts the management and control frameworks with a flood of forged *EAPOL* packets and sends them to the AP to delete the authentication state of an authenticated and associated user (Ahmad and Tadakamadla, 2011). For this type of attack the MDK3 tool (MDK3, 2022) was used;

3) *Beacon Flood*: Attack that causes damage to the management boards, and this through the issuance of range of packets with several fake SSIDS to the frequency spectrum of the network, thus bringing disorder to the user who tries to connect to the AP (Ahmad and Tadakamadla, 2011). For the use of this type of attack was used the tool MDK3 (MDK3, 2022).

Figure 1 - Topology of the *Wireless* Network (WPA2) contained in a segment at Hospital Nossa Senhora da Conceição (H.N.S.C).

Source: Authors (2023).

2.2 PRE-PROCESSING AND STANDARDIZATION

In the pre-processing phase that had the aid of the Wireshark tool (*Wireshark*, 2022), on the need of the offending station keeping only significant attributes to the MAC frame (*Protocol Version*, *Type*, To DS, From DS, More Fragment, *Retry,* Power Management*, More Data,* WEP *, Order,* Duration*, Transmitter* address, Destination address, Source *address, Receiver address, BSS Id* and *Sequence number*). This accomplishes a proper organization for the data collection proposed to this study. Pointing out, the identification of attacks that impact the operation of the IEEE 802.11 wireless network, to be used as a reference, in different approaches of various wireless ecosystems, and *with the need for an* Info attribute in *the easy identification of the type of attack* and / or said as normal, in addition to the representativeness in its quantity to (Table I).

Table I. Data sampling values of type Normal, *Deauthentication,* Eapol-logoff*,* and *Beacon Flood*.

Info	Number of samples
Normal	9134
Deauthentication	5094
Eapol-*logoff*	1428
Beacon Flood	1047
Total	16703

Source: Authors (2023).

However, the collection of peculiar samples requires the normalization that was performed through *Java (Java*[1] , 2022) in order to avoid noise about the actual data, specifically with the labels,

[1] The Java programming language is object-oriented, intended to run on any platform or even devices.

from Normal to the value 0 (zero), and the other attacks, Deauthentication *value corresponding to 1 (one),* Beacon Flood to *its value 2 (two) and* Eapol-logoff *respective to* 3 (three) . Facilitating the whole, with the balancing itself in the extraction of the fields for a normalized data set, in obtaining machine learning algorithms and the evaluations by performance metrics.

2.3 MACHINE LEARNING ALGORITHMS

To introduce effect to this project were related machine learning algorithms. With the categorization of the dataset, to a new observation to which it belongs, and it is on the mastery of the *Waikato Environment for Knowledge Analysis (*Weka*) tool (Frank, Hall and Witten, 2016)* and its functional principles that are a desirable feature for IDS (Scarfone and Mell, 2007). With the randomization of the *Application Programming Interface* (API) in *Java* that the study of anomaly associations to IEEE 802.11 were submitted.

1) *Logistic Regression*: The model based on logistic regression aims to create a direct dependence relationship between the class variable and the characteristics, seeking to bring values between 0 and 1, values that represent the probability of returning the value 1 for the linear expression:

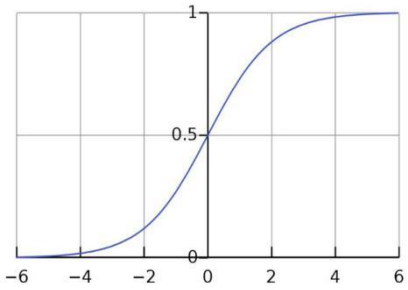
$$\theta x = \theta 0 + \theta 1 x1 + \theta 2 x2 + ... + \theta n xn \qquad \text{(Eq.1)}$$

The values of X and θ are vectors that feed the hypothesis function. Such a function has functionality of determining the value of θ so that it returns the expected y based on the input value of x. The hypothesis function is given by the sigmoid:

$$h = g(z) = \frac{1}{1+e^{-z}} \qquad \text{(Eq.2)}$$

The values generated by the function are represented in the graph below (see image I). Note that the function generates 2 asymptotes, 1 tending to 0 for negative values of z and another tending to 1 for positive values.

Figure 2. Sigmoid graph for the function g(z)(Eq.2)



Source: Authors (2023).

In order for the values presented in the function above to be approximated to the real model, it is necessary to use another equation for this, a function called the logarithmic loss function. Presented below:

$$\text{Custo}\,(h\theta(x), y) = -\log(h\theta(x)) \text{ se } y = 1 \text{ ou}$$
$$-\log(1 - h\theta(x)) \text{ se } y = 0 \qquad \text{(Eq.3)}$$

This model is more suitable for binary classification methods and can be used in multiclass functions with greater manpower. In *Weka,* logistic regression is the implementation of the *Logistic* algorithm and can be found by Cessie and Houwelingen (1995).

2) *J48*: This type of algorithm represents the form of binary decision trees, but with great stability between precision time and calculation, with less training effort on the algorithm for nonlinear classifiers. By *Weka*, the decision tree is the implementation of the *J48* algorithm. Details of this algorithm can be found in Quinlan (2014). Eventually the use of entropy on the degree of uncertainty of random elements and information gain are commonly the most used in this type of algorithm (Ravipati and Abualkibash, 2019). Despite this, the entropy will calculate the homogeneity of the detailed samples, so that the data analyzed completely as homogeneous will be respective to zero, if not the entropy with the perspective to 1 (Ravipati and Abualkibash, 2019) in the emission of the formula:

$$I(S = \sum_{x=1}^{x} -P_I \, log_2 \, P_x) \qquad \text{(Eq.4)}$$

The gain of information comes with the obstinacy of building a decision tree, to estimate the information about each attribute returning the greatest gain over the independent attribute. This induces, the gain of information (T,X) that applies the resource on the attribute X; however, Entropy(T) against the complete data set, and Entropy(T, X) with its applicability to the feature suffers a due

disadvantage, in relation to the fit of the model in the treatment of the division of strong training data, and considerably reduce the accuracy of the test (Ravipati and Abualkibash, 2019).

$$\text{Information Gain (T, X) (T, X= Entropia(T)} - \text{Entropia)} \quad \text{(Eq.5)}$$

3) *Naive Bayes*: Highly scalable algorithm, requiring a high number of linear variables (predictors) in a learning problem. Calculating the conditional probability of each attribute followed by a contained application of Bayes' theorem *(*see equation 6), to determine the relative probability on the characteristics of the attributes, in the sense of predicting the outcome (Aggarwal, 2014). Further details of the implementation of the *Naive Bayes* algorithm in *Weka* can be found by John and Langley (1995).

$$P(A|B= \frac{P(B|A) \times P(A)}{P(B)}) \quad \text{(Eq.6)}$$

2.4 PERFORMANCE METRICS

For this due situation, it occurs in the involvement of metrics, in *micro average* (Abracadabra, 2018), to be used as performance measures specific to the study. The approaches adopt the feasibility on sampling, and for this were suitable for the study:

1) True Positive indicates that the dataset is classified as an attack by the classification model.

2) True Negative predicts a non-(normal) response and is consistent with the data observed in the connection.

3) False Positive lists the number of instances classified as normal but being identified as anomalies by the classifier.

4) False Negative predicts an anomalous connection as no, but this one should be yes.

5) True positive rate (TPR, which is the type of sensitivity) and the probability of an actual test being positive. A (equation 7) defines the type of the measure.

$$TPR = \frac{VP}{VP+FN} \quad \text{(Eq.7)}$$

6) False positive rate. Let (equation 8) be in the determinance, that the FP are the numbers of false positives and the VN is the number of true negatives, with its probability being triggered when the value of the intrusion is true, but it is determined as negative.

$$FPR = \frac{\text{FP}}{\text{FP+VN}} \qquad \text{(Eq.8)}$$

7) Accuracy analyzes the ability to correctly classify a data object as normal or anomalous. It is defined with (equation 9).

$$Acurácia = \frac{\text{VP+VN}}{\text{FP+VP+VN+FN}} \qquad \text{(Eq.9)}$$

8) Accuracy evaluates the amount of positive ratings that are consistent with the data set. A (equation 10) defines the type of measure.

$$Precisão = \frac{\text{VP}}{\text{VP+FP}} \qquad \text{(Eq.10)}$$

9) *Recall* is the ratio of the classifier that could recognize the numbers of positive attacks. A (equation 11) defines the structure of the measure.

$$Recall = \frac{\text{VP}}{\text{VP+FN}} \qquad \text{(Eq.11)}$$

10) *F-Measure* considered as a classifier precision, in addition to defining a harmonic mean between the precision measure and *recall*. For its use the (equation 12) is defined.

$$F - Measure = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão+Recall}} \qquad \text{(Eq.12)}$$

11) False Alarm Rate only calculates the number of incorrect predictions of the classifier algorithm by the number of true positive. A (equation 13) defines the characteristic of its use.

$$Taxa\ de\ Alarme\ Falso = \frac{\text{FP + FN}}{\text{VP}} \qquad \text{(Eq.13)}$$

12) MCC or *Matthews Correlation Coefficient* is the coefficient that measures the quality of the classification (Liu et al., 2014). A (equation 14) is bounded for use.

$$MCC = \frac{(\text{VP}\times\text{VN}) * (\text{FP}\times\text{FN})}{\sqrt{(\text{VP+FP}) * (\text{VP+FN}) * (\text{VN+FP}) * (\text{VN+FN})}} \qquad \text{(Eq.14)}$$

13) ROC or *Receiver Operator Characteristic Curve*, being a measure of the accuracy area of the model that indicates the compensation between the TPR true positive rate *and the FPR false positive rate* $= \frac{\text{VP}}{\text{VP+FN}}$. $= \frac{\text{FP}}{\text{FP+VN}}$ This tells the model itself and the test suite to analyze the correct data and compensate for incorrect data.

14) Time that represents the construction of the classification model until the results.

2.5 CLASSIFICATION

What determines the dazzle of the data is dictated with the characteristics of (pseudocode I). Presenting in a simple way, the involved parts of data instances related to the algorithms (*Naive Bayes*, *J48* and *Logistic*), in a supervised classification based on patterns and associations of the data labeled to this study to the standard of 70% on top of the training base and 30% on the test base. In addition, cross-validation that breaks down the dataset into parts ensuring that they are in one form random of i=10 or $E= \frac{1}{10}\sum_{i=10}^{i} E_i$ i.e., isolating the Data group on track for training to estimate the models, while another party makes the relation to the test, validating each of the models. And, the pre-evaluated data then identified, in an instance of the respective and predictive class, are previously analyzed their results through performance metrics.

Pseudocode I: Processing Model
1. Reading from the *Dataset*
2. 70% data instance split into training
3. 30% test data instance split
4. Scroll through each indicated model (*Naive Bayes, J48* and *Logistic*)
5. Model (*Naive Bayes, J48* and *Logistic*) classify the training data
6. Evaluate the model with the test data
7. Perform cross-validation $E= \frac{1}{10}\sum_{i=10}^{i} E_i$
8. Cycle through valid attribute classes $\sum_{i=4}^{i} E_i$ (*Info*)
9. Analyze data through performance metrics
10. Return the time in *ms* and the total time of each model presented

Source: Authors (2023).

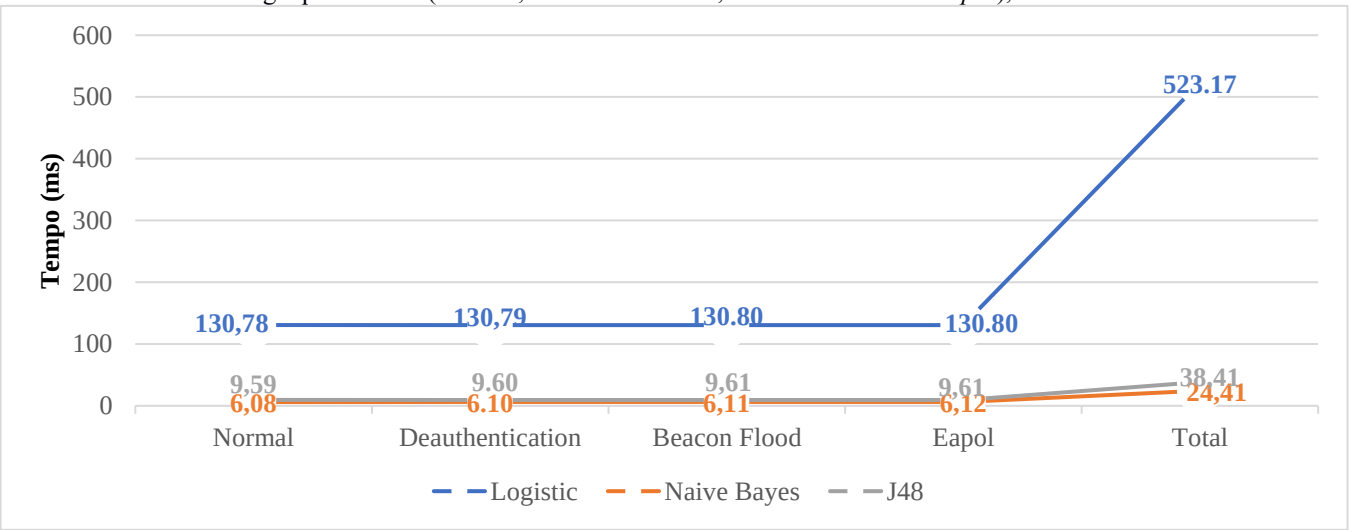
3 RESULTS AND DISCUSSION

In this section is basically the performance of the machine learning classification models that are evaluated on the Wireless network scenario in the context of IEEE802.11, in detection of the anomalies described in this study. As described, it evaluates the appropriate performance metrics and fluidity (or speed). And for the experiments we used 1 (one) Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz, 3401 MHz, 4 Core*(s), 8 Logical Processor(s), in addition to 8 (eight) GB's of* ram at *1333* MHz *and the* Windows 10 operating system. The software development environment was used, *IntelliJ* IDEA 2022.3.1 (*Community Edition*), *Apache Spark* API and *Weka API* in *Java*.

3.1 TIME OF ANALYSIS OF PERFORMANCE METRICS

In order to evaluate the performance metrics and even the computational total time calculation ratio of the classifiers proposed to the study, we obtained the processing speed in milliseconds (*Ms*) of 16,703 instances of the dataset (see Figure 3). The analyzed results indicate that, in the detection of the non-anomalous data (Normal), the classifier *Naive Bayes* obtained a cost of (6.08 *Ms*), already the type of attack *Deauthentication*, its value was respective to (6.10 *ms*), however, for the anomaly *Beacon Flood* The value presented is only (6.11 *ms*), already the consequent attack *Eapol* the value was his exact (6.12 *ms*) with a total of only (24.41 *ms*) being the best classifier, in computational cost in the verification of the aforementioned attacks. However, the *J48* regarding the type of data Normal, a computational cost of approximately (9.59 *ms*), in relation to the anomalous data *Deauthentication* the importance of (9.60 ms), following the other types of attacks, *Beacon Flood* and *Eapol* with the similarity of (9.61 *ms*), and already dealing with the total time, the *J48* had a performance not equal to its predecessor, but only (38.41 *ms*) accumulated. Being a range (*C*) of data, the *Logistic* and its respective and determined value of (130.78 *ms*) in the detection of normal data, very close to the type of attack *Deauthentication* with approximately (130.79 *ms*), however, the anomaly *Beacon Flood* and *Eapol* were obtained from similar values (130.80 *Ms*), with the consternation of the values reached on the algorithm *Logistic*, it is to be admired that its total value is the greatest of all, with its exact ones (523,17 *Ms*).

Figure 3. Computational consumption of the respective algorithms, *Logistic, Naive Bayes* and *J48.* Meaning by total values of each of the conforming equivalences (Normal, *Deauthentication*, *Beacon Flood* and *Eapol*), in *ms*.



Source: Authors (2023).

3.2 ANALYSIS OF PERFORMANCE METRICS ON THE CONJUNCTURE OF INTRUSIONS IN IEEE 802.11 NETWORKS WITH MACHINE LEARNING AT HOSPITAL N.S.C.

The analysis of *performance metrics* for the combination of attacks restricted to the study is analyzed in Table II, Table III, Table IV, Table V, presenting some of the results obtained, both in the training phase, as well as with regard to the test phase and the comparison with the proposed machine learning classifiers. All the results achieved were formed from a cross-validation of $E = \frac{1}{10}\sum_{i=10}^{i} E_i$ or 10 times. This is with the dictated presentations on study performance, which can be evaluated using some and/or several metrics of the confounding matrix: accuracy, precision and *recall* (TARCA et al., 2007), as well as the rates of true positives, true negatives, false positives, and false negatives of each class. However, a complementation of the data in sampling with *F-measure*, the false alarm rate, true positive rate, false positive rate, ROC and MCC.

3.3 ANALYSIS OF PERFORMANCE METRICS ON THE THEN NON-ANOMALOUS DATA TYPE - NORMAL

During the detection phase of the project, the non-anomalous data type (Normal) to the Naive *Bayes, Logistic* and *J48* classifier were well analyzed. Identifying, in table II, the rates of true positives between the marks of (8,249.0) to *Naive Bayes,* the indication of Logistic *presenting a number (8,246.0*), in addition to J48 and its (8,226.0) numbers. Have seen that, the numbers of similarities, in true negatives of number (7,569.0), false positives of number (0.0) and the rate of false positive with (0.00%) being well feasible for an intrusion detection system, in addition to an accuracy of exactly (100.00%) among all classifiers. Related to performance, in identifying that these values are coexisting with the type of data, as normal, in addition to inducing an accuracy of (94.70%) respective to the *Naive Bayes,* very approximate the *Logistic (94.68* %) and the classifier J48 (94.56 %).*Despite identifying the false negatives with (885.0) numbers on the* Naive Bayes *and* Logistic *with (888.0) numbers and just above (908.0) numbers on the* J48.

Table II. Performance metrics in evaluation of classifiers (*Naive Bayes*, *Logistic* and *J48*) by non-anomalous data type (Normal).

Classifier	VP	VN	FP	FN	A %	P %	R %	F-M %	TAF %	TVP %	TFP %	RC %	MCC %	T (*ms*)
Naive Bayes	8.249,0	7.569,0	0,0	885,0	94,70	100,00	90,31	94,91	10,73	90,31	0,00	96,70	89,92	6,08
Logistic	8.246,0	7.569,0	0,0	888,0	94,68	100,00	90,28	94,89	10,77	90,28	0,00	96,59	89,89	130,78
J48	8.226,0	7.569,0	0,0	908,0	94,56	100,00	90,06	94,77	11,04	90,06	0,00	96,47	89,67	9,59
Total samples: 16.703.0														

Caption: VP = True positive; VN = True negative; FP = False positive; FP = False negative; A = Accuracy; P = Precision; R = Recall; F-M = F-measure; TAF = False alarm rate; DVT = True positive rate; PFT = False positive rate; RC = ROC or *Receiver Operator Characteristic Curve*; MCC = *Matthews Correlation Coefficient*. Rows with data in their total volume of (VP, VN, FP and FN) differ from the due percentages (%) between treatments.

Source: Authors (2023).

Accordingly, cited to the study in the Table II, there are few divergences of values among the classifiers presented through the performance metrics. And this, is respective with the *recall* of *Naive Bayes* (90.31%), already approximated the *Logistic* (90.28%) and the classifier *J48* presents (90.06%). Despite this, the *F-measure* equals a somewhat similar value of *Naive Bayes* and *Logistic,* reaching (94.91%) and (94.89%), while the *J48* with the rate of (94.77%). The highest false alarm rate was presented by *J48* (11.04%), the *Logistic* with (10.77%) and the *Naive Bayes* with the lowest rate of (10.73%). Respectively, the *Naive Bayes* in detection of non-anomalous data in true positive rate was (90.31%), followed by the *Logistic* with a rate of (90.28%) and the *J48* and their respective (90.06%). The area of the ROC curve to the algorithm *Naive Bayes* (96.70%), slightly below the *Logistic* presenting (96.59%) and *J48* with the rate of (96.47%). When it comes to the quality of the classification according to the MCC, the *J48* obtained the lowest value (89.67%), followed by the *Logistic* demonstrating a rate of (89.89%) and the proportion of *Naive Bayes* being the highest among all classifiers (89.92%). However, the indication of computational time calculation is very relevant for the detection of the non-anomalous type, being the *Naive Bayes* representing (6.08 *ms*), followed by *J48*, and its perculiar time of (9.59 *ms*) and the last algorithm *Logistic* denoting (130.78 *ms*).

3.4 ANALYSIS OF PERFORMANCE METRICS ON THE THEN ANOMALOUS DATA TYPE - *DEAUTHENTICATION*

As for the search for the detection of anomalous data to the study, which is of the Deauthentication type. There is data on the performance metrics corresponding to Table III. To which the classifiers *Naive Bayes* and *Logistic* demonstrate that they filed the same metric of true positive, with its exact number (5,094.0), while the difference of J48 and its (5,085.0) numbers of true positives. The difference between the results is that the *Naive Bayes presented (10,724.0) true negatives and false positives (885.0), and however, the* Logistic classifier presented (10,722.0) true negatives and (887.0) false positives, however, *the J48* obtained a lower value of true negatives, with exact (10,715.0) numbers and a higher value of false positives (894.0) numbers. Nevertheless, it is worth

noting that *Naive Bayes and* Logistic *obtained a false negative rate merely equal to (0.0), and this comes to identify that* J48 obtained a practically higher number (9.0) of false negatives. However, due to the slight increase in numbers corresponding to true positive, true negative, false positive and false negative, *J48* presented a value slightly below accuracy with only (94.59%), while the *Naive Bayes and* Logistic *algorithms acquired a rate of (94.70%) and (94.69%).*

Table III. Performance metrics in evaluation of classifiers (*Naive Bayes*, *Logistic* and *J48*) by anomalous data type (*Deauthentication*).

Classifier	VP	VN	FP	FN	A %	P %	R %	F-M %	TAF %	TVP %	TFP %	RC %	MCC %	T (*ms)*
Naive Bayes	5.094,0	10.724,0	885,0	0,0	94,70	85,20	100,00	92,01	17,37	100,00	7,62	96,12	88,71	6,10
Logistic	5.094,0	10.722,0	887,0	0,0	94,69	85,17	100,00	91,99	17,41	100,00	7,64	96,03	88,69	130,79
J48	5.085,0	10.715,0	894,0	9,0	94,59	85,05	99,82	91,85	17,76	99,82	7,70	95,99	88,47	9,60
Total samples: 16.703.0														

Caption: VP = True positive; VN = True negative; FP = False positive; FP = False negative; A = Accuracy; P = Precision; R = Recall; F-M = F-Measure; TAF = False alarm rate; DVT = True positive rate; PFT = False positive rate; RC = ROC or *Receiver Operator Characteristic Curve*; MCC = *Matthews Correlation Coefficient*. Rows with data in their total volume of (VP, VN, FP and FN) differ from the due percentages (%) between treatments.
Source: Authors (2023).

However, with the high number of false positives (894.0), *J48* presented the worst accuracy with (85.05%). *Naive Bayes* and *Logistic* obtained a slightly higher margin (85.20%) and (85.17%). In order to have an attack completeness corresponding to *Deauthentication*, the *recall* of *J48* was (99.82%) and *Naive Bayes* and *Logistic* achieved (100.00%). With the harmonic mean between precision and *recall* was then obtained, the *F-measure* of the classifier *J48 (91.85%), followed by* Naive Bayes and *Logistic at values of (92.01%)* and *(91.99*%). In addition, in the false alarm rate the *J48* algorithm resulted in a basically higher value (17.76%), and respectively the *Naive Bayes and* Logistic obtained values with a lower rate presented (17.37%) and (17.41%). However, Table III helps to highlight the values assigned on the true positive rate to the *Naive Bayes* and *Logistic algorithm (*100.00%) and just below, the *J48* and its (99.82%). In addition, the false positive rate in *Naive Bayes is only (7.62%),* Logistic (7.64*%), but being resultant for* J48 *(*7.70%). However, the area of the ROC curve was of one opinion, in *Logistic* to its (96.03%), while to *Naive Bayes* a slight increase of (96.12%) and *J48* simply with the rate of (95.99%). Even so, the MCC quality coefficient on the *Naive Bayes* is respective to (88.71%), with the *Logistic* at a cost of (88.69%) and later the decision tree (*J48*) obtained a rate of only (88.47%). However, it is worth mentioning that the computational time of *Logistic* was the highest among all classifiers, followed by (130.79 ms), in the detection of the *Deauthentication* anomaly, *followed by the lowest value presented by* Naive Bayes (6.10 ms) and, *soon after, the* J48 *and its computational value of (9.60* ms*).*

3.5 ANALYSIS OF PERFORMANCE METRICS ON THE THEN ANOMALOUS DATA TYPE - *BEACON FLOOD*

By means of the classification following the *dataset* identified to this study, in the Hospital Nossa Senhora da Conceição (H.N.S.C.). The anomalous data type *Beacon Flood* can be analyzed by means of performance measures characterized to the classifiers interposed to the Table IV. However, there are a number of true positives (1,047.0), between, *Naive Bayes*, *Logistic* and *J48.* However, there is no similarity of data to the number of true negatives, with the *Naive Bayes* (15,656.0), soon after, the *Logistic* and its (15,655.0) numbers, and followed by the *J48* (15,633.0) numbers. In addition, there is a discrepancy of false positives, the *Naive Bayes* presents a number (0.0), then slightly above or rather irrelevant, the *Logistic* and its number (1.0), so the *J48* relevant (23.0) numbers. With the development on the numbers of false negatives where there was practically (0.0) or null on the classifiers presented.

Table IV. Performance metrics in evaluation of classifiers (*Naive Bayes*, *Logistic* and *J48*) by anomalous data type (*Beacon Flood*).

Classifier	VP	VN	FP	FN	A %	P %	R %	F-M %	TAF %	TVP %	TFP %	RC %	MCC %	T (*ms*)
Naive Bayes	1.047,0	15.656,0	0,0	0,0	100,00	100,00	100,00	100,00	0,00	100,00	0,00	100,00	100,00	6,11
Logistic	1.047,0	15.655,0	1,0	0,0	99,99	99,90	100,00	99,95	0,10	100,00	0,01	100,00	99,95	130,80
J48	1.047,0	15.633,0	23,0	0,0	99,86	97,85	100,00	98,91	2,20	100,00	0,15	99,95	98,85	9,61
Total samples: 16.703.0														

Caption: VP = True positive; VN = True negative; FP = False positive; FP = False negative; A = Accuracy; P = Precision; R = Recall; F-M = F-Measure; TAF = False alarm rate; DVT = True positive rate; PFT = False positive rate; RC = ROC or *Receiver Operator Characteristic Curve*; MCC = *Matthews Correlation Coefficient*. Rows with data in their total volume of (VP, VN, FP and FN) differ from the due percentages (%) between treatments.
Source: Authors (2023).

With the construction of rates to be analyzed by accuracy itself, there was a very good performance among all algorithms, with *Naive Bayes* presenting (100.00%) of assertiveness, *Logistic* with its rate of (99.90%), and just below J48, with its (99.86%). The accuracy of *Logistic* presented (99.90%) and then *Naive Bayes* (100.00%), while slightly below the *J48* (97.85%). Then the recall values were obtained, at a parity between all algorithms, with a rate of (100.00%). Similarly, table IV presents values, in *F-measure of (*100.00%) to Naive *Bayes,* in *Logistic the rate of (*99.95%) is visualized, with the completeness of the J48 algorithm seeks its value at a rate of (98.91 %). Despite this, the false alarm rate combined with the *J48* was the highest (2.20%), however, the rate of the *Naive Bayes* with the value of (0.00%) and *Logistic* (0.10%). However, the aforementioned algorithms obtained a similarity value of true positive rate (100.00%). The false positive rate for *Naive Bayes was exactly (0.00%) and* Logistic *(0.01%), and* J48 *(0.15* %). In arguing the false positive rate, the *Naive Bayes classifier obtained the lowest rate with its exact (0.00%), followed by the* Logistic classifier and the rate of (0.01%), and *concomitant* the J48 at a rate of (0.15%). To obtain the area of the ROC curve,

the algorithms cited in the study reached a value of (100.00%) between *Naive Bayes* and *Logistic*, so *J48* obtained the rate of (99.95%). The MCC to be presenting the quality of *the Naive Bayes* with the highest value (100.00%), *the Logistic* with (99.95%) and in pursuit with the value below the other algorithms, the J48 (99.85 %). And through, of a computational power, it was necessary to analyze in the table above, that the computational time of the *Naive Bayes* algorithm, was the most momentous, with (6.11 ms), then with a value a little above, the classifier *J48* displaying a rate of (9.61 ms*), in sequence, the highest value among all the classifiers presented by* Logistic *and the exact one (130.80* ms).

3.6 ANALYSIS OF PERFORMANCE METRICS ON THE TYPE OF ANOMALOUS DATA - *EAPOL*

Finally, for the detection of the *Eapol* attack, in a V frame. It points out, that the numbers interposed to each of the algorithms correlated to the study were very well presented, with a value of 100% on the respective rates before the metrics (Accuracy, Precision, *Recall,* F-measure*, False Alarm Rate, True Positive Rate,* False Positive Rate, ROC and MCC), in addition to having pleasant numbers of true positives (1,428.0) between, *Naive Bayes*, *Logistic* and *J48.* And followed by numbers of true negatives (15,275.0) among all designated classifiers, and with an optimal value of false negatives (0.0), then attracted to a number (0.0) of false positives.

Table V. Performance metrics in evaluation of classifiers (*Naive Bayes*, *Logistic* and *J48*) by anomalous data type (*Eapol*).

Classifier	VP	VN	FP	FN	A %	P %	R %	F-M %	TAF %	TVP %	TFP %	RC %	MCC%	T (*ms*)
Naive Bayes	1.428,0	15.275,0	0,0	0,0	100,00	100,00	100,00	100,00	0,00	100,00	0,000	100,00	100,00	6,12
Logistic	1.428,0	15.275,0	0,0	0,0	100,00	100,00	100,00	100,00	0,00	100,00	0,000	100,00	100,00	130,80
J48	1.428,0	15.275,0	0,0	0,0	100,00	100,00	100,00	100,00	0,00	100,00	0,000	100,00	100,00	9,61
Total samples: 16.703.0														

Caption: VP = True positive; VN = True negative; FP = False positive; FP = False negative; A = Accuracy; P = Precision; R = Recall; F-M = F-Measure; TAF = False alarm rate; DVT = True positive rate; PFT = False positive rate; RC = ROC or *Receiver Operator Characteristic Curve*; MCC = *Matthews Correlation Coefficient*. Rows with data in their total volume of (VP, VN, FP and FN) differ from the due percentages (%) between treatments.
Source: Authors (2023).

As there was no loss, but gain of information to what indicates the Table V, the accuracy data rates in all classifiers presented to the study was a ratio of (100.00%). Then, then the same (*Naive Bayes*, *Logistic* and *J48*) added more details to their own, with an accuracy of (100.00%) and a *recall* absurdity of (100.00%). The system presents a very related idea, for the test of the metric *F-measure*, with a detail presenting (100.00%) of all classifiers, and with the predictions themselves to the study were very relevant so that the false alarm rate did not exceed the range of (0.00%) among all algorithms. However, the true positive rate was quite high, with a rate of (100.00%) among all classifiers. That said, the false positive rate was from its relevant to *Naive Bayes*, *Logistic* and *J48*, of

only (0.00%), while the area of the ROC curve presented a significance of (100.00%) between the *Naive Bayes*, *Logistic* and *J48*, and it is observed that all classifiers were well qualified at a rate of (100.00%) on the CCM. However, it stresses that the computational cost presented by Table V, comes to point out, that the *Naive Bayes* identifies a lower value than the other classifiers, with the accurate use of (6,12 *ms*), as a result of the detection of the attack *Eapol*, the classifier *J48* presents a value slightly above (9.61 *Ms*), and with regard to the classifier *Logistic*, displays the highest value (130.80 *Ms*).

3.7 DISCUSSION

However, an IDS to analyze, process and classify the information in intrusion or normal is paramount for decision making that may occur on the *wireless* network. This primacy sought a great efficiency, due to the obtaining of results on a large number of true positives. However, the number of false positives was satisfactory, being small or even zero, and this study showed it. Nevertheless, to relate the structure of the project, the model of Aminanto et al. (2022) used the *Aegean Wi-Fi* Intrusion Dataset *2 (AWID2) dataset that has been corroborated in several studies and the convolutional neural network (CNN) in the classification of attacks on the Wi-Fi network with the perspective on the F1-Score evaluation metric with a score of (99.73%) in anomaly detection.*

Thus, in the WSN-DS database, Quincozes and Kazienko (2020) addressed the J48 (*Decision Tree)* classifier, in addition to Naive Bayes, REP *Tree, Random Tree, and Random Forest, processing in the best categorization of "gray holes" and "black holes" (i.e., similar attacks on DoS in wireless sensor networks), and while using* Random Tree (Random Tree) comes to categorize the detection of "floods" better, but evaluating the data for accuracy. Noting during the detection of "black hole", that *J48 achieved (97.88%) accuracy, while* REP Tree *(97.89%)*, being the most accurate algorithms. However, Naive *Bayes* (97.47%), Random Forest *and Random Tree* had a rate of (97.71%) and *(97.72%)*. However, regarding the detection of "gray hole", *J48* and *REP Tree* filed the same accuracy (98.11%). Despite this, the *Naive Bayes presented the lowest accuracy (97.50%), and between* the Random Forest and *Random Tree obtained (98.06%) and (98.07* %). Subsequently, the authors detected "flood" attacks, obtaining the Random *Forest* with a value of (99.13%) of accuracy, with a slightly lower accuracy (99.11%) to the *Random Tree,* and the other algorithms presented data variations.

And in the way of this, in Qin et.al's approach. (2018) have been using the *Aegean* Wi-Fi *Intrusion Dataset* (AWID) with the selection of 18 useful attributes instead of 154, for a *performance* in improving the accuracy of anomaly detection through support vector machine (SVM) with the approximation of 89.18%, 87.34% and 99.88%, in "flood" attacks, "injection" attacks and normal data. Other promising results were also obtained, when Patil and Agarkhed (2020) used the paradigm of the

Radial Bias function, for the best anomaly detection rate in WSN (Wireless Sensor Networks) and to meet the low amount of false positives, based on the knowledge of decision tree techniques in the accuracy of 98.00% over the *Sleep Deprivation* and *Sinkhole attacks*.

Thus the benefits of machine learning through the proposed environment, in intrusion detection in wireless networks, are realized through the increase of true positives and a low range of false positives. But based on the results just seen above, the most concise algorithm will depend on the type of attack. For, in view of the type of attack *Eapol*, the classifiers *Naive Bayes*, *Logistic* and *J48* obtained excellent results, among all the evaluation metrics. And, in our findings reveal that the detection of *Beacon Flood*, were obtained an optimal performance in evaluation metrics, mainly between *Naive Bayes* and *Logistic*, indicating higher accurate results, instead of *J48*. And in general, because it is a time in (*ms*), the *Naive Bayes* algorithm has been shown to be faster. Already dealing with the type *of attack Deauthentication* and its simultaneous sending of unrealistic frames. It succeeded designated false positive values over an average (888.6) numbers in the *Naive Bayes,* Logistic *and* J48 *algorithms,* but add that the designated rates is formidable. However, not performing as well as the other attacks, the EAPOL-Logoff, *Beacon Flood*, and the consequent non-anomalous type (Normal), that such representation of the algorithms provide significant results.

Therefore, the use of machine learning proves to be a deep technology and with several existing discoveries about the MAC 802.11 framework, in the way of correlative attacks on *wireless networks*.

4 CONCLUSION

In this study we proposed the release of anomalies in a corporate network contained in WPA2, in the promotion of the MAC framework 802.11 and its attributes (*Protocol Version,* Type*, To DS*, *From DS*, More Fragment*, Retry*, Power Management*, More Data, WEP,* Order , *Duration,* Transmitter *Address, Destination Address, Source* Address, Receiver Address*, BSS Id* and *Sequence Number*). In addition, presented by a supervised learning technique, in random sub-sampling of the majority class applied to the use of filtering to a certain percentage selection (30%), for random oversampling of the minority class, in addition to the condition of a cross-validation and the classifiers predicting a categorical class label, to an instance attributing the characteristic (i.e., *Info*) of information to the type of attack corresponding to the study.

With a real data set. As far as we go, we know that data pre-processing was quite important, thus introducing better performance, relative to the types of attacks identified in the study. Despite this, the contained metrics reaching values above the average respectively, which is comparable and observed in existing studies. And yet, we operate in better detail, on the algorithms cited (*Naive Bayes*, *Logistic* and *J48*). In the future we intend to improve the data set and the processing model in order

to apply methods (i.e., algorithms to classify in real time), as well as recognize other types of attacks as semi-automatic or automatic applied in machine learning.

REFERENCES

Abracadabra (2018). https://tomaxent.com/2018/04/27/Micro-and-Macro-average-of-Precision-Recall-and-F-Score/.

Aggarwal C. C., Data Classification: Algorithms and Applications, 1st ed. Chapman & Hall/CRC, 2014.

Ahmad Md. S. and Tadakamadla S. 2011. Short paper: security evaluation of IEEE 802.11w specification. In Proceedings of the fourth ACM conference on Wireless network security (WiSec '11). Association for Computing Machinery, New York, NY, USA, 53–58. https://www.jstor.org/stable/2347628?origin=crossref.
Aircrack-NG (2022). http://www.aircrack-ng.org/doku.php.

Aminanto M. E., Wicaksono R. S. H. , Aminanto A. E., Tanuwidjaja H. C., Yola L. and Kim K., "Multi-Class Intrusion Detection Using Two-Channel Color Mapping in IEEE 802.11 Wireless Network," in IEEE Access, vol. 10, pp. 36791-36801, 2022. https://ieeexplore.ieee.org/document/9745910.

Arasaki, A. M.; Della Flora, J. C. L. Teste de intrusão em redes sem fio padrão 802.11. 2012. 63p. Monografia - Curso de Pós-Graduação em Redes de Computadores e Segurança de Dados. Centro Universitário Filadélfia de Londrina - UniFil, Londrina, 2012.

Barford P., Kline J., Plonka D., and Ron A. , “A signal analysis of network traffic anomalies,” in Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement, pp. 71–82, ACM, 2002. https://dl.acm.org/doi/10.1145/637201.637210.

Cessie L. S. and Houwelingen V. J.C. (1992), Ridge Estimators in Logistic Regression. Journal of the Royal Statistical Society: Series C (Applied Statistics), 41: 191-201. https://doi.org/10.2307/2347628 Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Feng P., "Wireless LAN security issues and solutions," 2012 IEEE Symposium on Robotics and Applications (ISRA), 2012, pp. 921-924. https://ieeexplore.ieee.org/document/6219343.

IEEE Standard 802.11 (1999). IEEE Standards for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. https://ieeexplore.ieee.org/document/1389197.

IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Protected Management Frames, in IEEE Std 802.11w-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, and IEEE Std 802.11y-2008) , vol., no., pp.1-111, 30 Sept. 2009. https://ieeexplore.ieee.org/document/5278657.

IEEE Standard for information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 6: Medium Access Control (MAC) Security Enhancements, in IEEE Std 802.11i-2004 , vol., no., pp.1-190, 24 July 2004. https://ieeexplore.ieee.org/document/1318903.

Java (2022). https://www.java.com.

John G. H. and Langley P., Estimating continuous distributions in bayesian classifiers, in Proceedings of the Eleventh conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1995, pp. 338–345.

Linhares, A.G.; Gonçalves, P. A. da S. Uma análise dos mecanismos de segurança de redes IEEE 802.11: WEP, WPA, WPA2 e IEEE 802.11 w. Recife: UFPE, 2012.

Liu Z., Cheng J., Yan C., Wu X., and Chen F., "Research on the matthews correlation coefficients metrics of personalized recommendation algorithm evaluation," Int. J. Hybrid Inf. Technol, vol. 8, no. 1, pp.163–172, 2015. https://gvpress.com/journals/IJHIT/vol8_no1/14.pdf.

Mdk3 (2022). https://en.kali.tools/?p=34.

Mitchell, T. Machine learning. New York: McGraw-Hill, 1997.

Morimoto, C. E.; Redes, Guia Prático. Porto Alegre; Sul Editores; 2008.

Patil B. and Agarkhed J., "An Exploratory Machine Learning Technique for Investigating Intrusion in Wireless Sensor Networks," 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), 2020, pp. 1-6. https://ieeexplore.ieee.org/document/9297969.

Qin Y. , Li B., Yang M. and Yan Z., "Attack Detection for Wireless Enterprise Network: a Machine Learning Approach," 2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2018, pp. 1-6. https://ieeexplore.ieee.org/document/8567797.

Quincozes S. E. and Kazienko J. F., "Machine Learning Methods Assessment for Denial of Service Detection in Wireless Sensor Networks," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020, pp. 1-6. https://ieeexplore.ieee.org/document/9221146.

Quinlan J. R., C4. 5: programs for machine learning. Elsevier, 2014.

Ravipati R. D. and Abualkibash M., Intrusion Detection System Classification Using Different Machine Learning Algorithms on KDD-99 and NSL-KDD Datasets - A Review Paper (June 2019). International Journal of Computer Science & Information Technology (IJCSIT) Vol 11, No 3, June 2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3428211.

Scarfone, K. and Mell, P. (2007). Guide to intrusion detection and prevention systems (IDPS), volume 800. NIST.

TARCA, A.L. et al. Machine learning and its applications to biology. PLoS Computational Biology, California, v. 3, n. 6,p. 953-963, 2007.

Tews, E. (2007). Attacks on the WEP Protocol. Cryptology ePrint Archive, Report 2007/471.

Wi-Fi Alliance (2003). Wi-Fi Protected Access: Strong, Standards-based, Interoperable Security for Today's Wi-Fi Networks.

Wireshark (2022). https://www.wireshark.org/.