Journal of Engineering and Applied Sciences 11 (12): 2729-2739, 2016
ISSN: 1816-949X

Evaluation of Efficiency Techniques for Spectrum Decision Making in Cognitive Radio Wireless Networks

Angela Patricia Garcia, Danilo Lopez Sarmiento and Edwin Trujillo
Faculty of Engineering, Distrital Francisco Jose de Caldas University, Bogota, Colombia

Abstract: The level of success of spectral decision in cognitive radio networks depends on how good the prediction model of channel usage is in both licensed and unlicensed bands. In this research, three different techniques are explored to predict channel state from the point of view of the use given by the primary user and optimization of the best technique is proposed to achieve a closer prediction to reality from the use of actual data of a WiFi network as validation method. The article concludes that our algorithm is optimal for short history windows.

Key words: Cognitive radio, spectrum, primary user, secondary user, correlation, linear regression, autocorrelation

INTRODUCTION

Cognitive Radio (CR) may be defined as an intelligent wireless communication system that aims to provide highly reliable data transmission through the efficient use of the spectrum; this is possible by having awareness of the surrounding environment, learning from it and adapting to statistical variations in the input stimulus (Haykin, 2005). Three stages could be defined with different functions in a cognitive cycle (Fig. 1):

- Detection: Search unused spaces of spectrum
- Analysis: Identification of specific characteristics of the unused spaces found in the detection stage
- Decision: Selecting the frequency band at a given time and specific location that achieves meet the requirements of Quality of Service (QoS) of the user, without causing interference to other users

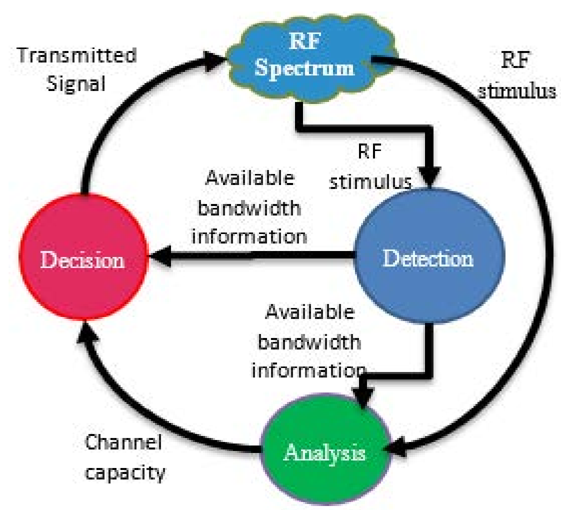

Fig. 1: Cognitive cycle (Lopen and Sanchez, 2015)

Once defined the frequency band to be used, communication can start; it tries to maintain communication and quality of service provided, even if it is necessary to change the frequency band due to some novelty as increased traffic, arrival of other licensed users in the band, etc.; this process refers to the ability of mobility. Fair access to the shared medium is the ability to share and cognitive cycle with mobility capacity and sharing are intelligent radio, cognitive radio (Akyildiz *et al*., 2006).

Decision stage: In CR two types of users are distinguished: the Primary User (UP) who has the license for the use of particular frequency band and the Secondary User (US) which has access and (opportunistic) possibility of using this frequency band without causing interference to UPs, so it looks to US use the band at the moment that the UP not be using it (Zhao and Swami, 2007).

Three tasks in order to ensure that the decision taken for the secondary user is the most appropriate are performed: The characterization allows to identify the properties of each frequency band based on the historical performance of the primary user and the current

Corresponding Author: Angela Patricia Garcia, Faculty of Engineering, Distrital Francisco Jose de Caldas University, Bogota, Colombia

conditions of the spectral band. Once the characterization process is complete, it was proceed to the selection of the most appropriate band according to the requirements of quality of service which cannot always be met. Finally, the reconfiguration of the transmission parameters to achieve a communication in the selected spectral band (Masonta *et al*., 2013).

MATERIALS AND METHODS

The spectrum is a limited resource, for this reason CR emerged to use it efficiently. This study aims to evaluate (with real data) three predictive techniques for decision making in the spectrum of CR wireless networks (WiFi), the techniques to be evaluated do not correspond to a mathematical model but an algorithm that seeks to reproduce the behavior channel to decide whether or not to transmit in the channel according to the predictions obtained, also it seeks to optimize the results obtained making changes to the algorithm with the best performance.

The methodology used throughout the research is shown in Fig. 2. To conduct the study it was necessary to use WiFi Analyzer software that would identify Access Points (AP) and WiFi channels, plus allowing to display the number of clients connected to each of the APs; the software used for the first phase of the study is Acrylic WiFi Professional of Tarlogic (free software, student licensed) that offers the monitoring mode, i.e., for viewing, capturing data and flowing WiFi traffic, regardless of whether it is directed to the device. Acrylic WiFi Professional allowed to export to a CSV file (Comma-Separated Values) data manually taken every 5 min, subsequently to join in a consolidated covering 12 h-145 measures.

Data fitting to statistical models attempts to obtain an accurate prediction for a future time; at this stage the data collected were taken and WiFi channel behavior was described in terms of busy (1) and idle (0); then statistical techniques of correlation, linear regression and autocorrelation are included; these algorithms were proposed by Uyanik. The implementation phase in algorithms software was possible using MATLAB, the flowchart for each of the algorithms was generated and was taken to a code. Each one of the algorithms was applied to different time intervals to verify the effectiveness of each of them. The last phase of the study was the analysis and validation of the results was obtained. The comparison of the prediction obtained was performed with the channel state corresponding to the time for which the prediction was made and concluded which technique offers the most accurate prediction.

Data adquisition and processing: The 2.4 GHz band offers a Bandwidth (BW) of 72 MHz, in Colombia the US standard is used, having 11 channels available each with a BW of 5 MHz; for data transmission in wireless networks often uses bandwidths of 20 and 40 MHz as 5MHz is insufficient. If each AP that transmits occupies 20 MHz, then the only way to avoid interference between channels is at a distance of at least 5 channels, that is why channels that are used most often for WiFi networks are 1, 6 and 11. The database that was used throughout the study was created with a collection of measurements generated with Acrylic WiFi Software. It was found that the maximum number of users in the interval from 9-0 to 21-720 min is 11 and the minimum number of users found in the interval is 4 (Fig. 3).

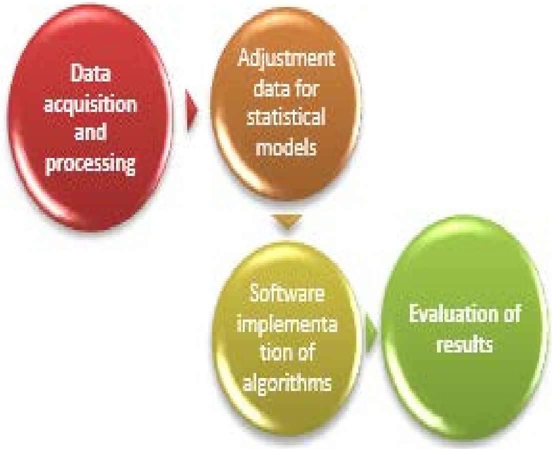

Fig. 2: Phases the research included

Channel 6 as it is default is the most popular, this fact is reaffirmed to find that channel 6 hosts most of the users, being always in the busy state; channels such as 2, 3, 5, 8 and 10 remain occupied <10% of the sampling time, this may be due to its proximity to the most used channels (1, 6 and 11).

Because each channel has a particular behavior, all channels and the number of users were taken to determine whether the channel was occupied or not at a given time (Fig. 4) allowing to identify the channels that have a favorable behavior for a secondary user, undoubtedly the channels that best fit would be those that remain unemployed the most part of the time, with channel 3 being ideal assuming a bandwidth of 20 MHz.

Data adjustment to statical models and implementation in algorithms software: To choose the appropriate spectral band, secondary users monitor the spectrum and obtain predictions based on history window (historical behavior of the spectrum). The number of time intervals for which the prediction is obtained is called prediction window and

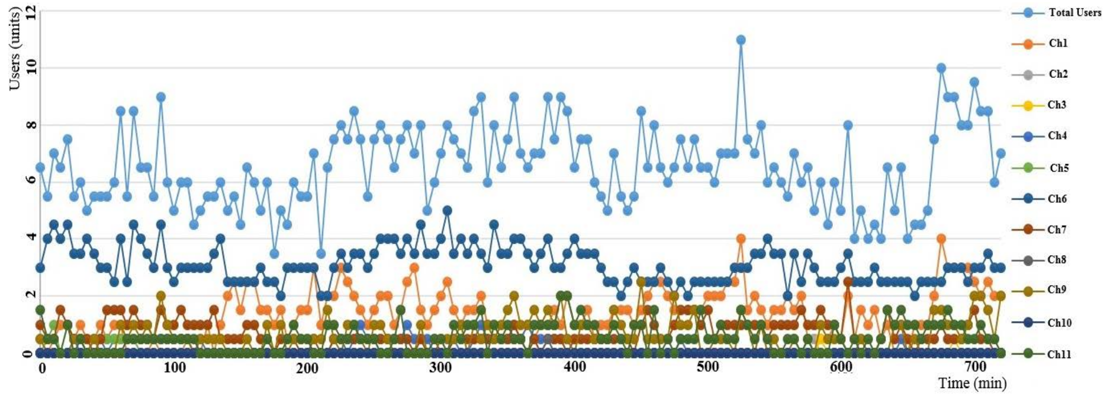

Fig. 3: Number of users

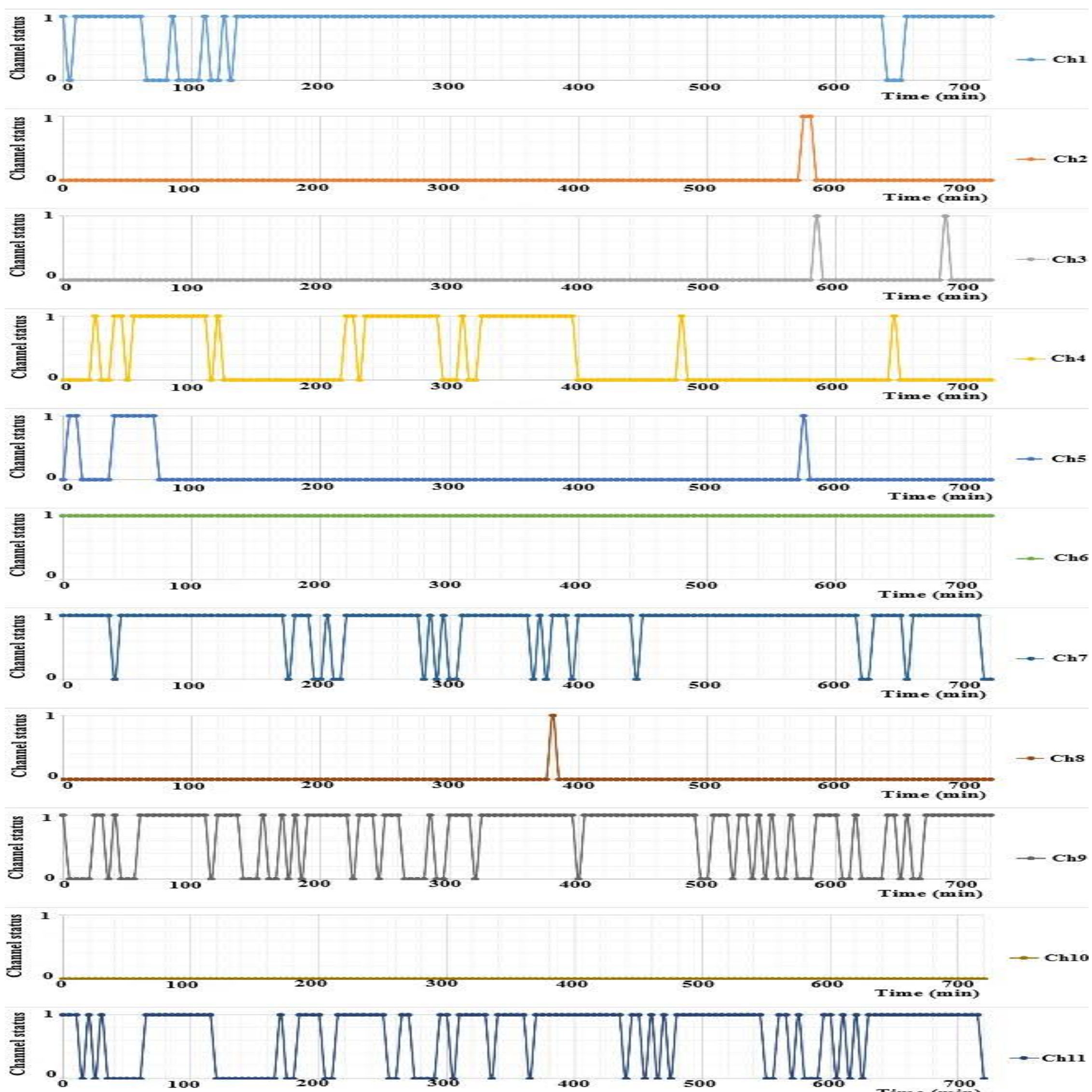

Fig. 4: Channel occupancy level as a function of time

Table 1: Channel occupation 12 h interval

Channels	Busy state	Idle state
1	90	10
2	1	99
3	1	99
4	33	67
5	7	93
6	100	0
7	87	13
8	1	99
9	69	31
10	0	100
11	69	31

may consist of one or more time intervals. Three algorithms proposed by Uyanik were used such algorithms represent the behavior of an independent channel of the other three channels and apply statistical techniques to generate a model to predict the behavior of the spectrum. The following variables are defined:

WH: History Window (Length vector i)
H_i: State of the spectrum captured in the history window at position i
[WH]: History window size
X: Vector index ($X = [1, 2,..., |W_H|]$)
WP: Prediction Window
δ_c: Correlation threshold

History Window W_H: Because, the algorithms are applied to one channel at a time, the channels history data were used with which channel occupancy percentages were obtained during the 12 h interval (Table 1) and it was found that data channel 4 would be used due to its variable behavior with wide idle time and busy intervals. The data processing of a chosen channel was conducted in Matlab, applying each algorithm to the history window, that is n = |WH| and interval corresponds to the number of predictions that want to be generated.

Threshold value δc: To set the threshold value used for all algorithms it is taken into account that the Pearson's correlation coefficient should be not <0.1 as a lower coefficient would indicate no correlation and should not exceed 0.5 as these values indicate that there is an absolute correlation and all 3 algorithms would have the same result as with real data there would not be sufficient correlation. Tests are performed simulating WP = 0.1 and WP = 0.3, finding that with $\delta_c = 0.1$ predictions are more adjusted to the actual data, this is because the signal is not periodic.

Algorithm_1; correlation:

majResult→Majority(W_H)
if | corrCoeff(X, W_H) |>δ_c then
W_P← [$H_{|WH|}$ ... $H_{|WH|}$]
else
W_P←[majResult ... majResult]
end if
return W_P

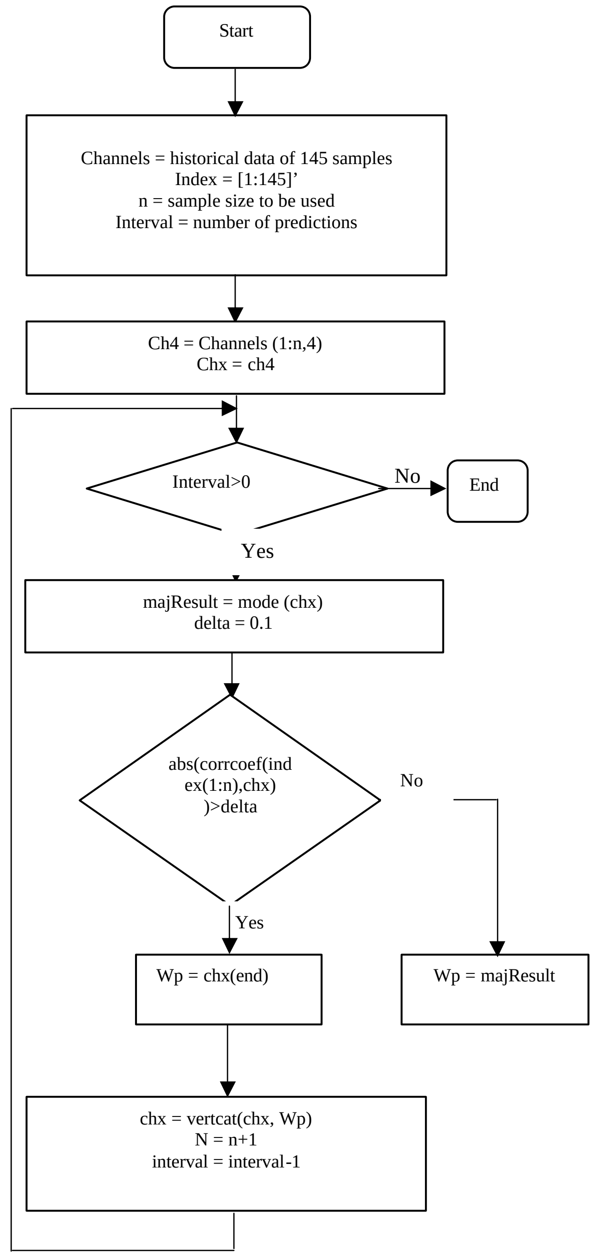

Fig. 5: Flowchart algorithm_1

For this algorithm the Pearson correlation coefficient is measured and compared with the threshold value ($\delta_c = 0.1$), if the coefficient value exceeds the threshold then the prediction window will settle with the last sample of the prediction window. If there is not enough correlation, then the window prediction is filled with the result of the majority. Below is the flowchart for the implemented code (Fig. 5). Data from the occupation of the eleven channels were introduced to Matlab as matrix (Channels) plus the index vector (index) was defined, the sample size which is the size of the history window (n) and the number predictions to be obtained (intervel). Once these data are obtained, the sample is drawn in a new vector (CHX) and is given to start a while loop that will be repeated until completing the total number of predictions defined above by adding a position at the end of the CHX vector. In the while loop, the process to find the channel state following the logic explained above is found.

Algorithm_2; correlation and linear regression:

majResult←Majority(W_H)
if | corrCoeff(X, W_H) | > δ_c then
W_P ←Points calculated by linear regression with binary values.
else
W_P ← [majResult ... majResult]
end if
return W_P

The Pearson correlation coefficient quantifies the degree of relationship between two variables and can be used to generate a prediction algorithm with linear regression.

The Pearson correlation coefficient is measured and compared with the threshold value (δ_c = 0.1), if the coefficient value exceeds the threshold, then the prediction window will settle with points calculated using linear regression (taking as reference 0.5, lower values would approximate to 0 and greater values to 1). If not sufficient correlation is found, then the prediction window is filled with the result of the majority.

The flowchart for the implemented code (Fig. 6) shows that as in the above algorithm, the data is taken and the variable “index”, “n”, “intervel” and “chx” are defined to start the “while” loop under the logic of the algorithm previously explained.

Algorithm _3. Autocorrelation:

majResult←Majority (W_H)
Calculating autocorrelation coefficients with a maximum delay |W_H|/2
maxCoeff ? max(coefficients)
if | corrCoeff(X, W_H) | > δ_c then
Periodicity ? delay de maxCoeff
W_P←Points calculated with the periodicity of the signal.
else
W_P← [majResult ... majResult]
end if
return W_P

This algorithm calculates the coefficients of the autocorrelation with a maximum delay of half the magnitude of the history window, from these coefficients the highest one is chosen and the delay corresponding to highest coefficient defines the periodicity of the signal.

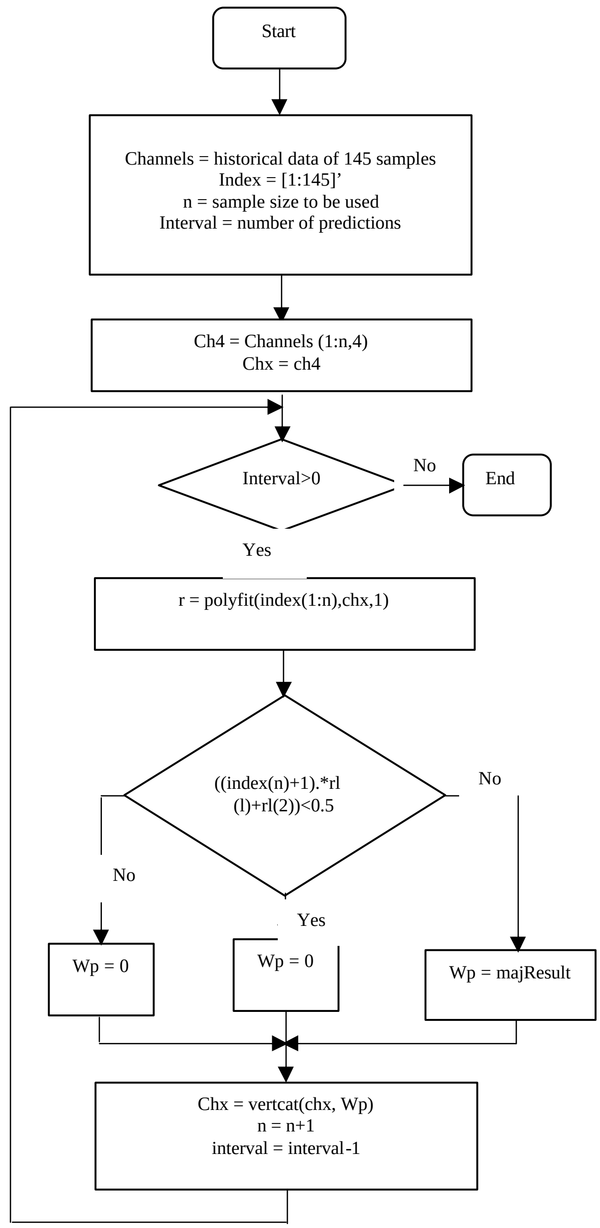

Fig. 6: Flowchart algorithm_6

Whenever, the Pearson correlation coefficient is greater than the threshold δ_c = 0.1), the result of the prediction is calculated according to the periodicity; otherwise, the prediction window settles for the last sample of the prediction window. Below is the flowchart for the implemented code (Fig. 7) as in the above algorithms, the data is taken and the variable “index”, “n, “interval” and “chx” are defined to start the “while” loop under the logic of the algorithm previously explained.

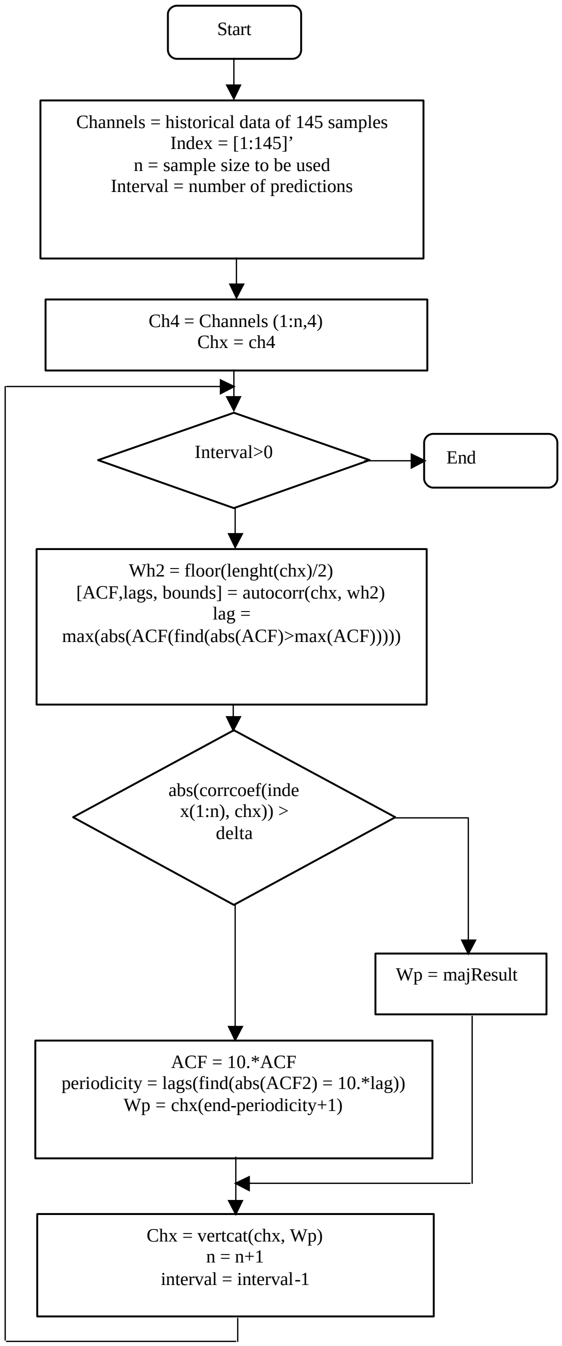

Fig. 7: Flowchart algorithm_3

RESULTS AND DISCUSSION

The analysis focused on the use of channel 4 regardless of other channels and the relationship between them since, the proposed algorithms define it so, it is noteworthy that the APs mostly have a fixed channel assigned and the data obtained with Acrylic WiFi Professional indicate that there was never >1 user in channel 4, unlike channel 6 which was always busy and has 3 users on average.

Below are the results of three algorithms (correlation, correlation and linear regression, autocorrelation) for 5 different time intervals: 41, 61, 81, 101 and 121 samples; graphs show 2 series, the reddish color with points demarcated with "+" are the actual data and blue with points demarcated with "o" are the interval taken and predictions in the last places.

Interval of 41 samples: The simulation is performed to predict 20 intervals and the same prediction is obtained with the 3 algorithms (Fig. 8), this is because in all cycles, the Pearson coefficient found is greater than the threshold of 0.1, this in the case of algorithm_1, means that will replicate the last sample in all predictions; in the case of algorithm_2 the line obtained always has a negative slope, making the prediction 0 for all time slots; in the case of algorithm_3, it tries to find periodicity in the data signal and since the signal is not periodic the periodicity is minimal and brings the data near the end of the chain having 0 for all predictions.

Interval of 61 samples: Different results for each algorithm are found; Algorithms 1 and 2 show an opposite behavior; by increasing the number of predictions for algorithm_1, its accuracy decreases (Fig. 9), in this case the value of the Pearson coefficient does not always exceed the threshold and since the number of samples is small, the result of the majority matches the one of the last sample, always getting "0" for all predictions.

Algorithm_2 manages to increase accuracy with the number of predictions (Fig. 10), the Pearson coefficient exceeds the threshold for all predictions and when evaluated with the equation it results always in a straight line with a positive slope, obtaining "1" for all predictions. Algorithm_3 is inaccurate for any number of predictions since the signal is not periodic (Fig. 11).

Interval of 81 samples: Different results for each algorithm are found; algorithm_1 and algorithm _3 present the same result, the accuracy decreases with increasing number of predictions (Fig. 12).

Algorithm_2 is not reliable for this interval, since the predictions obtained are not correct, only one from 20 possible predictions is successful (Fig. 13).

Interval of 101 samples: The same result for the three algorithms, matching the prediction with the data actually obtained (Fig. 14).

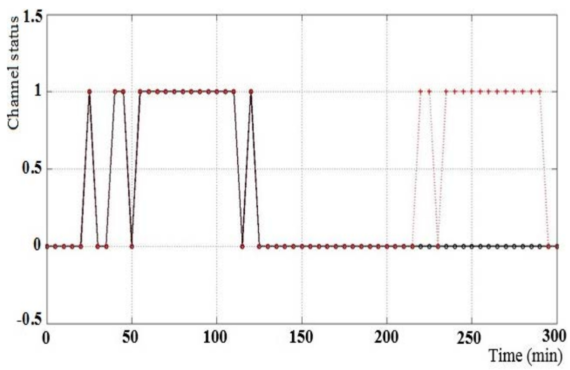

Fig. 8: Results for the interval of 121 samples

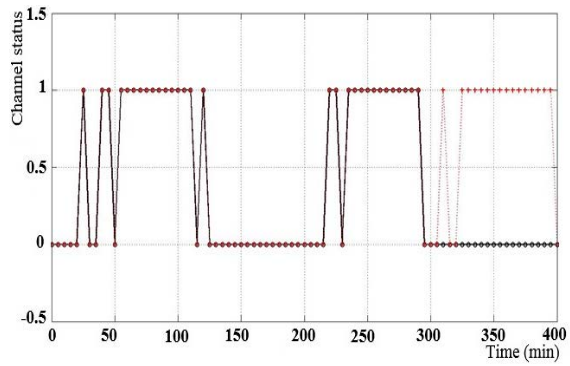

Fig. 9: Results for the interval of 61 samples using algorithm_1

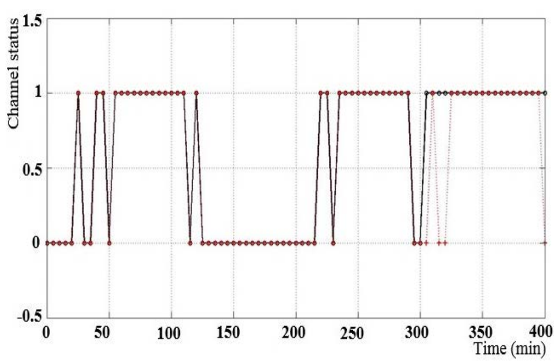

Fig. 10: Results for the interval of 61 samples using algorithm _2

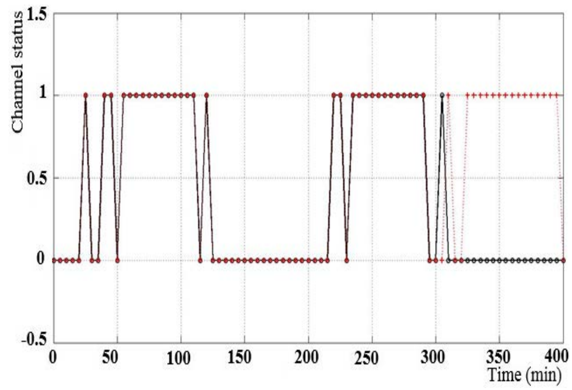

Fig. 11: Results for the interval of 61 samples using algorithm _3

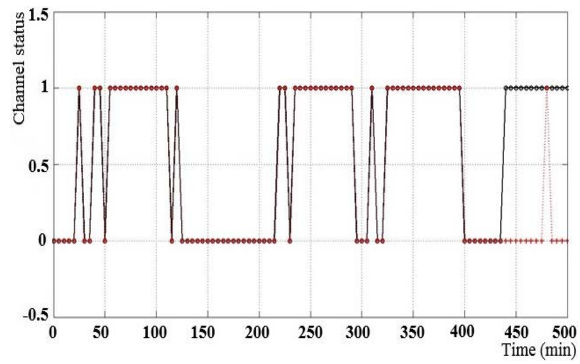

Fig. 12: Results for the interval of 81 samples using algorithm_1/algorithm_3

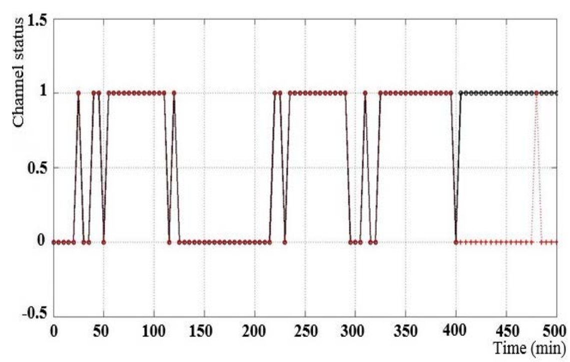

Fig. 13: Results for the interval of 81 samples using algorithm_2

Interval of 121 samples: The same result for the three algorithms, matching the prediction with the data obtained in reality at all points except one (Fig. 15). Precision for prediction windows is tabulated below (Table 2).

Discussion and optimization of channel state prediction: Since, the results for the first intervals are unsatisfactory, it is necessary to verify that the channel was the right choice, so then it is decided to explore the channel decision in further. To obtain the candidates for the channel in which the secondary user will transmit intervals with the same number of samples as those used for evaluation of the algorithms were taken and

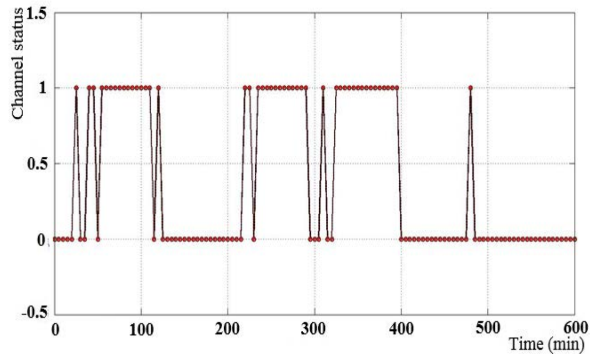

Fig. 14: Results for the interval of 101 samples

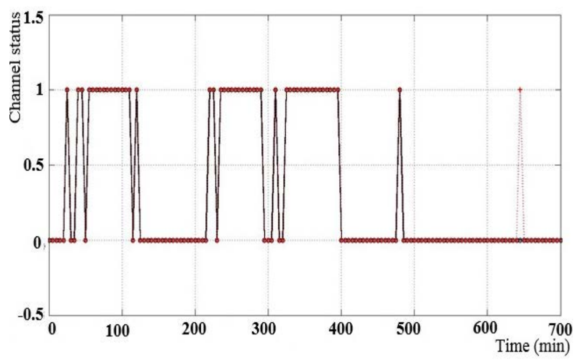

Fig. 15: Results for the interval of 121 samples

Table 2: Precision for prediction windows

		Predictions		
Samples	Variables	3 (%)	10 (%)	20 (%)
41	Alg 1	100	40	30
	Alg 2	100	40	30
	Alg 3	100	40	30
61	Alg 1	66.67	30	20
	Alg 2	33.33	70	80
	Alg 3	33.33	20	15
81	Alg 1	100	70	40
	Alg 2	0.00	0	5
	Alg 3	100	70	40
	Alg 1	100	100	100
101	Alg 2	100	100	100
	Alg 3	100	100	100
121	Alg 1	100	90	95
	Alg 2	100	90	95
	Alg 3	100	90	95

considering that the channels that are most interval occupied would not be a candidate suitable for a US to transmit, such channels are excluded; assuming a bandwidth of 20 MHz for transmission and considering that it is intended not to cause any interference to UP is also necessary to disqualify the immediate neighbors of these channels. The data obtained are shown in Fig. 16 which shows in dark blue color the channel which is most

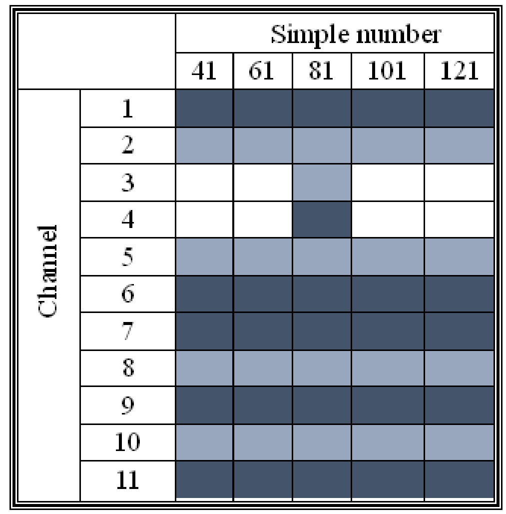

Fig. 16: Channel eligibility

of the time busy, light blue its immediate neighbors and blank the channels that up until to that sample were idle most of the time.

These data clearly show that it was not wise to transmit in the interval of 81 samples, in which the worst performances were obtained when algorithm_2 only matches to a position of 20.

Since the accuracy of the algorithms for shorter intervals (41 and 61 samples) is not very good, it was sought to modify the algorithms to achieve greater precision, so algorithm_1 was taken to be the one that offers greater efficiency and is modified by combining it with algorithm_2 (Fig. 17) which is the one more offering greater precision with increasing number of predictions in the interval of 61 samples.

Below are the results obtained for the interval of 61 samples (Fig. 18) which presents an improvement to algorithm_1 with an accuracy of 66.67% for 3 predictions, 80% for 10 predictions and 85% for 20 predictions, these results are undoubtedly better than those obtained with other algorithms; for the remaining intervals the same result was obtained.

Performance evaluation with the best algorithms in terms of performance: Tests were also conducted with data obtained with the energy detection method using a spectrum analyzer; the data sequence comprising 1000 samples spaced 290 msec equivalent to an interval of 4.8333 min (Fig. 19); it should be noted that measurements were taken during a week in December 2015 and correspond to a research by the University Francisco Jose de Caldas.

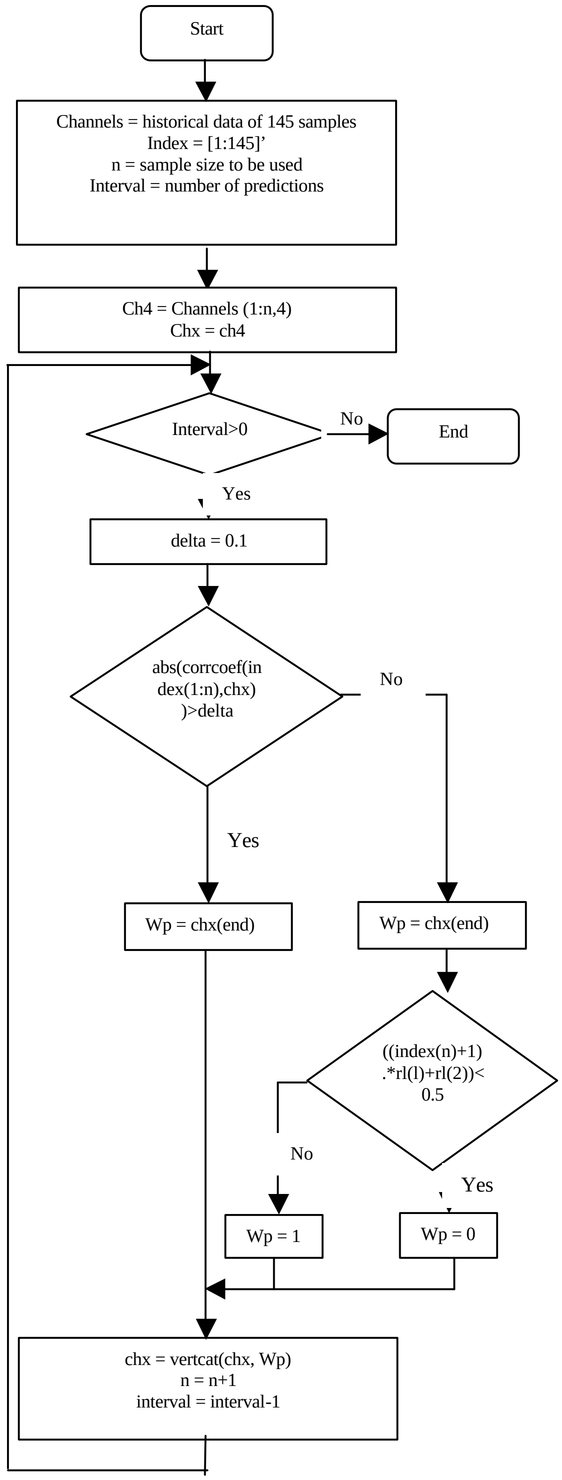

Fig. 17: Flowchart for modified algorithm_1

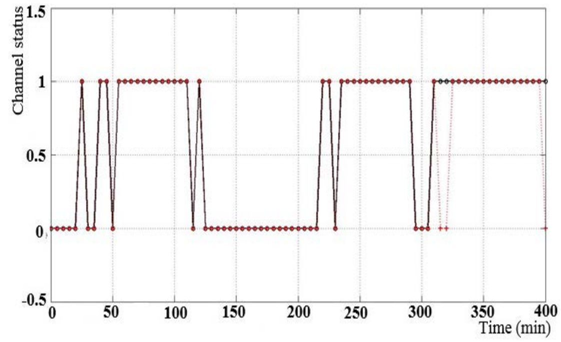

Fig. 18: Results for the interval of 61 samples

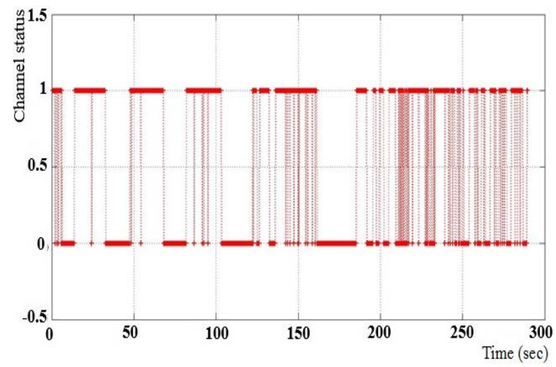

Fig. 19: Sequence of 1000 samples, 290 sec

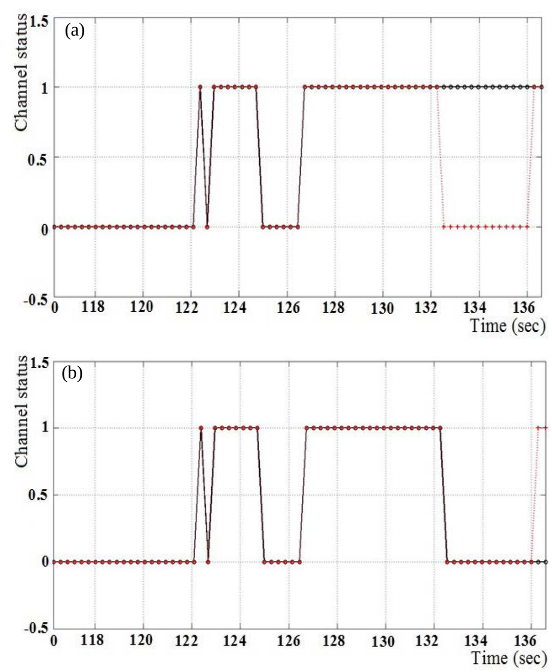

Fig. 20: a, b) Results for the interval of 451 samples using algorithms: 1 (Top) and 1-modified (Bottom)

This test is performed in order to check the operation of the algorithms with better performance (algorithm_1 and algorithm_1-modified) (Fig. 20) using a more robust

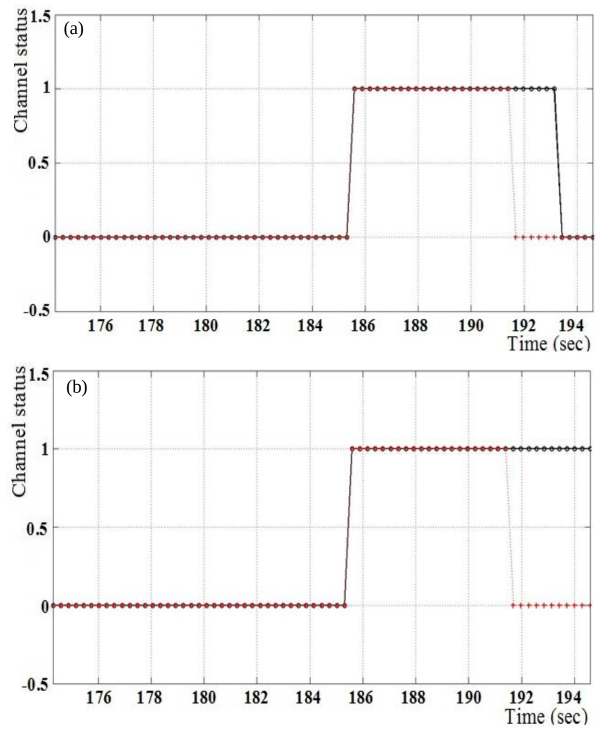

Fig. 21: a, b) Results for the interval of 651 samples using algorithms: 1 (Top) and 1-modified (Bottom)

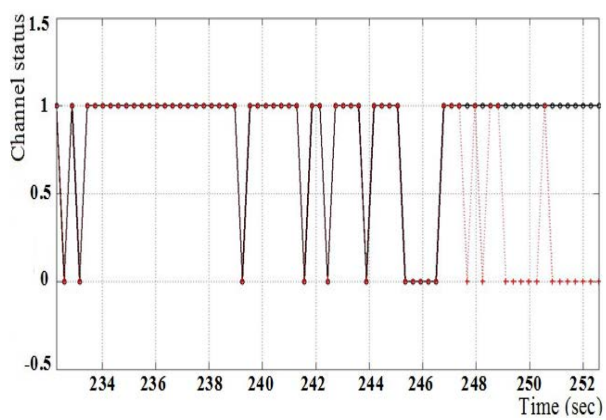

Fig. 22: Results for the interval of 851 samples using algorithms: 1 and 1-modified

database and taking historical of 451, 651, 851, 951 and 979 samples. Below there are graphics with the results obtained (Fig. 20-24) using algorithms 1 and 1 modified, it can be noted that as the history window grows, the accuracy of prediction obtained with algorithms decreases to the point that when the size of history window is too large the steps that follow the algorithm are the same for each iteration as the Pearson correlation coefficient never exceeds the threshold value ($\delta_c = 0.1$). It is also noteworthy that as in previous tests the best performance corresponds to algorithm_1 modified, presenting a different behavior from algorithm_1 for history windows with a fewer number of samples. Precision for windows prediction is tabulated in Table 3.

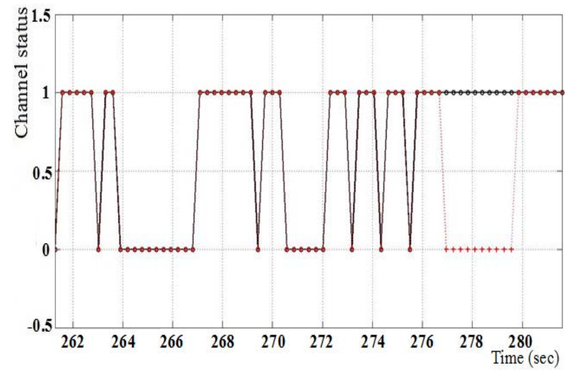

Fig. 23: Results for the interval of 951 samples using algorithms: 1 and 1-modified

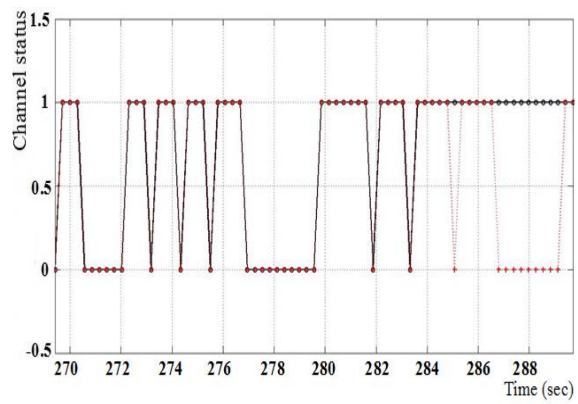

Fig. 24: Results for the interval of 979 samples using algorithms: 1 and 1-modified

Table 3: Precision for prediction windows

		Predictions		
Samples	Variables	3	10	20
451	Alg 1	100	50	35
	Alg 1m	100	100	90
651	Alg 1	100	90	45
	Alg 1m	100	90	70
851	Alg 1	66,67	50	30
	Alg 1m	66,67	50	30
951	Alg 1	100	30	50
	Alg 1m	100	30	50
979	Alg 1	100	80	50
	Alg 1m	100	80	50

CONCLUSION

According to the results obtained for prediction windows, the algorithm providing the most accuracy regardless of the length of the history window is algorithm_1 however, the combination of algorithm _1 and algorithm _2 provides substantial improvement for shorter history windows; it should be noted that long-term predictions are not reliable because there are no exact patterns that define the use/disuse of channels.

Although, the analysis was performed for a single channel its results are reliable since others (Kone *et al*., 2012) have shown little or no correlation between channel use patterns. For a reliable prediction the best option is to obtain predictions for few time slots and constantly feeding back the database with actual occupancy data as time passes.

REFERENCES

Akyildiz, I.F., W.Y. Lee, M.C. Vuran and S. Mohanty, 2006. Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Comput. Networks, 50: 2127-2159.

Haykin, S., 2005. Cognitive radio: Brain-empowered wireless communications. Selected Areas Commun. J., 23: 201-220.

Kone, V., L. Yang, X. Yang, B.Y. Zhao and H. Zheng, 2012. The effectiveness of opportunistic spectrum access: A measurement study. IEEE./ACM. Trans. Networking, 20: 2005-2016.

Lopez, R.B. and M.S. Sanchez, 2015. La radio cognitiva y su impacto en el uso eficiente del espectro de radio. Ingenieriia Electron. Autom. Comunicaciones, 36: 42-55.

Masonta, M.T., M. Mzyece and N. Ntlatlapa, 2013. Spectrum decision in cognitive radio networks: A survey. IEEE. Commun. Surv. Tutorials, 15: 1088-1107.

Zhao, Q. and A. Swami, 2007. A survey of dynamic spectrum access: Signal processing and networking perspectives. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 4, April 15-20, 2007, Honolulu, HI., USA., pp: IV-1349-IV-1352.