A Machine Learning Model to Resource Allocation Service for Access Point on Wireless Network

Davi Militani[1], Samuel Vieira[1], Everthon Valadão[2], Kátia Neles[3], Renata Rosa[1], Demóstenes Z.Rodríguez[1]
[1]Department of Computer Science, Federal University of Lavras, MG, Brazil
[2]IFMG Campus Formiga, Formiga, MG, Brazil
[3]Faculdade Martha Falcão, Manaus, Brazil
{militani;samuel}@ufla.br, everthon@ifmg.br, katia.neles@gmail.com,{demostenes.zegarra;renata.rosa}@ufla.br

***Abstract*—Currently, an access point (AP) is usually selected based on the signal strength parameter. However, the signal strength is not a guarantee of a good quality of service (QoS). Machine learning algorithms are used to automatically learn and improve some tasks and based on a network device characteristics is possible to select the most important input for a better network coverage. Thus, in this paper is implemented a Resource Allocation service for wireless networks based on machine learning algorithms. In this research, the Random Forest algorithm was implemented to automatically determine the AP selection strategy (SS). The results of the RF algorithm applied to heterogeneous network technologies showed an improvement of the channel condition, in relation to the throughput. In the validation tests phase, the experimental results demonstrated that our proposed AP SS based on Random Forest algorithm outperforms an existing AP SS based on signal strength.**

***Index Terms*—Random Forest, access points, wireless network, QoS**

I. Introduction

The access point (AP) selection usually does not have a centralized control and many times there are multiple layers of coverage, in which a user receives a high level of receiving signal strength (RSS) from more than one AP [1]. The selection problem in Variable channel-width WLANs can also occur and algorithms can be used for the problem treatment [2]. Additionally, devices that are produced by more than one vendor present difficult to use the network provided by another vendor. These kind of problems promote the development of high level solutions, which can be accepted by different equipments and devices [3]–[5].

Nowadays, the heterogeneous network interaction has been studied in many studies [6], mainly in the case of an AP selection [7], [8]. The signal strength strategy (SSS), which is performed in AP equipments sometimes is not related to a good quality of service (QoS) [9], [10] to users in heterogeneous network with an unbalanced load scenario [2]. The solutions developed for AP selection strategy (SS) [3]–[5] looking for a good connectivity performance. However, the AP SS in heterogeneous networks has not yet been deeply investigated and the solutions do not consider that new future attributes can be aggregated in the AP SS, working in the majority with fixed metrics.

The machine learning algorithms [5], [11] including neural networks [12] have been used in general applications [13]–[15]. Some studies applied the Bayesian analysis technique [16], in which it presents computational simplicity; however, it achieves a low accuracy for SS classification. The Decision Tree-based scheme is used in [17] presenting low time complexity in the testing process, with good classification accuracy; however, it requires a high time complexity in the training process. In this research, in order to choose the resource allocation service for access point on wireless network, different machine learning algorithms were tested, and the Random Forest reached the best performance with good accuracy and low complexity in training and testing phase of classification of SS.

The remainder of this article is structured as follows. Section II presents the related works about the AP SS and the machine learning algorithms applied to the increasing of throughput in networks. In section III presents the proposed SS model. Section IV presents the results about the scenario of network emulation. Finally, the conclusions and future works are presented in section V.

II. Related Works and Main Concepts

In this section, the main works about AP SS and machine learning algorithms are treated.

A. Access Point Selection Strategy

The SS is used to select the AP that will deliver the maximum benefit in channel condition for the user. In the majority, the SS considers throughput and packet loss. The SS can occur randomly, in which an algorithm chooses an AP at random [18]; or the SS can choose the AP with the strongest signal strength; the SS also can choose an omniscient way, in which an algorithm simulates AP probes with the aim of choosing the AP with the best bandwidth [19].

Solutions using fuzzy logic are used in [1] to design an AP selection method for visible light communication (VLC) networks. The study [1] determines the users that should be connected to the system, and it assigns the remaining users in a stand-alone VLC.

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grant 2015/24496-0 and Grant 2018/26455-8

Another study [19], uses an algorithm to test every candidate AP and calculate its expected bandwidth by communicating with a server. In [20], every candidate AP is probed, and its bandwidth is computed according to the time spent exchanging the probing frames. However, these schemes present additional latency and signaling overhead to analyze every AP, becoming critical in dense WLANs.

Some conditions of APs are analyzed in [21], in which mobile users among the APs are distributed. This solution requires considerable modifications at the APs to balance the AP loads. Another scheme [22] collects historical data of the AP performance. However, [22] requires a server to save Wifi-Reports; additionally, the mobile devices consume additional battery power to submit reports to the server.

The software-defined networking (SDN) also is used in other solutions [23], in which the SDN controller considers the QoS requirement of a user going to a network. Nevertheless, in this solution substantial signaling overhead is applied to track the flows connected to the network.

Differently of the cited works, the solution in this work does not consume extra resources of the mobile devices or AP, not affecting the power consumption of the network device. Additionally, in the proposed model, in case of new characteristics or attributes of an AP, it is only necessary to generate another model automatically. Thus, the solution can be adapted to complex and dynamic customized rules.

B. Machine Learning Applied on Network Problems

Machine learning algorithms are useful to work with large amount of information, providing results with high precision [24]. The algorithms commonly used are Decision Trees (J48), support vectors (SMO) or artificial neural networks (ANN) [25].

The increasing use of machine learning algorithms in different areas has been happening in the recent years, such as in the security scenarios [26] and speech recognition [27].

In [28] a mathematical model of the multi-objective mapping for virtualized resources is established for heterogeneous radio access network. Thus, the dynamic differential evolutionary algorithm is used to solve the multi-objective model. The weight values of objective function are adjusted by machines learning algorithm. In [28] only a phase is built by machine learning algorithm and the details about the machine learning phase are not explained.

Some studies [16], [17] use machine learning for the network selection problem; however they present an accuracy about 80%. Additionally, in [17] a high time complexity in the training process is presented. It is important to note that there is scarce works using machine learning algorithms to resolve the problem of resource allocation service on heterogeneous network and with an accuracy higher than 80%.

The Random Forest is a supervised learning algorithm in which it creates a forest in a random way. The created forest is a combination (ensemble) of Decision Trees, according is shown in Fig. 1, in most cases trained with the bagging method and the algorithm has presented good accuracy in many studies [29], [30]. The Random Forest aggregates many Decision Trees to limit overfitting as well as error due to bias and therefore yield useful results [30].

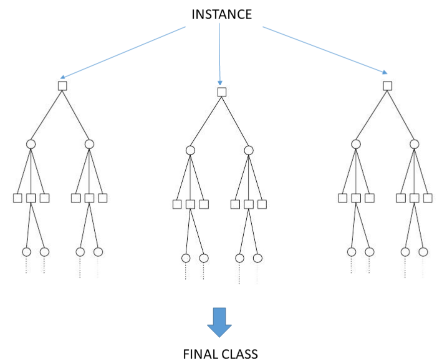

Fig. 1. Random Forest algorithm topology.

A Random Forest is a predictor, containing a collection of M randomized regression trees. In the case of a j tree, the predicted value is denoted by $m_n(\mathbf{q}; \Theta_j, \mathcal{D}_n)$ at the query point q. The $\Theta_1, \ldots, \Theta_M$ are independent random variables. The k_{th} tree estimate is represented by:

$$\mathrm{m}_n(\mathbf{x}; \Theta_k, \mathcal{D}_n) = \sum_{i \in \mathcal{D}_n^\star(\Theta_k)} \frac{1_{\mathbf{x}_i \in A_n(\mathbf{x}; \Theta_k, \mathcal{D}_n)} Y_i}{N_n(\mathbf{x}; \Theta_k, \mathcal{D}_n)} \quad (1)$$

where $\mathcal{D}_n^\star(\Theta_k)$ is the set of data selected before the tree construction, $A_n(\mathbf{x}; \Theta_k, \mathcal{D}_n)$ represents the cell containing x , $N_n(\mathbf{x}; \Theta_j, \mathcal{D}_n)$ represents the number of points that fall into $A_n(\mathbf{x}; \Theta_j, \mathcal{D}_n)$.

Among the various reasons for adopting machine learning algorithms to the problem of resource allocation service for AP, the main reason is the flexibility. Through a machine learning algorithm the system can learn new knowledge of a particular domain, customizing automatically new rules. The solution can perform network management tasks without intervention from a human, adapting itself to changes of the environment. The main proposal of this work is present a machine learning solution that presents an accuracy higher than 80% with low complexity in training and testing phase of classification of SS.

In the present work, the Random Forest implementation was developed in the Waikato Environment for Knowledge Analysis (Weka) package, which is a collection of machine learning algorithms. The processor used is an Intel Xeon with 4 physical cores and a frequency of 2.4 GHz and 8 GB of RAM.

In this work, different machine learning algorithms were tested, such as Naive Bayes, decision trees and Random Forest. However, the Random Forest reached the best performance.

III. Proposed Method for Access Point Selection Strategy

This section presents the proposed method of AP SS based on machine learning model. The Fig. 2 presents the steps for the AP SS, in which the AP characteristics are extracted and directed to the machine learning model.

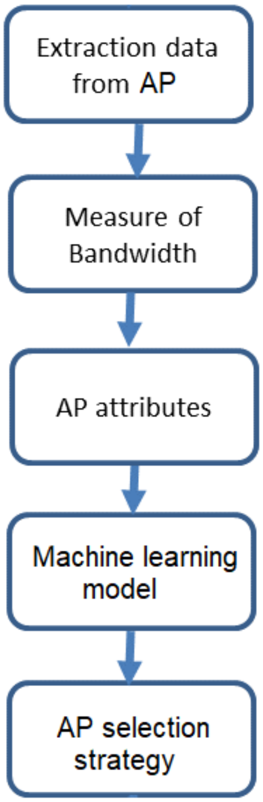

Fig. 2. High-level architecture of the proposed model for AP SS.

A. Machine Learning Model

Different machine learning algorithms were tested and the throughput was measured. The algorithms tested were: Naive Bayes, Decision Trees and Random Forest. However, because the Random Forest reached the best performance, with an average accuracy of 87% in the testing phase, it was chosen in the final tests. The algorithm Naive Bayes presented an average accuracy of 79% and the Decision Tree obtained an accuracy of 83%.

There are three kind of Naive Bayes models, in this work was tested the Gaussian distribution with standard configuration; the decision tree algorithm used in this work was the C4.5 (J48) with confidenceFactor of 0.25 and numFolds of 3. In the Random Forest algorithm, two completely-random tree forests and two Random Forests were used. In which, each completely random tree forest contained the total of 500 completely-random trees and each Random Forest contained 500 trees. This topology was selected after many tests to discover the better topology. The forest produced an estimate of class distribution and then averaging across all trees in the same forest.

The main attributes used in the machine learning algorithms were the ranges: distance location, channel number, frequency, signal range, transmit power, antenna gain, antenna height, stations number in range, stations associated and bandwidth.

These attributes were chosen according to the previous tests and to studies [31], [32] about wireless network characteristics.

Firstly, a model is obtained for the following steps, in which the model is performed in a server to not overhead the stations:

- the wireless characteristics are extracted by an API and send to a server;
- a script joins the attributes to serve as input to the Random Forest;
- the Random Forest runs and the model is obtained; the model is carried as input of any AP device.

After the model is obtained, it is not necessary obtain frequently new models, only in case of a new device with new characteristics is detected by an automatic script. Thus, the machine learning model is used for the classification of the better resource allocation Service, following the steps:

- All available access points in range are selected.
- The bandwidth between the target station and the available AP is calculated. This information is saved in the machine learning model.
- Other characteristics of the AP that are influenced by the movements are captured.
- The bandwidth between the station and the APs are calculated. This information is also saved in the machine learning model.
- By the end, the Random Forest choose the best AP.

The machine learning algorithms, used in this work, have its parameters all adjusted, such as, learning rate, momentum rate, number of times, training method and number of layers and neurons.

B. Network emulation

The network emulation in this work was performed by the Emulator for Software-Defined Wireless Networks, the Mininte-WiFi.

A scenario of network with hybrid access technologies of WiFi, satellite and LTE was emulated. The network is composed by 3 WiFi APs, 1 satellite AP, 1 LTE AP and 35 stations in an area of 300mx300m and the simulation time was 1000s. The parameters used in the emulation are the following:

- WiFi - IEEE 802.11b - AP range 75m - Channel Bandwidth 20 Mbits/s
- Satellite - AP range 1000m - Channel Bandwidth 10 Mbits/s
- LTE AP - range 1000m - Channel Bandwidth 10 Mbits/s
- Mobile Stations - range 75m
- Packet Loss 10%
- Mean Mobile Stations speed 5 m/second

Initially, the mobile stations are distributed randomly, with a random speed. The stations are connected in a server “s1” via

WiFi APs according to the SSS by default. The “s1” processes multimedia content and text from the mobile stations. The iperf tool runs on each one of the stations in the beginning of the experiment, to capture the bandwidth value between themselves and the “s1” in real time. The bandwidth values are recorded for analysis. The mobile stations began to gather after a determined time of 300 second, in which an incident occurs and the devices toward the incident. The mobile stations take different positions and they settled in an area near the incident.

It is important to note that the different scenarios of movement were recorded to use the same scenarios, but using the AP SS based on signal strength. This was performed for comparison of the AP SS performed by the machine learning model and the AP SS without the machine learning model.

IV. Results

In this section, we describe the experimental results of the current AP selection method and the proposed method.

A. Machine Learning Topology

Previous network emulation tests were performed with another machine learning algorithms and Table I presents the values of average throughput gain compared to the SS technique based on signal strength.

TABLE I
INCREASE OF THE AVERAGE THROUGHPUT OF THE MACHINE LEARNING ALGORITHMS COMPARED TO THE SS TECHNIQUE BASED ON SIGNAL STRENGTH

Algorithm	Average Throughput
Naive Bayes	33.22 to 54.17 %
Decision Tree	89.11 to 91.08 %
Random Forest	123.15 to 250.34 %

B. Throughput

The throughput for some of the mobile stations using the Random Forest algorithm is shown in Fig. 3 and Fig. 4.

In Fig. 3 and 4 the x-axis represents the time elapsed on the experiment, and the y-axis represents the end-to-end throughput detected by the iperf software. The time when the machine learning is executed is shown by the black dash line. After the machine learning algorithm is executed, a significant raise can be observed in every mobile station. In the Fig. 3 and Fig. 4 are shown only the mobile devices 3, 11, 17 and 30. However, all the mobile devices present similar behavior.

The average throughput gain value of each mobile station varies between 123% until 250%, according is shown in Fig. 5.

Fig. 5 presents the comparison of both AP selection methods, the Random Forest for the SS and the SS based on signal strength, in terms of throughput, in which were selected 14 mobile stations.

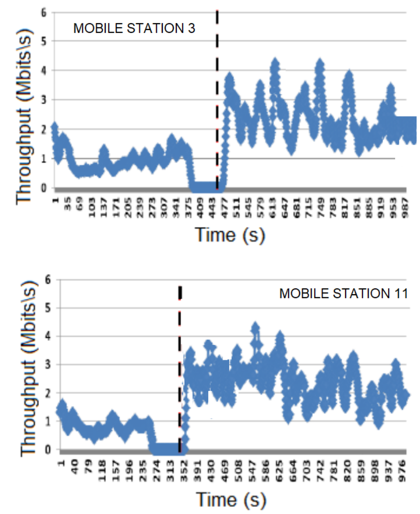

Fig. 3. End-to-end throughput of mobile stations in which the machine learning algorithm is applied - mobile stations 3 and 11.

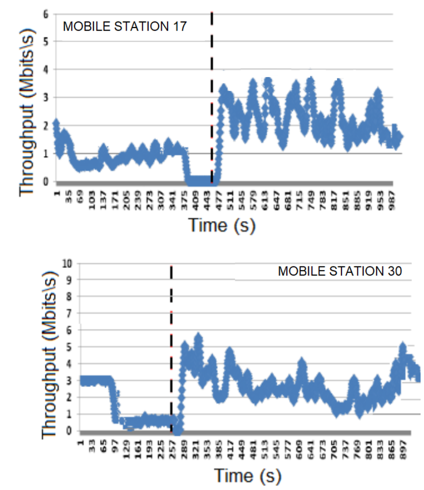

Fig. 4. End-to-end throughput of mobile stations in which the machine learning algorithm is applied - mobile stations 17 and 30.

C. Overhead

In this work the overhead was also measured in terms of execution time. This occurs since the start of execution of the machine learning model until the AP new selection. The overhead is about 3.1s and the AP SS based on signal strength is about 9.5s. This occurred due to the fact that the machine learning algorithm after the model is performed do not causes a relevant overhead independent of the scale and complexity of the network.

V. Conclusion and Future Works

Currently, the combination of various network technologies is very common and is necessary new techniques for AP SS. The experiments implemented in this study have shown that the AP SS based only on signal strength is not more an

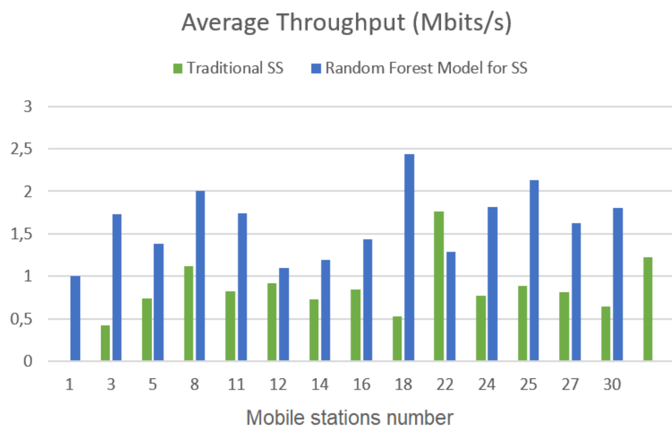

Fig. 5. Comparison of the both selection methods in terms of throughput.

efficient parameter; if the QoS is involved, then a higher level of performance is required. Thus, the research presented in this work searched for a solution based on machine learning algorithm, mainly based on Random Forest algorithm and the results show its efficiency for treating this kind of problem. The Random Forest model was compared to the SS based on signal strength and the evaluation results presented a throughput gain of 123% until 250%, also presenting low values of overhead. Thus, the Random Forest has shown an increase of the average throughput on mobile stations with a low overhead, with an average accuracy of 87%. In future works, other WiFi protocols will be used in the network emulation and we will consider varying the parameters used on simulator.

REFERENCES

[1] X. Wu, D. Basnayaka, M. Safari, and H. Haas, "Two-stage access point selection for hybrid vlc and rf networks," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2016, pp. 1–6.

[2] X. Chen, W. Yuan, W. Cheng, W. Liu, and H. Leung, "Access point selection under qos requirements in variable channel-width wlans," *IEEE Wireless Communications Letters*, vol. 2, no. 1, pp. 114–117, February 2013.

[3] Zied Ben Houidi, "A knowledge-based systems approach to reason about networking," in *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*, New York, NY, USA, 2016, pp. 22–28, ACM.

[4] Q. Zhou, C. Wang, S. McLaughlin, and X. Zhou, "Network virtualization and resource description in software-defined wireless networks," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 110–117, November 2015.

[5] D. Lee and C. S. Hong, "Access point selection algorithm for providing optimal ap in sdn-based wireless network," in *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Sep. 2017, pp. 362–365.

[6] E. Gioia, P. Passaro, and M. Petracca, "Amber: An advanced gateway solution to support heterogeneous iot technologies," in *2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, Sep. 2016, pp. 1–5.

[7] X. Wu, M. Safari, and H. Haas, "Access point selection for hybrid li-fi and wi-fi networks," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5375–5385, Dec 2017.

[8] I. Ben Dhia, M. Bouhtou, T. En-Najjary, S. Lahoud, and X. Lagrange, "Optimization of access points selection and resource allocation in heterogeneous wireless network," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.

[9] D. Z. Rodríguez and S. Möller, "Speech quality parametric model that considers wireless network characteristics," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, June 2019, pp. 1–6.

[10] D. Z. Rodríguez, R. L. Rosa, F. L. Almeida, G. Mittag, and S. Möller, "Speech quality assessment in wireless communications with mimo systems using a parametric model," *IEEE Access*, vol. 7, pp. 35719–35730, 2019.

[11] E. T. Affonso, R. L. Rosa, and D. Z. Rodríguez, "Speech quality assessment over lossy transmission channels using deep belief networks," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 70–74, Jan 2018.

[12] Xuewen Qian and Marco Di Renzo, "Receiver design in molecular communications: An approach based on artificial neural networks," in *International Symposium on Wireless Communication Systems (ISWCS)*, 08 2018, pp. 1–5.

[13] P. Vallabh, R. Malekian, N. Ye, and D. C. Bogatinoska, "Fall detection using machine learning algorithms," in *2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, Sep. 2016, pp. 1–9.

[14] Y. Cheng, H. Chou, and R. Y. Chang, "Machine-learning indoor localization with access point selection and signal strength reconstruction," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5.

[15] E. T. Affonso, R. D. Nunes, R. L. Rosa, G. F. Pivaro, and D. Z. Rodríguez, "Speech quality assessment in wireless voip communication using deep belief network," *IEEE Access*, vol. 6, pp. 77022–77032, 2018.

[16] Chunjing Song, Jian Wang, and Guan Yuan, "Hidden naive bayes indoor fingerprinting localization based on best-discriminating ap selection," *ISPRS International Journal of Geo-Information*, vol. 5, no. 10, 2016.

[17] Wang-hong Li and Zhu Qi, "Network selection algorithm based on decision tree in heterogeneous wireless networks," *MATEC Web of Conferences*, vol. 189, pp. 04010, 01 2018.

[18] M. Lu and J. Wu, "Localized access point selection in infrastructure wireless lan," in *MILCOM 2007 - IEEE Military Communications Conference*, Oct 2007, pp. 1–7.

[19] Anthony J. Nicholson, Yatin Chawathe, Mike Y. Chen, Brian D. Noble, and David Wetherall, "Improved access point selection," in *Proceedings of the 4th International Conference on Mobile Systems, Applications and Services*, 2006, MobiSys '06, pp. 233–245.

[20] T. Sun, Y. Zhang, and W. Trappe, "Improving access point association protocols through channel utilization and adaptive probing," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1157–1167, May 2016.

[21] T. Lei, X. Wen, Z. Lu, and Y. Li, "A semi-matching based load balancing scheme for dense ieee 802.11 wlans," *IEEE Access*, vol. 5, pp. 15332–15339, 2017.

[22] J. Pang, B. Greenstein, M. Kaminsky, D. McCoy, and S. Seshan, "Wifi-reports: Improving wireless network selection with collaboration," *IEEE Transactions on Mobile Computing*, vol. 9, no. 12, pp. 1713–1731, Dec 2010.

[23] A. Raschellà, F. Bouhafs, M. Seyedebrahimi, M. Mackay, and Q. Shi, "Quality of service oriented access point selection framework for large wi-fi networks," *IEEE Transactions on Network and Service Management*, vol. 14, no. 2, pp. 441–455, June 2017.

[24] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, April 2017.

[25] B. Balaguer, G. Erinc, and S. Carpin, "Combining classification and regression for wifi localization of heterogeneous robot teams in unknown environments," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2012, pp. 3496–3503.

[26] F. Jiang, Y. Fu, B. B. Gupta, F. Lou, S. Rho, F. Meng, and Z. Tian, "Deep learning based multi-channel intelligent attack detection for data security," *IEEE Transactions on Sustainable Computing*, pp. 1–1, 2018.

[27] F. L. de Almeida, R. L. Rosa, and D. Z. Rodriguez, "Voice quality assessment in communication services using deep learning," in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2018, pp. 1–6.

[28] S. Zou, Y. Zhao, and Y. Tang, "A novel algorithm of virtual resource allocation in heterogeneous radio access networks," in *2016 11th International Conference on Computer Science Education (ICCSE)*, Aug 2016, pp. 401–404.

[29] Te Han, Dongxiang Jiang, Qi Zhao, Lei Wang, and Kai Yin, "Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 8, pp. 2681–2693, 2018.

[30] Susan Athey, Julie Tibshirani, and Stefan Wager, "Generalized random forests," *Ann. Statist.*, vol. 47, no. 2, pp. 1148–1178, 04 2019.

[31] Andrea Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[32] S. Aust, R. V. Prasad, and I. G. M. M. Niemegeers, "Outdoor long-range wlans: A lesson for ieee 802.11ah," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1761–1775, 2015.