International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:14, No:11, 2020

Reinforcement Learning-Based Coexistence Interference Management in Wireless Body Area Networks

Izaz Ahmad, Farhatullah, Shahbaz Ali, Farhad Ali, Faiza, Hazrat Junaid, Farhan Zaid

Abstract**—**Current trends in remote health monitoring to monetize on the Internet of Things applications have been raised in efficient and interference free communications in Wireless Body Area Network (WBAN) scenario. Co-existence interference in WBANs have aggravates the over-congested radio bands, thereby requiring efficient Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) strategies and improve interference management. Existing solutions utilize simplistic heuristics to approach interference problems. The scope of this research article is to investigate reinforcement learning for efficient interference management under co-existing scenarios with an emphasis on homogenous interferences. The aim of this paper is to suggest a smart CSMA/CA mechanism based on reinforcement learning called QIM-MAC that effectively uses sense slots with minimal interference. Simulation results are analyzed based on scenarios which show that the proposed approach maximized Average Network Throughput and Packet Delivery Ratio and minimized Packet Loss Ratio, Energy Consumption and Average Delay.

Keywords**—**WBAN, IEEE 802.15.4 Standard, CAP Super-frame, Q-Learning.

I. Introduction

NEW technologies have been developed in wireless sensor networks (WSNs) and transmitting to perform different functions. One of them is WBAN. WBAN consists of lightweight devices (sensors) that can be placed on the body or around the body [3]. These sensors are short-range and low power. Sensors are connected over a wireless channel. Sensors are short-range wireless communicating devices, so the collected data are transferred from sensors to the coordinator, and then forwarded to the health monitoring center for further processing through the internet [4]. Task Group 6 (TG6) was recognized in November 2007 [5] by IEEE 802.15 to standardize WBAN and complete a baseline document in February 2012 [6]. Two standards are used in WBAN: IEEE 802.15.4 and IEEE 802.15.6 [1]. The IEEE 802.15.4 standard defines the physical and MAC specification for low rate WBAN, and IEEE 802.15.6 standard defines the physical and MAC layer for WBANs to communicate with high data rate but in short range. The quality of service (QoS) is an initial problem of WBAN [2].

The human body is subject to mobility due to which a WBAN may come closer to another WBAN. In such conditions, WBANs may create inter-WBAN interference and intra-WBAN interference which results in performance degradation. If two sensors or more sensors of a WBAN wants to transmit data at the same time to a coordinator then it creates Intra-WBAN interference. LRPWAN standards for on-body communication are IEEE 802.15.4 [11], IEEE 802.15.6, and Bluetooth which are subjected to intra-WBAN interference. WBANs are designed for monitoring e-health applications like temperature, blood pressure, heart rate, Electro Encephalography (EEG). WBAN is now used for daily routine and medical conditions [1].

Two WBANs coexist where nodes are fixed on the body and connected to their coordinator but are in the range of another WBAN so there is a need to careful access medium protocol to avoid collisions and interference that belongs to other WBAN [7].

Many systems (channel hopping, Beacon shifting, and active super-frame interleaving) are used to resolve the coexistence of WBANs. However, homogeneous coexistence interference is common and periodic to resolve the coexistence of WBANs. Collisions probability is high in some slots as compared to other slots and therefore we can manage these slots by Reinforcement Learning.

II. Literature Review

Coexistence interference among WBANs can be mitigated by swap the current channel to other channels or arranging of super-frame. MAC protocols are proposed by different authors who aim to maximize throughput and minimize transmission power while mitigating interference.

In [8] and [9] Dynamic Coexistence Management (DCM) is designed to manage the coexistence of IEEE 802.15.4 standard WBANs. DCM works on beacon replacement and channel switching. Candidate channels are scanned by the coordinator for possible channel switching. Two possibilities for scanning candidate channel are (i) full Beacon Interval

Izaz Ahmad is Faculty Member of Abasyn University Peshawar with specialization wireless and WBAN at Bahria University Islamabad Campus, Pakistan (e-mail: izazahmad445@gmail.com)

Farhatullah, Faiza and Farhan Zaid are currently studying at Abasyn University Peshawar, Pakistan with specialization Machine Learning, Deep Learning, and WBAN (e-mail: farhatkhan8398@gmail.com, faizahabibi858@gmail.com, farhanzaidkhan@gmail.com).

Shahbaz Ali completed last degree from national college of business administration and economic Lahore with of Machine Learning, cyber security, and WBAN (e-mail: ranaalipk7@gmail.com).

Farhad Ali completed master degree from Federal Urdu university Islamabad with of Machine Learning, cyber security, and WBAN (e-mail: farhadms13@gmail.com).

Hazrat Junaid is currently studying at university of Malakand (UOM) Dir Pakistan with specialization at Machine Learning, Deep Learning, WBAN and Wireless Ad-hoc (e-mail: abidj3692@gmail.com)

(BI) scan and (ii) inactive period scan. In full BI scan, coordinator receives information of preexisting WBANs, while in inactive period coordinator finds the gap between the active periods without missing a beacon.

In interference aware channel switching algorithm, coordinator monitors the signal to interference ratio (SIR) [10]. If SIR is below the threshold value, user switches to a new channel. Three values of threshold have been used, i.e. low, medium, and high.

Dynamic Resource Allocation (DRA) depends on spatial reuse to avoid interference between multiple WBANs. Every coordinator creates a list of interfering sensors of other WBANs that are in its transmission range and then broadcasts it to its neighbors. Every coordinator takes its own decision for channel assignment to avoid interfering with sensor nodes [11].

In [12], authors propose an Interference Avoidance Algorithm (IAA) to mitigate the co-channel interference in WBAN. IAA uses CSMA/CA and FTDMA scheme, CSMA/ CA is used between the sender and relays, while FTDMA is used between relays and coordinator. Two possible options of using base channel are; with Back-Off time and without Back-Off time. In IAA, super-frame consists of two parts: Contention Access Period (CAP) and Time Division Multiple Access (TDMA). CAP is separated by two parts: CAP1, and CAP2. CAP1 is further divided into CAP1A and CAP1B sub parts. CAP1A is for those sensors that use base channel, CAP1B is for those sensors that use base channel with Contention Window (CW), and CAP2 is for those sensors that use reserved channel. A source will use CAP1A, if it experiences no interference and has Signal to Interference and Noise Ratio (SINR) greater or equal to threshold. A source will use CAP1B, if it experiences interference and has SINR greater or equal to threshold, will also use base channel after finishing its CW. A source which has SINR is less than threshold will use CAP2 part.

In [13], authors proposed Smart Channel Assignment (SCA) for a static scenario between coexisting WBANs. SCA is an adaptive interference minimizing system for coexisting WBANs. The mobility of nodes within each WBAN is considered in the proposed scheme. The SCA minimizes the interference and power consumption of sensors in a WBAN. SCA works in four phases: (i). Orthogonal Transmission: The shared channel is divided into Number of coexisting WBANs (Nc) time slots, and more divided into Number of sensors (Ns) in each WBAN portion. (ii). Interference Set Formation: Each coordinator creates a table energy power of each sensor of all coexisting WBANs. Each coordinator compares energy power of its own WBAN sensors with the received power of other WBAN sensors. A sensor that has greater received power than threshold will be considered as an interference sensor. (iii). Information Exchange: interference lists are broadcasted by every coordinator. Coordinators create an Interference Set of those sensors which have high level of interference than the interference threshold value. (iv). SCA: Sensors that exist in the Interference Set (IS) are allotted with the same time slots as previously, while the remaining sensors of all coexisting WBANs that are not in the IS allocated equal time slots.

In [14], the authors propose a QoS-based scheduling for multiple coexisting WBANs. Coordinators of all interfering WBANs share information before data transmission, thus coordinator will decide which sensor can access the channel. The channel is divided into beacon periods. Network traffic is arranged on priorities based by coordinator at each WBAN.

A highest priority sensor will get first. In the first beacon, coordinator broadcasts its WBAN identification and timer information. If other WBANs are within its transmission range and listen to the first beacon, coordinators save the first beacon information in its tables. In case of overlapping frequency range and slots, priority level of slots is compared in interference range, lower priority slot will be delayed to transmit later. The coordinator broadcasts the traffic information of all sensors in the next beacon.

In [15], authors propose a Distributed Active Channel Reservation Scheme (DACROS) to mitigate coexistence interference using ready to send (RTS) and clear to send (CTS) to reserve a channel for super-frame. When coordinator broadcasts a beacon that consists of time slots for the super-frame of WBAN, all other networks, i.e., IEEE 802.11, and ZigBee, keep silent and do not transmit in that time to avoid collisions.

Two-layer MAC (2L-MAC) was proposed for the interference mitigation of WBANs [16]. The coordinator of a WBAN always checks for the idle channel before polling its sensor nodes. The proposed scheme has two phases: (i). polling with Back-Off mechanism: The coordinator uses carrier-sensing to check the channel is free or not. If coordinator finds the channel free, then it sends a polling frame to all sensors. If it finds the channel busy, coordinator performs Back-Off procedure and waits until the channel became free. The coordinator knows the priority of each node and therefore sets the CW based on priority. Back-Off Time (BT) is decreased in finding the channel free. When BT becomes zero, coordinator sends a polling frame to sensors and waits to transmit data from sensors. (ii). Channel switching mechanism: Each sensor waits for a delay time (Total-delay), that is the time from its activating wakeup event to receiving polling frame from coordinator. If the Total delay is greater than threshold, both the sender and coordinator switch from current channel to backup channel.

In [17], the author suggests a hybrid channel access system called the HCA scheme to minimize coexisting interference between IEEE 802.15.4-based WBANs. The HCA scheme has two phases. In the first phase, the CFP shift and the Mini slot assignment are used to prevent collisions between the CFP portions of the super-frame. When the first step was added, the PRR is still smaller than the threshold value. WBAN switches the control channel mode from CFP to CAP since the limit is more manageable against homogeneous interference than the CFP.

A. Reinforcement Learning Figure

Reinforcement learning [18] is a type of machine learning where an agent figures out how to communicate on states

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:14, No:11, 2020

keeping in mind to get maximum rewards. The reward on a state is the feedback of that state where agent takes actions on different states. The agent figures out how to act and upgrade its behavior from the response of its activities [19], which implies, that for the first time, if an agent makes action and gets a poor reward for it. The agent decides not to pick a similar activity in next state from where it learns a poor reward. For example, a child burned by fire will not go again near fire because he learned and also got a reward from previous action. Fig. 1 illustrates the flow of reinforcement learning. Reinforcement learning consists of agents and environment. Agent takes an action on a state from environment. The agent takes reward from environment; a negative reward will be obtained by an agent if action is not successfully performed whereas a positive reward will give to agent in case of not successful action.

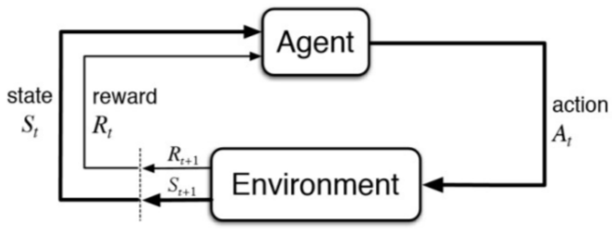

Fig. 1 Reinforcement learning

B. Q Learning

Q-learning [20] is a generally utilized free model RL technique where the q value of each action A in a given state S is kept up in a table. Another side of the statement, it takes the action-value function which gives the expected utility of making action in a given state and from that point following the best policy. The joining of Q learning to the best action values has been demonstrated under specific presumptions [21]. The values for each state and activity are updated each time by an update function. Thus, the learning is online and there is no requirement for an obvious transition model.

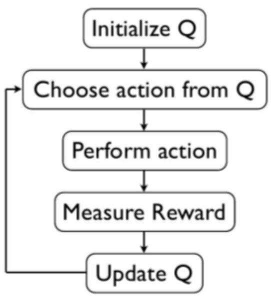

Fig. 2 Q-learning

III. Proposed Method

In QIM-MAC, we use a fixed star topology for each WBAN consisting of a fixed number of nodes and a coordinator. All nodes are static with no mobility. We adopt one-hop communication scheme and consider a beacon-enabled slotted CSMA/CA for first Super Fame, and subsequent transmission based on Q-learning. Fig. 3 shows the network model, we have used for QIM-MAC. It shows coordinator and sensor nodes. We have used several coexisting WBANs, each with star topology.

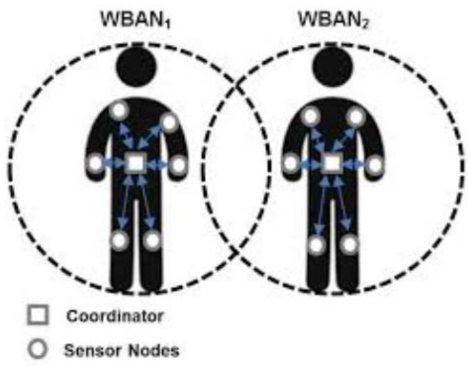

Fig. 3 Network model for WBANs

A. Reinforcement Learning Based Interference Mitigation

QIM-MAC is based on reinforcement learning and is a model-free reinforcement learning technique based on receiving scalar rewards from the environment. It assigns Q-values to each sensor on an action representing the approximate goodness of the action. Each node selects and performs one operation during the learning process to change the Q value [22]-[24]. Over time the node learns the real action values and can select the most appropriate. Our proposed scheme called QIM-MAC consists of five elements as discussed below:

1) In QIM-MAC, each sensor node corresponds to an agent.
2) All time slots of CAP of a Super Frame represent states, i.e. either success or failure of transmission in any particular slot due to interference.
3) In QIM-MAC, a node decides which slot it should select to sense to minimize interference as a possible action, i.e. a slot with highest success ratio to sense and thus can be used for transmission.
4) Reward indicates, whether carrier sense in the selected time slot according to Q value benefits as success or results in failure.
5) Q value is the rewarded value of each node. Initially Q value set to zero for each and every node of WBAN. Q function contains k number of slots in each Super Frame, and i number of nodes in a WBAN. Node i has its own Q values for each k slot. Q values are represented by Qt(i, k), where i represents number of node that take an action on slot k. Q value is updated after the learning rate for each node according to (1):

$$Qt+1(i, k) = Qt(i, k) + \_(r - Qt(i, k)) \quad (1)$$

Equation (1) updates the value of Q value for each node i for a particular slot k while Îs is the learning rate mostly used greater than zero and less than one (0 < _ < 1), and r is a reward. *Qt* (*i, k*) is the previous Q value of node i on slot k while *Qt*+1(*i, k*) is the Q value for next super-frame.

B. Learning Rate

Learning rate is the time after which node updates its Q value. Learning rate will be fixed in the range of 0 to 1. We take learning rate as 0.01 for QIM-MAC which results in a smaller number of packets drops in case of collisions, if node i receive collisions on slot j so i will change j after less time (learning rate), high learning rate will result high number of packets drops in case of collisions, if node i receive collisions on slot j and will use j till completion of learning rate.

C. Q-learning Interference Mitigation MAC (QIM-MAC)

In our proposed scheme, first super-frame transmissions are based on CSMA/CA/mechanism of IEEE 802.15.4. Each node learns collisions in first super-frame, which are due to interference. Each node updates Q value by (1) 1 after learning rate. Node learns/experiences slots on which nodes transmit. If sensor (Ix) receives acknowledgement packet (ACKx) from coordinator for data packet (Px) on time slot (Kx), it updates its Q value and uses that time slot for next super-frame with the highest Q value. QIM-MAC algorithm first finds the CAP length from beacon broadcasted by coordinator. If a node has data to send, it will use CSMA/CA for transmission. Initially each node has zero Q value. Node i sends data on time slot k and receives acknowledgment for that packet, it updates its Q value with plus 1, and if do not receives acknowledgment for that packet, it updates its Q value with minus 1. The operation of QIM-MAC is illustrated in Algorithm 1.

Algorithm. 1 QIM-MAC at each sensor node:

```
Input: No of nodes n, reward r [+1, -1], learning rate, Q value.
Output: Q-value of each sensor q.
Repeat for each sensor 1 to n
Q value = 0 initialization
while i do
Br = Received beacon Extract l CAP, i.e. CAP information from Br
if first super-frame then
    if node i has data P to transmit then
        sense using CSMA/CA on slot k
      if ACK received for P then
          update (1) 1 with a reward +1 on k to
    minimize interference
            end
      if timeout/no ACK for packet P then
        update Equation 1 with a
        reward -1 for k
  end
end
end
else
      sense the time slot j with highest Q value to
  minimize interference transmit data P in slot j
      if ACK received for P then
    update Equation 1 with a reward +1
    on k to minimize interference
  end
  if timeout/no ACK for packet P then
    update (1) with a reward -1 for k
  end
end
end of while
```

In the Algorithm 1, first super-frame is based on CSMA/CA. Each node i has Q value which is initially set to zero. Coordinator broadcasts beacon which has information about CAP length. Each node extracts CAP length l cap from beacon and starts transmission using CSMA/CA. When node i receives acknowledgment packet ACK for data packet P, it updates its Q value with +1 using (1) and if not receive ACK for P then updates its Q value with -1 using (1).

All nodes i will use that slot j in next super-frame, at which i has highest Q value to minimize interference and use the same process for updating Q value.

D. Numerical Example of QIM-MAC Operation

In this section, we explain the operation of QIM-MAC with the help of an example. Let's suppose we have 5 number of nodes and 7 slots for CAP part of super-frame, and use (1) to update Q value. Node 1 accesses slot number 4, transmits successfully, and gets a reward of +1, calculates Q value Q(1,4) = 0.01. Node 2 got access to slot 1 and 2. Transmission fails at slot 1 by which got a reward of -1 and calculated Q value Q(2,1) = -0.01, but transmits successfully two times on slot 2, and calculating Q value Q(2,2) = 0.009901. Node 3 transmits data three times on slot 5 and receives acknowledgment for each data packet and has Q value Q(3,5) = 0.029701. Node 4 uses slot 6 and 3 to transmit but did not receive acknowledgment so Q-value for node 3 became Q(4,6) = -0.01. Node 4 successfully transmits two times by which got Q value of Q(4,3) = 0.009901. Node 5 transmit data on slot 7 and get a reward of +1, Q value of node 5 became 0.01. Node 5 will use that slot of CAP for next super-frame for which node has high Q value.

TABLE I
Q-LEARNING NUMERICAL EXAMPLE

No of Node	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7
1				0.01			
2	-0.01	0.009901					
3					0.029701		
4			0.009901			-0.01	
5							0.01

Table I shows that node 1 will use slot 4, node 2 will use slot 2 because Q value at slot 2 is higher than slot 1. Node 3 will use slot 5, whereas node 4 will use slot 3 because node 4 has greater Q value on slot 3 than slot 6. Node 5 will use slot 7. Q value associated with each time slot is illustrated in Fig. 4.

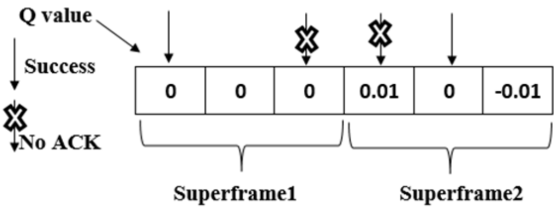

Fig. 4 Q-value for time slot on two super-frames

Fig. 4 shows three slots of Super Frame1 at which three nodes have 0 Q value. Nodes learn collisions on slot in Super Frame1. As we see in Fig. 4 that a node successfully transmits data at first slot of Super Frame1 due to which it is awarded by +1 and its updated Q value is shown in Super Frame2. Slot 3 of Super Frame1 shows transmission failure and its Q value is updated by -1.

E. QIM-MAC Performance Evaluation

We evaluate the performance of QIM-MAC and compare with HCA scheme [18] under two different scenarios including varying number of nodes, and varying number of WBANs. For scenario1 we take 18 nodes, and for scenario2, we take 30 nodes.

F. Scenario 1

In this scenario, we simulate QIM-MAC scheme and HCA scheme [18] using star topology of two coexisting WBANs with 18 number of sensor nodes. Simulation area is of 1300 x 550 m in size. We analyze average throughput, delay, energy consumption, and packet loss ratio varying number of nodes increases.

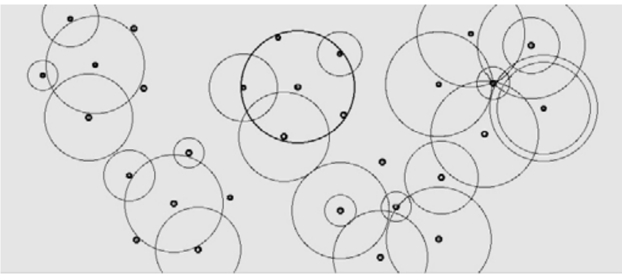

Fig. 5 NAM of scenario 1

G. Parameters

We simulate two coexisting WBANs with star topology consisting of 18 sensor nodes in the area of 1300 x 550 meters with the data rate of 250 kbps for ten minutes. Each sensor has a transmission range of 20 meters with an interference range of 550 meters.

TABLE II
SIMULATION PARAMETERS

Parameter	Count
Data rate	250 kbps
Simulation time	10 minutes
Super-frame Duration	122 micro sec
Packet Size	40 bytes + Header
ACK packet	28 bytes + Header
Energy power	10 J
Traffic	CBR
Interference range	20 m
Simulation area for scenario	1300 x 550 m
No of nodes for scenario 1	18

1) Fig. 6 shows the average network throughput of scenario 1. In QIM-MAC, collision reduced due to intelligent carrier sensing based on most optimal time slot, which results in high network throughput as compared to HCA scheme [18]. HCA scheme [18] does not take into account suitable time slots for transmission which may experience higher number of collisions due to coexisting WBANs. As Fig. 6 shows, QIM-MAC increases as increasing number of nodes, HCA scheme [18] also shows an improvement to node 6, but as number of nodes exceeds 6, HCA scheme [18] shows no improvement however QIM-MAC throughput increases. We observe that as QIM-MAC is better than HCA scheme [18] which network tends to be dense.

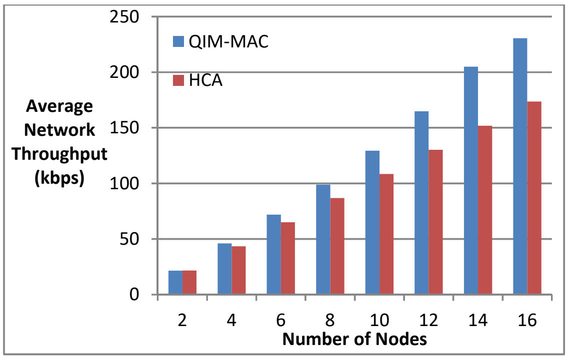

Fig. 6 Average Network Throughput vs Number of Nodes

2) Fig. 7 shows energy consumption with varying number of nodes. As nodes increase, the energy consumption of HCA scheme [18] increases. HCA scheme [18] and QIM-MAC consume equal energy at starting, when number of nodes are 2, however as number of nodes increases to 12, HCA scheme [18] consumes 2.8 joule due to the high number of retransmission of data packets due to next slot selection for carrier sensing, whereas QIM-MAC decreases 1.8 joule due to a smaller number of retransmissions.

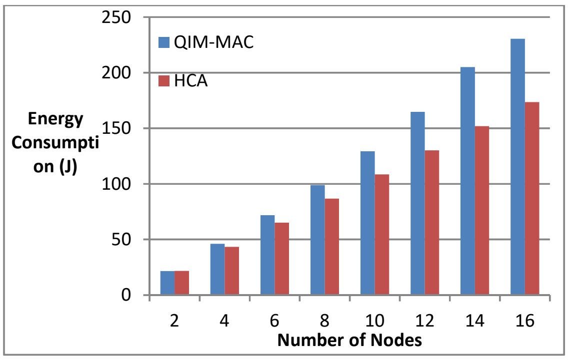

Fig. 7 Energy Consumption vs Number of Nodes

3) In scenario 1, average delay of nodes with varying number of nodes is as shown in Fig. 8. The average delay was calculated in milliseconds. HCA scheme [18] experiences higher delay than QIM-MAC from node 2. When the number of nodes exceeds 2, both schemes gradually decrease, but when the number of nodes goes to 6, HCA scheme [18] remains constant till the end of simulation, however QIM-MAC keeps decreasing as the

number of nodes increases till the simulation ends due to a smaller number of collisions which results in a smaller number of retransmissions.

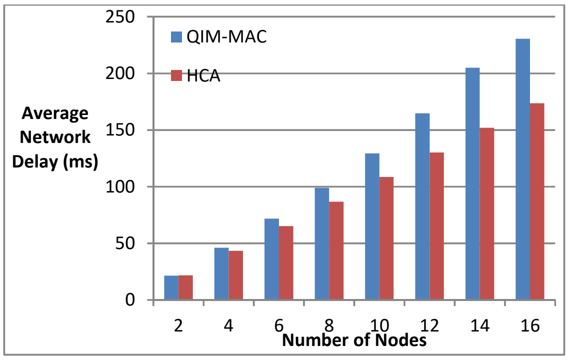

Fig. 8 Average Network Delay vs Number of Nodes

4) Packet loss ratio with number of nodes is evaluated in this scenario for 8 minutes. In Fig. 9 the number of nodes is simulated which gives different packet loss ratio. For node 2, HCA scheme [18] experiences 6.5% packet loss however QIM-MAC experiences 6% packet loss. When number of nodes exceeds 2, packet loss ratio drops for both schemes. As the number of nodes increases, HCA scheme [18] increases packet loss ratio however QIM-MAC decreases packet loss ratio due to learning carrier sensing. Packet loss ratio for QIM-MAC decreases to 2% whereas HCA scheme [18] decreases to 3.8% packet loss ratio as compared to QIM-MAC.

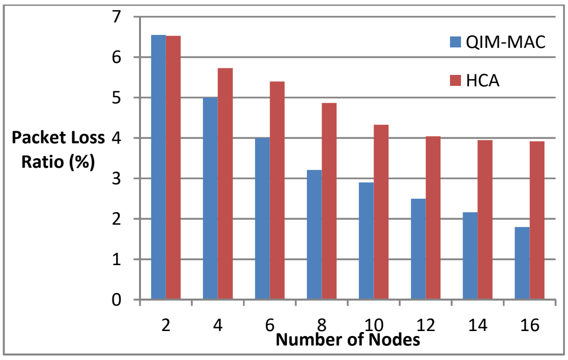

Fig. 9 Packet Loss Ratio vs Number of Nodes

H. Scenario 2

We simulate QIM-MAC and HCA scheme [18] under five coexisting WBANs using star topology. Each WBAN consists of 5 sensor nodes and 1 coordinator. We simulate 30 nodes placed in the area of 2500 x 550 meters. Simulation time is 8 minutes, and evaluates average network throughput, average network delay, energy consumption, and packet loss ratio.

1) Fig. 11 shows the performance of average network throughput of scenario 2. In this scenario 5 WBANs are simulated which results in high throughput using QIM-MAC against HCA scheme [18]. We calculate separate throughput for two to five coexisting WBANs. We observe that there is less difference between both schemes with two WBANs however as the number of WBANs increases QIM-MAC achieves greater throughput than HCA scheme [18]. QIM-MAC uses learning based CSMA/CA as instead of blind CSMA/CA which results in high network throughput as compared to HCA [18].

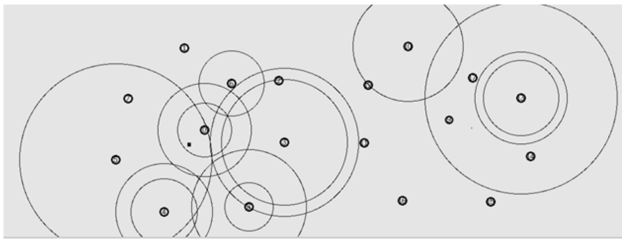

Fig. 10 NAM of scenario 2

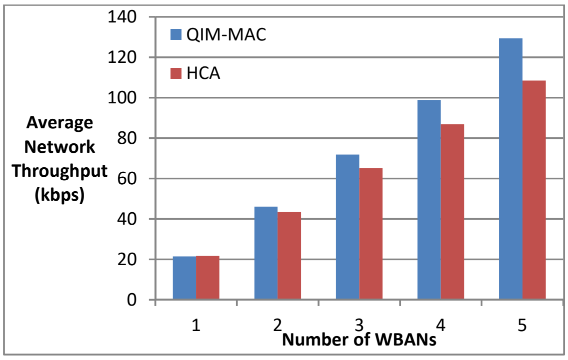

Fig. 11 Average Network Throughput vs Number of WBANs

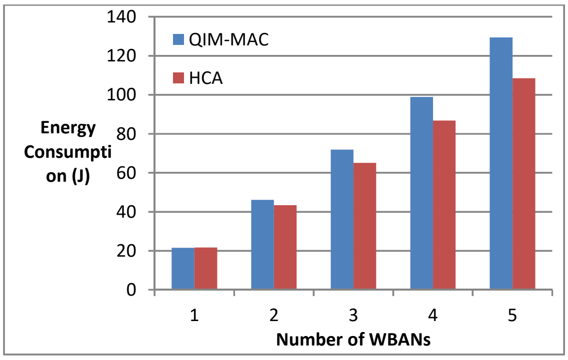

Fig. 12 Energy Consumption vs Number of WBANs

2) In this scenario, we calculate energy consumption for each WBAN. Topology change the number of WBANs from two to five and other parameters are the same as shown in Table II. Due to a smaller number of retransmissions in QIM-MAC scheme which is mainly attributed to Q-based carrier sensing, it consumes less energy. HCA scheme [18] experiences a high number of retransmissions due to blind carrier sensing and transmission in the next time slot, which results in higher energy consumption. Fig. 12 shows energy consumption of each WBAN. Energy consumption in QIM-MAC

decreases with an increase in the number of WBANs increased, however HCA [18] energy consumption increases as the number of WBANs increased.

3) Fig. 13 illustrates average delay experienced by each WBAN. When simulation setup has two coexisting WBANs, results of both schemes remain the same. When number of WBANs is increased to three, four, and five, HCA [18] shows high increase in average delay; however delay of QIM-MAC shows very less increase.

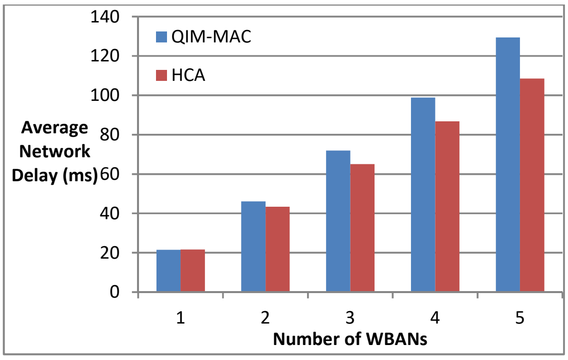

Fig. 13 Average Network Delay vs Number of WBANs

4) Fig. 14 shows that QIM-MAC experiences less packet loss ratio than HCA scheme [18]. This difference is negligible for two WBANs. However, when the number of WBANs is increased to three, four, and five, the difference becomes more prominent than HCA [18]. QIM-MAC experiences less increase in packet loss ratio.

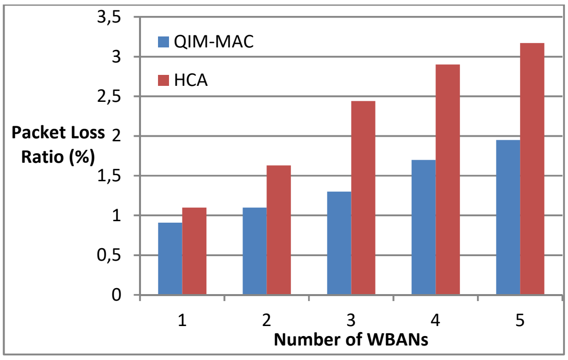

Fig. 14 Packet Loss Ratio vs Number of WBANs

IV. Conclusion and Future Work

Homogeneous interference is very common and periodic in WBANs, which may affect the performance of health monitoring and other applications. In our work, we focus on interference management among coexisting WBANs. Existing literature proposes different heuristics to manage the interference; however, these schemes do not perform an intelligent carried sensing based on the most optimal time slot learned through the past experiences.

Unlike existing schemes, we propose a reinforcement learning approach to manage the coexistence interference between WBANs. Q-learning based CSMA/CA senses the time slots based on learning and therefore selects the slots with minimum probability. We have performed the simulations in a WBAN based on star topology and have compared with most recent work, HCA. We have performed simulations in three different scenarios. In the first scenario, two WBANs are simulated with varying times. Second scenario also has two WBANs but the results are taken with varying number of nodes, an increasing number of WBANs is simulated in third scenario. Simulation results show that our scheme outperforms HCA [18] scheme in all three scenarios. Simulation results show that QIM-MAC has significant improvement in network throughput while decreasing energy consumption, average delay and packet loss ratio.

Future work is to investigate performance of QIM-MAC in peer to peer and mesh topologies. The second step is to apply Q-learning on IEEE 802.15.6 by taking into account both the collaborative and non-collaborative approaches. It will be interesting to investigate the performance of QIM-MAC under heterogeneous interference as well.

References

[1] A. Astrin, "IEEE Standard for Local and Metropolitan Area Networks, Wireless Body Area Networks," IEEE, 29 Feb 2012.
[2] B. B. Benoit Latrie, Ingrid Moerman, Chris Blondia, and Piet Demeester, "A survey on wireless body area networks," Wireless Networks, pp. pp 1-18, 2010.
[3] J.-H. C. Beomseok Kim, and Jinsung Cho, Member, "A Hybrid Channel Access Scheme for Coexistence Mitigation in IEEE 802.15.4-Based WBAN," IEEE, vol. Vol. 17, NO. 21, 1 November 2017.
[4] R. M. Cavallari, F. Rosini, R. Buratti, C. Verdone, "A Survey on Wireless Body Area Networks: Technologies and Design Challenges," IEEE Communication Survey, 2014.
[5] G. T. Chen, Chen, W.T, Shen, "2L-MAC: A MAC Protocol with Two-Layer Interference Mitigation in Wireless Body Area Networks for Medical Applications in Proceeding of the IEEE International Conference on Communications (ICC), Sydney, Australia," June 2014.
[6] M. N. J. Deylami, "E. A Distributed Scheme to Manage the Dynamic Coexistence of IEEE 802.15.4-Based Health-Monitoring WBANs," IEEE, 2013.
[7] M. M. Fumihiro Inoue, Takayuki Nishio, Koji Yamamoto, Fusao Nuno, Takatoshi Sugiyama, "Novel coexistence scheme between wireless sensor network and wireless LAN for HEMS. Smart Grid Communication," IEEE, pp. pp 271-276, 2013.
[8] A. M. Jamthe, A. Agrawal,, "Scheduling Schemes for Interference Suppression in Healthcare Sensor Networks," In Proceeding of the IEEE International Conference on Communications (ICC), Sydney, Australia, June 2014.
[9] L. P. Kaelbling, Littman, M. L, "Reinforcement learning: A survey. Journal of Artificial Intelligence Research," pp. pp 237 - 285,, 1996.
[10] B. B. Latrfe, B. Moerman, I. Blondia, C. De-meester, "A survey on wireless body area networks. Wireless, Network," IEEE, 17 Jan 2011.
[11] J. M. Mahapatro, S. Manjunatha, M. Islam, "Interference-Aware Channel Switching for Use in WBAN with Human-Sensor Interface," n Proceeding of the 4th International Conference on Intelligent Human Computer Interaction (IHCI), Kharagpur, India, December 2012.
[12] H. M. Mohamad Jaafar Ali, Ahmed Mehaoua, "Interference Avoidance Algorithm (IAA) for Multi-hop Wireless Body Area Network Communication" arXiv:1602.08658v2 (cs.NI), 2 November 2016.
[13] S. A. Movassaghi, M. Lipman, J. Smith, D. Jamalipour, "A Wireless Body Area Networks: A Survey," IEEE, 2014.
[14] S. A. Movassaghi, M. Smith, "Smart Spectrum Allocation for Interference Mitigation in Wireless Body Area Networks in Proceeding of the IEEE International Conference on Communications (ICC), Sydney, Australia," IEEE, June 2014.
[15] C. M. Otto, A. Sanders, C. Jovanov, "System Architecture of a Wireless Body Area Sensor Network for Ubiquitous Health Monitoring," pp. pp

307 - 326, 2006.
[16] M. W. A. Patel, "Challenges, and Prospective in Emerging Body Area Networking Technologies," IEEE, pp. pp 80-88, 2010.
[17] T. J. Ratko Magjarevic, Peter Kramar, and Anze Zupanic, Home care technologies for ambient assisted living, in 11th Mediterranean Conference on Medical and Biomedical Engineering and Computing. 2007.
[18] A. M. Samaneh Movassaghii, Abbas Jamalipour, David Smith, Mehran Abolhasan, "Enabling Interference-Aware and Energy-Efficient Coexistence of Multiple Wireless Body Area Networks with Unknown Dynamics. Digital Object Identifier 10.1109/ACCESS," 2016.
[19] I. Standard, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR - WPANs)," IEEE, 2006.
[20] R. S. Sutton, and Barto, "Reinforcement learning: An introduction. MIT Press," 1998.
[21] C. J. Watkins, Dayan, "P. Q-learning. Machine Learning," pp. pp 279 - 292, 1992.
[22] C. J. C. H. Watkins, "Learning from delayed rewards. PhD thesis, University of Cambridge, England," 1989.
[23] A.E. Forster, A.L. Murphy “FORMS: Feedbacks Routing for Optimization Multiple Sinks in WSN with Reinforcement Learning” in 3rd International Conference on Intelligent Sensors, Sensor Networks and Information. Melbourne, Qld., Australia. IEEE, dec 3-6. 2007
[24] A. Förster and A. Murphy, "Froms: A failure tolerant and mobility enabled multicast routing paradigm with reinforcement learning for WSNs", Ad Hoc Networks, vol. 9, no. 5, pp. 940-965, 2011. Available: 10.1016/j.adhoc.2010.11.006.