ML-based estimation of the number of devices in industrial networks using unlicensed bands

Oluwatobi Baiyekusi*, Haeyoung Lee†, Klaus Moessner†‡
*Department of Computer & Information Sciences, University of Strathclyde, Glasgow, UK
*5GIC & 6GIC, Institute for Communication Systems (ICS), University of Surrey, Guildford, UK
†Faculty of Electronics and information Technology, University of Technology Chemnitz, Germany
Email: oluwatobi.baiyekusi@strath.ac.uk, Haeyoung.Lee@surrey.ac.uk, klaus.moessner@etit.tu-chemnitz.de

***Abstract*—Advanced automation is being adopted by manufacturing facilities and wireless technologies are set to be a key component in driving the factories of the future. It is expected that private cellular networks and WLAN technologies would be deployed for smart factory operations. Since both wireless technologies can operate on the same channel in unlicensed bands, then efficient resource sharing becomes important. When multiple devices compete for the resource, the estimation of number of devices contending for the channel resource can help the design of an efficient resource sharing scheme. This paper aims to address the challenge of estimating the number of factory devices contending to transmit over the unlicensed channel. We adopt three machine learning (ML) techniques and develop a novel device number estimation system by collating and analysing the idle-time interval between transmission across the channel. By using NS-3 simulation, the performance of the proposed estimation approach is evaluated. The results presented reveal the significance of the chosen features and performance of each ML algorithm used.**

***Index Terms*—Machine learning, smart factory, number of device estimation, unlicensed band.**

I. Introduction

The anticipated industrial revolution 4.0 has begun and wireless communication technologies are expected to play a pivotal part in actualising its potential. As part of that revolution, factory automation is enhanced and one of the key use cases for 5G networks has been factory automation via private networks [1]. While factory automation in the past was limited (e.g. robotic arms tethered to control units via cables, etc.), that limitation can be taken away with wireless technologies. 5G private networks could bring multiple advantages in improved coverage, better control and enhanced security for factory automation. Furthermore, with access to unlicensed spectrum, private networks can be more affordable and widespread. While unlicensed bands are open to wireless networks coexisting together, Wireless Local Area Networks (WLANs) such as Wi-Fi have been so far dominant. However, cellular networks are expected to also operated in the unlicensed bands. In fact, standards have been ratified by the 3GPP for cellular networks to operate over unlicensed bands.

One of the main challenges that has arisen from the entrant of cellular technology into unlicensed bands has been the potential impact it may have on the performance of existing technologies operating in unlicensed bands (e.g. Wi-Fi) [2]. Different feasibility studies were conducted to evaluate the performance of LTE-LAA (License Assisted Access) in unlicensed bands but also its potential impact on technologies such as Wi-Fi [3], [4]. LTE-LAA and 5G NR-U (New Radio Unlicensed) adopted the fairly similar channel access procedures to the IEEE 802.11 standard, which Wi-Fi use, with the aim of achieving fair or harmonious coexistence with technologies using IEEE 802.11 standards. But one important aspect to fair, efficient and optimal use of the spectrum is determining the number of devices operating over the same channel. This information can be very useful in designing an effective channel access protocol that considers all users operating on the same channel from a fairness perspective. In the private network for the factory automation use case, estimating the number of factory devices (e.g. robotic arms, control units, smart cameras) can be crucial for implementing a solution which meets the requirements of time-sensitive applications.

Node number estimation has been studied in the literature widely [5]–[8]. The authors in [5] proposed a listen-before-talk (LBT) mechanism allowing estimation of the number of WLANs nodes by determining the collision probability. The estimation of number of nodes by calculating the collision probability has been the foundation for numerous estimation methods. One problem using the collision probability is the impact of the contention window (CW) size at which the collisions are measured, i.e. measuring the collision probability within a different CW size will give a different estimate of node numbers. Another problem is the different estimation accuracy level for the different node numbers. According to [6], as the node number increases, the estimation becomes less accurate. The authors in [6] proposed the extended Kalman filter based estimation coupled with a change detection mechanism, which can estimate the number of Wi-Fi nodes with high accuracy. However, the high accuracy reported was for large number of nodes without considering smaller number of nodes. In [7], the authors approached the estimation problem using batch and sequential bayesian techniques to reduce the computational burden. They proposed maximum a posteriori algorithm, but it traded-off accuracy to reduce computational complexity. The authors in [8] proposed an estimation technique by using the average idle slot interval. Three thresholds are determined to track the variation in the node numbers on the network. Based on empirical data, the average idle slots are obtained and a formula is provided. The node number estimation is performed based on the measured average idle slot and the threshold it falls under. The estimate of node numbers in [8] comes with a large variance in the number of nodes and remains insensitive to smaller increases in node numbers on the network.

While aforementioned works are non-ML based problems, in this paper, we present a *machine learning (ML) based node number estimation* approach. The approach exploits the capability of devices operating over the unlicensed bands to sense the channel before transmission, the LBT concept. The periodic but varying idle-time over the channel can be associated to a statistical distribution. The mean and standard deviation of this idle-time distribution can be characterised to the number of nodes actively contending over the channel. The dataset acquired from observing the idle-time can be used to train ML models to perform number of device estimations for contending factory devices operating over unlicensed bands. By using NS-3 system simulations, we show that the proposed ML-based number of device estimation approach, can reveal relationship between the idle-time distribution and the node number, but more importantly make predictions on the number of devices actively contending over the channel. The remainder of the paper is organized as follows. Section II describes the considered scenario and the number of device estimation problem. In Section III, the proposed machine-learning based prediction approach is elaborated. The performance validation are explained in Section IV to show the effectiveness of our proposed approach. Finally, we draw important conclusions in Section V.

II. Scenario Setup and Problem Formulation

In this section, we explain the scenario setup including network model of the factory automation use case and the MAC models of Wi-Fi and LTE-LAA.

A. Network Model

We consider the 3GPP indoor scenario consisting of a private LTE-LAA network coexisting with Wi-Fi, deployed on the same channel [9]. Fig. 1 illustrates our scenario setting, the layout of the network and factory devices. Four small cells are operated on each network, aligned and centred along the longer dimension of the building. The separation of the Access Points (APs) and Base Stations (BSs) are uniform across nodes from the same operator. The Wi-Fi connected devices and LAA connected devices are randomly dropped within the coverage area inside the building. User device association is based on the proximity to the access node. We consider the downlink transmission but this work could be extended to the uplink.

The transmission power for the access nodes are the same given as P_L for LTE-LAA and P_W for Wi-Fi. For the path loss model, the ITU Indoor hotspot (InH) model [10] is used. For the line of sight which is the case in our network model, the model (1) is applicable for 10 m $< d < d_{BP}$ [11] where d denotes the distance between the access node and the user device and d_{BP} is the breaking point distance. fc indicates the centre frequency of the channel.

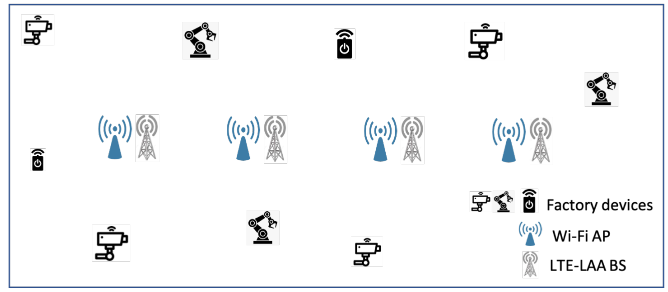

Fig. 1. The system model for an indoor factory environment

$$PL(d) = 28.0 + 22\log_{10}(d) + 20\log_{10}(f_c). \tag{1}$$

B. Wi-Fi MAC Model

The Wi-Fi system operates the Distributed Coordination Function (DCF) mechanism based on the IEEE 802.11 standard. In order to transmit over the unlicensed channel, the Wi-Fi device contends for the channel by selecting an integer randomly within a contention window (CW) which is known as the backoff (BO) time. The slot time is 9 μs. During the BO period if transmission is detected over the channel, the BO counter is paused until the channel is sensed to be idle again and then continue decreasing the BO counter. The node senses the channel for a set period known as the DCF Inter-Frame Space (DIFS) before transmitting, after which it begins the arbitration process of selecting a new BO number. The contention begins with the smallest CW size and increments exponentially whenever a collision is detected which is represented as $CW_s = (0 - 2^s \cdot CW_{\text{min}})$. CW_{min} is the minimum contention window size. The maximum contention window $CW_{\text{max}} = 2^s \cdot CW_{\text{min}}$ when $s = m$ where s represents the exponential backoff stage and m is the maximum backoff stage. The stationary transmission probability τ_w under saturated traffic conditions for a Wi-Fi device is given in [12] as

$$\tau_w = \frac{2(1-2p)}{(1-2p)(2^s CW_{\text{min}}+1) + p2^s CW_{\text{min}}(1-(2p))^s}, \tag{2}$$

where p is the independent collision probability. The channel idle probability P_q influenced by the DCF mechanism is a function of all nodes and devices not transmitting over the channel which can be written as

$$P_q = \prod_{n_w=1}^{N_w} (1 - \tau_{n_w}), \tag{3}$$

where N_w are the number of Wi-Fi nodes.

C. LTE-LAA MAC Model

For the downlink channel access procedures for LTE-LAA [13], on selection of a carrier for transmission of the physical downlink shared channel (PDSCH), the BS senses the channel to be idle for the defer duration T_d. Then a uniformly distributed random number n_{rd} is chosen from the CW size. If the channel is sensed idle the number is decremented by $n_{rd} - 1$ after each slot period T_{sl}. The BS may transmit over the channel when $n_{rd} = 0$. The time T_d and T_{sl} are set to 16 μs and 9 μs, respectively. The channel access procedures are governed by priority classes which define the min and max CW sizes and the channel occupancy time (COT). These parameters and their corresponding values are given in Table I. The CW adjustment is based on detecting 80% HARQ-ACK as NACKs in the reference subframe. The next higher CW size is chosen within the priority class when a NACK occurs. The maximum energy detection threshold $l_{th_{\max}}$ when coexisting with other systems such as Wi-Fi is governed by the equation

$$l_{th_{\max}} = \max \begin{cases} -72 + 10\log_{10}(\frac{BW}{20}) \\ T_{\max} \\ T_{\max} - T_A + (P_H + 10\log_{10}(\frac{BW}{20}) - P_{TX}) \end{cases} \quad (4)$$

where T_A could be 10 dB or 5 dB for transmission of PDSCH or discovery signal transmission respectively. P_H is 23 dBm and P_{TX} is maximum output power of the BS. BW is the channel bandwidth of the MHz unit and $T_{\max}$ is calculated as follows.

$$T_{\max} = 10\log_{10}(3.16228.10^{-8} \cdot BW). \quad (5)$$

TABLE I
CHANNEL ACCESS PRIORITY CLASS OF LTE-LAA [11]

Priority Class	$CW_{(min,p)}$	$CW_{(max,p)}$	$T_{(mcot,p)}$	Allowed CW_p sizes
1	3	7	2 ms	3, 7
2	7	15	3 ms	7, 15
3	15	63	8 or 10 ms	15, 31, 63
4	15	1023	8 or 10 ms	15, 31, 63, 127, 255, 511, 1023

D. Data Acquisition and Preparation

As aforementioned, LTE-LAA and Wi-Fi perform channel sensing to execute its opportunistic channel access protocols in order to operate in unlicensed bands. Since both networks can sense idle periods and transmission over the channel, makes analysing the periodic idle-time interval, a suitable parameter for predicting the number of devices serving in the factory environment. The randomly selected BO numbers are chosen using a uniform distribution as a collision avoidance mechanism. Each contender for the channel independently selects a number within a contention window, where differences in the selected backoff numbers represents a new dependent variable. Hence, every randomly selected BO number by each device contending using either the LTE-LAA channel access procedures or the 802.11 MAC protocol, represents an independent variable with a uniform distribution. The resulting numeric difference in the variables creates a new distribution. This new distribution can be obtained by counting the idle times between transmissions of all contending nodes. For the purpose of carrying out the number of device estimation, the time unit of the idle-time interval will be slots. Each slot time is 9 μs. The intervals for different device numbers were collected over simulations campaigns conducted with different node numbers. The number of nodes considered in the simulations were $4, 8, 10, 16, 20, 24, 30, 36, 40$.

$$Y = X_1 - X_2 - \cdots - X_n. \quad (6)$$

In (6), $X_1, \cdots, X_n$ are independent variables, randomly selecting values within a CW via a uniform distribution. The difference between all variables gives the new distribution Y.

The collated idle-time data from the simulation are represented in microseconds. However, these are converted to number of slots by the equation:

$$N_{sl} = \frac{T_{idle}}{T_{sl}}, \quad (7)$$

where N_{sl} represents the number of slots, T_{idle} is the measured idle-time. Obtaining the number of slots for every idle-time period provides the basic dataset for our predictive ML models. An average of about 16,000 idle-time measurements where collated for each simulation performed for each device number across the multiple randomised seed. These dataset were split into multiple samples which formed the features of the data. The features are the mean and standard deviation of the samples from the collated data. These features became the data points in training the ML model. A set of means and their standard deviations were labelled with the associated number of devices.

III. PROPOSED MACHINE LEARNING PREDICTION DESIGN

In this section, we explore three ML algorithms [14] in performing the number of factory device estimation or prediction. The selection of these algorithms were based on capability for multiclass classification. In the data prepared from the measured idle-time intervals, two data variables are chosen, which are the mean and standard deviations of subsets (features) of the overall dataset.

A. Data Organization for ML Models

To train the three ML models selected for predicting the number of factory devices actively contending over the channel, the data structure adopted was a $[m \text{ x } n]$ matrix where n being 3 columns with input data $P \in [m \text{ x } n-1]$ and output vector $Q \in m$. A consists of the mean and standard deviation which are the features of the samples from the full dataset. B comprises the labels for each pair of A entries in the matrix. The training dataset P_{tr}, Q_{tr} and test dataset P_{test}, Q_{test} were used to train and evaluate the ML models respectively.

B. Multilinear Regression

Multiple linear regression is a prediction model which takes multiple independent variables and make predictions on a dependent variable. The mean and standard deviation of the idle-time intervals constitutes the independent variables to predict the node numbers. The independent variable are training dataset P_{tr} and Q_{tr} while m_1 and m_2 are the estimated regression coefficients as shown in (8). Q_{tr} represents the dependent variable through which the prediction is made. The coefficients in the multilinear equation represent the rate of change of Q_{tr} with respect to dependent variables m_1 and m_2. The challenge with multilinear regression is obtaining the best fit to the data points. This will consequently affect the accuracy of the model. Each label (number of devices) represents a class for which an output prediction is made.

$$Q_{tr} = b + m_1(P_{tr} \in [m \times n_{j-2}]) + m_2(P_{tr} \in [m \times n_{j-1}]) \quad (8)$$

where j is the number of columns in the matrix.

C. k-Nearest Neighbour (k-NN)

k-NN is a type of supervised learning ML technique. Our k-NN model is trained on labelled data representing the number of devices. The concept is to provide predictions on unlabelled data based on the proximity of the unlabelled input data to the labelled training data provided to the model. k-NN is known to give good accuracy because no assumptions are made. In the case of the number of device estimation proposed in this paper, the k-NN algorithm performs classification by calculating the distance between the input mean and standard deviation features to the ones used to train. The distance calculation depends on the type of data being trained. Euclidean and Manhattan distance are used for continuous data while Hamming distance are used for categorical data. In this work, euclidean distance was used for distance calculations as the data processed was continuous data. The prediction is made based on the class with the highest proximity to the input data when compared to the distance calculated. The k represents the 'number of neighbours' from the trained data to the input data to be considered in making the classification. Selecting the optimal 'k' number is crucial to the accuracy of the model. A very low k value could cause overfitting to the model and negatively influence the prediction, while too high k value can lead to underfitting and higher computational cost in calculating the distance for all the points. The Euclidean distance for the model is determined by the equation below

$$d(P_{tr}, Q_{tr}) = \sqrt{\sum_{i=1}^{n} (p_{tx} - q_{test})^2} \quad (9)$$

D. Random Forest (RF)

RF, also a supervised learning algorithm, works by building an ensemble of decision trees, with the aim to improve prediction accuracy. Decision trees make up the component parts in the RF. The decision trees in the model we trained essentially contains branches which provides a possible decision or occurrence based on the distinctions in the data features presented to the model. Our dataset exhibits higher entropy (level of randomness) based on the number of classes trained for prediction. Hence, the decision tress splits the data into smaller samples to reduce the entropy. This enables better decision making by the decision nodes and root nodes. The efficiency of the split dataset is influenced by the conditions in making the split. The goal will be to reduce the entropy to zero as much as possible. In the model we trained, the classifier had a minimum sample of two with no limit to the maximum samples. The classifier in the RF combines these decision trees defined by the specific class conditions built into the branches to produce a class prediction. RF has a lesser training time when compared to other ML models and the risk of overfitting is significantly reduced due to the use of multiple trees. The low correlation between each decision trees actually produces more accurate predictions.

TABLE II
NETWORK PARAMETERS USED IN THE SIMULATION

Parameter	**LTE-LAA**	**Wi-Fi (802.11)**
Slot time	9 μs	9 μs
SIFS	N/A	16 μs
Defer Time/DIFS	43 μs	34 μs
Tx Power BSs/APs	18 dBm	18 dBm
Bandwidth	20 MHz	20 MHz
Total Data SB	1200	56
Min & Max CW	15 & 1023	15 & 1023

IV. RESULT DISCUSSION

In this section, we present our findings based on the analysis performed on the channel idle-time interval dataset acquired for LTE-LAA and Wi-Fi from the simulation campaign on ns3. Some network parameters used in the simulation is given in Table II. As aforementioned, three multiclass classification methods, Multilinear Regression, k-NN, and Random Forest, are evaluated. The models where trained, tested and evaluated using python libraries Pandas, Numpy and Sklearn. The dataset is split with a 70% and 30% for model training and testing respectively. We evaluate the performance of the models for different level of granularity of device numbers. The first models was trained on a lower granularity (LG) dataset of 10, 20, 30, 40 devices. The second models was trained with higher granularity (HG) of 4, 8, 10, 16, 20, 24, 30, 36, and 40 devices. We compare the performance of both models based on granularity and obtain the result shown in Table III.

We first embark on gaining insight to the level of association between the features of the dataset and thereby appreciate their relevance to the intended task of estimating the number of devices. In Table III, the correlation coefficients shown, reveal a high negative correlation with the number of devices. This is important to understand the relevance of the chosen parameters to estimating the number of devices. As expected, the dataset with LG show a higher linear correlation between the mean, standard deviation and the number of devices, when compared

TABLE III
CORRELATION OUTPUT FOR DATASET FEATURES

Number of Device Category	Mean	Standard Deviation	Number of Device
10-20-30-40 (LG)	-0.92765	-0.95028	1.0000
4-8-10-16-20-24-30-36-40 (HG)	-0.70462	-0.74271	1.0000

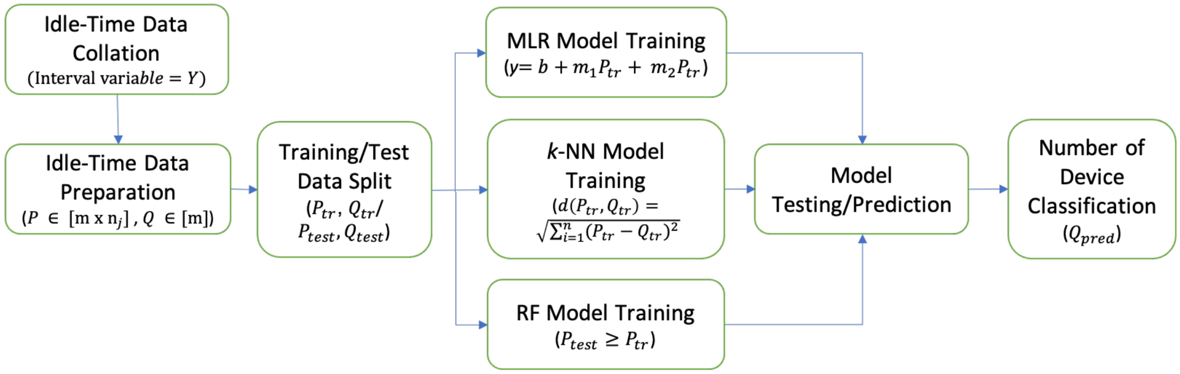

Fig. 2. The operational procedures of proposed ML-based node number estimation

to the ones with HG. However, the dataset with HG similarly shows good linear correlation to perform number of device estimations.

Different evaluation metrics are applied to ML models. However, for multiclass classification problems the F1 score is mostly used, particularly in k-NN and RF models. The F1 score is mainly good because it offers a result based on the harmonic mean of the precision and recall metrics. The F1 score are presented per class, i.e. per number of devices for both LG and HG. However, we adopt the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) for Multilinear Regression. The MAE and RMSE are more widely used for evaluating linear regression predictions.

A. Multilinear Regression Results

Table IV shows the MAE and RMSE evaluation results for the Multilinear regression model. The results for LG shows a good prediction performance with a low MAE and RMSE. This is expected as the separation between the features for each number of devices is sufficient to provide accurate predictions. The results for HG also reveals relatively good performance with respect to an increase in specificity of number of actively contending devices over the channel. Considering the HG classifies more than twice the size of number of devices, when compared to the LG; the MAE and RMSE for HG category, reveals comparable performance to the LG. Also, these results reveal with multiple random placement of devices, good number of device estimation can be attained using the Multilinear regression algorithm.

TABLE IV
PERFORMANCE RESULTS FOR MULTILINEAR REGRESSION

Device Number Category	MAE	RMSE
10-20-30-40 (LG)	2.9275	3.5658
4-8-10-16-20-24-30-36-40 (HG)	6.4999	7.8399

B. k-Nearest Neighbour Results

Its already mentioned above the k-NN model performs classification based on the proximity of the input data to the trained data. The 'k' parameter is crucial to the performance

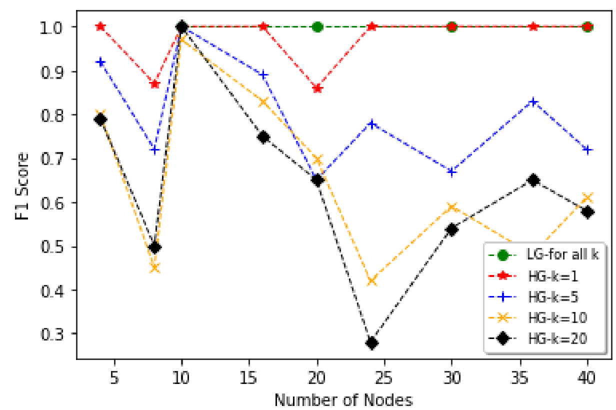

Fig. 3. k-NN performance based on different k values

of the model because it determines how many data points to consider in the proximity of the input data to make a classification output. The error rate for the value of k is a good way of finding the optimal k value. The k value with the lowest error becomes the optimal k. In Fig. 3, the F1 score of the different number of devices are plotted for both LG and HG category with respect to their k value. The classification score for LG shows the highest classification performance. This is largely due to the dataset features being sufficiently spaced to allow such high accuracy classifications. In the HG category, it is observed the curve of F1 score varies across the different k values and number of devices. The lower performance in classification with higher k values are due to higher error rates. It is therefore clear that best performance is achieved when k=1. This show good classification is achievable using the k-NN models.

C. Random Forest Results

In the case of the RF model, the decision tree ensemble plays the critical role in the final classification output. The RF model shows best performance based on the F1 score for different number of devices in Fig. 4. The number of estimators is a hyperparameter for RF models in making classification outputs. Again, the RF model gives excellent

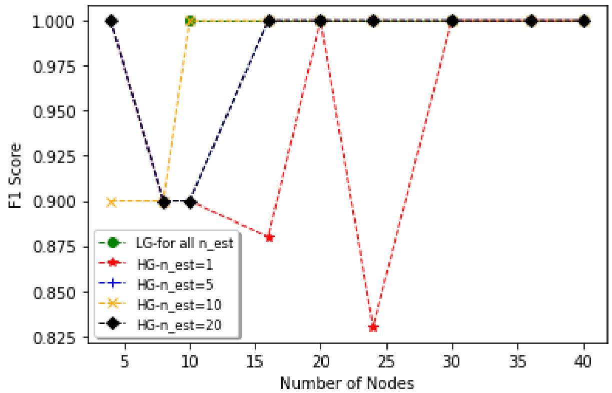

Fig. 4. Random Forest performance for different n estimators

performance with the LG category with a consistent F1 score for all classes and all chosen estimators used in training the model. We evaluate the performance for 3, 5, 10 and 20 estimators. The minimum F1 score measured from all classes with 3 estimators came at 0.83. It can be observed that from 5 estimators, a consistent F1 score is measured. It will then be adequate to use 5 estimators for the classification problem because using any higher will increase the computational cost in performing the classification tasks. From these results for HG category, we clearly see the superiority of the RF algorithm in making accurate classifications above k-NN and Multilinear regression.

V. Conclusions

In this paper, we embarked on developing ML algorithms to predict and classify the number of devices operating in unlicensed bands. We used a novel approach, where the statistical distribution of idle-time periods over a channel is collated and analysed to enable the model training and testing. In performance evaluation of the three ML models which are multilinear regression, k-NN, and RF, the RF based algorithm outperforms the other two algorithms. It could be analysed that RF provides best predictions in complicated scenarios while k-NN and Multilinear regression provide benefits in computational cost. As a preliminary work for efficient resource sharing, this work focused on the performance analysis of different ML mechanisms to estimate the number of devices operating in unlicensed bands. The findings could be utilized in our future work, the design of the contention-based MAC protocol operating in unlicensed bands.

Acknowledgment

This work was partly supported by the European Union's Horizon 2020 research and innovation programme in project 5G-HEART under grant agreement No. 857034 and project DEDICAT 6G project under Grant Agreement No. 101016499. We also would like to acknowledge the support of the 5GIC/6GIC members for this work.

References

[1] G. Brown, "Private 5G Mobile Networks for Industrial IoT," *White paper produced for Qualcomm*, July 2019.

[2] O. Baiyekusi, S. Vahid, and K. Moessner, "A Time-Based Fairness Approach for Coexisting 5G Networks in Unlicensed Bands," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017, pp. 1–5.

[3] A. Mukherjee, J.-F. Cheng, S. Falahati, H. Koorapaty, D. H. Kang, R. Karaki, L. Falconetti, and D. Larsson, "Licensed-Assisted Access LTE: coexistence with IEEE 802.11 and the evolution toward 5G," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 50–57, June 2016.

[4] M. Mehrnoush, S. Roy, V. Sathya, and M. Ghosh, "On the Fairness of Wi-Fi and LTE-LAA Coexistence," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 735–748, Dec. 2018.

[5] H. Ko, J. Lee, and S. Pack, "A Fair Listen-Before-Talk Algorithm for Coexistence of LTE-U and WLAN," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 116–10 120, Feb. 2016.

[6] G. Bianchi and I. Tinnirello, "Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, Apr. 2003.

[7] T. Vercauteren, A. L. Toledo, and X. Wang, "Batch and Sequential Bayesian Estimators of the Number of Active Terminals in an IEEE 802.11 Network," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 437–450, Feb. 2007.

[8] S. Chun, D. Xianhua, L. Pingyuan, and Z. Han, "Adaptive Access Mechanism with Optimal Contention Window Based on Node Number Estimation Using Multiple Thresholds," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2046–2055, Apr. 2012.

[9] 3GPP TR 36.889, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on Licensed-Assisted Access to Unlicensed Spectrum (Rel. 13)," Tech. Rep., v.13.0.0, 2015.

[10] ITU-R M.2135, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," Tech. Rep., June v.1, 2009.

[11] 3GPP TR 36.814, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Further advancements for E-UTRA physical layer aspects (Release 9)," Tech. Rep., v9.2.0, 2017.

[12] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[13] 3GPP TR 36.213, "3rd Generation Partnership Project (3GPP), "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Rel. 14)," Tech. Rep., v.14.2.0, 2019.

[14] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions," *IEEE Access*, vol. 7, pp. 137 184–137 206, 2019.