An Efficient Spectrum Utilization Scheme Using Incremental Neural Network

Dissertation Submitted in partial fulfilment of the requirements for the award of the degree

Of

MASTER OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

Prashant Kumar Bharti

Admission No. 18MT0323

Under the Guidance of

Dr. Ansuman Bhattacharya

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology (Indian School of Mines), Dhanbad-826004

May - 2020

Indian Institute of Technology (Indian School of Mines), Dhanbad-826004

(Declared as Deemed-to-be-University U/S of the UGC Act, 1956

Vide Notification, F11-4/67-U3, Dated 18.09.1967 of Govt. of India)

DECLARATION

The Dissertation titled **An Efficient Spectrum Utilization Scheme Using Incremental Neural Network** is a presentation of my original research work and is not copied or reproduced or imitated from any other person's published or un-published work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions, as may be applicable. Every effort is made to give proper citation to the published/unpublished work of others, if it is referred to in the Dissertation.

To eliminate the scope of academic misconduct and plagiarism, I declare that I have read and understood the UGC (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations, 2018. These Regulations have been notified in the Official Gazette of India on 31st July, 2018.

I confirm that this Dissertation has been checked with the online plagiarism detector tool Turnitin provided by IIT (ISM) Dhanbad and a copy of the summary report/report, showing Similarities in content and its potential source (if any), generated online through Turnitin is enclosed at the end of the Dissertation. I hereby declare that the Dissertation shows less than 10% similarity as per the report generated by Turnitin and meets the standards as per MHRD/UGC Regulations and rules of the Institute regarding plagiarism.

I further state that no part of the Dissertation and its data will be published without the consent of my guide. I also confirm that this Dissertation work, carried out under the guidance of Prof. Ansuman Bhattacharya, Assistant Professor, Department of CSE, has not been previously submitted for assessment for the purpose of award of a Degree either at IIT (ISM) Dhanbad or elsewhere to the best of my knowledge and belief.

(Prashant Kumar Bharti)

Admission No. - 18MT0323
M.Tech (Computer Science and Engineering – Specialized in Information Security)
Department of CSE
IIT (ISM), Dhanbad

(Ansuman Bhattacharya)

Supervisor
Department of CSE
IIT (ISM), Dhanbad

CERTIFICATE

This is to certify that Mr. Prashant Kumar Bharti (Admission No. 18MT0323), a student of M.Tech. (Computer Science and Engineering - Specialized in Information Security), Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad has worked under my guidance and completed his Dissertation entitled **An Efficient Spectrum Utilization Scheme Using Incremental Neural Network** in partial fulfilment of the requirement for award of degree of M.Tech. in Department of Computer Science and Engineering from Indian Institute of Technology (Indian School of Mines), Dhanbad.

This work has not been submitted for any other degree, award, or distinction elsewhere to the best of my knowledge and belief. He is solely responsible for the technical data and information provided in this work.

(Ansuman Bhattacharya)

Assistant Professor and Guide
Department of CSE
IIT (ISM), Dhanbad

FORWARDED BY:

(Haider Banka)
Head of the Department
Department of CSE
IIT (ISM), Dhanbad

Acknowledgement

I am deeply obliged to my supervisor Prof. Ansuman Bhattacharya, Assistant Professor, Department of Computer Science and Engineering Indian Institute of Technology (Indian School of Mines), Dhanbad for his treasure's guidance. I am always grateful for the moral support. I always regard him towards his continuous assistance and motivation that derived me to move ahead in the research activity.

I express my gratitude towards Prof. Haider Banka, Head of the Department of Computer Science Engineering, Indian Institute of Technology (Indian School of Mines) Dhanbad for his valuable support while working on my project work. I am grateful towards all academic members for guiding me with their tremendous knowledge. I appreciate each and every member of Computer Science Engineering department for helping me during my research work.

I am thankful to my seniors and beloved mates for their support.

I am gratitude to my family members believing in me. They stood by me every walk of life. I sincerely thank to my parents for their tremendous support.

(Prashant Kumar Bharti)

M.Tech. (Computer Science and Engineering –
Specialized in Information Security)
Computer Science and Engineering
Admission No. 18MT0323

Date: *May 2, 2020*

Place: IIT(ISM) Dhanbad

Abstract

Rapid progress in radio technology has culminated in the production of something like a smart system classified as Cognitive Radio (CR). This form of radio harnesses popular Artificial Intelligence tactics to create an outstanding dynamic behaviour. The cognitive radio network (CRN) is regarded to be amongst the potential ways to tackle the challenge of spectrum shortage and optimal use of spectrum. Within a CRN the Secondary User (SU) is permitted to own the spectrum that the Primary User (PU) is currently not utilising. The CRN is a wireless communication network of the next era that enables SUs to own the underused or unoccupied spectrum, defined as white spaces, in authorized spectrum with minimized interference to PUs. This thesis contributes a method that implements cognitive radio using a machine learning technique, that is, an incremental neural network combined with reinforcement learning. By this method we could achieve efficient utilization of spectrum without any interference to PUs.

Contents

List of Tables

List of Figures

List of Equations

Chapter 1

1 Introduction

As the development in wireless communication has increased, the use of mobile devices in modern times has also increased tremendously. With the rapid development of IoT (Internet of Things), substantial growth of smart devices is anticipated in the coming years. This rising number of smart devices requires a huge proportion of frequency to help. But the accessible spectrum is a sparse resource. If we search the existing spectrum allocation table, it is very difficult to find available spectrum to accommodate the incoming wireless system volumes [1] and cellular data congestion [3].

There comes Cognitive Radio (CR), which is a smart system that can be dynamically designed. It is a wireless communication system of the next generation that permits free users to access and own unused authorised spectrum or white spaces owned by authorised users to optimise the use of the spectrum [4]. CR is a system launched to invade the coming issue of spectrum jam. CR users are unlicensed users who intelligently find underused licensed spectrum for their use, without interfering with licensed users. The CR technology improves bandwidth availability at each unlicensed user, thereby enhancing the performance of the unlicensed user network. Few of CR's current methods are either complicated, requiring high computing power to determine the idle spectrum or failing to focus on real-time spectrum resources [1].

1.1 Related work

Several conventional machine learning approaches are implemented in CR for spectrum sensing, spectrum sharing, spectrum decision, and spectrum mobility [5]. Methods based on different artificial intelligence approaches like genetic algorithms [12], support vector machine [4], game theory [9], decision tree, fuzzy logic, multi-agent systems [3], Bayesian, case-based reasoning, neural networks [5], Markov model [2] [3], artificial bee colony algorithm and reinforcement learning [1] [2] which can increase the efficiency and accuracy in all applications of CR [10].

These types of traditional schemes are useful when we know about the environment and hence can apply supervised learning algorithms. In our case, we are considering unsupervised learning to focus on unknown environments, and hence such an algorithm is needed, which can do better sensing and efficient spectrum utilization even if the environment is not known. Currently, CRN based on unsupervised learning is one of the very active research areas. A number of researchers have previously worked in this field [7] [12] and have proposed a Q-learning algorithm, a reinforcement learning algorithm, for spectrum sensing where an optimal matrix is calculated by Q-learning algorithm showing the strength of the PUs signal. The Q-Learning algorithm generates an exact matrix for the working agent to optimize its reward in the long run [12]. Though this approach in itself is not wrong, this is reasonable only for small environments and gradually loses its feasibility as the number of states and actions in environmental rises. The solution of this problem arises from the fact that the values in the matrix are of relative significance only, i.e., the values only have importance with respect to the other values. Thus, this thinking leads us to the integration of IANN combined with Q-Learning, which uses an artificial neural network to approximate/optimize the values. This optimization of values does not hurt as long as it maintains relative significance. This thesis proposes a method, which leads to giving more accurate results on spectrum utilization.

1.2 Objective of the Thesis

In this dissertation, a new approach proposed for spectrum sensing to enhance signal detection and efficiently utilize the under-utilized spectrum or whitespaces. The approach uses IANN combined with Q-learning, where IANN does sensing on the spectrum. The input is fetched to the IANN incrementally, i.e., one by one manner and detection is performed. As our environment is unpredictable, we need to have an unsupervised machine learning algorithm, and so we need incremental learning.

The term “Incremental” refers to the processing of the data incrementally. So, by incremental neural network means sensing the environment and feeding data into neural network one by one. The method is discussed in Chapter 3 in detail.

1.3 Organization of the Thesis

The remaining thesis is conveyed according to the arrangement given below:

In chapter 2: We discuss the literature survey which is useful in our work.

In chapter 3: We discuss the detailed process of our proposed method of utilization of spectrum using IANN.

In chapter 4: We implement our work, and the results are shown.

In chapter 5: Conclusion of the thesis is discussed.

Chapter 2

2 Review of Literature

In this chapter, first we discuss how spectrum utilization happens in CR. Next, we discuss how the Q-learning algorithm, which comes under a reinforcement learning technique, is applied in CR to sense whitespaces. At last, we discuss ANN and what is Incremental ANN, which is useful for the enhancement of signal detection.

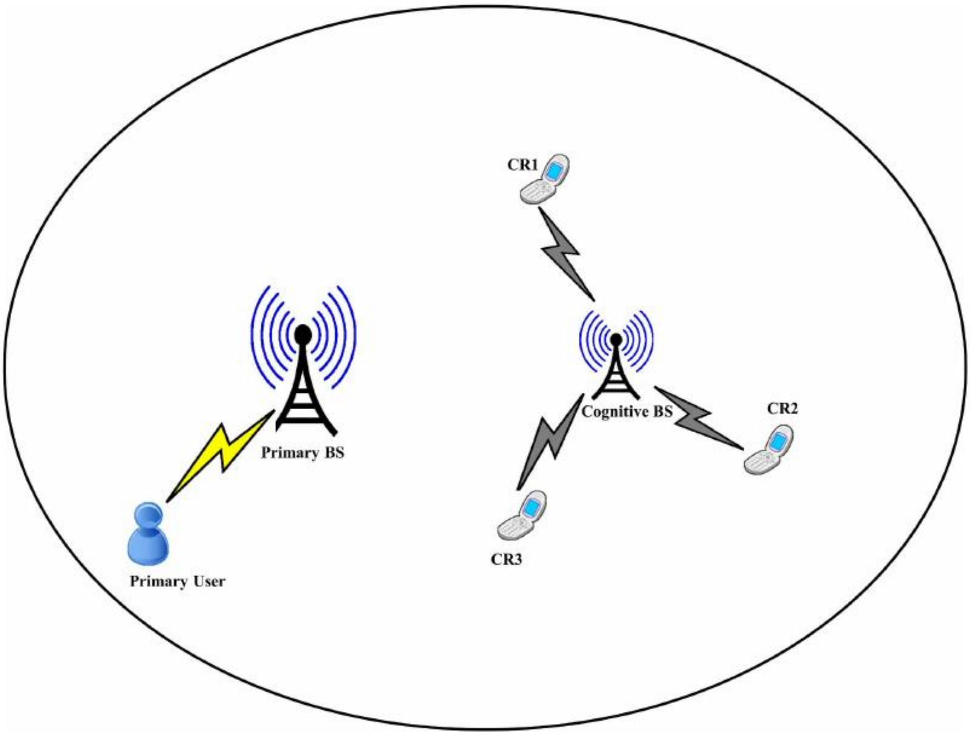
2.1 Cognitive Radio Technology

Cognitive radio is a radio that can be dynamically programmed and designed to use the optimised wireless networks in its area to prevent interference and disturbance between users. This radio identifies channels available in the radio spectrum, and adjusts its parameters of propagation or transmission to enable more concurrent radio communications at one location in a provided spectrum range [4].

While previously considered cognitive radio as a software-defined radio implication, most research tends to focus on spectrum-sensing. The key challenge in cognitive radio spectrum detection is the design of high-end hardware and methodologies to start sharing spectrum sensing data among networks. A basic energy detector has been shown to be unable to guarantee correct observation of signal existence, requiring for more advanced spectrum sensing methods and needing frequent exchange of spectrum sensing data among nodes [3] [14].

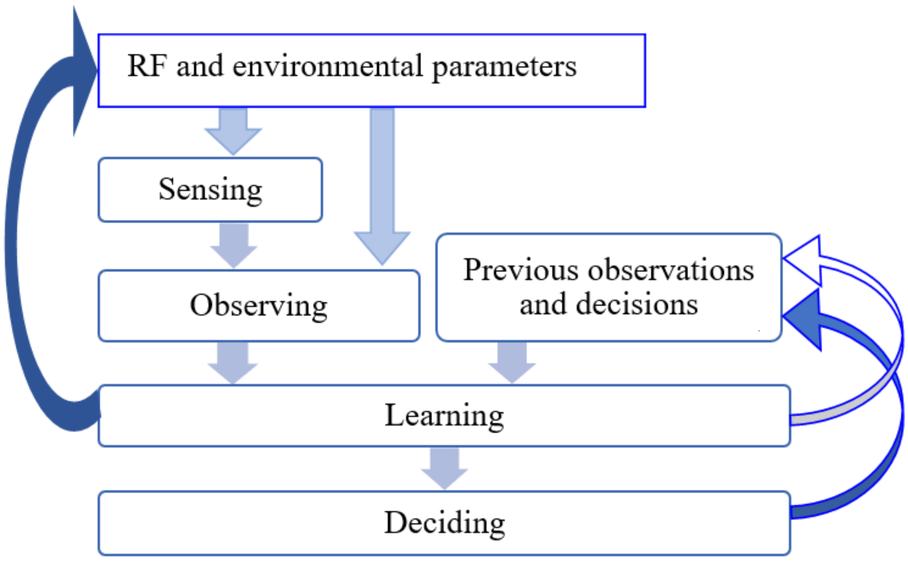
The idea of CR is a promising technology from a technical perspective to achieve effective use of RFS. In 1999, Joseph Mitola presented the principles for CR models [10]. In these models, operators allowed to use different frequency bands are known as Primary Users (PUs). Whereas, Secondary Users (SUs) are known as participants who are not approved. The CR model is based on the assumption that leaves parts of its spectrum unchanged, and the PU may not make full use of its permitted bands. Such unoccupied white spaces, or gaps, are connected in terms of frequency, time, and space [6]. An SU can use such holes in contrast to the unlicensed bands usually used by it [13].

Figure 2-1: Cognitive Radio System.



CR provides an insight into the radio network to ensure highly secure contact with effective radio spectrum use. In order to make CR truly intelligent, the learning ability must also be present. Learning means that the real behaviour would be based on past and present environmental observations [10]. Artificial intelligence technology like Reinforcement Learning (RL), Neural networks, and many more are applied to CR, which can increase efficiency in the diverse quality of CR applications like spectrum detection and spectrum utilization. RL is a machine learning field that deals with how software agents can take action in an unknown environment to optimize some cumulative reward notion. With the RL approach, many representations may be needed, including state and action, discounted and delayed rewards [4].

Figure 2-2: Cognitive cycle.



Whereas, an Artificial Neural Network (ANN) functions like how the brain of humans interprets data, which includes a large number of neural networks that function together for information processing [8] [11]. On the basis of CR, this thesis represents an enormous enhancement of these representations through the Machine learning algorithms. Q-learning can be paired with Incremental Artificial Neural Network (IANN) to further improve signal detection for effective spectrum utilization.

2.2 Introduction to Artificial Intelligence

Artificial intelligence seeks to allow systems to perform tasks similar to a specialist. The smart system perceives its atmosphere and takes actions to optimize its functionality. The core issues in artificial intelligence involve inference, logic, solution, representation of information, and able to learn [10].

2.2.1 Reinforcement Learning

Reinforcement learning is an impressive concept for unsupervised machine learning, which has attracted more and more attention in both academia and industry. With RL, agents can learn what actions should be taken to deliver maximum reward without depending on descriptors, right actions acknowledged by authoritative external supervisor. Rather, agents need to carry out a range of actions to gain reward information. Moreover, every action must be tried multiple times to acquire rewarding information related to the various states. By utilizing cumulative information of reward, agents take the actions required to receive maximum rewards [4]. RL is a useful tool for addressing quality improvisation issues. By quality improvisation, we mean the question of knowing the finest action to optimize some objective function in the state visited by the program. Usually, RL is used if the network has a massive number of states (more than 1000) and a robust probabilistic design that is not appropriate for analysing the closed type. If networks have a limited number of states, and the fundamental natural design is fairly simple, then dynamic computing can be used [2].

Q-learning is a form of RL, frequently used in different applications because of its model-free structure. The model-free structure allows agents to learn optimal policy of action directly by communicating with environments rather than analysing environmental structures, such as the likelihood of change [7]. Even though Q-learning employs an incremental method to adjust each state's Q-value and every action's Q-value, the trade-off among exploitation and exploration is a major issue that needs to be resolved.

2.2.2 Q-Learning Working Principle

The core concept in Q-Learning is to identify or learn the optimal action through trial and error in any state the program encounters (also called an optimal policy). Throughout the actual-world environment (usually observed in artificial intelligence) or in a simulation (usually shown in managerial science / industrial engineering); the trial and error process may be applied [7].

The agent takes action, collects feedback, *f*, and uses feedback to update their repository. The agent preserves a so-called Q-factor to its repository for every state-action pair. If the feedback for choosing an action in a state is favourable, the value of the associated

Q-factor is raised whereas if the feedback is unfavourable the value is diminished. The feedback is the direct gain or reward plus the next-state value [12].

Figure 2-3: Flow diagram of the connected component labelling algorithm.

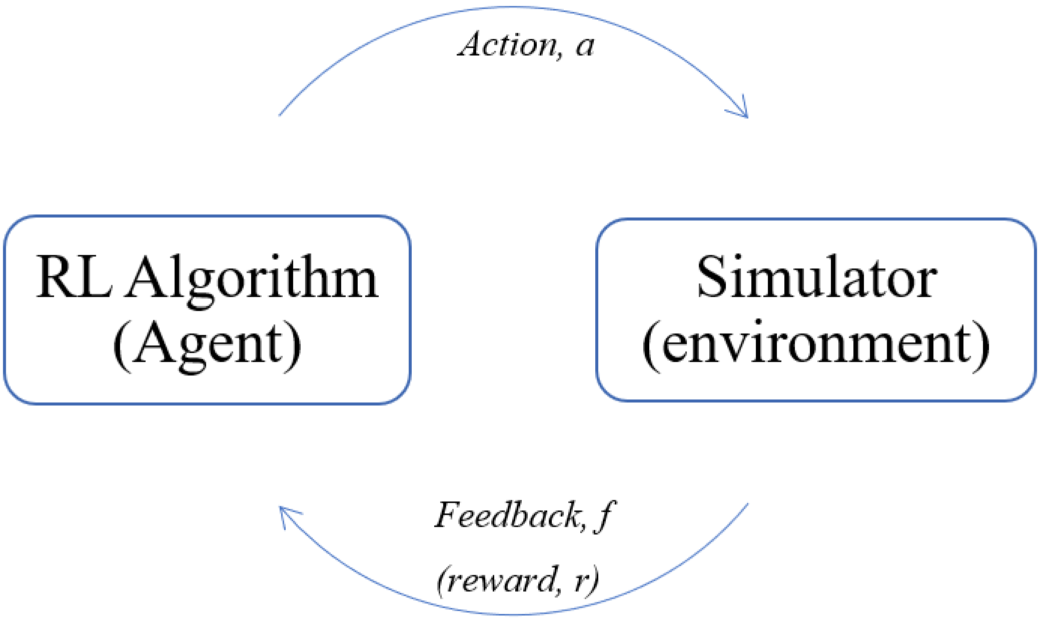


Fig. 2.3 shows RL process of trial and error. The RL Agent selects an action and is supplied to the simulation. The action is then simulated by the simulator and fed the resulting feedback back into the base of knowledge of the agent (Q-factors). The agent upgrades its base of knowledge using the RL algorithm, grows way more intelligent in the trial phase and afterwards selects a better action [10].

Here, we denote reward, *r*, as the instant reward, $r\ (i, j, a)$ where *i* becomes the current state, *a* is the selected action of the current state, and *j* would be the next state. In that state, the value of every state is given by the highest Q-factor. So, if two actions come into effect in any state, the state value is the maximum of the two Q-factors in that case.

Equation 2-1

$$f = r\ (i, j, a) + \lambda\ (\max Q\ [j,b]),$$

where the discount factor is λ.

The equation for Q-factor update used by Q-Learning algorithm is shown here:

Equation 2-2

$$Q\,[i, a] \mathrel{+}= \{1+ \alpha\}\; Q\,[i, a] + \alpha\,\{\, f\,\},$$

$$\Rightarrow \qquad Q\,[i, a] \mathrel{+}= \{1+ \alpha\}\; Q\,[i, a] + \alpha\,\{r\,(i, j, a) + \lambda\,(\max Q\,[j,b])\}$$

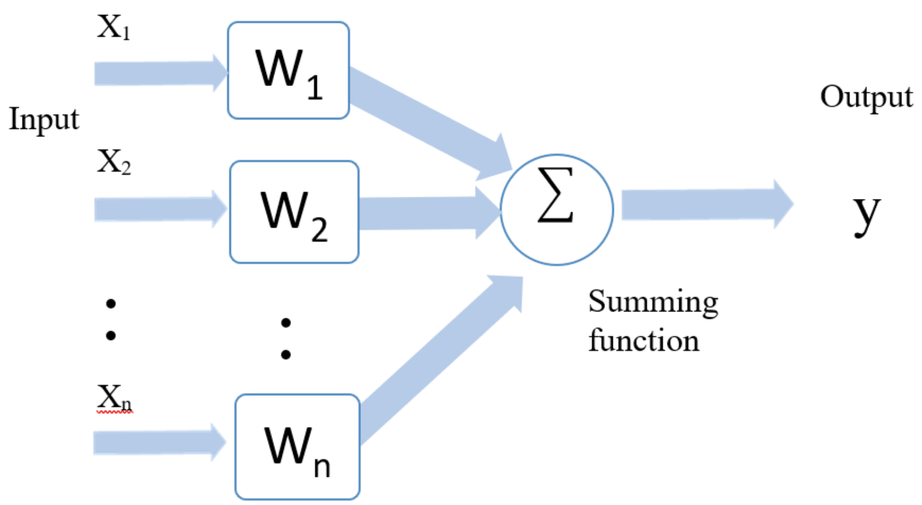
where the learning rate or the step size is α.

2.2.3 Artificial Neural Network

The biological nervous system in the human body, which structures the learning behaviour of the human brain, is the concept that is behind the formulation of ANN. By using the data and learning information, ANN models try to identify patterns and relations. Billions of neurons in the human brain gather knowledge to help the brain learn from the collected information. Similarly, the function of ANN is to simulate lots of interconnected artificial neurons in a computer so that the model can learn things, make decisions, and give predictions following the human way. It can learn from data and make decisions, reproduce, and model nonlinear processes. ANN technology has attracted researchers and scientists from a wide range of disciplines for predictions and decision making in various applications [8].

A general structure of ANN with multiple inputs and one output is shown in Figure 2.3.

Figure 2-4: Neuron structure.



With the help of Summing function (transfer function) all the weighted inputs are summed up. If we denote transfer function i.e. output as y,

Equation 2-3

$$y = \sum_{i=0}^{n} w(i)\, x(i),$$

where the i^{th} weight of neuron is $w(i)$, and the i^{th} input is $x(i)$.

There are two types of algorithms primarily available for ANNs, Feed Forward Propagation and Backpropagation. The backpropagation algorithm aims at minimizing the error by updating the weights and influences the predicted output to get closer to the target value of the output.

In our study, we use the backpropagation algorithm with an Incremental neural network. In the next chapter, we discuss this in detail.

Chapter 3

3 Proposed Method

Neural networks are really effective for learning non-linear functions, and could adjust to non-linear structures of primary signals. Keeping that in mind, the proposed sensing scheme utilizes an ANN to improve spectrum sensing. We have seen that there are several schemes for spectrum sensing using Q-learning and ANN approach. In our approach we have combined these two methods and it was seen this approach enhances the primary signals better. Now we see how we have done this.

3.1 Proposed Scheme

This segment involves discussion of the working of each phase of the proposed scheme where we apply Incremental Artificial Neural Network (IANN) combined with Q-learning in CRNs, which detect the signal strength of the primary signals more efficiently.

With the Q-learning approach, a Q-factor is calculated for each state-action pair. In this approach, if the number of state-action pairs are massive, it's not possible to separately store each Q-factor. It is then sensible to store the Q-factors in a single neural network for a given action. When a Q-factor is required, it is extracted from its neural network. After upgrading a Q-factor, the latest Q-factor would be used to update the neural network, i.e., backpropagation. *Q (i, a)* denotes the state function *i*, for any given action, *a*. So, in what follows, we'll call it a Q-function.

Neural networks typically consist of two kinds: incremental updating or batch updating. The incremental neural networks take one part of data at one time, while batch updating involves all of the data in one go. We need incremental neural networks for RL because each moment the agent gets feedback, we get a new part of data which should be used to upgrade other neural networks.

Neurons are generally used for linear fitting models. Back-propagation is an approximation of universal function, and should preferably fit any of the Q-function. Neurons can also be used for part-wise fitting of the Q-function, where a linear fit is applied to each part of the data.

3.2 Algorithms for our Scheme

The objective of our algorithm is to have enhanced detection of PUs and potential availability of the spectrum, i.e., whitespaces, for SU. This would help to achieve efficient utilization of spectrum.

The following algorithms are our proposed algorithms which is divided into three parts:

- **Algorithm 1 Policy:** It denotes the policy function, which decides next the state to be taken and reward is calculated by given state-action pair. Here, the agent has two choices of action: either occupy the channel (*taken*) or not-occupy the channel (*not-taken*). The reward is calculated based on an action taken by the agent.
- **Algorithm 2 Q-update:** It denotes the updating equation, which is the core of the Q-Learning algorithm, as we have already discussed in chapter 2, *Equation 2-1* and *Equation 2-2*.
- **Algorithm 3 Incremental Neural Network:** It denotes the IANN function, which returns the Q-matrix denoting the availability of whitespaces and the detection of signal strength of PUs.

Our main algorithm is **Algorithm 3**. The spectrum is divided into the number of channels, which is given as *input* to this algorithm in the form of a square matrix. In this algorithm, weights of the neural network are initialized, and then transfer function, *y*, is calculated which is actually *output* here. A function call to **Algorithm 2** gives weight update, which is done by the Q-learning algorithm. A function call to **Algorithm 1** gives policy.

Algorithm 1 Policy:

Input: State and action for a channel /**function call from* ***Algorithm 3****/

Output: Next state, Reward /*reward is calculated based on given inputs*/

```
1.  function policy (state, action)        /*action have two options: taken or not-taken*/
2.      if state > 0
3.          s = state
4.      else
5.          s = -1
6.      end if
7.      if state > 0 and action == taken
8.          reward = maximum               /*maximum reward is given*/
9.      else state > 0 and action == not-taken
10.         reward = minimum               /*minimum reward is given*/
11.     else
12.         reward = 0
13.     end if
14. return s, reward
```

Algorithm 2 Q-update:

Input: Q_{old}, Q_{next}, Reward /**function call from* ***Algorithm 3****/

Output: Q-factor /**Returns updated Q-factor i.e.* Q_{new}*/

1. **function** Q_update (Q_{old}, Q_{next}, reward)
2. **return** $(1- \alpha) * Q_{old} + \alpha * (reward + (\lambda * Q_{next})$

where the discount factor is λ and the learning rate is α.

Algorithm 3 Incremental Neuron Network:

Input: State and action for all channels *`/*action is either take or not-take*/`*

Output: Q-matrix *`/*the matrix denoting available whitespaces and signal strength of PUs*/`*

```
function SensingWithIANN(state, action)
    Initialize weights of the neuron to small random numbers
    Set the number of iterations as Kmax
    while Kmax != 0
        Take i as the state and j as the next state
        Randomly choose a as an action
        for pair (i, a), calculate Q            /*Q-factor*/
            Q_old = (w[1][a] + w[2][a])i
        for j associated to each a, calculate Q            /*Q-factor*/
            Q_next1 = (w[1][1] + w[2][1])j
            Q_next2 = (w[1][2] + w[2][2])j
        Get Q_next = maximumof {Q_next1, Q_next2}            /*Evaluate maximum Qnext*/
        [st, reward] = policy(state[i], action[a]) /*Get next state, reward from Algorithm 1*/
        if st == -1
            Q_new = 0
        else
            Q_new ← Q_update (Q_old, Q_next, reward)       /*Get Q_new from Algorithm 2*/
        end if
        while Mmax != 0
            for s = 0 to length of state
                w[s][a] ← μ*(Qnew – Qold) * state[s]
            end for
            Decrement M_max by 1
        end while
        Increment i by 1
        Decrement K_max by 1
    end while
    Q-matrix = w.max (1)
return Q-matrix
```

Chapter 4

4 Implementation and Results

Simulations for this research work has been performed on Spyder 4.0.1 with Anaconda Navigator 1.9.7 on the python programming language to give clear picture of the feasibility of the work we are proposing. It was shown that the proposed approach achieves higher spectrum usage efficiency than the conventional approach (Refer to table 4.1). Here, the spectrum is divided into some number of channels in such a way that a square matrix is created where every element represents a signal strength of a PU. Elements, where strength is greater than zero indicates the presence of PUs and elements where strength is smaller than or equal to zero indicates no signal strength or no activity and that could be used for SUs. Our algorithm was run from 2 channels (1x1 dimensions) to 100 channels (10x10 dimensions). However, we would compare our work with the conventional approach with 64 channels (8x8 dimensions). The discount factor λ and the learning rate α are selected as 0.6 and 0.7, respectively. The results of the implementation are provided in Table 4.1 and are reasonably acceptable.

For better understanding the resultant tables, see the respective values of input to the output in matrices. The zeros in the output matrix indicate the whitespaces, whereas the positive values are enhanced to a higher frequency status in order to reflect the signal and can be detected easily. When a signal is properly interpreted, it can be detected without fault.

Thus, the findings from tables indicate that all possible CR signals can be detected using the IANN algorithm combined with the Q-learning model.

Table 4-1: Comparing Q-learning technique and proposed work.

Input in the form of square matrix of the spectrum (channels)	-3	-9	2	-100	0	-300	0	-1
	-20	-50	-60	7	-4	7	10	-6
	-70	-20	-1	100	-40	50	8	-9
	-3	7	3	-9	0	-50	10	-30
	9	-30	-22	70	-55	10	7	-20
	-44	40	-55	-88	60	10	60	-43
	-12	-32	33	7	-77	50	10	-66
	50	0	0	10	78	-22	80	55

Output of the Q-matrix (whitespaces) formed by traditional Q-learning algorithm in CRN	0	0	21	0	14	0	12	0
	0	0	0	11	0	20	21	0
	0	0	0	100	0	62	20	0
	0	11	22	0	14	0	21	0
	12	0	0	71	0	23	19	0
	0	42	0	0	71	23	69	0
	0	0	51	11	0	62	21	0
	52	4	20	14	88	0	88	70

Output of the Q-matrix (whitespaces) formed by proposed algorithm in CRN	0	0	402	0	0	0	0	0
	0	0	0	217	0	139	295	0
	0	0	0	1727	0	980	158	0
	0	441	277	0	0	0	578	0
	330	0	0	789	0	434	451	0
	0	291	0	0	671	547	883	0
	0	0	726	348	0	579	437	0
	543	0	0	255	360	0	983	387

Chapter 5

5 Conclusions

This thesis concludes that the Incremental Artificial Neural Network (IANN) algorithm, when combined with the Q-learning, optimizes the efficiency of the Cognitive Radio (CR) in signal detection of Primary Users (PUs). The algorithm, *Q*-learning, learns from the data sent to the program using the reward and penalty concept, which creates a *Q*-factor, and ANN uses that for updating weights by incremental means. By this method, we get the enhanced signal strength of PUs, and hence we get the available whitespaces correctly. Then we can allocate these to the Secondary Users (SUs) without having any interferences from PUs. This scheme utilizes the spectrum very efficiently.

This research work helps to detect the presence of PUs successfully. The future scope of the current thesis would include visual examples and practical implementations of our proposed approach.

References

[1] Cottis, A. V. (2016). "A reinforcement-learning based cognitive scheme for opportunistic spectrum access". *Wireless Personal Communications, 86*(2), 751-769.

[2] Gosavi, A. (2009). "Reinforcement Learning: A Tutorial Survey and Recent Advances". *INFORMS Journal on Computing*, 178-192.

[3] J. Lundén, S. R. (2013). "Multiagent Reinforcement Learning Based Spectrum". *IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, 7*(5), 1932-4553.

[4] K.-L. A. Yau, G.-S. P.-R. (2014). "Application of Reinforcement Learning in Cognitive Radio Networks: Models and Algorithms". *The Scientific World Journal*.

[5] Kansal, R. S. (2016). "ARTIFICIAL NEURAL NETWORK BASED SPECTRUM RECOGNITION IN COGNITIVE RADIO". *IEEE*.

[6] M. KULIN, T. K. (2018). "End-to-End Learning From Spectrum Data: A Deep Learning Approach for Wireless Signal Identification in Spectrum Monitoring Applications". *IEEE*, *6*, 2169-3536.

[7] M. Li, Y. X. (2009). "A Q-Learning Based Sensing Task Selection Scheme for Cognitive Radio Networks". *IEEE*.

[8] M. R. Vyas, D. P.-B. (2017). "Artificial Neural Network Based Hybrid Spectrum Sensing Scheme for Cognitive Radio". *IEEE*.

[9] Muthukkumar, R., Manimegalai, D., & Santhiya, A. S. (2015). "Game-theoretic approach to detect selfish attacker in cognitive radio ad-hoc networks". *3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, 1-5.

[10] N. Abbas, Y. N. (2015). "Recent advances on artificial intelligence and learning techniques in cognitive radio networks". *EURASIP Journal on Wireless Communications and Networking, 2015*(1), 174.

[11] Ravinder, R. G. (2018). "Artificial Neural Network Based Approach for Spectrum Sensing in Cognitive Radio". *IEEE*.

[12] Reddy, Y. B. (2008). "Detecting Primary Signals for Efficient Utilization of Spectrum Using Q-Learning". *in Fifth International Conference on Information Technology: New Generations.* Grambling, LA.

[13] S. Arunthavanathan, S. K. (2013). "Reinforcement Learning based Secondary User Transmissions in Cognitive Radio Networks". *in Globecom 2013 Workshop - Broadband Wireless Access.* Melbourne, Australia.

[14] Y.-J. Tang, Q.-Y. Z. (2010). "Artificial Neural Network Based Spectrum Sensing Method for Cognitive Radio". *IEEE*.