DYNAMIC SPECTRUM ACCESS USING DEEP REINFORCEMENT LEARNING IN DISTRIBUTED SYSTEM

Enrolment no -- **9916103181, 9916103206**

Name of Students -- **Harshit Jain, Shivam Jolly**

Name of Supervisor -- **Mr. Himanshu Agrawal**

May – 2020

Submitted in partial fulfilment of the Degree of

Bachelor of Technology

In

Computer Science Engineering

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING & INFORMATION TECHNOLOGY

JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY, NOIDA

TABLE OF CONTENTS

DECLARATIONS

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Place: Noida
Date: May 20, 2020

Signature:
Name: Harshit Jain
Enrollment No: 9916103181

Signature:
Name: Shivam Jolly
Enrollment No: 9916103206

CERTIFICATE

This is to certify that the work titled "**Dynamic Spectrum Access using Deep Reinforcement Learning in Distributed System**" submitted by "**Harshit Jain, Shivam Jolly**" in partial fulfillment for the award of the degree of B. TECH of Jaypee Institute of Information Technology, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor :

Name of Supervisor: Mr. Himanshu Agrawal

Designation: Mentor

Date: May 2020

Acknowledgment

We would like to take this opportunity to thank everyone who enabled us to create and execute our minor project. We would like to show our gratitude to **Mr. Himanshu Agrawal** for giving us a good guideline for the project throughout numerous consultations. We would also like to expand our deepest gratitude to all those who have directly and indirectly guided us in this project.

We would like to extend our thanks to our classmates and team members itself, for their insightful comments and constructive suggestions to improve the quality of this project work.

Harshit Jain (9916103181)

Shivam Jolly (9916103206)

DATE: May 20, 2020

SUMMARY

Dynamic spectrum access (DSA) is an effective tool and a powerful technology when it comes to sharing radio spectrums among different networks. For a secondary user i.e. someone who utilizes the band for free, the DSA device will face the problem of preventing interference with the primary user and thus reducing collisions. This problem becomes even more significant in a distributed DSA because there is no centralized controller present in this type of network to control SU. In this project, we devise a model for communication in a distributed DSA network when spectrum sensing errors are present. We apply a machine learning approach i.e. deep reinforcement learning (DRL) so that SU can sense the spectrum band and collect information given that it has no clue of the underlying system settings. Furthermore we utilize a type of recurrent neural network i.e. reservoir computing (RC) to realize DRL by using the present temporal correlations in the network. Using our model the SU can devise strategy relying solely on its past and present sensing outcome. Our experimental results suggest that the RC based spectrum access model can help the SU to reduce collision significantly and converge faster than other models when the channel number is large.

LIST OF FIGURES

LIST OF ABBREVIATIONS

S. No.	Name	Full-Form
1	DSA	Dynamic Spectrum Access
2	PU	Primary User
3	SU	Secondary User
4	DRL	Deep Reinforcement Learning
5	MLP	Multi-Layer Perceptron
6	DQN	Deep Q-Network
7	RC	Reservoir Computing
8	ESN	Echo State Network

1. INTRODUCTION

1.1. General Introduction

As years are cruising by, innovations are being made. Moreover, consistently a great many of these gadgets are being included in the media transmission world. These remote gadgets take a shot at the premise of range groups. The range is not the same for every nation. Be that as it may, as the gadget check is expanding, numerous issues are being identified with them such as inadequate usage of data transfer capacities, impact between clients, and so on. For a decade, numerous new techniques have been created to handle these issues, and research is going on to identify the best way to expand range productivity. There is gigantic undiscovered potential in this field. Many nations are presently focussing on this area for financial development. Numerous researches have shown that if the spectrum utilized in a much better way by the economy, it can boost the GDP of the economy of a country. Various technologies are being introduced in this area to help meet this objective. At present day, 5G has been introduced by the companies that provide its users faster speed as compared to its predecessor. But it is still not feasible by many countries due to factors like location, underdeveloped towers, or spectrum efficiency. As, these technologies being introduced, the system is becoming more and more complex. Due to this, spectrum efficiency is decreasing; hence it becomes very important to utilize reinforcement learning in this scenario.

1.2. Problem Statement

For an unlicensed user, the device in the spectrum band will face the critical problem of avoiding interference with primary users and transmit traffic according to the state of the bandit senses. This occurs because there is no centralized controller in the system. So in this project we work on a new strategy to help a SU reduce collision and improve convergence.

1.3. Significance of the Problem

The worldwide portable information traffic has encountered a rapid increase which is likely to continue shortly. Spectrum expansion is required to curb this growth. In any case, the radio spectrum is an expensive and rare asset and is even harder to access due to the restrictions placed by the government. Also trial tests and examinations from third party organizations show that the static spectrum portion arrangement from the Federal Communications Commission (FCC) causes the underutilization of allotted authorized groups. This has persuaded them to re-evaluate the present static spectrum assignment arrangement and utilize dynamic spectrum get to (DSA) to advance spectrum uses.

1.4. Brief Description of the solution approach

In this project, we use machine learning to obtain results for the DSA network. Deep reinforcement learning (DRL) is used to reduce the control overload to enable SU to learn in a distributed way assuming no awareness of the underlying system. Also reservoir computing (RC) is used to perform DRL by exploiting the temporal correlations within the DSA network. To be precise a DSA scheme utilizing DRL and RC is developed to help SU access the channel while protecting PUs from interference and collisions.

1.5. Approach to the problem in terms of Technology

1.5.1 Machine Learning
1.5.2 Reinforcement Learning
1.5.3 Deep Reinforcement Learning
1.5.4 Reservoir Computing

1.6. Comparison of existing approaches to the problem framed

- Many performance evaluations of the introduced system have been conducted to show that our approach can learn PU patterns quickly and reduce collisions significantly in the created environment.
- Our scheme converges faster than the Q-learning scheme for a larger number of channels.
- Compared with DRL+MLP, our DRL+RC strategy utilizes an underlying temporal correlation to yield substantial performance improvement.

2. Literature Survey

2.1. Summary of papers studied

Paper 1

Title of the paper	**Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks**
Authors	Yiding Yu, Taotao Wang, Soung Chang Liew
Year of Publishing	2018
Web Link	https://ieeexplore.ieee.org/document/8422168
Summary	This paper works on state-of-the-art findings of deep reinforcement learning and integrates it with the advances of the latest Multiple Access Control (MAC) protocols. This work is executed with the inspiration that a time will come when distinctive networks will share a spectrum dynamically primarily based on instant supply and demand. This layout will assist in the usage of assets related to time more optimally with the help of numerous observations and moves, without mainly understanding the mechanisms of different networks. In this take a look at, MAC protocol which is referred to as deep – reinforcement mastering a couple of getting admission to (DLMA) protocol and the node running DMLA is taken into consideration as a DRL agent. In this, the problem of sharing time slots is taken among diverse wireless channels. DLMA can acquire more than one objectives and the first one we're focussing is maximizing the summation of throughputs of all networks. The results display that, DLMA can generate close to – top of the line answer while as simultaneously coexisting with TDMA network, ALOHA network, and a mix of them without having previous understanding. The study additionally suggests that why DRL as opposed to its old counterparts gives speedy convergence and robustness. Robustness is important as the machine does no longer has the expertise of the kind of networks present in

	the system. MAIN CONTRIBUTIONS • The use of DRL for DLMA, which is a MAC protocol for wireless networking. The effects indicate to us that it offers us close to ideal answers while not having any earlier expertise of the community i.e. whether or not it is ALOHA or TDMA. • This additionally reveals the advantages of using DRL in networking compared to conventional RL methods. Thus, it gives extra convergence closer to the most appropriate answer and supplies us with a much better machine against not optimal parameters. These two are very essential properties for the machine. • In this, a brand new approach is found which helps in generalizing Q – learning so that extra objectives can be met. In general, this indicates that we want to separate q – learning characteristic and goal feature which is not in conventional q – learning.

Paper 2

Title of the paper	**Self-Tuning Sectorization: Deep Reinforcement Learning Meets Broadcast Beam Optimization**
Authors	Rubaiyat Shafin, Hao Chen, Young Han Nam, Sooyoung Hur, Jeongho Park, Jianzhong (Charlie) Zhang, Jeffrey Reed, and Lingjia Liu
Year of Publishing	2019
Web Link	https://arxiv.org/abs/1906.06021
Summary	Cellular information is increasing day by day. This is because cellular

	devices are growing at a very fast pace. This has precipitated an exponential growth in mobile traffic. Moreover, Cisco Visual Networking Index (VNI) has foretold that worldwide IP traffic will boom three-fold by 2022. To offer a high-quality experience to customers, complexity is also growing of wireless cellular networks at an alarming rate. To mitigate this issue, self-organizing networks (SON) are brought which affords functionalities which include self – optimization, self – configuration. Self – optimization approach helps in the self-tuning of parameters to reap the most beneficial solutions. Multiple Input Multiple Output (MIMO) is one of the fundamental systems for the contemporary day mobile community. Large MIMO, is wherein several antennas are present at a base station (BS) is considered as a device for 5G systems. Beamforming is an approach used by MIMO to combine numerous alerts produced antennas present in MIMO array. Sectorization is a system where a wide beam is used to cover a region of a cell site, providing the coverage of networks. At present, broadcast beam parameters are modified manually, where a network engineer visits the site and make changes. Thus, every so often these parameters are not modified for years. This provides us with a suboptimal solution. To cope with this trouble, now researchers are imposing DRL based solutions. DRL predicts accurate q – values from the action – state pairs and helps in enhancing network performance. In this paper, DRL primarily based technique is introduced for MIMO broadcasting on the way to assist in covering a big location. MAIN CONTRIBUTIONS • A double deep – Q Network primarily based machine is proposed which can optimize sector – particular beam optimally. It can also autonomously track its parameters based totally on a person's mobility and pattern. • A new factorization algorithm has been delivered for each single and multi-sector environment. In the case of multi-regional, a singular framework for computing Q – values have been added which can help in computing Q – values for distinctively selected broadcast beams.

Paper 3

Title of the paper	**Dynamic Spectrum Access in Cognitive Radio**
Authors	Zelikow Tabakovic and Sonja Grgic
Year of Publishing	2009
Web Link	https://www.researchgate.net/publication/224085427_Dynamic_spectrum_access_in_cognitive_radio
Summary	According to the author in this paper Dynamic Spectrum Strategies can be divided into three models are Dynamic exclusive use model, Spectrum commons modal, and hierarchical model. Dynamic exclusive use model keeps the fundamental structure of the current spectrum rule system consistent where spectrum bunches are approved to the organizations for a particular use. Spectrum commons model uses open sharing among peer customers without need allocation to the organization or class of customers. The hierarchical access model receives a hierarchical radio spectrum access structure with essential and optional clients where essential clients have authorized access whereas optional clients have restricted access. In this paper Cognitive Radio (CR) approach is used for two essential jobs that are exceptionally solid correspondences and compelling utilization of radio range. Steps of the cognitive cycle are spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility. Spectrum sensing: CR screens its conditions and utilization measurements for different clients to decide its null areas. Spectrum decision: CR studies the received information to decide its operational settings and parameters to transfer as much information as possible without causing interference.

	Spectrum sharing: CR achieves a balance between the amounts of information and resources transferred within the rules of the environment. Spectrum mobility: CR stops or changes the spectrum frequency of a secondary user when a primary user starts to operate to avoid interference in real-time. To secure the essential and auxiliary clients, CR needs to apply proper limitations to the authority levels. These are called Spectrum Usage Rights (SUR) and essentially of two sorts Block Edge Mask (BEM) and Power Flux Density (PFD). BEM controls authority by creating an envelope type framework in which radio discharge must persist. PFD system offers a success ratio by defining the degree of interference the neighbors may experience. The essential difference with the BEM approach is that the guideline is given based on the total power received by the user. PROS AND CONS • CR helps in achieving highly reliable communication and efficient utilization of spectrum bands than its predecessors. • BEM approach is simpler to execute but it does not consider the obstructions faced by the client. PFD approach is more helpful for client security but is difficult to implement in real life.

Paper 4

Title of the paper	**Q-learning based dynamic spectrum access in cognitive Industrial Internet of Things**
Authors	Feng Li, Kwok Yan Lam, Zhenggou Shang, Li Wang, Kanglian Zhao, and Xinggang Zhang
Year of Publishing	2018
Web Link	https://www.researchgate.net/publication/327581054_Q-Learning-Based_Dynamic_Spectrum_Access_in_Cognitive_Industrial_Internet_of_Things
Summary	Here the author has proposed a Q-learning based model for dynamic spectrum access using IoT. The heterogeneous qualities of the remote sensor system are used to improve the effectiveness of the spectrum band and collision rate. As there are many channels available at the same time, the author has proposed Q-learning to calculate the data and choose the channel with the most Q value through self-training and learning. In this paper, the key contributions presented are: • On account of the complex multitier structure present in the network deep learning is used. • To reduce the number of conflicts, a distributed DSA strategy is raised. • Numerical results show that the solution provided has better performance than the previous solutions. PROS AND CONS • Numerical results prove that the proposed algorithm has better channel accessing effects compared with traditional simplified self-access protocol and aloha method. • This method's complexity is high.

Paper 5

Title of the paper	**Dynamic Primary-Secondary Spectrum Sharing with Cellular Systems**
Authors	J.M. Peha, Rathapon Saruthirathanaworakun
Year of Publishing	2009
Web Link	https://ieeexplore.ieee.org/document/5577706
Summary	This paper considers the primary-secondary spectrum sharing network where the primary is a phone framework and sharing is permitted only if the primary can withstand the shared channel. We consider situations when primary and secondary systems cooperate and when they don't. When they share, the secondary decides whether it can transmit data without causing harmful interference. A sensor network works to ensure that the transmission would be tolerable to the primary user. The experiments show that the accuracy of the model increases as the number of primary users in the network increase. PROS AND CONS • Results show potential value in this model when the signal strength of the primary user is strong. • Also, this work shows the value of the sharing network wherein the secondary user has information about the primary user despite not cooperating with it.

Paper 6

Title of the paper	**Echo State Networks for Proactive Coaching in Cloud-based Radio Access Networks with Mobile Users**
Authors	M. Chen, W. Saad, C. Yin, and M. Debbah,
Year of Publishing	2017
Web Link	https://ieeexplore.ieee.org/document/7880663
Summary	In this paper, proactive data storing is perused for cloud radio access networks (CRANs). In the concentrated model, the baseband units (BBUs) can calculate the demands for each of the required resources and reserve them at remote radio heads. To deal with this issue, a model is defined using echo state networks. Using ESNs, the BBUs can calculate each user's resource requirements while having very little data on the state of either the system or the client. To predict each customer's needs the ESN network's memory needs are determined based on the situations. This memory need is limited to calculate the most optimum data requirement of the client using sub-linear algorithms. PROS AND CONS • Experimental results show that the proposed method yields significant results and the limit ranges up to 27.8% and 30% based on two different standard calculations. • The choice of a reasonable size of the planning dataset and the required number of ESN units for the network are the two important parameters that influence the accuracy of the forecast. We calculate the errors in the prediction for each client using two parameters that are sureness model δ and the sensible blunder type $. As δ and $ increase, the probability of error increases. This is because as δ and $ increase, the quantity of content required for forecast decreases, and thus the number of tests required to calculate error decreases. We can similarly see that the deviation between the proposed reserve and the actual requirement increases as the quantity of clients changes.

Paper 7

Title of the paper	**Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks**
Authors	S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari,
Year of Publishing	2018
Web Link	https://www.researchgate.net/publication/323302666_Deep_Reinforement_Learning_for_Dynamic_Multichannel_Access_in_Wireless_Networks
Summary	We consider a dynamic multichannel issue, where different channels seek after a joint Markov model. A customer at each time picks a channel to transmit data and gets a reward based on the result of transmission. The objective is to find a pattern that provides a reward for a long time. The issue is combined with a part discernible Markov decision process (POMDP) with dark structure components. To vanquish the challenges of these system components such as prohibitive computation, we apply DQN to manage the massive data provided that we have no earlier information. The author gives an insightful explanation of this model to show that it performs better than other methods by utilizing Whittle Index-based heuristic. It shows that the model shows a relation close to the ideal case as the complexity of the environment increases. PROS AND CONS • This DQN approach can manage a huge framework with no prior knowledge. It can discover close and ideal arrangements by studying the relationship between different framework elements. Also, as the likelihood increases, the reward increase, and thus there is more assurance in the framework. • During channel access, a client can generally see the conditions of his channel at each vacancy.

Paper 8

<table>
<tr><td>Title of the paper</td><td>Traffic Signal Timing via Deep Reinforcement Learning</td></tr>
<tr><td>Authors</td><td>Li Li, YishengLv, Fei-Yue Wang</td></tr>
<tr><td>Year of Publishing</td><td>2016</td></tr>
<tr><td>Web Link</td><td>https://ieeexplore.ieee.org/document/7508798</td></tr>
<tr><td>Summary</td><td>In this article, a large number of parameters are proposed to plan signals using deep reinforcement learning. Traffic control has always been a difficult issue for architects due to problems of physical model creation and optimization of the said model. There are generally two sorts of approaches to manage these difficulties. One such method is the flow-based model in which one portrays the visible traffic stream for different areas. But when the traffic situations turn complex the cost of demonstration and traffic update of the model increases significantly. The other method is the recreation-based method in which future traffic is predicted using AI or simulations but this method is tedious. In this paper, the Reinforcement learning approach is utilized to plan traffic signals. This approach models the elements of a complex framework of the environment and learns the resultant traffic changes. At that point, it searches for an optimum signal pattern based on input-output data. The trouble in this model lies in the exponentially growing complex data. So deep Q-network is utilized to take care of demonstration and prediction issues simultaneously.
PROS AND CONS
By utilizing DRL it is found that normal traffic can be reduced by up to 14%. In morning top hours, a vehicle may spend about 13s less to pass. Additionally, the quantity of completely halted vehicles is enormously decreased. The technique creates a fair line for every one of the bearings and the line length is likewise decreased.
One impediment of this model is that it can’t be applied physically on a large scale. The model doesn't think about the crisis circumstances that may emerge during the day.</td></tr>
</table>

Paper 9

Title of the paper	**Traffic Prediction with Reservoir Computing for Mobile Networks**
Authors	Peng Yu, Wang Jian-min, Peng Xi-yuan.
Year of Publishing	2009
Web Link	https://ieeexplore.ieee.org/document/5366168
Summary	The specific traffic model and its prediction have a significant role in network planning. The author in this paper concerns with traffic prediction of the mobile network when it is nonlinear and dynamic. The specific model of the traffic of this compact framework can be barely acquired. Thus a prediction model based on the past input states has been devised using reservoir computing which utilizes the relations between these inputs to predict the future output. It is used to predict traffic flow according to the time at that instant. The traffic from different mobile devices may have entirely different characteristics based on the location from where they are transmitting the traffic. Using RC a topology is generated to depict the various parameters of the area being studied. At that point a model using ESN is generated to predict future traffic. PROS AND CONS • The simulations of the proposed model show that the ESN model predicts the traffic forecast within the error estimation. • One drawback of this research paper is that the technique is not orderly and the information indexes for confirmation of the model are difficult to get. This model would not work on a large data set. Also not all of the parameters influencing the traffic are considered.

Paper 10

Title of the paper	**Deep Multi-User Reinforcement Learning for Distributed Dynamic Spectrum Access**
Authors	O. Naparstek and K. Cohen,
Year of Publishing	2017
Web Link	https://ieeexplore.ieee.org/document/8254101
Summary	In this paper the author considers the issue of dynamic range access in multichannel remote systems. At the start of each scheduled vacancy, a client chooses a channel and transmits it with a certain probability. After this each client gets a response based on the results of the transmission. The goal of this paper is to create a technique to increase the scope of successful transmission for all the clients without having to generate contact between them. To handle this issue a dispersed model is created to access the multichannel using a calculated algorithm. PROS AND CONS • We see from the numerical tests that the clients learn about the channel network just based on the acknowledgment response they receive. Also the model increases the system throughput. • It has to notice that the required equipment quality for the AI calculations is largely unavailable. Hence to work more in this field it is required to create more imaginative types of equipment for proposed calculations.

Paper 11

Title of the paper	**An Overview of Dynamic Spectrum Sharing: Ongoing Initiatives, Challenges, and a Roadmap for Future Research**
Authors	S. Bhattarai, J.-M. J. Park, B. Gao, K. Bian, and W. Lehr
Year of Publishing	2016
Web Link	https://ieeexplore.ieee.org/document/7516641
Summary	Radio spectrum is undergoing a tremendous change in the present time and its importance is going to increase in the coming time. The use of the radio spectrum has a significant impact on a country's economy. One way of managing it is a dynamic sharing spectrum in which we observe the sharing of spectrum timely among clients. In 2010, the first country to bring major changes in the usage of spectrum sharing was the USA. Following this, many countries like Canada, China, and Singapore made likewise changes in their strategies. One of the major changes that are being utilized in this field is Psychological radio. To effectively use this, it requires information about radio conditions, and security also takes an important part in this. To get radio conditions, the interloper is used to get the spectrum client's information such as area and other subtleties. Despite the changes being made in the system, we are still far from getting great results. There are sectors like security, observing obstruction, and observing spectrum where the considerable improvement is yet to be seen. Many nations are still using old strategies and are not changing much but on the other hand there are other nations too that are accommodating changes in their existing systems. CONCLUSION This study provided us with the number of examples, changes of authoritative exercise, and troubles in researches that are coming in the advancement of technology and how are they helping to manage and efficiently utilize radio range. Most of the old systems used throughout the

	world are undeniably static and unchanging making it inconvenient, or making it impossible to use the range to its limit in a robust manner. The assignment that was submitted described that old systems of radio usage were making it difficult to utilize changes that will help in regulating range and as such blocking its improvement in organizations and advances. A progressively unique model holds the key for the future for range sharing or Dynamic Spectrum Access (DSA). In this, the range is shared based on various parameters such as time, space, or course and among various user settings like business, legacy, and government and different classes of right holders. To understand DSA, specific changes have to made in managerial rights, in-game plans and it requires improvement on how and who can use range rights, changes are required in instruments, on static zones, uniting database propels, etc.

Paper 12

Title of the paper	**Reservoir Computing Properties of Neural Dynamics in Prefrontal Cortex**
Authors	P. Enel, E. Procyk, R. Quilodran, and P. F. Dominey
Year of Publishing	2016
Web Link	https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004967
Summary	The prefrontal cortex (PFC) is used for making decisions on complex choices, arrangements, etc. PFC assumes a major job in creatures where they can encode sudden circumstances similar to RC which makes complex decisions based on past and present states. RC is a sort of repetitive neural network with fixed associations based on neurons. In this paper, the author presents the proof of overlap present in RC and PFC. He conducted a trial in which 2 monkeys needed to find out which 2 of the 4 boxes contained juice. A monkey picked a box and the result was saved. This procedure was carried on back and forth which divided the experiment into two phases. The search stage included what the monkey picked in the first place and recorded it to check whether it was right or wrong. The second phase concerned with the repetition of the process. It was observed that the monkeys stayed away from the wrong boxes and always chose the right ones after some time. An RNN model was 5thus created using RC to make predictions and matched with the experimental results. The reservoir is created using 1000 neurons where input weights which are present between input and reservoir neurons are selected using a uniform distribution with the interval [-1, 1] and the probability of 0.1. Internal weights are chosen among neurons with a probability of 0.1. Here, it can be seen that RC is capable of making complex decisions dynamically.

Paper 13

Title of the paper	**Reservoir Computing Smart Grids: Attack Detection Using Delayed Feedback Networks**
Authors	K. Hamedani, L. Liu, R. Atat, J. Wu, and Y. Yi
Year of Publishing	2018
Web Link	https://ieeexplore.ieee.org/document/8094025
Summary	In this paper the author talks about the ways to harness wind energy. With the increase in the demands for renewable sources of energy new ways are being found to harvest this type of energy. The benefits offered in this field have offset all the investment costs. Here the author presents ways to generate energy for smart grids using wind turbines. Another significant concern in this field is cybersecurity. As these grids rely heavily on computational systems, cyber assaults can happen like false data injection (FDI). This assault adds wrong parameters into the system which can lead to wrong estimations and can result in large damages such as blackouts over a large area. This prompts the use of attack detection over the network. Expanding the use of FDI discovery can directly improve the network output. Subsequently, RC comes into play which can help to detect attacks on smart grids. Main contributions – • An RC based approach is used to improve accuracy and increase insensitive towards attack variations such as the scale of the attack and compromised meters. • MLP is introduced to deal with nonlinear data and task classification. There are steps to this process. First, convert the produced spikes into analogy currents and then produce the states of virtual nodes. The output of these nodes is then multiplied with the feedback gained to preserve the recurrent nature of DFN. The results show that this approach has much better results than its counterparts like SVM, SVE, etc.

Paper 14

Title of the paper	**Reservoir Computing Based Echo State Networks for Ventricular Heart Beat Classification**
Authors	Qurat-ul-ain Mastoi, Teh Ying Wah and Ram Gopal Raj
Year of Publishing	2019
Web Link	https://www.researchgate.net/publication/331205589_Reservoir_Computing_Based_Echo_State_Networks_for_Ventricular_Heart_Beat_Classification
Summary	In the lower part of the heart called ventricular, if there is not normal conduction in cardiovascular development, then it can affect in unexpected passing. In this research, a reservoir computing based method has been proposed for ventricular heartbeat gathering. The proposed system was effectively planned for IoT in the medical industry, for example wearing remote devices for watching heartbeat, etc. this study is based on two datasets called AHA and MIT – BIH –SVDM. The accuracy it got was 98.96%. CONCLUSION The estimation proposed is considered to be helpful in wearable gadgets.

Paper 15

Title of the paper	**Nonlinear Prediction of Speech by Echo State Networks**
Authors	Ziyue Zhao, Huijun Liu, Tim Fingscheidt
Year of Publishing	2018
Web Link	https://ieeexplore.ieee.org/document/8553190
Summary	Speech forecast assumes a key job in numerous speech signal preparation and speech specialized strategies. In this paper, a nonlinear speech forecast is coordinated by a unique kind of repetitive neural framework not requiring any preparation, the echo state network, which automatically updates its loads. Simulations show that it executes better than the other approaches by up to 8 dB. A nonlinear indicator is established using a basic ESN and is applied for speech forecast. Loads of the ESN are updated using an RLS algorithm while the information and neuron loads remain constant. No preparation is required beforehand. This model of speech indicator can be applied in any setting where a forecast is required in the present world.

Paper 16

Title of the paper	**On the Statistical Challenges of Echo State Networks and Some Potential Remedies**
Authors	Qiuyi Wu1, Ernest Fokoue1, Dhireesha Kudithipudi1
Year of Publishing	2018
Web Link	https://www.researchgate.net/publication/323335094_On_the_Statistical_Challenges_of_Echo_State_Networks_and_Some_Potential_Remedies
Summary	In this paper the author has talked about the benefits and problems with Echo State Network (ESN). ESN is a type of recurrent network with some challenges. They are shaky and insecure which makes finding the path toward a tolerable ESN for a dataset exceptionally hard. They are not particularly accurate. Additionally, ESN can show present results for the most dataset, yet it fails as time goes on. Only results that are backed by huge data storage help ESN to give accurate predictions. Considering this situation, to further develop ESN we add it up to huge data sets to balance out its framework and bootstrap the input data. CONS • The execution of the model takes a lot of parameters and settings calculation at the initial stage. • The framework would fail if the dataset contains a lot of noise. • ESN can just track dynamic situations and is incapable of estimating in the long run.

2.2 Integrated Summary

There is a great deal of undiscovered potential in spectrum usage due to which numerous nations are focussing their assets to use the spectrum to their fullest capabilities. They are likewise making arrangements to suit more spectrum groups. Research is going on in this field at a quick rate and numerous strategies have been created to cope up with issues like cognitive radio. Neural network models like RNN are also being utilized. Reservoir computing is one such system that helps the user to make complex decisions when network conditions are not known. It can likewise be compared with the prefrontal cortex present in mind. RC is exceptionally useful, especially in Dynamic Spectrum Access where there are numerous clients and each client needs a separate spectrum band for its use.

3. REQUIREMENT ANALYSIS AND SOLUTION

3.1. Overall description of the project

For an unlicensed user (SU), the device in the spectrum band will face the critical problem of avoiding interference with primary users and transmit traffic according to the state of the bandit senses. This occurs because there is no centralized controller in the system. So in this project we work on a new strategy to help a SU reduce collision and improve convergence. In this venture, we research correspondence techniques of a distributive DSA arrange under the nearness of spectrum detecting mistakes.

3.2. Requirement Analysis

3.2.1 Hardware Requirement

- Processor – i5, 2.40 GHz
- Hard Disk – 500 GB
- Memory – 4 GB RAM

3.2.2 Software Requirement

- PyCharm (Python Implementation)
- Windows 10 (64 bit)

3.3 Solution Approach

We apply deep reinforcement learning (DRL), for SU to learn spectrum strategies in a dispersed manner provided that it no prior knowledge of the system. Besides, reservoir computing (RC) is used to acknowledge DRL by exploiting the temporal relations within the DSA network. Utilizing this system, SU could settle on a spectrum channel solely depending on their present and past spectrum detecting results. Through various tests, we identify that the RC based spectrum access technique can push the SU to fundamentally decrease the odds of a collision with other PUs. We additionally show that our method outperforms the present systems being utilized.

4. MODELING AND IMPLEMENTATION

4.1. Design Diagrams

4.1.1 Learning Process

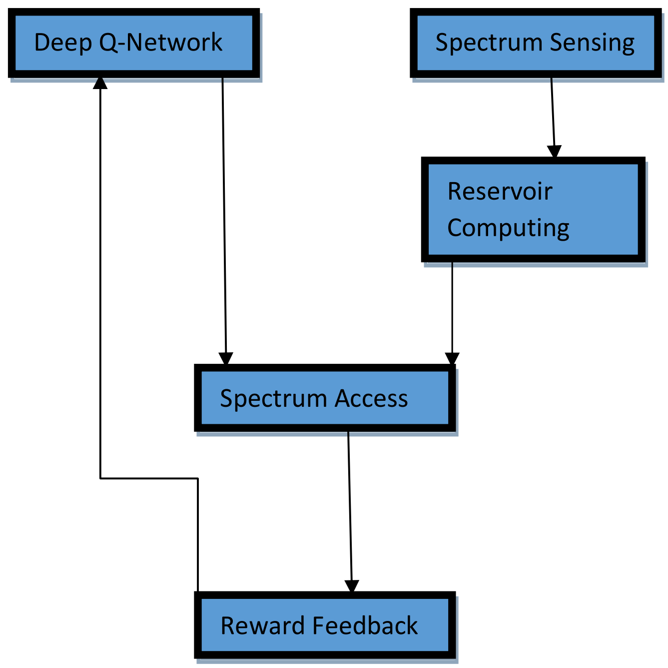

Fig 1: Learning Process in DSA

4.1.2 State Changes

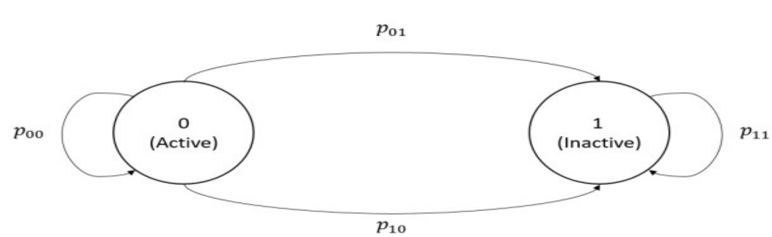

Fig 2: Markov States

4.2. Implementation details and Issues

We consider a DSA network comprising of N-PU's existing together with a DSA optional network with a SU. It is expected that there are a total of N remote channels such that every PU transmits on one unique remote channel to ensure no obstruction among PU's. Moreover, we expect that every PU will communicate a caution signal containing data so that the SU can identify which channel is being utilized by the PU and become aware of the corresponding collision.

Each channel is occupied by a PU and each PU may be in one of the two states:

- Inactive (1) – It means that SU is allowed to access the corresponding channel.
- Active (0) - It means SU cannot access the corresponding channel because the PU is using it.

The dynamics of each PU's activity is described as a two-state Markov chain as shown in Fig.2. The transition probability denoted as:

$$P_n = [P^n_{00}\ \ P^n_{01}\ \ P^n_{10}\ \ P^n_{11}]$$

As an outcome, the reward function of the SU on the n-th channel can be expressed as

$$r^l(t+1) = \begin{cases} -C & \text{Interference with PU} \\ \log_2(1 + SINR^l_n/T) & \text{otherwise} \end{cases}$$

SU doesn't have a clue about the progress probabilities of channel states and the probabilities of detecting errors. It utilizes the SINR received from sensing the channel to develop strategies to maximize its own total discounted reward given by:

$$R^l = \sum\infty_{t=1} \gamma^{t-1}\, r^l(t+1)$$

The dynamics of PUs are demonstrated as autonomous two-state Markov chains with states: Inactive (1) or Active (0). To start with we haphazardly pick p11 and p00 from a uniform appropriation over [0:7; 1] and [0; 0:3] individually for each channel. At that point p10 = 1 -

p11 and p01 = 1 - p00 can be determined as need be insuring that the conceivable estimation of p00 is low and the conceivable estimation of p11 is high.

To show the adequacy of our solution we compare our results with those of different methods such as myopic, Q-learning, etc. We also use these techniques as a baseline to show that our model provides more information and converges faster than its predecessors.

5. TESTING

5.1. For single SU

In this we have taken 22 PU's, 1 SU, and 22 channels. Here, each PU can utilize only 1 channel at a time. To guarantee a reasonable examination, we train the DQN+RC and Q-learning with a similar learning pace of 0.01. Although Q-learning and DQN+RC have comparative execution when in a similar environment, it tends to be seen that Q-learning is much slower than DQN+RC when the quantity of channels is enormous. This is because Q-learning needs serious Q-table updates, and DQN+RC approximates the Q-value utilizing RC, so DQN+RC unites quicker if the estimation is precise. What's more, the outcomes show that DQN+RC and Q-learning can beat the myopic technique even though the myopic strategy treats change probabilities of channels and probabilities of detecting blunder as known data.

To summarize, our proposed DQN+RC has the upsides of quicker union speed when the quantity of channels is huge and preferable execution over the myopic technique. Also the model utilizes the temporal correlations within the network to derive more useful information.

5.2. Example Dataset

We are using Pycharm for our implementation. This data set was generated using random integer property of the random package. In this 0 means empty and 1 means occupied. We have used random functions to generate different input sequences at different times. This will help to provide timely changes that can be seen in real – life. This data set is fed to SU where this is compared to a different sequence of 0,1 present in the environment for further work.

```
C:\Users\deepak\AppData\Local\Programs\Python\Python35\
python.exe C:/Users/deepak/PycharmProjects/8major/data.py
[0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0
, 1, 1]
[0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 1
, 0, 1]
[0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0
, 1, 1]
[1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1
, 1, 0]
[0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1
, 0, 0]
[1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 1
, 1, 1]
[1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1
, 0, 1]
[0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0
, 0, 1]
[0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1
, 0, 1]
[1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1
, 1, 0]
[0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0
, 0, 1]
[0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1
, 1, 1]
[0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1
, 1, 1]
[1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0
, 1, 0]
[1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1
, 0, 1]
[1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1
, 0, 1]
[0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1
, 0, 0]
[0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1
, 1, 0]
[1, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0
, 1, 1]
[0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0
, 1, 1]
[0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1
, 0, 0]
[1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1
, 1, 0]
```

```
24 [1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0
   , 1, 0]
25 [1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1
   , 0, 0]
26 [1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0
   , 1, 1]
27 [0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0
   , 0, 1]
28 [0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1
   , 1, 1]
29 [1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0
   , 0, 0]
30 [1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1
   , 0, 1]
31 [1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1
   , 0, 0]
32 [1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0
   , 0, 0]
33 [1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1
   , 0, 1]
34 [0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1
   , 0, 1]
35 [1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1
   , 1, 1]
36 [0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0
   , 0, 1]
37 [1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1
   , 1, 0]
38 [1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0
   , 0, 0]
39 [1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0
   , 0, 0]
40 [0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0
   , 1, 1]
41 [1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1
   , 1, 0]
42 [0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0
   , 1, 0]
43 [0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1
   , 0, 0]
44 [0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0
   , 0, 0]
45 [1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1
   , 0, 0]
46 [0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1
   , 1, 1]
```

```
47 [1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1
   , 1, 1]
48 [0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1
   , 1, 1]
49 [0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1
   , 0, 1]
50 [1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0
   , 0, 1]
51 [0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1
   , 1, 1]
52 [0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0
   , 0, 1]
53 [0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1
   , 0, 1]
54 [0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1
   , 1, 0]
55 [1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1
   , 0, 1]
56 [0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1
   , 0, 1]
57 [1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0
   , 0, 0]
58 [1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0
   , 0, 1]
59 [0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1
   , 0, 1]
60 [0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0
   , 1, 1]
61 [1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1
   , 0, 0]
62 [1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0
   , 1, 1]
63 [1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0
   , 0, 1]
64 [0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0
   , 1, 1]
65 [1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1
   , 1, 0]
66 [0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1
   , 1, 0]
67 [0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0
   , 0, 1]
68 [1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1
   , 0, 0]
69 [0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0
   , 1, 0]
```

```
[1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0,
 0, 1, 0]
[1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1,
 1, 0, 1]
[1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 1,
 0, 0, 1]
[0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0,
 1, 1, 0]
[0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0,
 1, 0, 1]
[1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1,
 1, 1, 0]
[1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0,
 0, 0, 1]
[1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1,
 0, 1, 1]
[0, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1,
 1, 1, 0]
[1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0,
 1, 0, 0]
[0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1,
 1, 1, 0]
[0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1,
 0, 1, 0]
[0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0,
 0, 1, 0]
[0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1,
 1, 0, 0]
[0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1,
 0, 1, 0]
[0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1,
 0, 0, 1]
[1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0,
 0, 1, 0]
[1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0,
 0, 0, 0]
[0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1,
 0, 0, 1]
[0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0,
 1, 0, 1]
[0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1,
 1, 1, 0]
[1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0,
 0, 1, 0]
[1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1,
 1, 1, 1]
```

```
[1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0,
 1, 0, 0]
[1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0,
 1, 0, 1]
[0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0,
 1, 0, 0]
[1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0,
 1, 0, 0]
[1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1,
 1, 1, 1]
[1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0,
 1, 0, 0]
[1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0,
 0, 1, 0]
[1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0,
 0, 0, 1]
[0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1,
 1, 0, 0]

Process finished with exit code 0

```

6. FINDINGS AND CONCLUSIONS

6.1. Results and Findings

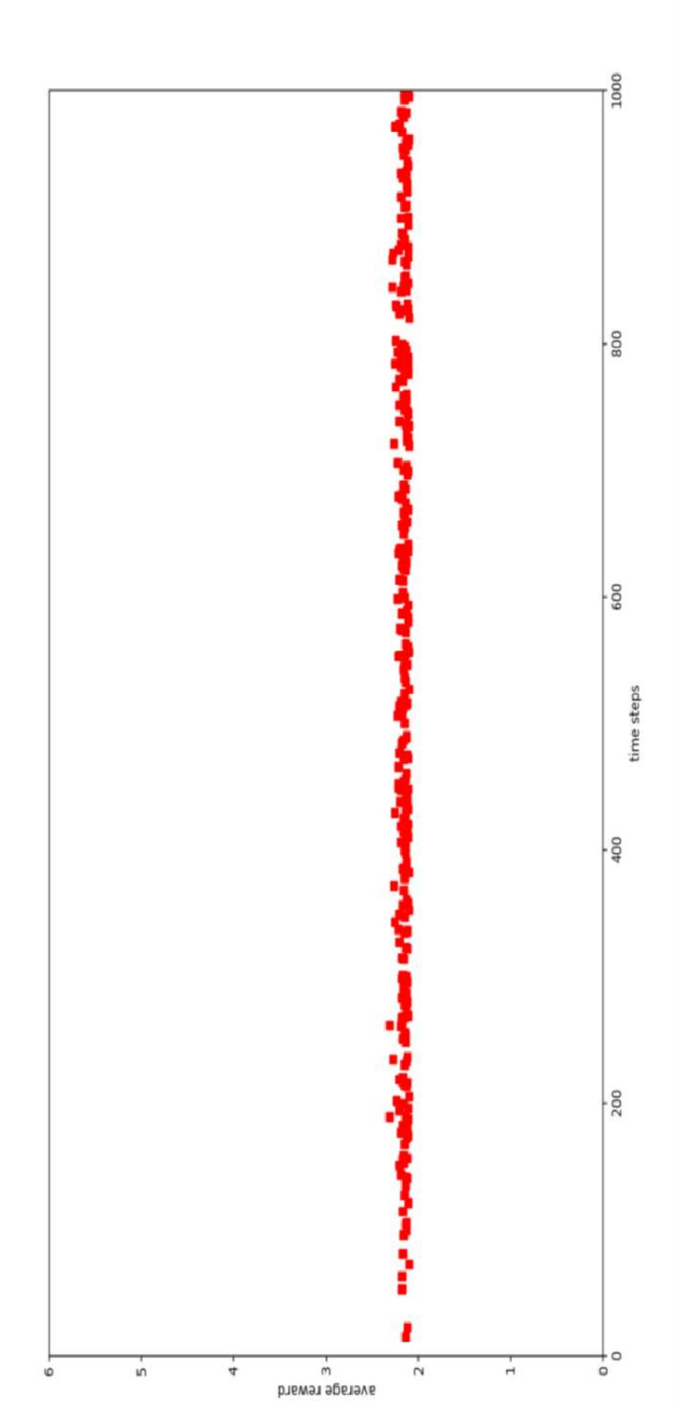

Fig 3. The graph between average reward and time steps

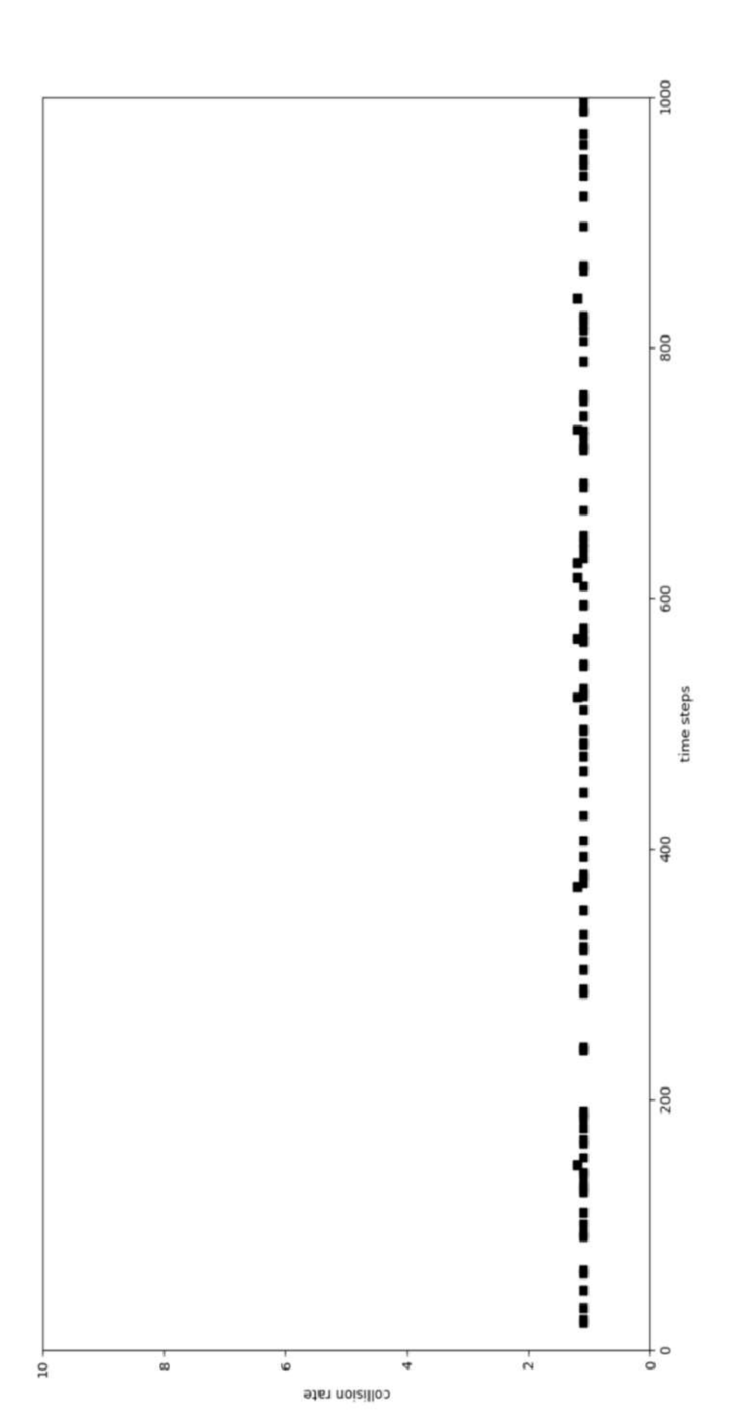

Fig 4. The graph between collision rate and time steps

- Extensive execution evaluation of the introduced dynamic spectrum framework shows that our technique could quickly learn about PUs' activities and decrease the chance of contact with PUs when accessing the channel.

- Based on our understanding, we wrote code and got linear graphs but it shouldn't be. In the case of fig 3, the graph should be increased after some time steps and then be linear. In the case of fig 4, the graph should be decreasing after some time steps and then be linear.
- Theoretically, this method overshadows the myopic strategy and converges faster than Q-technique when channel quantity is large.
- As compared to DRL+MLP, our DRL+RC strategy can misuse the concealed time-based relations present while detecting the results and yield more appropriate information for faster execution.

6.2. Limitations of the Solution

- The random values assigned in the reservoir for input generate instability within the system and increase variance.
- The learning system requires complicated adjustments within the given parameters which require tremendous efforts.

6.3. Conclusion

In this project we study spectrum sensing for the DSA network when there is no centralized system and sense is imperfect. Our strategy combining DQN and RC is introduced in which DRL is used to learn the state of the channel and RC is utilized to realize DQN by studying the temporal correlations present in the network. Relying on our strategy the SU can take proper decisions relying on the limited information gained from PUs at each sensing step. Experiments have been performed to verify the performance of our method. These results suggest that our method provides a higher transmission rate and lower collision frequency as opposed to myopic methods. It converges faster and provides faster information collection and execution through its recurrent structure.

7. REFERENCES

1) Yiding Yu, Taotao Wang, Soung Chang Liew, “Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks”, IEEE 2018.

2) Rubaiyat Shafin, Hao Chen, Young Han Nam, Sooyoung Hur, Jeongho Park, Jianzhong (Charlie) Zhang, Jeffrey Reed, and Lingjia Liu, "Self-Tuning Sectorization: Deep Reinforcement Learning Meets Broadcast Beam Optimization", IEEE June 2019.

3) Hao-Hsuan Chang, Hao Song, Yang Yi, Jianzhong (Charlie) Zhang, Haibo He, and Lingjia Liu, "Distributive Dynamic Spectrum Access through Deep Reinforcement Learning: A Reservoir Computing Based Approach", IEEE 2018.

4) Eliko Tabakovic and Sonja Grgic, “Dynamic Spectrum Access in Cognitive Radio” 51st International Symposium ELMAR-2009, 28-30 September 2009, Zadar, Croatia.

5) Feng Li, Kwok Yan Lam, Zhenggou Shang, Li Wang, Kanglian Zhao, and Xinggang Zhang. (2018) “Q-learning-based dynamic spectrum access in cognitive Industrial Internet of Things. Mobile Networks and Applications”, 23 (6). pp. 1636-1644.

6) J.M. Peha, Rathapon Saruthirathanaworakun "Dynamic Primary-Secondary Spectrum Sharing with Cellular Systems” Proc. IEEE Special Issue on Cognitive Radio, vol.97, no.4, pp.708-19, Apr. 2009.

7) M. Chen, W. Saad, C. Yin, and M. Debbah, “Echo state networks for proactive caching in cloud-based radio access networks with mobile users,” IEEE Trans. Wireless Commun., vol. 16, no. 6, pp. 3520–3535, Jun. 2017.

8) S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari, “Deep reinforcement learning for dynamic multichannel access in wireless networks, IEEE Trans. on Cogn. Commun. Netw., vol. 4, no. 2, pp. 257–265, Jun. 2018.

9) Li Li, YishengLv, Fei-Yue Wang." Traffic Signal Timing via Deep Reinforcement Learning" IEEE/CAA Journal of Automatica Sinica (Volume: 3, Issue:3, July 10 2016).

10) Peng Yu, Wang Jian-min, Peng Xi-yuan. " Traffic Prediction with Reservoir Computing for Mobile Networks " 2009 Fifth International Conference on Natural Computation, IEEE.

11) O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for dynamic spectrum access in multichannel wireless networks," in GLOBECOM, Dec. 2017, pp. 1–7.

12) S. Bhattarai, J.-M. J. Park, B. Gao, K. Bian, and W. Lehr, "An overview of dynamic spectrum sharing: Ongoing initiatives, challenges, and a roadmap for future research," IEEE Trans. on Cogn. Commun. Netw., vol. 2, no. 2, pp. 110–128, Jun. 2016.

13) P. Enel, E. Procyk, R. Quilodran, and P. F. Dominey, "Reservoir computing properties of neural dynamics in the prefrontal cortex," PLoS computational biology, vol. 12, no. 6, p. e1004967, Jun. 2016.

14) K. Hamedani, L. Liu, R. Atat, J. Wu, and Y. Yi, "Reservoir computing meets smart grids: Attack detection using delayed feedback networks," IEEE Trans. Ind. Informat., vol. 14, no. 2, pp. 734–743, Feb. 2018.

15) Qurat-ul-ain Mastoi, Teh Ying Wah, and Ram Gopal Raj, "Reservoir Computing Based Echo State Networks for Ventricular Heart Beat Classification" February 2019.

16) Ziyue Zhao, Huijun Liu, Tim Fingscheidt, "Nonlinear Prediction of Speech by Echo State Networks", September 2018.

17) Qiuyi Wu1, Ernest Fokoue1, Dhireesha Kudithipudi1, "On the Statistical Challenges of Echo State Networks and Some Potential Remedies", September 2018.

9916103181_DYNAMIC SPECTRUM ACCESS USING DEEP REINFORCEMENT LEARNING IN DISTRIBUTED SYSTEM

ORIGINALITY REPORT

17%	11%	11%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Hao-Hsuan Chang, Hao Song, Yang Yi, Jianzhong Zhang, Haibo He, Lingjia Liu. "Distributive Dynamic Spectrum Access Through Deep Reinforcement Learning: A Reservoir Computing-Based Approach", IEEE Internet of Things Journal, 2019 Publication	4%
2	Submitted to Indian Institute of Technology Guwahati Student Paper	2%
3	export.arxiv.org Internet Source	1%
4	www.vcl.fer.hr Internet Source	1%
5	ohmisskate.com Internet Source	1%
6	www.studymode.com Internet Source	1%

7	Ziyue Zhao, Huijun Liu, Tim Fingscheidt. "Nonlinear Prediction of Speech by Echo State Networks", 2018 26th European Signal Processing Conference (EUSIPCO), 2018 Publication	1%
8	users.ece.cmu.edu Internet Source	1%
9	Mingzhe Chen, Walid Saad, Changchuan Yin, Merouane Debbah. "Echo State Networks for Proactive Caching in Cloud-Based Radio Access Networks With Mobile Users", IEEE Transactions on Wireless Communications, 2017 Publication	1%
10	www.mdpi.com Internet Source	1%
11	link.springer.com Internet Source	1%
12	documents.mx Internet Source	1%
13	Kian Hamedani, Lingjia Liu, Rachad Atat, Jinsong Wu, Yang Yi. "Reservoir Computing Meets Smart Grids: Attack Detection Using Delayed Feedback Networks", IEEE Transactions on Industrial Informatics, 2018 Publication	<1%

14	dblp.uni-trier.de Internet Source	<1%
15	Shangxing Wang, Hanpeng Liu, Pedro Henrique Gomes, Bhaskar Krishnamachari. "Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks", IEEE Transactions on Cognitive Communications and Networking, 2018 Publication	<1%
16	Enel, Pierre, Emmanuel Procyk, René Quilodran, and Peter Ford Dominey. "Reservoir Computing Properties of Neural Dynamics in Prefrontal Cortex", PLoS Computational Biology, 2016. Publication	<1%
17	engine.scichina.com Internet Source	<1%
18	Submitted to Higher Education Commission Pakistan Student Paper	<1%
19	Yiding Yu, Taotao Wang, Soung Chang Liew. "Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks", IEEE Journal on Selected Areas in Communications, 2019 Publication	<1%

20	eprints.utm.my Internet Source	<1%
21	"Table of Contents - Volume 2", 2009 Fifth International Conference on Natural Computation, 2009 Publication	<1%
22	Submitted to University of Hull Student Paper	<1%
23	Long Hu, Daxin Tian, Kai Lin. "Editorial: Cognitive Industrial Internet of Things", Mobile Networks and Applications, 2018 Publication	<1%

Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On		

9916103181_DYNAMIC SPECTRUM ACCESS USING DEEP REINFORCEMENT LEARNING IN DISTRIBUTED SYSTEM

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41

PAGE 42

PAGE 43

PAGE 44