Secondary User QoE Enhancement Through Learning Based Predictive Spectrum Access in Cognitive Radio Networks

Anirudh Agarwal(✉), Shivangi Dubey, Ranjan Gangopadhyay, and Soumitra Debnath

Department of ECE, The LNM Institute of Information Technology, Jaipur, India
alanirudh@gmail.com, shivangidbl@gmail.com, ranjan_iitkgp@yahoo.com, soumitra_deb@yahoo.com

Abstract. Quality of experience (QoE) of a secondary spectrum user is mainly governed by its spectrum utilization, the energy consumption in spectrum sensing and the impact of channel switching in a cognitive radio network. It can be enhanced by prediction of spectrum availability of different channels in the form of their idle times through historical information of primary users' activity. Based on a reliable prediction scheme, the secondary user chooses the channel with the longest idle time for transmission of its data. In contrast to the existing method of statistical prediction, the use and applicability of supervised learning based prediction in various traffic scenarios have been studied in this paper. Prediction accuracy is investigated for three machine learning techniques, artificial neural network based Multilayer Perceptron (MLP), Support Vector Machines (SVM) with Linear Kernel and SVM with Gaussian Kernel, among which, the best one is chosen for prediction based opportunistic spectrum access. The results highlight the analysis of the learning techniques with respect to the traffic intensity. Moreover, a significant improvement in spectrum utilization of the secondary user with reduction in sensing energy and channel switching has been found in case of predictive dynamic channel allocation as compared to random channel selection.

Keywords: Machine learning · Dynamic spectrum access · Prediction · Spectrum utilization · Channel switching

1 Introduction

With the ever-increasing need for spectral resources, it becomes necessary for a secondary user (SU) to smartly and efficiently access the resources of idle channel primary radio systems without creating harmful interference to the licensed users, which is possible through Cognitive Radio (CR) technology. However, random spectrum sensing by an SU can result in a bad channel selection, as the channel might be heavily used by the existing primary user (PU) during sensing. Moreover, it might result into multiple unnecessary channel switch which would create delay in the SU data transmission [1]. Therefore, conservation of spectrum sensing energy and reduction in channel switching become really important for efficient dynamic spectrum access

D. Noguet et al. (Eds.): CROWNCOM 2016, LNICST 172, pp. 166–178, 2016.
DOI: 10.1007/978-3-319-40352-6_14

(DSA) and better QoE of SU, thereby improving the spectrum utilization of a CR user (terms, CR user and SU, are used interchangeably in this paper). This in turn suggests the need of spectrum prediction before sensing and channel allocation, where SU predicts the primary activity and sense only if the primary predicted state is idle. In this way, SU would perform prediction based sensing, starting with the channel having longest idle time and if the predicted state is busy, would switch to the next longer idle time channel for improved opportunistic spectrum access.

Machine learning (ML) proves to be a powerful tool for a CR system opening up versatile DSA applications viz. spectrum sensing, spectrum occupancy modeling, spectrum prediction, traffic pattern prediction, spectrum scheduling etc. [2]. The primary advantage of ML over other statistical models is that it does not require a-priori knowledge of the distributions under consideration. In the context of CR, ML techniques are generally used for signal classification, feature extraction, spectrum prediction [3–5] etc. For CR, mainly artificial neural networks (ANN) and Support Vector Machines (SVM) have been investigated in case of supervised ML [6]. But application specific work in reference to DSA and channel allocation based on learning has not been explored in sufficient detail in the existing literature. Moreover, most of the mentioned papers are restricted to one possible traffic model only.

Tumuluru *et al.* [7] has done the spectrum prediction based on MLP in Poisson traffic and has shown some improvement in SU spectrum utilization and reduction in sensing energy but the essence of channel switching has not been considered. In this paper, we have analyzed the performance of three supervised ML techniques e.g. artificial neural network based Multilayer Perceptron, Support Vector Machines with Linear Kernel (LSVM) and SVM with Gaussian Kernel (GSVM), for the prediction of primary activity as governed by several well known network traffic models namely, Poisson, Interrupted Poisson (IP) and Self-similar (SS) traffic. These traffic models reasonably capture the traffic characteristics that exist in most of the types of the wireless networks. The performance analysis of the prediction techniques is done in accordance with the statistical variation of the primary user data traffic. ML technique with highest prediction accuracy in estimating the average length of OFF period of primary in a single channel, is chosen for predicting the primary activity in multiple channel scenario. SU spectrum utilization, spectrum sensing energy and channel switching have been taken into account for claiming better QoE of CR user.

The paper is organized as follows: Sect. 2 discusses about the system model and methodology. A brief description of traffic models and various ML techniques are provided in Sect. 3. The performance analysis and the results are discussed in Sect. 4. Finally, Sect. 5 concludes the paper.

2 System Model and Methodology

In our model, for simplicity, we assume initially, one SU, targeting a single channel of PU whose channel state information (CSI) is used by the system for learning based prediction of future primary activity with three ML techniques in different network traffics. This model is subsequently used in the scenario of multiple PU channels where the CR would utilize the information of the average length of the OFF period $\left(\overline{L_{OFF}}\right)$ of

PU activity, predicted by an accurate ML technique. This process is repeated for the available channels and the channel having longest $\overline{L_{OFF}}$ is chosen, thereby resulting in an improved prediction based CR sensing-transmission strategy for DSA. It is also assumed that the sensing information is highly accurate. The channel allocation methodology has been explained by the flow graph in Fig. 1.

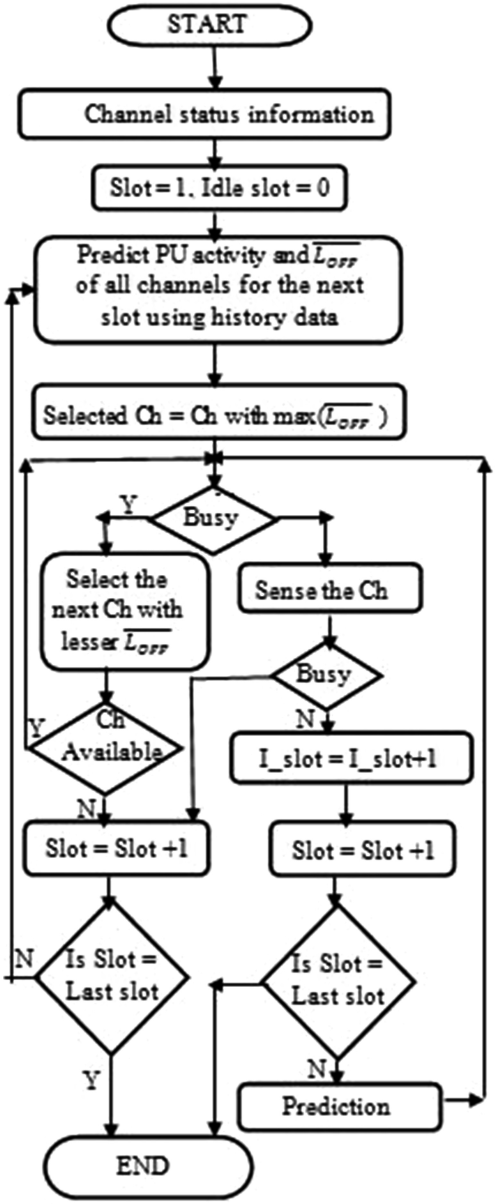

Fig. 1. Channel allocation methodology

3 Data Traffic Models and ML Prediction Techniques

3.1 Data Traffic Models

In this work, we have considered three different traffic models, i.e. Poisson traffic, Interrupted Poisson (IP) traffic and Self-similar (SS) traffic, for characterizing the statistics of a PU channel. Poisson traffic is one of the most widely used traffic model used to model the traditional voice data. The IP traffic is a good representation of data found in computers, e-mails, etc., i.e., there is heavy traffic for some time and then no traffic for some time. Self-similarity is a well known feature in the Internet traffic. SS traffic is characterized by long range dependence of traffic, burstiness and high correlation over varying time scales.

3.2 ML Prediction Techniques

In this sub-section, a brief description of three ML prediction techniques used in the present study is provided.

3.2.1 Multilayer Perceptron Neural Network Based Prediction

An MLP is a feed-forward network of simple neurons called *perceptrons*. It consists of three or more layers (an input and an output layer with one or more hidden layers) of nodes in a directed graph. Each node excluding the nodes at the input layer is a computing unit i.e. perceptron. The perceptron computes single output from multiple real-valued inputs by forming a linear combination of their input weights and then putting the output through some nonlinear activation function. Mathematically, this can be written as:

$$\boldsymbol{y} = \varphi\left(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x} + \boldsymbol{b}\right) \tag{1}$$

where $\boldsymbol{y}$ is the output vector, $\boldsymbol{\varphi}$ is the activation function, $\boldsymbol{w}$ is the weight vector, $\boldsymbol{x}$ is the input vector and $\boldsymbol{b}$ is the bias.

The activation function is often chosen to be the logistic sigmoid or tangent hyperbolic. The number of hidden layers and the number of neurons in each layer vary according to the application [7].

MLP networks are typically used in supervised learning problems, that can be further solved by the *back-propagation algorithm (BPA)* [8]. The signal flow graph for MLP is presented in Fig. 2, where the circles in blue, excluding the input layer, denote neurons.

For the implementation of the MLP algorithm, we have used MATLAB Neural Network Toolbox with 4 inputs in the input layer, two hidden layers consisting of 15 and 20 neurons respectively and one neuron in the output layer. We have used Tangent Sigmoid and Purelin as the activation functions respectively for the hidden layer and the output layer neurons. Both learning rate and momentum constant for the gradient descent method in BPA are taken as 0.2 and 0.9 respectively. These values are chosen after rigorous cross-verification for optimality.

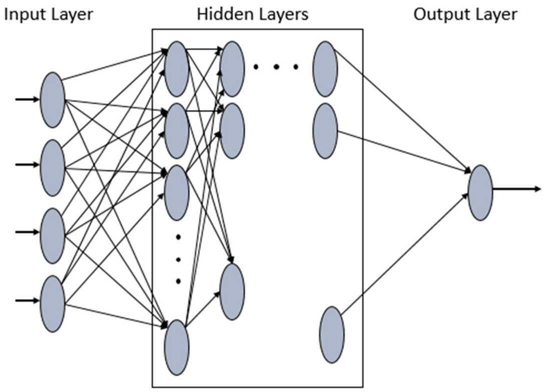

Fig. 2. Signal flow graph for MLP

3.2.2 Linear and Gaussian Support Vector Machines Based Prediction

SVM is a maximum margin discriminative classifier which means that, it learns a decision boundary that maximizes the distance between samples of the two classes, given a kernel. The distance between a sample and the learned decision boundary can be used to make the SVM a "soft" classifier. In the present implementation, we have used linear kernel based SVM and Gaussian kernel based SVM. The training feature and response vectors can be represented as $\boldsymbol{Z} = (\boldsymbol{T_i}, \boldsymbol{x_i})$ where $\boldsymbol{T_i} \in \{-1, 1\}$. The two classes are separated by a hyperplane denoted as H: $\boldsymbol{w^T x} + \boldsymbol{b} = \boldsymbol{\varepsilon}$, where $\boldsymbol{w}$ is the weight vector, $\boldsymbol{b}$ is the bias and $\boldsymbol{\varepsilon} = \sum_{i=1}^{m} \boldsymbol{\varepsilon_i}$ is a slack variable vector whose 1-norm is the penalty term. The hyperplanes which separate the two classes are given by:

$$T_i = \left\{ \begin{array}{c} 1 \ when \ \boldsymbol{w^T x_i} + b > 1 - \varepsilon \\ -1 \ when \ \boldsymbol{w^T x_i} + b < -1 + \varepsilon \end{array} \right\} \tag{2}$$

With a soft margin, the optimization problem for the SVM can be defined as follows:

$$\min_{(w,b,\varepsilon) \in \mathrm{R}^{n+1+m}} \left(\frac{1}{2} \|w\|^2 + C \sum_{\boldsymbol{i}=1}^{\boldsymbol{m}} \boldsymbol{\varepsilon}_{\boldsymbol{i}}^2 \right) \tag{3}$$

$$s.t. T_i(\boldsymbol{w^T x_i} + \boldsymbol{b}) > 1 - \boldsymbol{\varepsilon},$$
$$fori = 1, 2, \ldots, m.$$

where $C > 0$ is a regularization parameter that balances the weight of the penalty term $\sum_{i=1}^{m} \boldsymbol{\varepsilon_i}$ and the margin maximization term $\frac{1}{2}\|w\|^2$ [9].

For training and testing purposes, we have utilized the widely used software tool i.e. LIBSVM [10], integrated and compiled in MATLAB, where the algorithm is iterated until the minimum tolerance value (taken as 0.0001 in this work) is achieved.

4 Performance Evaluation, Results and Discussion

For evaluating the performance of the prediction based technique for the reliable prediction of $\overline{L_{OFF}}$ estimate for different traffic scenarios, we have used two performance measures i.e. probability of error in predicting the busy state, *Prob*$_{err}$*(busy state)*, and the mean square error in predicting the average length of the OFF duration of primary activity, *MSE of* $\overline{L_{OFF}}$. The significance of *Prob*$_{err}$*(busy state)* lies in the interference caused by the CR user to PU. More *Prob*$_{err}$*(busy state)* would lead to more tendency of CR user to violate the interference constraint. However, the overall utilization efficiency of SU with better channel allocation strategy, is governed by *MSE of* $\overline{L_{OFF}}$. Moreover, the prediction accuracy and *MSE of* $\overline{L_{OFF}}$ for all the ML techniques are calculated and depicted in Table 1. The traffic intensity for an ON-OFF random data traffic is defined as:

$$\rho = \frac{T_{ON}}{T_{ON} + T_{OFF}} \tag{4}$$

where T_{ON} and T_{OFF} are respectively the average time for which the primary user is busy and idle.

Table 1. Comparison of ML Techniques for different training lengths and traffic intensities in 3 traffic scenarios.

Percentage of training data	Traffic intensity	ML technique	Poisson		IP		SS	
			MSE	PA	MSE	PA	MSE	PA
30 %	0.5	MLP	2.032e-1	91.558	2.034e-1	87.040	1.195e-1	85.104
		LSVM	8.828e-7	92.850	2.233e-6	87.470	5.817e-7	85.398
		GSVM	2.012e-1	91.890	2.055e-1	87.290	6.597e-7	85.368
	0.8	MLP	5.360e-4	92.645	5.223e-3	89.950	2.140	86.746
		LSVM	7.382e-8	92.931	1e-7	90.003	4.082e-7	87.068
		GSVM	1.088e-7	92.919	7.330e-7	89.988	8.850e-7	87.054

The mean-square error in predicting the average length of the OFF duration of primary is calculated as:

$$MSE\ of\ L_{OFF} = \frac{1}{N}\sum_{n=1}^{N}\left(\overline{L_{OFF-pre}}(n) - \overline{L_{OFF-org}}(n)\right)^2 \tag{5}$$

where N is the total number of simulation intervals, $\overline{L_{OFF-pre}}(n)$ is the average length of OFF duration of PU activity in predicted data at the n^{th} simulation interval and similarly, $\overline{L_{OFF-org}}(n)$ is the average length of OFF duration of PU activity in original data at the n^{th} simulation interval. The other performance metric, the prediction accuracy (PA), is defined as:

$$PA = (1 - P_e) \times 100 \tag{6}$$

where P_e is the overall probability of error, i.e. when the busy state is predicted as idle and vice versa, in the prediction by an ML technique. For prediction analysis, 30 % of the primary user data is used for training while the rest of the data is utilized for testing the trained model. The total number of traffic slots for primary data is taken as 50000 in this work. Moreover, as the characteristics of the particular data traffic might change with time, we have evaluated the performance of all the considered parameters after averaging over sufficient number (50 in this work) of simulation intervals.

In a multi-channel scenario, two cases have been analyzed for the performance, i.e. random sensing, where CR user randomly senses the channels and check the status at every slot, and prediction based sensing, where CR user predicts the channel status of all the channels with the help of slot history information. In the second case, the channels are prioritized in the decreasing order of their $\overline{L_{OFF}}$, then the CR user starts priority based sensing of the channel among the channels with predicted idle status. Moreover, we have considered 10 licensed channels of different traffic characteristics, as described in Table 2, and a single SU in the slotted-time mode. It is assumed that at every slot, SU has the sensing information of all the channels.

Table 2. Different Channels for Primary User System

Channel number	Mean inter arrival time	Traffic intensity
1	22	0.760
2	20	0.500
3	18	0.700
4	16	0.625
5	12	0.600
6	10	0.700
7	10	0.600
8	20	0.500
9	25	0.525
10	22	0.500

The QoE of SU is characterized using three performance measures:

1. *Spectrum Utilization Improvement (α_{imp})*
   The spectrum utilization (α) is defined as the fraction of slots in the system over a finite duration of time, for which number of slots are detected as idle by the CR user. So, spectrum utilization improvement (in %) due to prediction is given by

$$\alpha_{imp} = \frac{\alpha_{PS} - \alpha_{RS}}{\alpha_{RS}} \times 100 \tag{7}$$

   where α_{RS} and α_{PS} are respectively the spectrum utilizations by CR in the case of random sensing and prediction based sensing.

2 *Reduction in Channel Switching (β_{red})*
The percentage reduction in channel switching due to spectrum prediction can be expressed as

$$\beta_{red} = \frac{\beta_{RS} - \beta_{PS}}{\beta_{RS}} \times 100 \tag{8}$$

where β_{RS} and β_{PS} are number of times CR has to switch the channel in case of random sensing and prediction based sensing respectively.

3 *Reduction in Sensing Energy (ξ_{red})*
Sensing energy is reduced in case of prediction based sensing, because in random sensing, CR user has to sense all the slots while through prediction, it senses only when the state of the channel is predicted to be idle. Here, we have assumed that one unit of sensing energy is required to sense one slot. The percentage reduction in sensing energy is given by

$$\xi_{red} = \frac{\xi_{RS} - \xi_{PS}}{\xi_{RS}} \times 100 \tag{9}$$

where ξ_{RS} and ξ_{PS} denote the product of unit sensing energy and corresponding number of idle slots sensed by CR in case of random sensing and prediction based sensing.

Figures 3 and 4 depict the probability of error in predicting the busy state of the primary for different traffic types utilizing the three learning schemes. It can be observed that for all types of data traffics, the probability of error in predicting the busy state decreases as we increase the traffic intensity. As ρ increases, the number of times

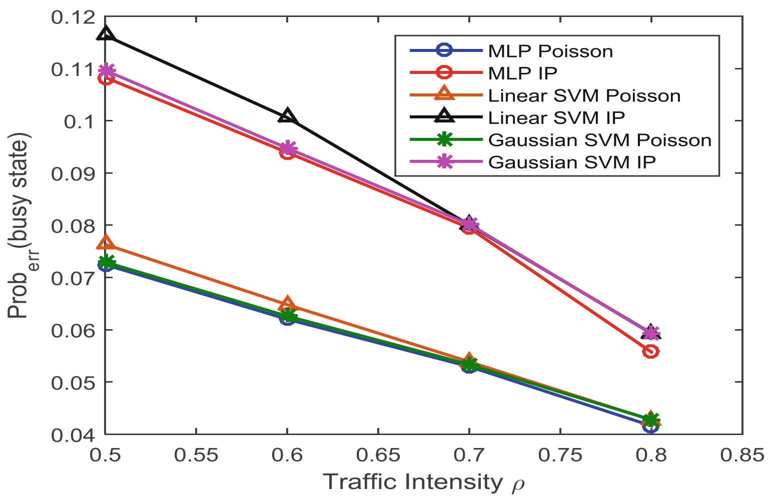

Fig. 3. Prob$_{err}$(busy state) vs. ρ for different data traffic using different ML techniques.

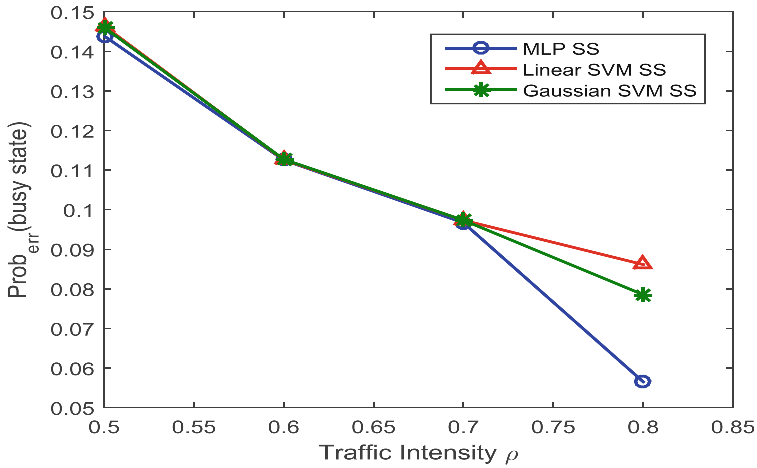

Fig. 4. Prob$_{err}$(busy state) vs. ρ for SS traffic using different ML techniques.

the channel being idle tends to decrease. It may be noted that MLP prediction technique performs slightly better than the other two in this case.

Figure 5 shows the variation of *MSE of* $\overline{L_{OFF}}$, for Poisson and IP traffic against the traffic intensity, ρ. The decreasing nature is attributed from the fact that, there are less number of transitions from busy state to idle and vice-versa with increase in ρ. This leads to more dependency of future states on the present and previous states, thereby suggesting a decrease in prediction error and an improvement in the prediction

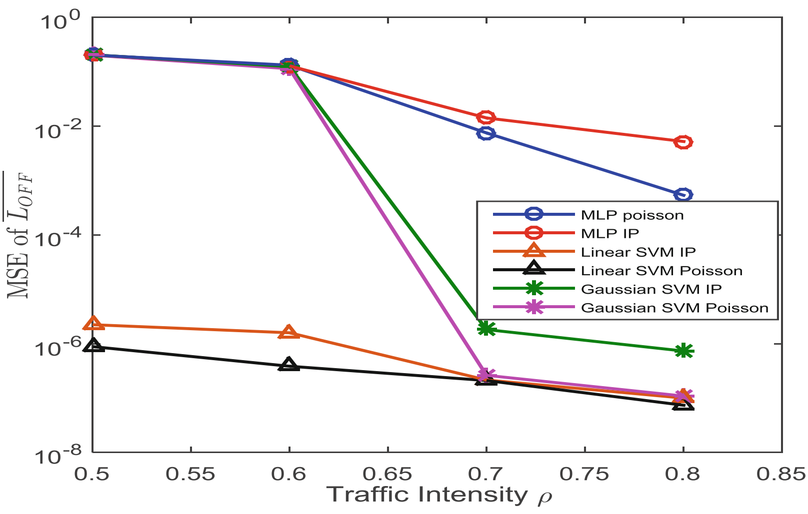

Fig. 5. MSE of $\overline{L_{OFF}}$ vs. ρ for different traffic using different ML techniques.

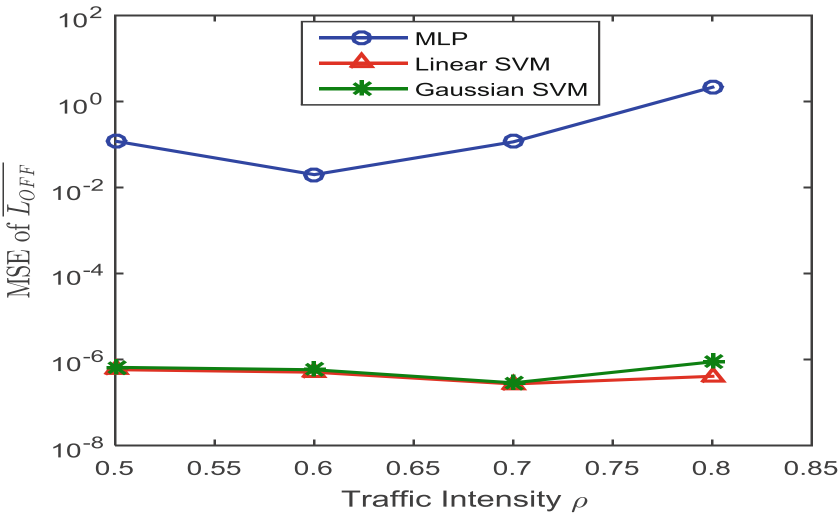

Fig. 6. MSE of $\overline{L_{OFF}}$ vs. ρ for SS traffic using different ML techniques.

accuracy. However, in this case, LSVM based prediction turns out to be the best in terms of $\overline{L_{OFF}}$.

With the same reasoning, similar pattern was expected for bursty SS traffic in Fig. 6. But it is observed that the prediction is not so accurate for high traffic intensity. This may be due to the heavy burstiness and strong OFF period correlation in the data traffic. Nevertheless, LSVM is found to perform uniformly and reliably for the SS traffic.

Table 1 provides a comparison of the performance of ML techniques under various simulated conditions in terms of the MSE of $\overline{L_{OFF}}$ and prediction accuracy of the algorithm under consideration. It is clearly observed that for any data traffic, LSVM outperforms both the other techniques.

Figures 7 and 8 show the comparison of random sensing and prediction based sensing on the basis of improvement in spectrum utilization and reduction in channel switching respectively. It can be seen in Fig. 7 that the number of times a channel used by a CR user for a duration of time, is far more in the case of prediction based sensing than that in random sensing. This is due to the fact that through prediction, a CR can find more idle slots in a channel thereby utilizing it optimally. Moreover, the percentage of channel switching would also be decreased as shown in Fig. 8. In prediction based sensing, CR remains in the best predicted channel, i.e. the channel with longest $\overline{L_{OFF}}$, for more time and switches to the next prioritized channel only when the state of the present channel is predicted to be busy, unlike in random sensing where CR senses any channel randomly and is supposed to switch, if it is found to be busy in sensed slot.

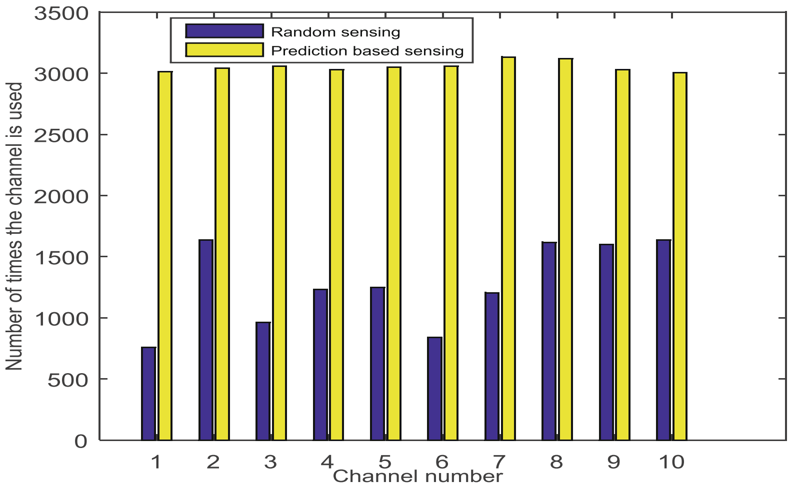

Fig. 7. SU utilization of a channel in both cases of sensing (Color figure online)

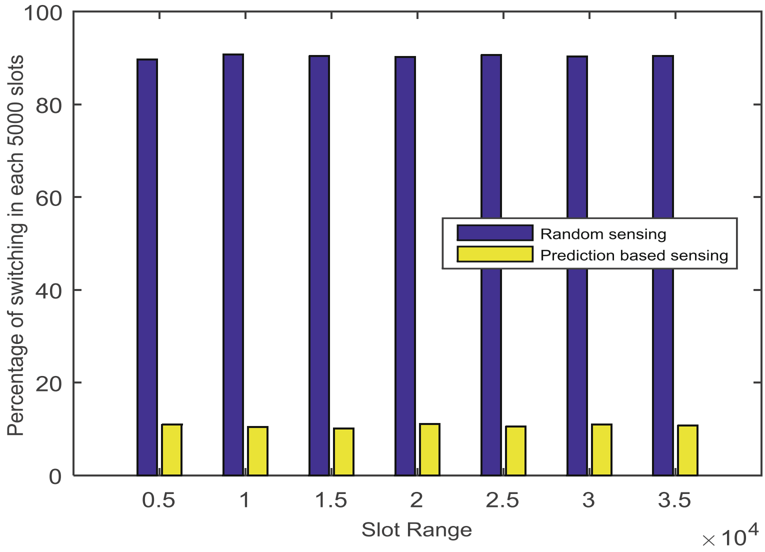

Fig. 8. Comparison of % channel switching in both cases of sensing (Color figure online)

In Table 3, an improvement in spectrum utilization due to prediction ranging from around 88 % to 146 % considering different number of channels, has been found. Channel switching is reduced by more than 83 %. Finally, Table 4 shows the reduction in sensing energy for a channel with $\rho = 0.5$ and for different mean inter-arrival times. Energy is saved in the case of prediction because there is no sensing operation when the predicted status of the slot is busy.

Table 3. Comparison of α_{imp} and β_{red} due to prediction for different number of channels

Number of channels	α_{imp} (%)	β_{red} (%)
3	88.5394	87.4291
5	124.0066	85.1706
7	146.3669	83.9832
9	145.1946	86.7339
10	139.7345	88.1905

Table 4. ξ_{red} in a channel due to prediction for $\rho = 0.5$ and different mean inter-arrival time

Mean inter-arrival time	ξ_{red} (%)
10	57.1265
12	55.9721
14	55.0063
16	54.2976
18	54.1090
20	53.5461
22	53.2832

5 Conclusion and Future Work

In this work, the importance of machine learning spectrum prediction is highlighted in the context of CR for efficient DSA. LSVM is found to consistently predict the primary $\overline{L_{OFF}}$ in data traffics with high accuracy. This technique is further used in multi-channel CR scenario and prediction based sensing is done over the channels which are prioritized on the basis of decreasing $\overline{L_{OFF}}$. A significant improvement in spectrum utilization, decrease in channel switching frequency and reduction in spectrum sensing energy have been found thereby providing improved quality of experience to CR user for opportunistic spectrum access.

However, this preliminary work needs further extension for multiple SU's without assuming perfect detection. Moreover, some advanced prediction algorithms based on deep belief need to be exploited for further QoE enhancement of the SU in multiple channel radio access in a CR system.

Acknowledgement. This work has been carried out as a part of the project "Mobile Broadband Service Support over Cognitive Radio Networks", sponsored by Information Technology Research Academy (ITRA), Department of Electronics and Information Technology (DeitY), Ministry of Communications & IT, Govt. of India.

References

1. Hoyhtya, M., Pollin, S.: Improving the performance of cognitive radios through classification, learning, and predictive channel selection. Adv. Electron. Telecommun. **2**, 28–38 (2011)
2. Azmat, F., Chen, Y., Member, S., Stocks, N., Mar, N.I.: Analysis of Spectrum Occupancy Using Machine Learning Algorithms, 0–22
3. Xing, X., Jing, T., Cheng, W., Huo, Y., Cheng, X.: Spectrum prediction in cognitive radio networks. IEEE Wirel. Commun. **20**, 90–96 (2013)
4. Lee, J., Park, H.-K.: Channel prediction-based channel allocation scheme for multichannel cognitive radio networks. J. Commun. Networks. **16**, 209–216 (2014)
5. He, A., Member, S., Bae, K.K., Newman, T.R., Gaeddert, J., Member, S., Kim, K., Menon, R., Morales-tirado, L., Neel, J.J., Zhao, Y., Reed, J.H., Tranter, W.H., Fellow, L.: Survey of artificial intelligence for cognitive radios. IEEE Trans. Vech. Tech. **59**, 1578–1592 (2010)
6. Bkassiny, M., Member, S., Li, Y., Member, S.: A Survey on Machine-Learning Techniques in Cognitive Radios, pp. 1–53
7. Tumuluru, V.K., Wang, P.W.P., Niyato, D.: A Neural Network Based Spectrum Prediction Scheme for Cognitive Radio (2010)
8. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Cog. mod. **5**, 3 (1988)
9. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Mining Knowl. Discov. **2**(2), 121–167 (1998)
10. http://www.csie.ntu.edu.tw/~cjlin/libsvm