

# Predictive learning model for Cognitive Radio using Reinforcement Learning

Mithra Venkatesan<sup>1</sup>, Anju.V.Kulkarni<sup>2</sup>, Radhika Menon<sup>3</sup>, Sharadha Tubachi<sup>4</sup>, V.Ramakrishnan<sup>5</sup>

<sup>1</sup>Associate Professor, Department of E&TC, D.Y.Patil Institute of Technology, SPPU, Pune, India.

<sup>2</sup>Professor, Department of E&TC, D.Y.Patil Institute of Technology, SPPU, Pune, India.

<sup>3</sup>Professor, Department of Mathematics, D.Y.Patil Institute of Technology, SPPU, Pune, India.

<sup>4</sup> PG Student, Department of E&TC, D.Y.Patil Institute of Technology, SPPU, Pune, India.

<sup>5</sup>Software Architect, Siemens PLM, Pune, India.

ORCID: 0000-0002-6541-447X (Venkatesan), 0000-0002-3160-0450 (Kulkarni),

0000-0002-0389-9037 (Menon)

## Abstract

Intelligence is required to cope with the rapid advancement of wireless communications. Intelligence is required, particularly to manage and share the inadequate radio spectrum in the exceptionally current unstable environments. Cognitive Radio is a promising technology which is capable of handling different circumstances by using intellectual software packages which will impart their transceiver with radio-awareness, compliance and potential to learn. A Cognitive Radio is a system which continuously monitors the cognition cycle, in which it alters its operating parameters, looks at the results and, in due course the system takes actions, which helps us to decide operating behavior of radio configuration, desires to progress the radio. In practice, learning model make use of measurements sensed from the operating environment, collect skills and accumulate knowledge which will help us to make decision. In this paper we have studied a Reinforcement Learning based technique to predict the throughput for a Cognitive Radio System using the Continuous Actor Critic Learning Automation (CACLA) Algorithm and our proposed Modified Continuous Actor Critic Learning Automation Algorithm (MCACLA). The performance metrics used to assess the learning which have been used in the paper are execution time, prediction accuracy and prediction error. Further, comparison has been done with existing predictive models for learning to establish the improvements of the proposed methods.

**Keywords:** Cognitive Radio, Wireless communication, Reinforcement learning, Wireless Spectrum, Throughput prediction.

## I. INTRODUCTION

### A. Background

Due to swift growth of high data rate wireless applications such as Digital Video Broadcast (DVB), Digital Audio Broadcast (DAB), Internet, Wi-max, the scarcity of radio spectrum has become an alarming major issue. To get rid of this scarcity problem, Cognitive Radio comes into the picture to improve the spectrum efficiency [1]. For the most of the time the spectrum is assigned to some primary users who are licensed to radio spectrum by government agencies, but these primary users are not using the complete spectrum. Some of spectrum is wasted. In this case Cognitive Radio allows us to

use the underutilized spectrum and share this spectrum to some unlicensed secondary users who are really in need. Because of this massive spectrum insufficiency problem, we are paying high licensing fees for the available spectrum. In particular, if we want to scan radio spectrum in revenue urban rich areas, we will find some frequency slots are unoccupied for some of the time, many slots are moderately used and some slots are completely used. Hence the partial usage of radio spectrum is a major challenge to an emerging and future technology such as Cognitive Radio [1].

### B. Motivation

Cognitive Radio solves the partially utilized spectrum problems by sensing the radio environment in the following manner.

- i. Recognizing the frequency slots of the spectrum which are partially utilized by the licensed user.
- ii. Re-allotting those unused bands to unlicensed users [2,5].

The inspiration of Cognitive Radio is from making prolific usage of vacant frequency slots of the spectrum and having capacity to make human-like decisions to transmit without impediment. To accomplish the above said ability, Cognitive Radio works in reactive or proactive manner based on exterior ecological information, along with their aims, ethics, abilities, practice and knowledge. Hence prospect Cognitive Radio will have the potential, to choose the radio configuration on the fly, by in addition with the context of operation (device status and environment aspects), goals, policies, profiles and capabilities, and machine learning (for representing and managing knowledge and experience). Generally, the word radio configuration or configuration refers to a selected carrier frequency and a precise Radio Access Technology (RAT) however it can be extended to add other operating parameters such as transmit power, modulation type, etc.

The intelligence to the Cognitive Radio is imparted by the brain of the Cognitive Radio which is the Cognitive Engine [3, 4]. The main module in the Cognitive Engine is the learning module. The Cognitive Engine can be accomplished by variety of learning algorithms and techniques. So far, there have been many different learning techniques explored in literature for Cognitive Radio which includes Artificial Neural Networks, ANFIS, Evolutionary/ Genetic Algorithms, Hidden Markov Models, etc. Among the predictive learning models

throughput has been predicted using ANN based learning[26,27,28], Self-Organizing Map[30,31], ANFIS[29,30].

Reinforcement Learning is an unsupervised learning technique exclusively based on rewards which is the consequence of state-action pair, not on the empirical data or experimental results from the laboratory.

RL has been widely used in Cognitive Radio for spectrum sensing [6,7], spectrum analyzing[8,9,], spectrum management(which includes Dynamic Channel Selection, Opportunistic Spectrum Access, Dynamic Spectrum Access, Spectrum Assignment )[10,11,12,13,16], spectrum sharing in CR networks[20], RL has also been used for reactive jamming mitigation[17,18]. Further it has also been observed that RL can be used for security enhancement in CRN [19] etc.

However, the scopes of RL based techniques have never been explored in Cognitive Engine of Cognitive Radio. Learning algorithms or techniques or modules for Cognitive Radio is yet to be built using Reinforcement based techniques and this is a viable option considering the merits of Reinforcement based techniques. This has been a completely unexplored terrain and has huge scope for exploration. Hence, the main objective of this paper is in the realization of the learning capability in throughput prediction and modeling the predictive model using Reinforcement Learning based techniques. In this paper RL based throughput prediction based on CACLA algorithm and MCACLA algorithm has been developed and validated.

### C. Objective

The aim of present work is to investigate RL based CACLA and MCACLA algorithm and to study how the projected scheme is used to predict the throughput in Cognitive Radio Network.

### D. Organisation of the paper

The paper is organized into following sections. Section 2 proposes RL based learning scheme and its use for throughput prediction for existing CACLA method and proposed MCACLA method. Section 3 presents the results of existing CACLA algorithm and MCACLA algorithm, along with observations and inferences. The performance of CACLA and MCACLA methods are also compared in this section. Section 4 discusses conclusion and future work to be done.

The introduction of the paper should explain the nature of the problem, previous work, purpose, and the contribution of the paper. The contents of each section may be provided to understand easily about the paper.

## II. REINFORCEMENT LEARNING BASED THROUGHPUT PREDICTION

### A. Introduction

This section describes RL based throughput prediction for Cognitive Radio. RL is Artificial Intelligence technique, which combines best features of supervised and unsupervised learning.

Section covers basic overview of RL, RL based CACLA algorithm and MCACLA algorithm to predict the throughput. Finally we are comparing CACLA method and MCACLA methods.

### E. Reinforcement learning (An Overview)

Reinforcement Learning is an unsupervised learning technique which improves system performance using simple modeling. Since it is unsupervised learning, there is no supervision to watch over the learning process, and so, an agent learns information with reference to the uncertain operating environment by itself. An agent constitutes knowledge on the fly at the same time it performs its normal operation, although it is not using pragmatic data or experimental results from the laboratory.

### F. Methodology

The learning model uses different type of network to predict the throughput for a particular radio set-up. Towards accomplishing learning, a database is formed with 3 operating parameters namely RSSI, Data Rate and Throughput. The database is same as the database being developed in literature [30].

#### RL based CACLA algorithm

The algorithm is modified to predict the throughput based on RSSI values. Our goal is to maximize the throughput using RL.

State set S: Number of states which are equal to number of labels from the dataset such as 1, 6, 48 and 54. There are four states.

Action set A: Its equal to number of states, will return the label on which maximum throughput predicted.

V-table-initialized by passing values of RSSI and its throughput

#### Algorithm 1: CACLA (Continuous Actor Critic Learning Automation)

Step 1: Adopt MDP to act on which is made up of state set S, action set A and Reward function  $R(s_i, a_i)$ .

Step 2: Repeat

Step 3: Initialize V-table consisting of the RSSI and its throughput values each state (label)  $S_i$ .

Step 4: Selection of state band  $S_i$  randomly.

Step 5: Select an action based on  $\epsilon$ -greedy exploration strategy

Step 6: Calculate the reward  $R_{ijk}$  perceived by the user when taking action  $a_i$  using:

$$R_{ijk} = \begin{cases} \frac{v_i}{3} & \text{When } i = j = k \\ \frac{v_i}{2} & \text{When } i = j \neq k \\ & \text{or } i = k \neq j \\ v_i & \text{When } i \neq j \neq k \end{cases} \quad (1)$$

Step 7: Update V-table using:

$$v_{s_i}^E(t+1) = v_{s_i}^E(t) + \alpha [R_{s_i, a_i} - v_{s_i}^E(t)] \quad (2)$$

Step 8: if  $v_{s_i}^E(t+1) > v_{s_i}^E(t)$  (3)

then,

$$A_{s_i}^u = a_i \quad (4)$$

Maximum Throughput predicted;

else

Maximum Throughput predicted;

Step 9: if Decision center checks last label,

then go to step 10

else go to step 4

Step 10: END

### RL based MCACLA algorithm

The limitations such as worst performance of variance, problems with high-dimensional/continuous states and actions have been observed in the existing CACLA method.

Towards the same, the modified CACLA approach has been proposed in this paper, which is based on partially observable Markov decision processes (POMDPs). *This has been the salient incorporation and unique feature of the paper.*

Using POMDP approach this paper presents the below is MCACLA algorithm.

The algorithm is modified to predict the throughput based on RSSI values. Our goal is to maximize the throughput using RL.

The following steps describe the modified algorithm and the various parameters defined in the algorithm.

State set S: Number of states which are equal to number of labels from the dataset such as 1, 6, 48 and 54. There are four states.

Action set A: Its equal to number of states, will return the label on which maximum throughput predicted.

V-table= initialized by passing values of RSSI and its throughput.

### Algorithm 2: MCACLA (Modified Continuous Actor Critic Learning Automation)

Step 1: Adopt POMDP to act on which is made up of state set S, action set A, Reward function R (si, ai), Probabilistic state-action transitions P (si | a, sj), and Conditional observation probabilities O (o | si, ai).

Step 2: Repeat

Step 3: Initialize V-table consisting of the RSSI and its throughput values each state (label) Si and probabilistic state-action transitions initialization P (si | a, sj) and O (o | si, ai).

Step 4: Generate state transitions S as production of actions and probabilities using:

$$b'(s) = P(s' | a, o, b) = O(o | s', a, b) \cdot \frac{P(s' | a, b)}{O(o | a, b)}$$

$$O(o | a, b) = O(o | s')$$

$$P(s' | a, b) = \sum_{s \in S} P(s' | a, s) \cdot b(s)$$

$$O(o | a, b) = \sum_{s \in S} O(o | s') \cdot P(s' | a, b)$$

$$V(i, j) = P(s' | a, b) \cdot O(o | a, b) \quad (1)$$

Step 5: Selection of state band Si randomly

Step 6: Select an action based on  $\epsilon$ -greedy exploration strategy

Step 7: Calculate the reward Rijk perceived by the user when taking action ai using:

$$R_{ijk} = \begin{cases} \frac{v_i}{3} & \text{When } i = j = k \\ \frac{v_i}{2} & \text{When } i = j \neq k \\ & \text{or } i = k \neq j \\ v_i & \text{When } i \neq j \neq k \end{cases} \quad (2)$$

Step 8: Update V-table using:

$$v_{s_i}^E(t+1) = v_{s_i}^E(t) + \alpha [R_{s_i, a_i} - v_{s_i}^E(t)] \quad (3)$$

Step 8: if  $v_{s_i}^E(t+1) > v_{s_i}^E(t)$  , (4)

then,

$$A_{s_i}^u = a_i \quad (5)$$

Maximum Throughput predicted;

else

Maximum Throughput predicted;

Step 9: if Decision center checks last label,

then go to step 10

else go to step 4

Step 10: END

### III. RESULTS AND DISCUSSION

In this paper, we have designed two RL based algorithms for throughput prediction based on input dataset. For experimental study we have collected three different datasets with different sizes in terms of number of records. These datasets are collected from the real time wireless network communications. There are three main parameters such as RSSI value, data rate label value and throughput values are utilized by our algorithms CACLA and MCACLA. Table 1 shows the properties of three datasets used.

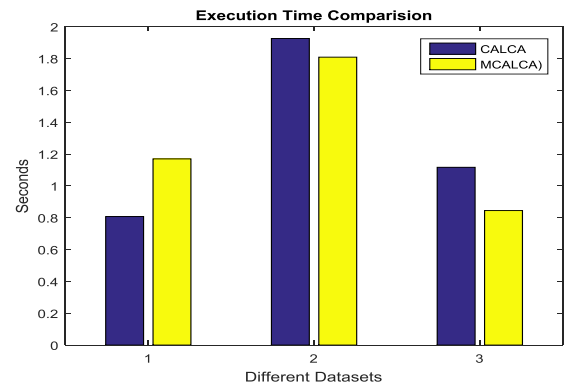
**Table 1.** Properties of Dataset Used

Dataset	Number of Records	RSSI Values	Data rate Values
D1	20	-18, -19, -31.	1, 6, 48, 54
D2	1084	-52 to -66, -69, -44, -42, -41, -49, -29, -30, -37, -39, -40.	1, 6, 48, 54
D3	312	-32 to -37	-32 to -37

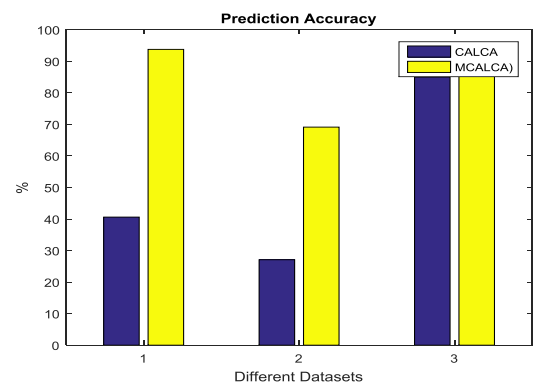
After applying both algorithms on three datasets, core performance metrics like execution time, error rate and accuracy rate are found. The results have been tabulated in table 2. Figures 1,2 and 3 graphically depict the execution time taken, percentage of correct predictions and incorrect predictions for different datasets.

**Table 2.** Results of CACLA and MCACLA methods

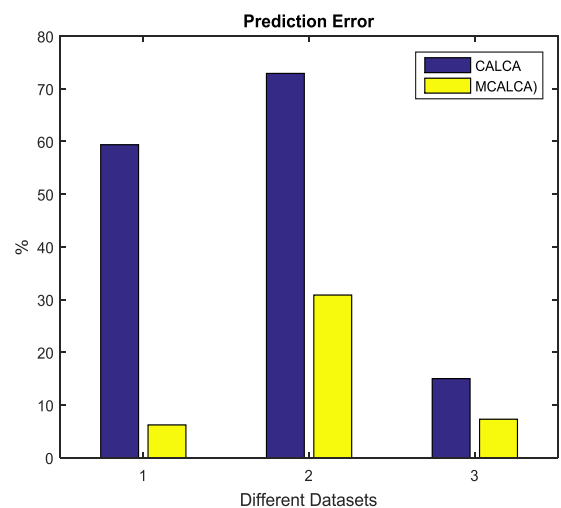
Algorithm used	Number of iterations/epochs	Prediction Accuracy(%)
Reinforcement learning using CACLA algorithm.	20	41.06
Reinforcement learning using CACLA algorithm.	1080	60.197
Reinforcement learning using CACLA algorithm.	298	79.8764
Reinforcement learning using MCACLA algorithm.	20	42.9536
Reinforcement learning using MCACLA algorithm.	1080	78.349
Reinforcement learning using MCACLA algorithm.	298	86.3748



**Fig 1.** Performance Analysis of Execution Time



**Fig.2.** Performance Analysis of correct Prediction



**Fig.3.** Performance Analysis of incorrect Prediction

Figure 1 illustrates the execution time of the CACLA and MCACLA algorithms. From the figure, we observed that the execution time of the MCACLA is small compare to CACLA method for the dataset 2 and 3. But MCACLA method takes more time to give results for dataset 1 because the MCACLA works on more volume of dataset. Since the dataset 1 is having only 20 values of data in it. Figure 2 illustrates the

prediction accuracy of the CACLA and MCACLA method, we can observe the prediction accuracy is doubled in MCACLA method compare to CACLA method for every dataset considered. Figure 3 illustrates the prediction error of the CACLA and MCACLA method, we can observe the prediction error is decreased half in MCACLA method compare to CACLA method for every datasets.

The major observation done is that the proposed MCACLA method outperforms CACLA method in terms of all performance metrics such as reduction in execution time, reduction in prediction error and improvement in prediction accuracy. The existing work is compared with the proposed work which is detailed in Table 3. Here we can observe that the proposed MCACLA model performs much better as compared to all previous CACLA method and we get prediction accuracy up to 78.349 %. The proposed MCACLA based RL system has improved prediction accuracy of 82.7% as compared with another unsupervised techniques like RL based CACLA, Feed forward neural networks [24,21,22] . Table 3 gives a comparative study of percentage of correct predictions for different methods. Though the prediction accuracy of Subtractive clustering based ANFIS Model is maximum compared to all existing supervised and unsupervised learning, the design complexity is huge. Hence, the RL based proposed MCACLA method exhibits less design complexity as well as moderate level of prediction accuracy.

**Table 3.** Comparative analysis of results various algorithms with proposed methods

Methodology/algorithm adopted	Number of iterations/epochs	Percentage of correct prediction (%)
Feed forward neural networks [21,22]	300	72
Elman networks [21,22]	300	81
Focused Time Delay neural networks[21,22]	300	80
Grid partition based adaptive Neuro Fuzzy Interference System [23]	100	89
Subtractive clustering based ANFIS Model [26]	100	97
Unsupervised technique [24]	10	78
Self organizing model [26]	10	82.7
Reinforcement learning using CACLA algorithm	1084	60.197
Reinforcement learning using MCACLA algorithm	1084	78.349

## V. CONCLUSION

This paper completely explores the usage of Reinforcement based learning techniques for building predictive models for Cognitive Radio which has been the unique feature of this paper. In this paper, two RL based algorithms for throughput prediction namely CACLA and modified version of CACLA

has been designed. The novel modified CACLA approach which uses partially observable Markov decision processes (POMDPs) has been presented. The performance of accuracy and error rate for the proposed MCACLA method shows great improvement compared to the existing method.

The proposed RL based system has the capability of predicting the throughput of a specific radio configuration and can hugely contribute to build predictive learning models for Cognitive Engine for Cognitive Radio. In fact, these learning and predictive models play a vital role in empowering the Cognitive Engine and make the Cognitive Radio truly astute and intelligent.

The capability of radio configuration includes throughput prediction, different access technology, modulation type, frame rate etc. Accordingly RL based method could be further modified to predict all these capability of radio configuration which could subsequently enhance the cognitive ability of the Cognitive Radio.

## REFERENCES

- [1] Kok-Lim Alvin Yau, Geong-Sen Poh, Su Fong Chien, and Hasan A. A. Al-Rawi "Application of Reinforcement Learning in Cognitive Radio Networks: Models and Algorithms" The Scientific World Journal Volume 2014 (2014), Article ID 209810, 23 pages Hindawi.
- [2] Yu Ren, Pawel Dmochowski, Peter Komisarczuk "Analysis and Implementation of Reinforcement Learning on a GNU Radio Cognitive Radio Platform" CROWNCOM2010.9170 Digital Object Identifier: 10.4108/ICST.
- [3] Kok-Lim Alvin Yau, Peter Komisarczuk and Paul D. Teal "Applications of Reinforcement Learning to Cognitive Radio Networks". Workshop on Cognitive Radio Interfaces and Signal Processing (CRISP) at the International Conference on Communications (ICC'10) IEEE, Cape Town, South Africa, May 2010.
- [4] Asma Amraoui, Badr Benmammar, Francine Krief, Fethi Tarik Bendimerad "Intelligent Wireless Communication System using Cognitive Radio" International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.2, March 2012.
- [5] Dhananjay Kumar, Pavithra Hari, Panbzhazagi Selvaraj and Sharavanti Baskaran "Efficient spectrum utilization in Cognitive Radio through reinforcement learning" ICTACT international journal on communication technology, september 2013, volume: 04, issue: 03.
- [6] Brandon F. Lo Ian F. Akyildiz "Reinforcement learning for Cooperative sensing gain in Cognitive Radio ad hoc networks" Springer Science+Business Media New York 2012.
- [7] Thesis by Nils Morozs Ph.D. University of York Electronics September 2015 "Accelerating

- Reinforcement Learning for Dynamic Spectrum Access in Cognitive Wireless Networks”.
- [8] Liljana Gavrilovska, Vladimir Atanasovski, Irene Macaluso, and Luiz DaSilva “Learning and Reasoning in Cognitive Radio Networks” January 2013 · IEEE Communications Surveys & Tutorials 01/2013.
- [9] Ana Galindo-Serrano, Lorenza Giupponi, Pol Blasco and Mischa Dohler “Learning from Experts in Cognitive Radio Networks: The Docitive Paradigm” In Proceedings of 5th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2010), 9-11 June 2010, Cannes (France).
- [10] Kok-Lim Alvin Yau, Peter Komisarczuk and Paul D. Teal “Context-Awareness and intelligence in Distributed Cognitive Radio Networks: A Reinforcement Learning Approach” 11th Australian Communications Theory Workshop (AusCTW'10) IEEE, Canberra, Australia, February 2010.
- [11] Francisco Bernardo, Ramon Agusti, Jordi P´erez-Romero and Oriol Sallent “Distributed Spectrum Management based on Reinforcement Learning” Proceedings Of The 4th International Conference On Crowncom 2009.
- [12] Wu Chun, Yin Mingyong, Ma Shaoliang, Jiang Hong “Multiagent Reinforcement Learning Dynamic Spectrum Access in Cognitive Radios” Sensors & Transducers, Vol.164, Issue 2, February 2014, pp. 170-175.
- [13] Pavithra Venkatraman, Bechir Hamdaoui, and Mohsen Guizani “Opportunistic Bandwidth Sharing Through Reinforcement Learning” IEEE Transactions on Vehicular Technology, Volume: 59 Issue: 6
- [14] T. Pavan Kumar , Suresh babu, Venkata Ramana. Sai Shashank “Survey: Routing Protocols in Cognitive Radio Mesh Networks” T. Pavan Kumar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (1) , 2015, 603-608.
- [15] Apurva Varade, Shriram Kulkarni “overview and literature survey on routing Protocols for Cognitive Radio networks” International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Volume 1 Issue 10 (November 2014)
- [16] Kok-Lim AlvinYau , PeterKomisarczuk , PaulD.Teal “Reinforcement learning for context awareness and intelligence in wireless networks: Review, new features and open issues” 25th International Conference on Advanced Information Networking and Applications (AINA'11) IEEE, Biopolis, Singapore, March 2011.
- [17] Marc Lichtman and Jeffrey H. Reed “Reinforcement Learning for Reactive Jamming mitigation” Journal of Cyber Security, Vol. 3 No. 2, 213–230 @2014 River Publishers.
- [18] Feten Slimenia, Bart Scheers, Zied Chtourou, Vincent Le Nir, Rabah Attia “Cognitive Radio Jamming Mitigation using Markov Decision Process and Reinforcement Learning” Procedia Computer Science 00 (2015) 000–000.
- [19] Mee Hong Linga, Kok-Lim Alvin Yau, Junaid Qadirb, Geong Sen Pohc, Qiang Nid “Application of reinforcement learning for security enhancement in Cognitive Radio Networks” M.H. Ling et al. Applied Soft Computing 37 (2015) 809–829 811
- [20] Tao Jiang, David Grace, Yiming Liu “Performance of Cognitive Radio Reinforcement Spectrum Sharing Using Different Weighting Factors” International Conference on Communications and Networking in China, 2008.
- [21] Nisha Kiran, Ravi Raj Alisha Khan “Maximization of Throughput in Cognitive Radio using non-persistent CSMA combined with Reinforcement Learning” International journal of innovative research in electrical, electronics, instrumentation and control engineering vol. 3, issue 5, may 2015.
- [22] Thesis by Tao Jiang Doctor of Philosophy (Ph.D.) University of York Department of Electronics September 2011 “Reinforcement Learning-based Spectrum Sharing for Cognitive Radio.
- [23] R.Urban and M. Steinbauer “Optimization of Machine Learning Parameters for Spectrum Survey Analysis” PIERS Proceedings, Guangzhou, China, August 25-28, 2014.
- [24] Xiaoshuang Xing, Tao Jing, Wei Cheng Yan Huo, Xiuzhen Cheng “Spectrum prediction in Cognitive Radio Networks” IEEE Wireless Communications > Volume: 20 Issue: 2.
- [25] Senthuran Arunthavanathan, Sithamparanathan Kandeepan, Robin J. Evans “Reinforcement Learning based Secondary User Transmissions in Cognitive Radio Networks” 2013 IEEE GLOBECOM workshops (gc wkshps)
- [26] Tsagkaris, K., Katidiotis, A., Demestichas, P. (2008). “Neural network-based learning schemes for Cognitive Radio systems” International Journal on Computer Communication, 31, 3394–3404.
- [27] Katidiotis, A., Tsagkaris, K., & Demestichas, P. (2010). Performance evaluation of artificial neural network-based learning schemes for Cognitive Radio systems. Computers & Electrical Engineering, 36(3), 518–535.
- [28] Mithra Venkatesan, Dr.Anju.V.Kulkarni and Dr.Radhika Menon “Artificial Neural Network based Learning in Cognitive Radio” Proc. of Int. Conf. on Advances in communication, Network, and Computing, 2013.
- [29] Hiremath, S. & Patra, S. K. (2011). Transmission rate prediction for Cognitive Radio using adaptive neural fuzzy inference system. Military Communications Conference 2011, MILCOM 2011, November 7–10.

- [30] Mithra Venkatesan, Anju Vijay Kumar Kulkarni  
“Comparative Performance Investigation of Supervised and Unsupervised Learning Outlines applied in Cognitive Radio Systems” *Wireless Personal Communication* December 2016, Volume 91, Issue, pp 1393–1417.
- [31] Tsagkaris, K., Bantouna, A., & Demestichas, P. (2012). Self-Organizing Maps for advanced learning in Cognitive Radio systems. Elsevier, *International Journal on Computers and Electrical Engineering*, 38, 862–881.