Multi-User Reinforcement Learning for the Deep Distributed Dynamic Spectrum Access

Dr.M.Monisha[1], Y. Kamala Kannan[2], B.Vignesh[3]
[1]*Assistant professor, Department of Electronics and Communication Engineering, VISTAS,*
[2,3]*UG scholar, Vels Institute of Science, Technology & Advanced Studies (VISTAS)*

***Abstract* - The problem in the DSA dynamic spectrum is the path of the distance between the nodes into the environment. The maximization in multichannel wireless networks. Where there is the problem which is it to take the shortest path by the node in the environment by the agent. This will make the message flow between the distance of the node. By this within it the certain attempt probability. From after into it time slot, Where the user of each node is delivered the packet successfully into the DSA. where into the distributed manner without online coordination or into it. The optimal rectification for this is make high cost due to this observe of the agent that states in the environment. So, for this to tackle this problem, so for that we developed a DSA algorithm which is shows the specific time slot into the environment by this we access the shortest action path and pass coordinate the message in o the environment. Which shows the principle for the implementation of the algorithm.**

***Index Terms* - Wireless networks, multi-user DSA, Agent, Environment, states and action, multi-agent learning, deep reinforcement learning.**

I.INTRODUCTION

Dynamic spectrum access (DSA) is a set of techniques where it is based on theoretical concepts in network information and game theory that is being researched also developed to improve to increase the performance of a communication network as a combine of a whole into it. The concept of DSM also draws principles from the of cross-layer, optimization, artificial intelligence, machine learning etc. It has been recently made possible by the availability of software radio due to development of fast enough processors both at servers and at terminals. For cooperative optimization. This can also be compared or related to optimization for the deep q-learning of the access where it is optimization. Where the DSA in the reinforcement learning and the performance optimized into the spectrum access. By this which is shows the environment is set the allowable action in to the DSA. And the certain states the environment and the current states associated with the performance of the channel into the reinforcement learning.

II.LITERATURE REVIEW

we provide the literature survey of the dynamic spectrum access techniques. Various approaches envisioned for dynamic spectrum access are broadly classified into under two models: dynamic exclusive use model, And the message that transfer into the dynamic spectrum access. Which is shoes the reinforcement learning, open sharing model, and hierarchical access model. Are based on this taxonomy,

[1] we provide a complete overview of the technical challenges and recent advances under each of to the model.MC Vuran et al.

[2] presented Dynamic Spectrum Access (DSA) and cognitive radio networks. The cognitive radio which is used to have a unlicenced band where the unused band are will be used into it and by this, we get a spectrum access that is moreover make the data or message to transfer from one place to another. Where this are the dynamic access spectrum were showed into the survey in the cognitive radio access in the reinforcement learning.

[3] proposed, the Dynamic spectrum access DSA are shows the performed of the spectrum access into the environment that receiving the spectrum and then transfer the dynamic plane access into the efficient of the frequency. Then this shows the DSA access into the multi-user reinforcement learning.

PROPOSED SYSTEM

Here the schematic proposed system of the diagram that is shown in fig.1.

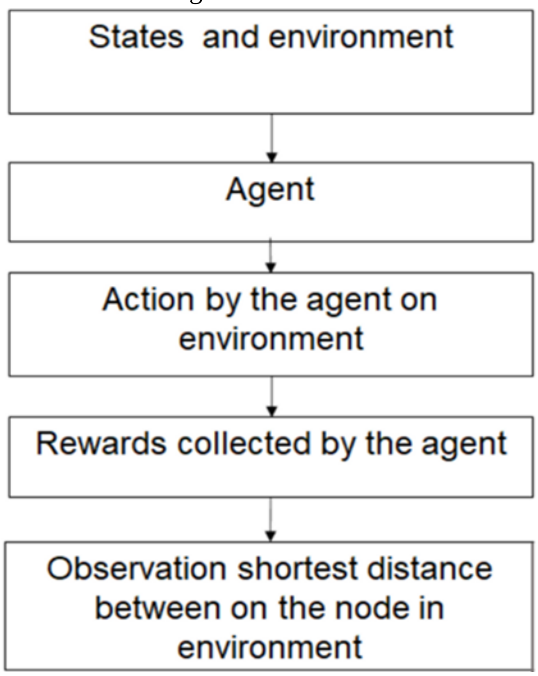


Figure:1.Block diagram of proposed system

A.STATES AND ENVIRONMENT

The reinforcement learning briefly is a paradigm of Learning Process in which the agent which is interact with the environment continuously into it. Where the environment which have a various obstacles and interaction into it. Which is learn by the agent in the state using reinforcement learning. The agent in the state while from a set of allowable actions within it. https://www.programmersought.com/article/5590537 2988/

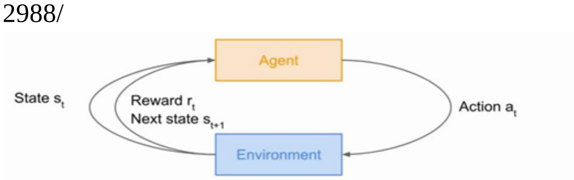


Figure:1a.Sample image

B.AGENT

The Agent in the reinforcement q-learning is used in the environment in which by using the reinforcement learning the with the state in environment make understand and utilize the environment of the current sates to increase and maximize the rewards. That this potential reward that expected the value of sum of weightage to all the rewards and the current state by the learning agent in the reinforcement learning.

C. ACTION ON ENVIRONMENT

The Q-Learning is to assign each Action-State pair a value. That the agent in the environment which is achieve to the maximum number of rewards to attain is in the certain states. So, by this if the which is make increase the potential rewards and action. Where the agent will get into in the state's s, where the highest rewards in the action by the environment to the agent is occur.

D. REWARDS

In the reinforcement q-learning is all about collecting rewards. The agent main goal and aim is to make the rewards maximization. The ultimate determine of the agent is to make the rewards colleting in the q-learning. The emphasis here is that the agent can also lose point due to the dangerous action. Which is the number of points by given to the current states in traffic within to does not focus on it.

Exploration – Exploration refers to the process of discovery the new generate product, then the resources, knowledge and opportunities, and it is associated to the performance of the radical changes and it make and create learning through experiment. Exploitation which is refers to the refinement and on to the existing products, resources, knowledge and competencies.

Exploitation – Exploitation is where the involves the activities such as make the refinement, by process increasing, selection, performance, to the execution on into it. The activities of exploitation are also different in terms of their on the various structures, process, cultures, and the capabilities on the changes make on the environment by the exploitation.

E. DISTANCE OF NODE IN ENVIRONMENT

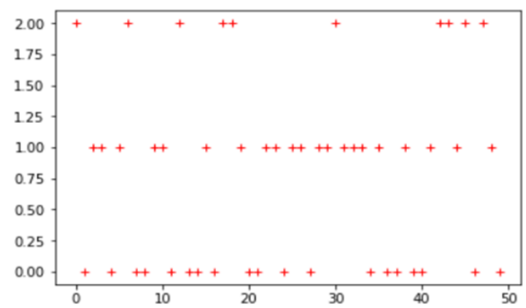
The time being is used on the Bellman–ford. If there will be no negative cycles, then we can solve it by the O (E + VLogV) time using Dijkstra's algorithm. Since that the graph is unweighted, so that we can solve this problem in O (V + E) time. The idea is to use a modified version of Breadth-first search in which we keep storing the predecessor of a given vertex while doing the breadth-first search. This algorithm will work even at the negative weight cycles are present in the graph. Now we get the length of the path from source to any changes vertex in O time from array d, and for printing the path from source to any vertex we can use length of array and that will take the O(V) time in worst case as V is the size of array of the search. So, most of the time of the algorithm is the Breadth-first search from a given source which we know takes O(V+E) time. Thus, this the time complexity of our algorithm is O(V+E)

F. OBSERVATION FORM CALCULATION

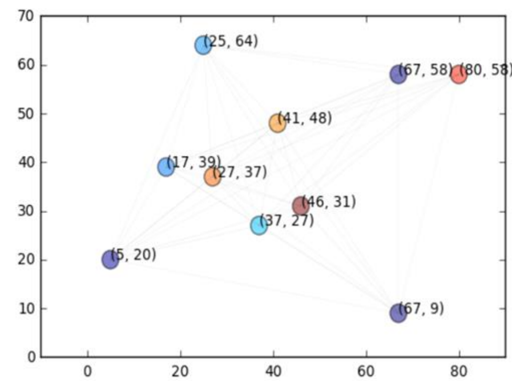
For action = [1,2,0] obs is [(1, 1.0), (1, 1.0), (0, 0.0), array ([0, 0], dtype=int32)]
The format of obs is [(ACK1, REWARD1), (ACK2, REWARD2), (ACK3, REWARD3), ..., (ACKn, REWARDn), (CAP_CHANNEL1, CAP_CHANNEL2,.., CAP_CHANNEL_k)].
When we pass actions to the environment, it takes these actions and returns the immediate reward as well as acknowledgement of the channel. Finally, it also returns the residual capacity of the channel (remaining capacity).
Where this are the 1,2,3 represents user which is user 1 and the user 2 then user 3 respectively for the first n of the tuples where n is number of users and k is the number of channels. Last element is an array [CAP_CHANNEL1, CAP_CHANNEL2, CAP_CHANNEL_k] denotes the remaining channel capacity or the fact that channel is available or not.
Since both channels were available at the beginning, user 1 and 2 allocates channel 1 and 2 respectively and user 3 remains idle. This can be concluded by the resulting output where there is (ACK, REWARD) pair as (1, 1.0) for user 1 and 2 and is (0, 0.0) for user 3. Both the channels are allocated by user 1 and 2 therefore last element is array ([0,0]).

III.RESULTS AND DISCUSSION

The Result of the graph shows total reward generated per time slot for 50 slots:
Here reward 0.0 means no user was able to send the packet and both the channels were free while reward 2.0 means both the channels were being used without collision and any one user was not sending the packets. Where the environment states and the virtual distance between the node and the formation of cluster are showed into the result.

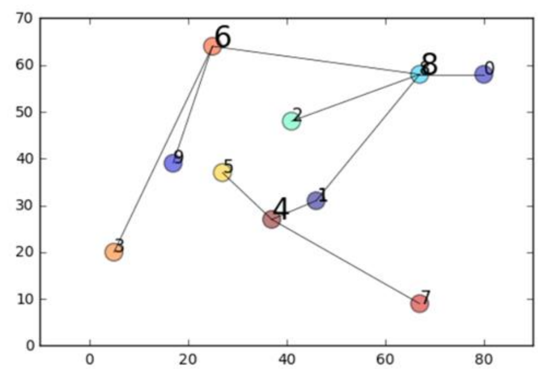


1.Action on environment by agent.



2.Generate nodes on random positions on canvas of width 'W'

The virtual distance to infinity for these nodes



CLUSTERING FOR WIRELESS SENSORY NODES FOR DISTRIBUTED COMPUTATION OF DRQN.

IV.CONCLUSION AND FUTURE WORK

The problem of dynamic spectrum access for network utility maximization in multichannel wireless networks was considered. For that we developed the novel distributed manners in which the dynamic spectrum access for the q-learning is used by the shortest path between the node in the environment is take and the agent which make by the q-learning is increases the reward maximization. By that the large space without message or online coordinates is be used due to the q-learning algorithm in which by the use of deep multi-user reinforcement learning.

The future implementation of the q-learning reinforcement is to be on 5G network where the unused band in the spectrum are to be used with the connected by the small receivers by the use of it take a shortest path on into it. Where is to be future work on the deep multi-user reinforcement learning.

REFERENCES

[1] U. Varshney, “Pervasive Healthcare,” IEEE Computer, vol. 36, no. 12, Dec. 2003, pp. 138–40.

[2] H. Furuhata, “Electromagnetic Interferences of Electric Medical Equipment from Hand-held Radiocommunication Equipment,” Int’l. Symp. Electromagnetic Compatibility, 1999, pp. 468–71.

[3] “National Standard of Canada CAN/CSA — C22.2 no. 60601-1-2:03 (IEC 60601-1-2:2001) Medical Electrical Equipment — Part 1-2: General Requirements for Safety — Collateral Standard: Electromagnetic Compatibility — Requirements and Tests,” 2003.

[4] H. Hong, Y. Ren, and C. Wang “Information Illuminating System for Healthcare Institution,” Proc. Int’l. Conf. Bioinformatics Biomed. Eng., May 16–18, 2008, pp. 801–4 .

[5] Kaushik R. Chowdhury, Cognitive Radio Applications in Healthcare and Vehicular Emergency Networks, IEEE Communication 2010

[6] U. Varshney, “Pervasive Healthcare,” IEEE Computer, vol. 36, no. 12, Dec. 2003, pp. 138–40.