IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 17, Issue 5, Ser. II (Sep. – Oct. 2022), PP 40-46
www.iosrjournals.org

Energy Efficient CR System: Multilayer DQN Approach

P Ravinder Kumar[1] Sandeep V M[2] Subhash Kulkarni[3]
[1]*Research Scholar, ECE Department, JNTUH, Hyderabad.*
[2] *Professor and HOD, CSE Department, Jayaprakash Narayan College of Engineering, Mahabubnagar.*
[3] *Principal and Professor, PESIT – Bangalore South Campus, Bengaluru*

Abstract:*This paper aims at increasing energy efficiency of CR system by minimizing sensing and interference time. With accurate PU behaviour information optimum system can be achieved. As first step through DQN synchronization of SU with PU reduces the interference. Secondly, SU will be synchronized at that instant of PU change. Thirdly, SU is synchronized with PU transitions through DQN timeseries analysis. Making SU to change the states in synchronization with PU, with low sensing and interference maximising energy efficiency.*
Key Word*:Cognitive Radio, Time series, Reinforcement Learning.*

I. Introduction

As the number of wireless users grows, spectrum scarcity and underutilization are serious issues. Static spectrum allocation methods are used earlier to allocate a large portion of the spectrum. The channel remains vacant when the primary user or the licenced user not using the channel, results in spectrum underutilization. A wireless communication technology that can intelligently using this available free spectrum improves spectrum utilisation. That is the goal of Cognitive radio (CR). CR allows the unlicensed or Secondary Users (SUs) to access licenced users' (also known as Primary Users (PUs)) spectrum during idle time. The SU performs spectrum sensing for efficient utilisation and reduces interference to the PU. By the conventional methods, frequent sensing is required. Due to false alarm the interference is more. There is no transmission during sensing and interference, so no throughput, resulting in energy wastage. This calls for is in terms of energy, time, and computation. Better utilization of the resources can be achieved by understanding PU behaviour through Machine Learning.

Machine Learning can be combined with CR to produce a potential solution for energy efficient resource allocation in CR networks[1], [2]. Machine learning and deep learning techniques are used to reduce these sensing and interference problems by knowing the exact behaviour (state) of primary. However, ML requires more data to analyse. Learning while observing the PU behaviour calls for Reinforcement Learning (RL). Here the PU with multiple behaviours (states) is the environment and the SU is an agent whose action is to predict the current state of PU. The correctness of this prediction result in throughput/interference as reward/punishment. For efficient time and space complexity the DQN algorithm is preferred [3].

The goal of our research is to develop an energy-efficient system with high-throughput. To reduce the interference, we need to know how the PU behaves and synchronize SU. Exploration and exploitation are used to learn the PU behaviour using reinforcement learning. If the PU's behaviour changes, the SU must resynchronize with it. In our research, the agent is structured as 3 layers. Layer-1 is in-charge of identifying PU behaviour. If we can determine how long this behaviour will continue, we can resynchronisation SU with new behaviour of PU at the instant of behaviour change. Timeseries models can help us better understand this behaviour. Layer-2 discovers the pattern of behaviour changes through RL. With this layer, energy wasted only during behavioural changes. Layer-1 provides the PU behaviour, and Layer-2 determines how long it will remain in that state and what the pattern is. Layer-3 is responsible in identifying the pattern of state changes. The pattern is identified through Reinforcement Timeseries Learning. Once layer-3 is active, layer-1 and layer-2 are relieved and SUs are directly synchronized with PU without sensing. The rest of the paper is structured as follows: The following section provides background information on CR and RL. Section 3 examines various timeseries models for identifying PU behaviour and achieving energy efficiency. The results and discussion in section 4 confirm our claims and Section 5 concludes the paper.

II. Background

Spectrum management is a critical issue in wireless communications today. Due to the phenomenal growth in wireless applications and services such as delay sensitive multimedia applications [4] and internet of things (IoT) [5] the spectrum for unlicensed secondary user (SU) remains overcrowded. Cognitive radio (CR) technology has emerged as a promising approach in resolving the issues of using spectrum resources [6], [7]. In CR technology, an SU, or CR user, uses the channel when the primary is inactive. This needs, accurate detection of the presence of PU for successful communication. The most difficult tasks in CR technology are reliable spectrum sensing and optimising network throughput. [8] investigates a sensing-throughput trade-off, estimating the optimal sensing time for maximising SU network throughput under a target detection probability. Accurate sensing improves spectral efficiency by reducing false alarms, maximising network throughput under energy constraints. Secondary Users (SUs) must adapt Dynamic Spectrum Allocation (DSA) to maximise the utilisation of idle licenced spectrum in an opportunistic manner to efficiently utilise the underutilised licenced spectrum [9]. When an SU wants to use a licenced channel, it will perform spectrum sensing to observe the state of the spectral occupancy (i.e., idle/busy). Frequent spectrum sensing, will shorten the network lifetime of wireless communication devices, which are typically battery powered [10]. Hence, creating an energy-efficient spectrum sensing method is critical. Qu et al. proposed a method to reduce energy consumption by selecting SUs with high detection accuracy for sensing and then distributing the results to neighbouring SUs [11]. This method can be improved further by incorporating machine learning techniques[12], [13]. Bhowmick et al. proposed a technique based on historical data in which SUs predict the future state of a channel and then either sense the idle channels or harvest energy from the busy channels [14],[15],[16].

Reinforcement learning [17], [18] is one of the most important machine learning research directions that has had a significant impact on the development of Artificial Intelligence (AI). Reinforcement learning is a learning process in which an agent can make decisions on a regular basis, observe the outcomes, and then automatically adjust its strategy to achieve the best policy. However, even though this learning process has been shown to converge, it takes a long time to arrive at the best policy because it must explore and gain knowledge of an entire system, making it unsuitable and inapplicable to large-scale networks. [18] provides an in-depth look at Deep RL for spectrum sensing in CR networks. For optimal detection performance, proposed a mathematical hypothetic model of Deep RL-based cooperative spectrum sensing and characteristics of the various Deep RL algorithms[17]–[19]. A time series is a sequence of data points measured typically at successive times and spaced at time intervals in statistics, signal processing, and many other fields [20]. Time series analysis refers to methods that attempt to comprehend such time series, frequently in order to comprehend the underlying context of the data points or to make forecasts [21]. In general, a time series model will reflect the fact that observations close together in time are more closely related than observations further apart. Furthermore, time series models frequently employ the natural one-way ordering of time, so that values in a series for a given time are expressed as deriving in some way from previous values[22].

III. Proposed Work

The experimental setup used in the current research is a channel, it’s PU and multiple Secondary Users (SU’s) competing for this channel. The goal is to optimize the Cognitive Radio (CR) system's energy efficiency with respect to throughput by reducing energy wastage. The sensing and interference periods are the main reasons for energy wastage as they never contribute to the throughput. At PU, interference periods and at SU, sensing and interference periods, causes this energy wastage. The energy wastage can be significantly reduced if the PU behaviour is known in advance so that the SU can be synchronized. This synchronization makes SU to act without sensing phase and yet no interference to the PU. The CR will have optimal efficiency as long as the PU behaviour is unchanged. Initially it is assumed that the PU behaviour is unchanged for some fixed duration.

When the PU behaviour is changed, the SU continues to be in sync with old PU behaviour for the period decided earlier. This desynchronizing of SU with PU adds to interference hence decrease in throughput. After identifying the change in behaviour, the SU must re-synchronize with the PU Behaviour. The interference is present in both cases. So, it can be reduced, if the PU behaviour is known, resynchronization is performed at that precise moment. Energy wastage is only during resync period and hence more efficient.

To achieve this, we need to know, the behaviours of PU at various times and their durations. The PU behaviour is time dependent and understanding this we need timeseries analysis. With no prior information available, learning is done through Reinforcement Timeseries learning (DQN).

Understanding PU behaviour amounts to large overhead in terms energy and resources. Most of the SU's have sparse resources and energy. This overhead is heavy burden to all SU's. Instead, if a single SU with sufficient resources and energy, is identified as central unit (CU) then all other SU's will be relieved from this burden. Hence the selected SU, called CU, now manages all the activities of secondary side of CR system. Whenever a SU wants the services, it requests the CU for the schedule.

The lack of prior PU behaviour data forces us to understand it in real time and this necessitates the use of DQN. The parameters ON and OFF periods of PU are chosen at random, and the overall performance of CR is measured by total interference to the PU and SU's inability to use the channel. Reinforcement Learning is used to update these parameters. Exploration is the process of randomly selecting the parameters, and exploitation is the process of fine tuning these parameters. It is encouraged to periodically explore and then continue with exploitation to reduce the interference period even further. It is assumed that this behaviour will continue for some fixed duration [18]. However, efficient CR system is when this period can be exactly predicted. For this the PU behaviour is studied through Reinforcement Timeseries Learning.

The literature recommends three time series techniques [3], [23]–[25]: ARIMA, LSTM, and Prophet. Our data is a time series behaviour of PU which may have trend and seasonality. This can be analyzed through ARIMA and appropriate model to predict can be built. ARIMA is stationary analysis model that uses time series data to understand the patterns in the data that will help in predicting the future. Time series data have components like trend and seasonality. ARIMA helps in accurate in model timeseries data having trend and seasonality. Along with trend and seasonality our data may also have predictable impulsive behaviour. ARIMA fails to model these impulses. Hence better model can be built through Prophet. More complicated behaviour can be easily learnt using LSTM. LSTM is a RNN which over comes long-term dependencies.

It may be possible that these state transitions may have some pattern if this pattern is known then SUs can be directly synchronized asper this information. This can result in relieving layer-1 and layer-2 of the agent and maximum energy can be saved. The pattern is identified using Reinforcement Timeseries Analysis in layer-3. Layer-2 outputs the sequence of states and their durations. This information helps layer-3 in learning the patterns of state transitions. Once the layer-3 learns the total behaviour of PU, state diagram is available with CU. So, whenever SU requests for service CU verifies this state diagram and informs the SU to work in which state and how long. With layer-3 all the resources of CU used for layer-1 and layer-2 can be released, the CU performance as SU will improve. Due to accurate synchronization of SUs with PU maximum throughput can be achieved for the energy spent. Hence, with layer-3 the CR system can be highly energy efficient.

IV. Results and Discussion

The experiment is conducted with PU having n states, each state representing behaviour of PU. This experiment is repeated for various values of n. The CU is a 3 layered agent that is attempting to predict the current state of PU. The layer-1 helps in identifying the current behaviour of primary user and synchronizes SU with this behaviour. However, just identifying current behaviour in synchronizing the SU is not effective unless the PU behaviour changes also be synchronized. The layer-2 of our experimentation uses the time series PU behaviour collected from layer-1 and analyses it. This time series is analysed using DQN through three techniques: ARIMA, LSTM and Prophet.

Figure 1 through figure 4 highlights the effect of using the second layer after learning through three techniques. Figure 1 shows the effect of layer-2 on interference to PU. It is found that without layer-2 to the PU variations are not synchronized and hence accounts to maximum interference. Here, the SU synchronized with PU for some predefined period. The central unit, is unaware of these changes till the end of this fixed period, causing interference to both PU and SU. At the end of this period, the CU realizes the problem and resynchronize to the new PU behaviour. This calls for another period of interference. Both these periods together cause heavy interference. This interference amounts to the penalty for the agent. LSTM, however, shows no significant improvement due to the very fact that large past data is required for better learning and learning in the present scenario on real-time basis.

ARIMA shows better performance as compared to LSTM, as small past data is sufficient for learning [4]. But ARIMA neglects impulsive behaviour of PU. The impulsive behaviour is very well understood by Prophet and hence the current study shows the Prophet is the best technique for understanding PU behaviour. With learning continues, sometime after very long period, LSTM may supersede these two. Till then Prophet is recommended and then can be switched over to LSTM. The study is performed over various behaviour changes of primary user.

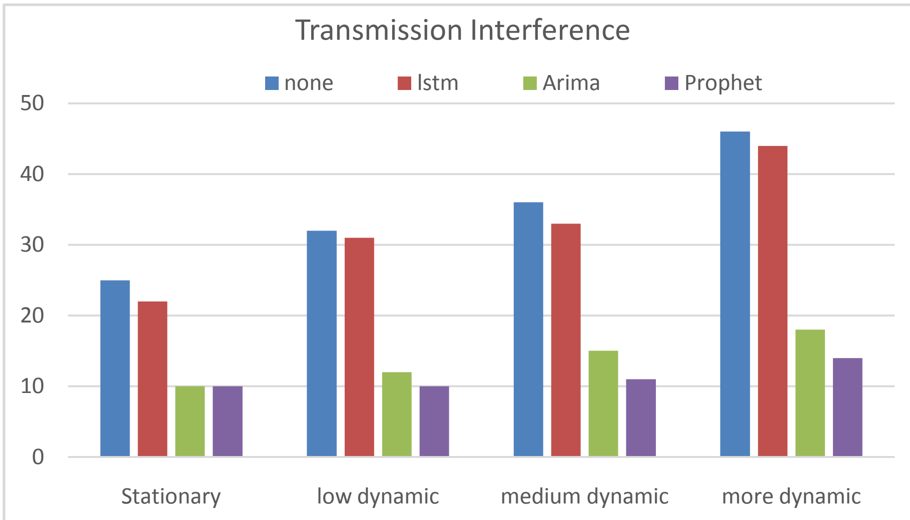

Figure.1: Transmission Interference comparison with different behaviours and methods.

When the PU behaviour is stationary, the synchronization of SU with PU needs to be done only once, hence, TI is small. From figure1 more observations can be made. When the PU behaviour is non changing the SU synchronizing with PU is done only once and hence overall interference is almost negligible. As the PU behaviour becomes dynamic the interference keeps increasing with dynamism as more and more resynchronization phases are needed. It is also observed that in all the cases Prophet technique is giving best performance.

Figure 2 shows the effect of layer-2 on transmission loss. All the issues discussed for interference also true for TL. However, it is far greater than interference as major goal of the research is to minimise the interference to PU. In this process, the effective utilization of channel by SU is scarified. Hence TL is more. Figure 2 endorses this effect.

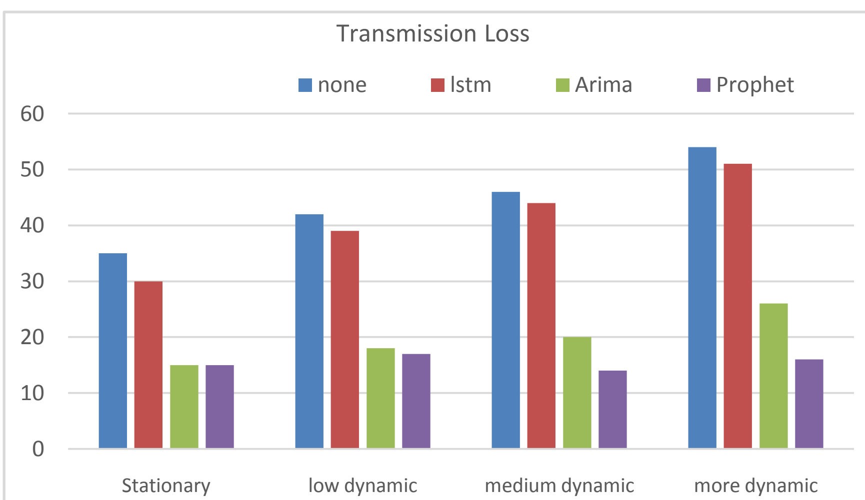

Figure.2: The Transmission loss for different models

Throughput of primary user, depends on interference and throughput of secondary user depends both on interference as well as transmission loss. The secondary user throughput is maximum whenever both TI and TL are at their minimum. The overall CR throughput is the total throughput of primary user and secondary user. The results shown in figure 3 have the same effect as of TI and TL in Figure 1 and 2.

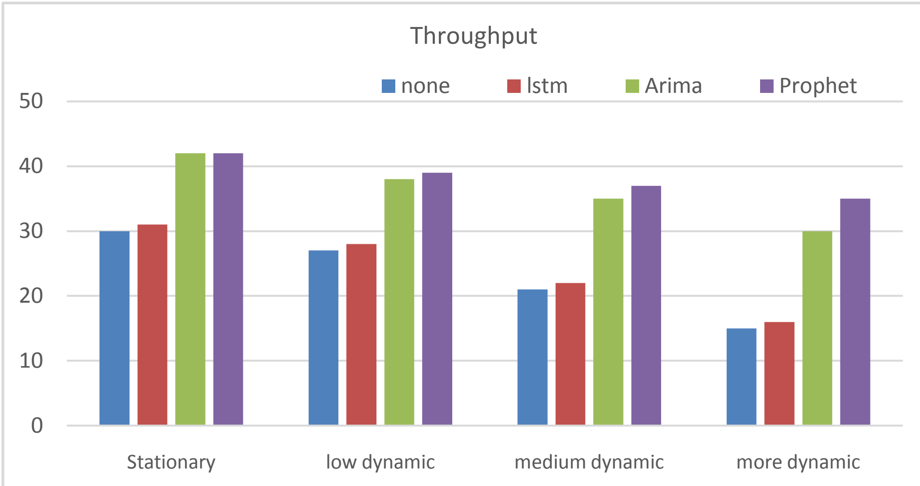

Figure.3: Throughput for different models

In our experiment, the Energy Efficiency is measured as throughput per energy spent. It is evident that the Prophet model has minimum interference, minimum transmission loss and hence higher throughput making it most efficient CR model. Similar explanations as in Figure 1 through figure 3 can infer the results in figure 4. Overall, the results claim that Prophet technique provides best Energy Efficiency in lesser time.

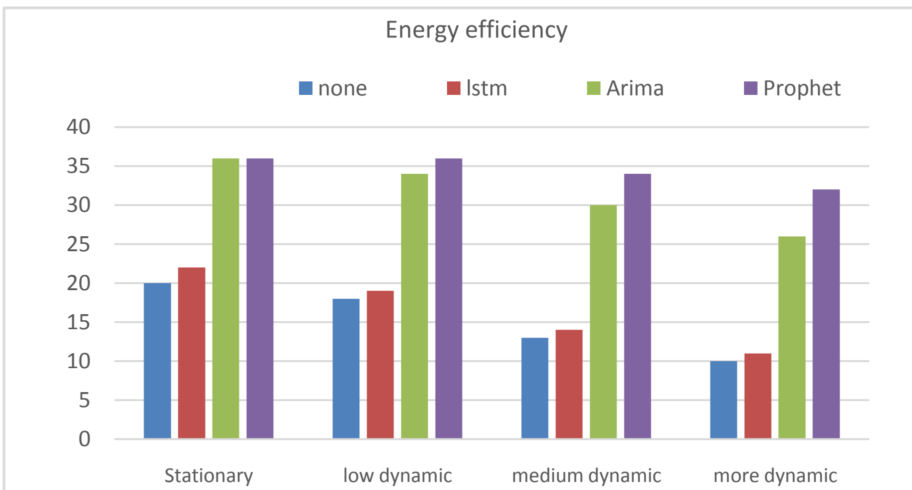

Figure 4: Energy efficiency for different models

The main reason for this Improvement in throughput and Energy Efficiency is layer-2 takes control of layer-1. In the absence of layer-2, layer-1 needs to work continuously for synchronizing SU with PU. With layer-2, the exact instant at which the PU changes its behaviour can be well predicted and hence layer-1 is operating only at these instants to resynchronize SU with PU. This contributes to maximum energy saving and hence higher Energy Efficiency.

With layer 3 complete behaviour of PU is learnt. This behaviour is stored as a lookup table and all the resources of CU are released. The CU now just uses this lookup table for guiding the SUs for perfect synchronization. Due to perfect synchronization highest performance is achieved as shown in figure 5.

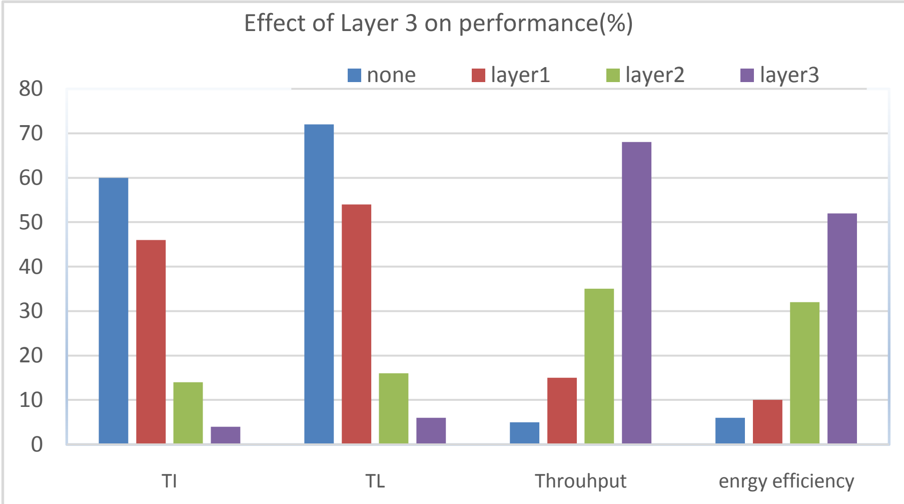

Figure.5: Effect of layer 3 on performance

V. Conclusion

In this paper, we proposed Reinforcement time series learning models for understanding and predicting PU behaviour. Three layered structure is used to measure the performance. Layer-1 knows the PU behaviour, the SU synchronizes with PU, thereby interference is reduced and hence the performance is improved. Layer-2 predicts the exact period of each PU state. This further improves the performance. Layer-3 is responsible for predicting pattern in state transitions of PU. Once layer-3 is operative, the CU works with bare minimum resources for scheduling the SUs achieving maximum throughput and hence energy efficiency.

References

[1] M. Bkassiny, S. Member, Y. Li, S. Member, S. K. Jayaweera, and S. Member, "A..A Survey on Machine-Learning Techniques in," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1136–1159, 2013.

[2] T. Chen, H. H. Chen, Z. Chang, and S. Mao, "Machine-Learning-Based Opportunistic Spectrum Access in Cognitive Radio Networks," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 6–8, 2020.

[3] S. Zhang, T. Wu, M. Pan, C. Zhang, and Y. Yu, "A-SARSA: A Predictive Container Auto-Scaling Algorithm Based on Reinforcement Learning," *Proceedings - 2020 IEEE 13th International Conference on Web Services, ICWS 2020*, pp. 489–497, 2020, doi: 10.1109/ICWS49710.2020.00072.

[4] T. Qiu, R. Qiao, and D. O. Wu, "EABS: An event-aware backpressure scheduling scheme for emergency internet of things," *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 72–84, 2018, doi: 10.1109/TMC.2017.2702670.

[5] T. Qiu, X. Wang, C. Chen, M. Atiquzzaman, and L. Liu, "TMED: A spider-web-like transmission mechanism for emergency data in vehicular Ad Hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8682–8694, 2018, doi: 10.1109/TVT.2018.2841348.

[6] B. Aneja, K. Sharma, and A. Rana, "Spectrum sensing techniques for a cognitive radio network," *Lecture Notes in Electrical Engineering*, vol. 509, pp. 133–144, 2019, doi: 10.1007/978-981-13-0665-5_11.

[7] I. Kakalou, D. Papadopoulou, T. Xifilidis, K. E. Psannis, K. Siakavara, and Y. Ishibashi, "A survey on spectrum sensing algorithms for cognitive radio networks," *2018 7th International Conference on Modern Circuits and Systems Technologies, MOCAST 2018*, pp. 1–4, 2018, doi: 10.1109/MOCAST.2018.8376562.

[8] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *Software Radio Technologies: Selected Readings*, pp. 413–418, 2001, doi: 10.1109/9780470546444.ch4.

[9] D. A. Guimarães, E. J. T. Pereira, A. M. Alberti, and J. V. Moreira, "Design guidelines for database-driven internet of things-enabled dynamic spectrum access," *Sensors*, vol. 21, no. 9, 2021, doi: 10.3390/s21093194.

[10] J. P., G. V., and T. Sudha, "Energy Efficient Spectrum Sensing Techniques for Cognitive Radio Networks: A Survey," *International Journal of Computer Applications*, vol. 160, no. 4, pp. 20–23, 2017, doi: 10.5120/ijca2017913033.

[11] Z. Qu, Y. Xu, and S. Yin, "A novel clustering-based spectrum sensing in cognitive radio wireless sensor networks," *CCIS 2014 - Proceedings of 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems*, pp. 695–699, 2014, doi: 10.1109/CCIS.2014.7175823.

[12] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 2018, 2018, doi: 10.1155/2018/5906097.

[13] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2209–2221, 2013, doi: 10.1109/JSAC.2013.131120.

[14] A. Bhowmick, K. Yadav, S. D. Roy, and S. Kundu, "Throughput of an Energy Harvesting Cognitive Radio Network Based on Prediction of Primary User," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8119–8128, 2017, doi: 10.1109/TVT.2017.2690675.

[15] A. Bhowmick, G. C. Das, S. D. Roy, S. Kundu, and S. P. Maity, "Allocation of optimal energy in an energy-harvesting cooperative multi-band cognitive radio network," *Wireless Networks*, vol. 26, no. 2, pp. 1033–1043, 2020, doi: 10.1007/s11276-018-1849-2.

[16] Z. Jin, K. Yao, B. Lee, J. Cho, and L. Zhang, "Channel status learning for cooperative spectrum sensing in energy-restricted cognitive radio networks," *IEEE Access*, vol. 7, pp. 64946–64954, 2019, doi: 10.1109/ACCESS.2019.2916065.

[17] N. C. Luong *et al.*, "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019, doi: 10.1109/COMST.2019.2916583.

[18] E. O. Felix Obite , Aliyu D. Usman, "An overview of deep reinforcement learning for spectrum sensing in cognitive radio networks," *Digital Signal Processing: A Review Journal*, vol. 113, no. 4, p. 103014, 2021, doi: 10.1016/j.dsp.2021.103014.

[19] S. Q. Jalil, S. Chalup, and M. H. Rehmani, "Cognitive Radio Spectrum Sensing and Prediction Using Deep Reinforcement Learning," in *Proceedings of the International Joint Conference on Neural Networks*, 2021, vol. 2021-July, pp. 1–8. doi: 10.1109/IJCNN52387.2021.9533497.

[20] Z. Wang and S. Salous, "Spectrum occupancy statistics and time series models for cognitive radio," *Journal of Signal Processing Systems*, vol. 62, no. 2, pp. 145–155, 2011, doi: 10.1007/s11265-009-0352-5.

[21] Z. Wang and S. Salous, "Time series arima model of spectrum occupancy for cognitive radio," *IET Seminar Digest*, vol. 2008, no. 12338, pp. 2–5, 2008, doi: 10.1049/ic:20080405.

[22] Z. Wang and S. Salous, "Spectrum occupancy statistics and time series models for cognitive radio," *Journal of Signal Processing Systems*, vol. 62, no. 2, pp. 145–155, 2011, doi: 10.1007/s11265-009-0352-5.

[23] Q. Dang, "Reinforcement Learning in Stock Trading To cite this version : HAL Id : hal-02306522 Reinforcement Learning in Stock Trading," 2019.

[24] S. W. Chang, C. L. Chang, L. T. Li, and S. W. Liao, "Reinforcement learning for improving the accuracy of PM2.5 pollution forecast under the neural network framework," *IEEE Access*, vol. 8, pp. 9864–9874, 2020, doi: 10.1109/ACCESS.2019.2932413.

[25] J. L. Hernandez Villapol, "Spectrum Analysis and Prediction Using Long Short-Term Memory Neural Networks (LSTMs) and Cognitive Radios," 2017.