Latency based Re-Enforcement Learning over Cognitive Software Defined 5G Networks

Baburao Kodavati[1], Madhu Ramarakula[2]
[1]JNTUK, Department of ECE, URCET,Telaprolu,AP, India - 521109.
[2]Department of ECE, JNTUK, Kakinada, Andhra Pradesh, India -533003.
[1]baburaokodavati@gmail.com, [2]madhu_ramarkula@jntucek.ac.in

2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9441725

Abstract - In present days, in software defined networks, cognitive network (CN) is the key control to enable Internet of Things (IoT) services, whenever CN plays a important task in future network internet applications used in different types of real time applications such as agriculture, monitoring healthcare and smart metering with different scenario's. Because of increasing popularity of variety software defined crowed applications; it follow on low transmission rate in communication, major challenge behind this is to improve efficiency of packet transmission in software defined network data transmission. So that in this paper, we propose A Novel Re-Enforcement Learning Approach (NRELA) to improve the transmission efficiency using cognitive radio software defined networks through multiple channel to increase network throughput. To increase the system ability in between nodes with respect to re-enforcement learning approach to find optimal data communication in software defined networks. Establish the connection between different nodes to accelerate the solution for efficient data transmission in software defined communication. An experimental result of proposed approach gives better and efficient latencyand energy results in data transmission in software defined networks.

Index Terms: Software defined networks, learning approach, cognitive networks, packet data transmission, multiple channel communication and network throughput.

I. Introduction

Software defined Sensor networks are relied upon to be coordinated into the IoTs [1-2], where the sensors that are reconfigurable, adaptable join the internet and are being used in different areas with the lot of scope[3-8], for example, enormous information applications, IoTs, E-business, medicinal gadget [9-10], computer generated experience and increased reality, and condition checking. The system condition additionally will in general become progressively entangled, and the correspondence assets become progressively rare. It is an extraordinary test to the remote sensor systems and IoTs. Adventitiously, the subjective system innovation can make up for these lacks [11-15]. Subjective hubs are wise remote gadgets that can detect the earth, watch the system changes, utilize the information gained from the past association with the system, and settle on canny choices to take advantage of the lucky breaks to transmit. The procedure of ceaselessly detecting the earth data, trading control data, learning data, choosing and executing a technique in the system can give the capacity of insight and flexibility to the remote sensor systems and the future IoTs. Consequently, the psychological radio innovation is a key correspondence approach for asset obliged remote sensor systems and future remote system [1-5]. At the point when psychological clients, i.e., sensors in remote sensor systems, get to the range, to viably utilize the system assets and fulfill the throughput interest for sight and sound applications, compelling instruments are required to organize the activities of the intellectual clients (tx power control, range get to, tx planning, et al.). With the quick increment in number of remote gadgets in the IoTs, more information will be put away in the system hubs. In this way, the strategy to quickly advance information with the constrained extra room and transfer speed is an extraordinary test for the present remote system of Internet of Things.

We propose A Novel Re-Enforcement Learning Approach (NRELA) to improve the transmission efficiency using cognitive radio software defined networks through multiple channel to increase network throughput. To increase the system ability in between nodes with respect to re-enforcement learning approach to find optimal data communication in software defined networks. Scheduler plays an essential job in cell systems, which makes it the "mind" of the entire framework [3]. Its fundamental capacity is to make choices about how radio assets are designated. Customary planning calculations for the most part take channel conditions what's more, QoS prerequisites into thought, and settle on choices as indicated by certain conditions. This deterministic method for booking experiences absence of adaptability.

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

Markov Decision Process (MDP) is introduced to address the issues present in the system by performing Dynamic Programming technique [4]. But this technique doesn't moderate the issue scale which is augmented. Hence, to allocate this issues, deep reinforcement learning (DRL), which joins Deep Neural Networks (DNNs) and Reinforcement Learning (RL) techniques are introduced. Multi-Measurement information will plays an elective role in finding the route of its capacity.

II. Related Work

At present, many existing written works (see in [5]-[15]) have considered the issue of system information transmission with obscure condition data in a cross-layer structure way. Among these written works, [5] and [6] have analyzed the versatile adjustment calculation in the information transmission arrange, while [7] concentrated on the dependable course revelation to lessen the time required for the information online. The adjustment of power and mindfulness schedulerQuality of service (QoS)for medium access control (MAC) layers at uplink and downlink to arrange the activity of the lower layers for asset proficiency was proposed in [8]. The throughput along with decency is concentrated in [9]. Contingent upon the inclination of the two highlights, calculations and techniques are proposed to dole out or plan clients to organize to boost throughput, augment decency or finding the proper harmony between the two. The better performance of system administration is obtained by using topology configuration, steering conventions and cross layer plan of MAC in [30]. Markov decision process (MDP) [1-3] will overcome the issues that are streaming from the remote systems. In any case, it is hard to take care of the MDP issue in light of the fact that the MDP has numerous factors. Thusly, the fortification learning technique can be acquainted with tackle it. The research for various applications and administrations in circulated remote systems is carried out in [33-37]. The performance parameters like Quality of experience (QoE) and Quality of Service (QoS) are related with booking instruments and the planning systems are considered in across-layer way.

This related research takes a shot at EH based frameworks that can be classified into two different classes dependent on the accessibility about the information of vitality entries. Basically, the methodologies that are disconnected with the stochastic frame work should consists of non-casual learning [11] [12] [13]. Specifically, the ideal uplink asset assignment was examined in [12] for the situation where two EH clients initially gathered vitality from the remote flag and after that agreeably sent data to the passageway. Additionally, the ideal packet planning over different access diverts was examined in [13], with the objective of limiting the time by which all parcels from the two clients are conveyed to the goal. The below average involves online methodologies [14] [15] [16]. Creators in [14] considered a multi-get to remote framework with EH transmitters, and the entrance issue was displayed as an in part recognizable Markov choice procedure (POMDP). In [15], the arrangements for controlling the power ideally for EH hubs in a multi-get to framework was studied and while performing procedure of EH will collect the elements from dam model.

In [16]at transmitter side few measurement learning aspects with respect to the dynamic framework is studied. In numerous reasonable applications, the total non-easygoing learning or even factual information of the framework elements (counting both the channel and vitality parts) probably won't be accessible, particularly when the EH procedures are non-stationary or from sources with obscure circulations. For instance, in a remote system with sunlight based EH hubs appropriated arbitrarily over a geological region, the qualities of the collected vitality at every hub rely upon the hub area, and change after some time in a non-stationary style [17]. It is very difficult to get the information which is coming from different elements from various sources.

III. Background Approach

CIoT that exists together with an authorized framework is shown in figure 1. In this we studied about the point ot guide transmission channels which are transfeered to the sin and area channels.

The parcels from its K neighbor hubs are accumulated by one hand-off, and at indistinguishable length Lthese packets are put away in K cushions. Thosepackets that originate from K neighbor hubs are accepted to have Poisson dissemination with indistinguishable entry rateλ.

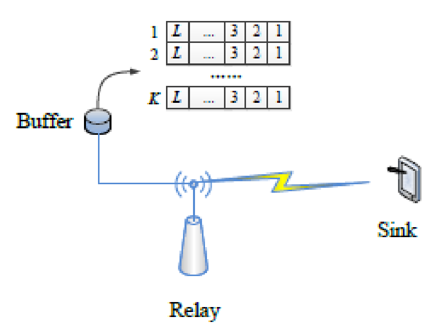

Fig 1 Channel transmission methodology with respect to node communication.

M channels are autonomous and indistinguishably dispersed. The transmission calendar is chosen by the hand-off. A channel is used by the hand-off process to transmit the hub to sink. The transfer does not include the transmission of bundle if the channel state is poor. At the point when a specific cradle is full, in the event that the transfer does not transmit parcel for it, at that point, the bundle is lost if parcels keep on landing in the following casing. Along these lines, and for the transmission process of bundles completely depends on the channel state and the support expresses that relate to the correspondence sets and transmission mode.

IV. Re-enforcement Learning with Multi Channel Access Procedure

In this section, we discuss about re-enforcement learning procedure with multi channel access control in software defined network communication. Basic problem formation described in figure 2.

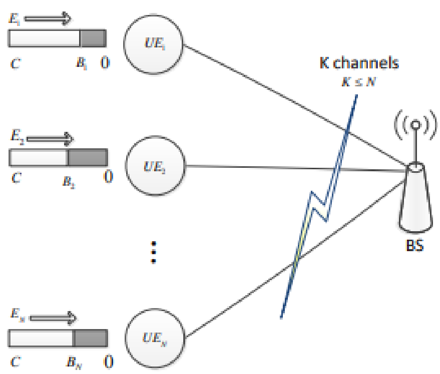

Fig 2 System implementation for k-channel access control

In this both BS and UEs are used for learning procedures. The main intent of BS is to learn about the current channel additions and in the same way the main intent of UEs is to check the states of battery. In this the system works depend on the channels performance. Generally, when framework enters into a new state then it changes the current position from BS to TS by using RL system. After this process, all the channels in the system will communicate with the UEs. Now, after some time, UEs will transfer the information to the BS by controlling the P individually. At last the information is executed and safely, transferred and saved into battery for future use.

4.1. Problem Description

Basically, to boost up the framework a control strategy is needed for the BS to expect the uplink aggregate rate. In this TS is divided into two sections mainly, present channel state and present UE battery state. The present channel state is represented as Ht = {H1t, · ,HNt} and the present UE battery state is represented as Bt = {B1t, · , BNt}. The combination of these two will give the St = {Ht, Bt} which is a UE choice decision. The condition P Kt ∈ A, with |Kt| = K, N i=1 Iit = K will be satisfied by using these strategies. The signal Rt is described as shown in below equation which is nothing but total rate.

$$R_t^{\lambda} = \sum_{k=t}^{\infty} \lambda^{k-t} R_{k+1}$$

$$= \sum_{k=t}^{\infty} \lambda^{k-t} \sum_{i \in K_{k+1}} z_{i(k+1)} F \log\left(1 + \frac{PH_{i(k+1)}}{\sigma^2}\right)$$

Main re-enforcement learning model maximize the cumulative discounted with access controly from starting state with optimization formulation for different channel

$$\max_{\pi} J_1(\pi)$$

4.2. Re-enforcement Model for Multi-Access Control

From figure 3 we can observe that the proposed learning system will overcome the issues that are coming from uplink. This will mainly control the access in effective way. At every starting point of TS, BS gets the data based on controller.

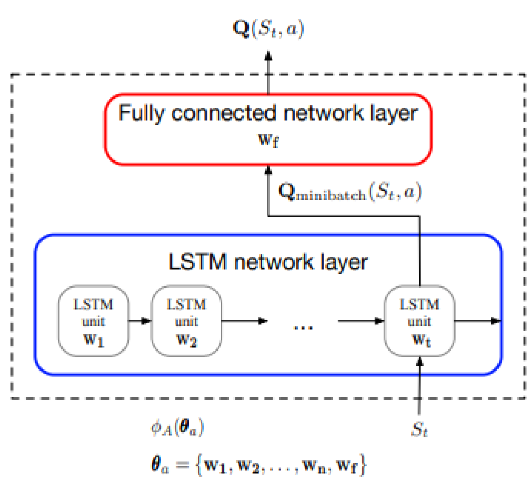

Fig 3 NRELA Network simulation model with access control

In this the layer which is available in the information is more smaller than the reinforcement model. Qminibatch(s, a) ∈ R batch size×L, is the present state model. Here the size of activity space is L based on the condition a ∈ A.

Now to alter the position vector for Q-esteem, a system layer is introduced ot used. The condition for this system layer is given as i.e., Q(s, a) ∈ R 1×L. from figure 3 we can observe that φA(θa) is nothing but an neural system which gives the output Q(s, a). In this LSTM layer parameters {w1, · ,wn}are quiet associated with the parameter wf. The quantity of LSTM units is donated by n. Q(St, an) is evaluated by φA(θa) when the learning schedule is opened.

1: Initialize the experience memory D,
2: Initialize the parameters of activity generator organize φA with irregular loads θa,
3: Initialize the absolute number of scenes Ep,
4: Initialize the earth and get beginning perception S1,
5: for t = 1, · , ∞ do
6: in the event that irregular() ≤ at that point 7: Select an arbitrary activity At ∈ A;
8: else
9: Compute Q(St, a) for all activities a ∈ An utilizing φA,
10: Select At = arg max a∈AQ(St, a).
11: end if
12: Execute At, watch compensate Rt and new state St+1,
13: Store progress (St, At, Rt, St+1) in D,
14: Sample irregular little clump of advances (˜s, a, ˜ r, ˜ s˜) from D,
15: Set yt = ˜r if t + 1 is the terminal advance of the scene (t + 1 = Ep); generally, yt = ˜r + γ maxa˜ Q(˜s, a˜; θ−a),
16: Perform stochastic slope drop venture on the misfortune work Lt(θa) = (yt − Q(˜s, a˜; θa))2 to refresh arrange parameters θa as indicated by (13).
17: end for

Algorithm 1 NRELA algorithm procedure with respect to access control.

At that accomplishes the most extreme Q(St, At), and the ideal approach is the insatiable strategy if Q(St, a) can be consummately assessed. Q(s, an) is enough in this approach to full fill the requirements. New activates will takes place when BS will investigate about the approach. The covetous approach; in investigation the BS takes activities arbitrarily with the point of finding better arrangements. The parity could be acknowledged by the -avaricious activity determination method(as given later in Algorithm 1) at each schedule vacancy, which either takes activities haphazardly to investigate with likelihood or pursues the insatiable arrangement to misuse with likelihood 1 − , where 0 << 1. In the wake of executing the chose activity At, the BS gets the remunerate Rt and the framework changes to the new state. We use experience replay to store the BS's encounters at every TS, which is indicated by tuple et = (St, At, Rt, St+1) in a dataset D = {e1, ..., et}. The value of L is set to replay memory size, which signifies that L experience tuples could be stored. Here, et is produced by the control approach π(a|s). In each TS, rather than refreshing θa dependent on advances from the present state, we arbitrarily test a tuple (˜s, a, ˜ r, ˜ s˜) from D. Refreshing system parameters along these lines to dodge issues brought about by solid connections among changes of a similar scene [22]. We parameterize a surmised worth capacity Q(s, a; θa) utilizing the proposed learning system with system parameters (loads) of θa as shown in figure 3. With the examined advances, yt = ˜r + γ maxa˜ Q(˜s, a˜; θ − a) is the objective Q-esteem with system loads θ −a acquired from past cycle. Finally multi channel access control with different attribute sequences in software defined network communication.

V. Experimental Simulations

In this section, we describeNovel Re-Enforcement Learning Approach (NRELA) to improve the transmission efficiency using cognitive radio software defined networks through multiple channel to increase network throughput and energy consumption with respect to different nodes. For efficient simulation setup, use NS3 with Ubuntu operating system with different node communication. Simulation parameters used in our implementation shown in table 1.

Parameter	Value
Area of network	1500*1500
Noodes with presented area	60
Time of Simulation	30S
Range of Transmission	250 m
Speed of Mobility	0-20m/sec
Number of Blackhole nodes	10
Check point nodes	4 nodes(Fixed)

Table 1 Basic simulation parameters.

Based on above simulation parameters, better and efficient results are achieved in the proposed structure with comparison of energy optimization with packet data loss with bit rates in between nodes in software defined communication.

Time comparison results in software defined networks with nodes communication with respect to time for packets droping in middle of data delivery by hop by hop communication. Table 2 shows analysis results with respect to time in data communication between nodes.

Number of Nodes	NRELA	CWA-CD
10	1.3	1.28
20	2.0.	2.7
30	3.2	2.8
40	4.2	4.5
50	5.6	4.9
60	5.4	5.5

Table 2: Node communication with time comparison results.

Time comparison results for different approaches with respect to different nodes with efficient communication.

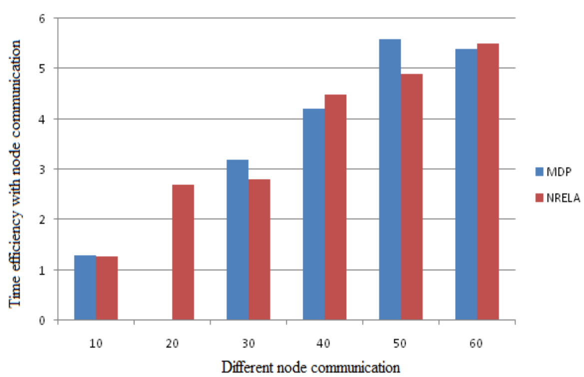

Fig 4 Comparison of time with respect to node communication

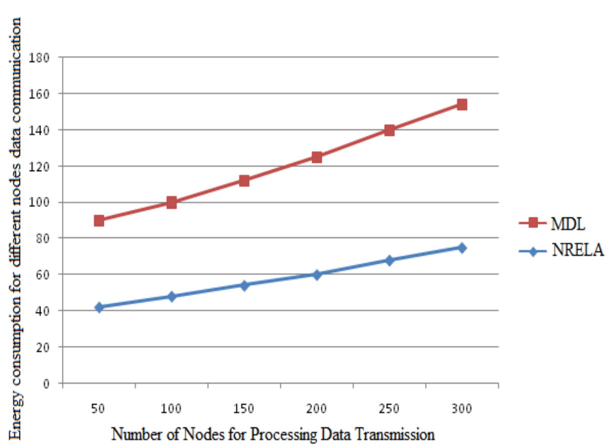

Fig 5 Performance of proposed approach in terms of energy optimization with different nodes.

As the number of nodes increased then the number of outcomes in real time data transmission of host to host communication energy consumption in our NRELA schema gives efficient communication without loss of data delivery in software defined network communication as shown in figure 4.5.

VI. Conclusion

In this paper, we present Novel Re-enforcement Learning model (NRELA) to provide solution for the user access control and also describe the battery prediction problems with respect to multi user energy sharing based communication system. The main intent of proposed system is to maximize the uplink sum rate which is driven by system instantaneous information in data sharing. Energy optimization is also discussed to minimize the packet loss. Simulation results of proposed approach are satisfied different conditions with increase effectiveness in terms of parameters like throughput, packet loss, latency and energy optimization.

References

[1]. Jiang Zhu, Yonghui Song, Dingde Jiang, " A New Deep-Q-Learning-Based Transmission Scheduling Mechanism for the Cognitive Internet of Things"2327-4662 (c) 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

[2]. Khalil N, Abid M R, Benhaddou D, et al. Software defined sensors networks for Internet of Things[C]// IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing. IEEE, 2014:1-6.

[3]. S. Jeschke, et al. (Eds.), "Industrial Internet of Things: Cyber-manufacturing Systems," ISBN 978-3-319-42559-7, Springer, Switzerland, 2016.

[4]. Yang J, He S, Lin Y, et al. Multimedia cloud transmission and storage system based on Internet of Things[J]. Multimedia Tools and Applications, 2015: 1-16.

[5]. Yuan B, Fu C, Chen D. Building a Large Scale Software defined Sensor Network for the Industrial Environment[C]// IEEE, International Conference on Embedded and Real-Time Computing Systems and Applications. IEEE Computer Society, 2016:96-96.

[6]. Khan A A, Rehmani M H, Rachedi A. When Cognitive Radio meets the Internet of Things?[C]// International Software defined Communications and Mobile Computing Conference. 2016:469-474.

[7]. Haustein, Thomas, et al. "Cognitive Software defined Communications–A paradigm shift in dealing with radio resources as a prerequisite for the software defined network of the future–An overview on the topic of cognitive software defined technologies." Frequenz 70.7-8 (2016): 281-288.

[8]. Otermat D T, Kostanic I, Otero C E. Analysis of the FM Radio Spectrum for Secondary Licensing of Low-Power Short-Range Cognitive Internet of Things Devices[J]. IEEE Access, 2016, PP(99):1-1.

[9]. Somov A, Dupont C, Giaffreda R. Supporting smart-city mobility with cognitive Internet of Things[C]// Future Network and Mobile Summit. 2013:1-10.

[10]. Nitti M, Murroni M, Fadda M, et al. Exploiting Social Internet of Things Features in Cognitive Radio[J]. IEEE Access, 2016, PP(99):1-1.

[11]. Man Chu, Hang Li, “Reinforcement Learning based Multi-Access Control and Battery Prediction with Energy Harvesting in IoT Systems”, arXiv:1805.05929v2 [eess.SP] 21 Sep 2018.
[12]. A. Haldorai and A. Ramu, “Security and channel noise management in cognitive radio networks,” Computers & Electrical Engineering, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784
[13]. A. Haldorai and A. Ramu, “Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability,” Neural Processing Letters, Aug. 2020. doi:10.1007/s11063-020-10327-3
[14]. A. Cammarano, C. Petrioli, and D. Spenza, “Online energy harvesting prediction in environmentally powered software defined sensor networks,” IEEE Sensors J., vol. 16, no. 17, pp. 6793–6804, Sep. 2016.
[15]. H. U. Yildiz, V. C. Gungor, and B. Tavli, “A hybrid energy harvesting framework for energy efficiency in software defined sensor networks based smart grid applications,” in 2018 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net). Capri, Italy: IEEE, June 2018, pp. 1–6.
[16]. A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, “Relaying protocols for software defined energy harvesting and information processing,” IEEE Trans. Software definedCommun., vol. 12, no. 7, pp. 3622–3636, Jul. 2013.
[17]. R. Zhang, J. Wang, Z. Zhong, C. Li, X. Du, and M. Guizani, “Energyefficient beamforming for 3.5 GHz 5G cellular networks based on 3D spatial channel characteristics,” Elsevier Comput. Commun., vol. 121, no. 5, pp. 59–70, Mar. 2018.
[18]. X. Zhou, B. Bai, and W. Chen, “Greedy relay antenna selection for sum rate maximization in amplify-and-forward mimo two-way relay channels under a holistic power model,” IEEE Trans. Software definedCommun., vol. 19, no. 9, pp. 1648–1651, Jun. 2015