Distributed Slot Allocation in Capillary Gateways for Internet of Things Networks

Fatima Hussain and Alexander Ferworn
NCART Lab, Dept. of Computer Science
Ryerson University, Toronto, Canada

***Abstract*—The applications and usage of the internet is expanding on a daily basis and the Internet of Things (IoT) is fast becoming the new approach for incorporating the internet into our personal, professional and social lives. IoT enables a wide variety of devices to inter-operate through the existing internet infrastructure. Capillary networks are proposed as a fundamental part of loT development, and will enable local sensor and devices to connect efficiently with other ubiquitous communication networks such as cellular systems. In this paper, we apply the Q-learning algorithm for the scheduling of capillary gateways for (M2M) communication in IoT networks. Q-learning algorithm is used to select conflict-free slot assignment for these gateways in a self-organizing manner. We analyze the performance of the proposed algorithm with respect to learning rates and rewards.**

Index Terms–Internet of Things, M2M Communication, Capillary Networks, Q-learning

I. INTRODUCTION

Machine to Machine (M2M) is the main building block for IoT networks and it is envisioned that massive numbers of these devices will be all around us. The major concerns in M2M networks are high traffic load and simultaneous massive access requests [1] by machine-type communication devices (MTCDs). Support for a massive number of MTCDs brings technical challenges affecting the performance of backhaul networks in terms of network congestion and efficient allocation of radio resources. Data aggregation and clustering [2], [3] is a technique to achieve efficient data communication and is recommended for M2M networks. It can alleviate network congestion and increase the energy efficiency of MTCDs [4], [5]. Also, with the advent of loT, we start to experience the proliferation of new constrained devices around us, which encompass a diverse group of new wireless technologies (e.g., Bluetooth, IEEE 802.11ah, IEEE 802.15.4) [6]. It is an important concern that deployment of these devices and enabling technologies, does not jeopardize the performance of the backhaul networks such as cellular, LTE, and 3GBPP systems. A major redesign of future networks is required and European Telecommunication Standards Institute (ETSIs) M2M technical committee has proposed a hybrid architecture, in which cellular enabled gateways act as traffic aggregation and protocol translation points for capillary (i.e. wireless sensor) M2M networks [7].

Capillary networks are composed of a potentially high number of devices (e.g. sensors) equipped with short-range radio interfaces. As a result, a vast range of constrained devices equipped with only short-range radios can utilize cellular network capabilities to gain global connectivity.

Q-learning algorithm is a well known reinforcement learning (RL) technique, which enables an agent (e.g., a sensor node) in M2M networks to learn by interacting with its environment. The agent will build up its experience through a number of trials and adopt the best action associated with a maximum reward [8]. A decentralized Q-learning technique is proposed in [9] to manage the interference generated by multiple wireless area networks. In [10], a RL-based base station selection algorithm is proposed that allows the MTCDs to choose the base station in a self-organizing fashion. Q-learning RACH access scheme (QL-RACH) is proposed in [11] by using ALOHA to control M2M traffic.

In this paper, we consider multiple capillary IoT networks using various short range communication techniques and, that are connected to cellular network through gateways. Capillary IoT network installations do not involve any network planning and these networks automatically configure themselves for easy deployment. Therefore, there are multiple capillary gateways within reach, offering communication services with different properties to different networks. These gateways may differ in properties like energy, latency and connectivity.

Gateways are responsible for data transmission/reception from/to eNB to/from different networks. These gateways aggregate data from various MTCDs from associated networks to eNB through an uplink channel; and receive aggregated data from eNB and relay the data to associated networks on downlink channels. We study how these gateways make distributive decisions about slot selection for transmitting data to devices/networks on downlinks. These decisions are influenced by the slot selection made by neighbouring gateways. If simultaneous data transmission is done by all the gateways on the same frequency band and time slot, a collision is caused and system performance is degraded. Here gateways are acting like agents having incomplete (or no) information about neighbours capabilities for decisions concerning the simultaneous transmission, this can be mapped onto a multi-agent system [12]. There is no central controller providing information to all the entities and for controlling interference. In this scenario, we propose a self-organizing slot allocation based on the paradigm of independent learning, where agents are unaware of the other agents' actions.

We utilize Q-learning [10], [11], [13] for scheduling of gateways, where Q-learning is run by gateways in a distributed manner avoiding collisions and inter-gateway interference. Also, this scheme allows multiple gateways to share the same channel being used by multiple devices, by transmitting one after the other, using different time slots. The major contribution of our work is the performance enhancement of capillary gateways through selecting the conflict-free time slots in congested scenarios using Q-learning.

This paper is organised as follows: Section II describe the system model and problem formulation. Simulation results are presented in section III and section IV concludes this paper.

II. System Model

We consider multiple capillary networks using short range communication technology and, require cellular network access for long-range communication. These networks are connected to eNB through gateways. Gateways connects MTCDs to the cellular backhaul network, transporting and receiving data to and from eNB. As a single eNB is not able to serve all the sensors and devices in each smaller networks, gateways help to reduce congestion [14], [15], and enables seamless integration of wireless sensor networks with cellular networks. These gateways are responsible for data aggregation and data transmission to/from associated devices/networks, and also are in communication with the eNB.

In this paper, we consider communication from gateways to associated devices/networks. Each gateway will send data (received from eNB) to its associated network in the downlink communication. We do not discuss how these networks make selections [16] of the best gateway for its communication, rather how resources such as slots are assigned to gateways. Let there be K gateways connecting N networks, having u devices in each, to eNB. Each gateway transmits data using time division multiple access (TDMA), over a frame having T slots, to its associated devices/network. TDMA allows several devices to share the same frequency band by dividing the signal into different time slots. In our case it allows multiple gateways to share the same channel used by multiple devices, by transmitting one after the other, using different time slots. Communication among devices within the network take place with short range communication technologies, and for simplicity of analysis, we assume that all networks use the same technology. We simulate two scenarios; firstly in which gateways transmit data on a randomly selected channel to associated devices/networks without considering link capacity and characteristics. While in the second scenario, gateways selects the best channel for the associated network and then transmit.

III. Problem Formulation and Solution

In the following we will discuss the slot selection by gateways employing the Q-learning technique.

A. Problem Formulation

Figure 1 illustrates a communication model between different capillary networks and capillary gateways using the same cellular network as backhaul. We consider distributed architecture in which gateways are not aware of their peers and there is no centralized entity responsible for scheduling of these gateways. The centralized architecture is easier to manage and is feasible for both the capillary and the cellular network however, load on the management server increases with increased network traffic. Also, scalability becomes an issue when unplanned capillary networks are added in the same area. Distributed architecture facilitates scalability and also ease of individual maintenance. A unique time slot is needed for data transmission by each gateway to an associated network. As gateways and associated networks are not scheduled for data transmission/reception, a careful distributed selection of channels is required to avoid collisions. Our task is to enable the gateways to select unique time slot for data transmission to its MTCD avoiding any inter-network collision.

Let MTCD i in j^{th} network be denoted by U_i^j. We first define a vector, $U = U_1^1, U_2^1, U_3^1, U_1^2, U_2^2, U_3^2, U_1^3, U_2^3, U_3^3$ that shows the i^{th} device in j^{th} network. We further define a matrix I, which has columns as the slot numbers and rows as the replicas of U at certain,

$$\boldsymbol{I} = \begin{bmatrix} U_1^1 & U_1^1 & U_1^1 \\ U_2^1 & U_2^1 & U_2^1 \\ . & . & . \\ . & . & . \\ U_3^3 & U_3^3 & U_3^3 \end{bmatrix} \tag{1}$$

The entry of this matrix is 1 if that device is assigned to a particular slot and 0 otherwise, i.e.,

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ . & . & . \\ . & . & . \\ 0 & 1 & 0 \end{bmatrix} \tag{2}$$

In this case, our problem can de defined as finding a matrix I of size $(u \times K) \times (T \times u)$ by observing the following constraints:

$$|I^i| = 1, \forall i = 1, 2...T \tag{3}$$

$$|I_j| \leq 1, \forall j = 1...u \times K \tag{4}$$

$$\sum_{k=1}^{u} U_l^k = 1, \forall l = 1, ..K \tag{5}$$

where I^i is the i^{th} column and I_j is the j^{th} row of matrix I. We define above constraints as:

- First constraint ensures that each slot is occupied.
- Second constraint ensures that each capillary network assigns at least one device to each slot
- Third constraint ensures that each network does not take more than one slot.

In order to find matrix I distributively with the above constraints, we resort to the reinforcement learning technique as discussed next.

B. Q-Learning Algorithm for Capillary Gateways

Q-learning is one of the most commonly used reinforcement learning algorithms. We use Q-learning algorithm for avoiding collisions among gateways by scheduling their packet transmission. Each gateway transmits data to its associated MTCD/network on a specific time slot. Each gateway tries to randomly transmit to associated devices, and if some other neighbouring gateway is simultaneously utilizing the same frequency and time slot for transmission, it will result in collisions. Every gateway learns from previous experience and schedules its MTCDs in such a way that no two MTCDs of neighbouring networks utilize the same slots in the same frequency. To achieve the learning experience, stateless Q-learning [17] is used in our scheme. Each gateway has individual Q values for each slot, and it is utilized in representing the preference of slot selections. Q values are denoted by $Q(j,T)$ and they represent that gateway j takes an action on slot T. The previous Q values and the current reward all contribute to the Q value update. The Q value is updated as below after the reward is returned:

$$Q_{t+1}(j,T) = Q_t(j,T) + \alpha(R - Q_t(j,T)) \tag{6}$$

- Agent: Gateways $j, \forall 1 \leq j \leq K$. Gateways are the agents running the Q-learning algorithm to determine the best slot selection.
- Action:$A(t) = a^{j,T}(t), T\epsilon[T_1, ...T_k]$. $a^{j,T}(t)$ is defined as the action of the j^{th} gateway at time t and is to choose time slot T_j out of T, where T_k is the maximum number of time slots in a frame. Gateways will transmit data on the specific slots according to some information obtained in the previous slots.
- Reward:R is defined as:

$$R = \begin{cases} 0 & T_j = \phi, \\ +1 & T_j \text{ is unique}, \\ -1 & T_j = T_{\hat{j}}. \end{cases} \tag{7}$$

where T_j is the slot chosen by the j^{th} gateway and $T_{\hat{j}}$ is the same slot chosen by any other gateway. Reward is 0 if any other slot except T_j is chosen.

Each gateway uses Q-learning algorithm to transmit data randomly in time slots. Gateways use the ACK/NACK protocol to know about the success/failure of the data transmission in a specific slot. If the transmission is successful then ACK is received, when no two gateways select the same slot. Thus the gateway will assign a positive reward value to that slot. If two (or more) gateways send data in the same time slot, a collision occurs and NACK is received by both (or more). This results in transmission failure and a negative reward is assigned to that slot. In this way, gateways will learn by experience which slot is suitable for data transmission by comparing their Q values. Slots with higher Q values will always be preferred, i.e.,

$$I^j = \max_{T,j} Q_t(j,T) \tag{8}$$

All the Q values are initialised to 0 at the very beginning. If multiple slots have the same Q value, gateways will randomly select one of them.

Another method of channel assignment is based on the best channel for each MTCD and gateway, as employed in channel based allocation (CBA) with Q-learning. The proposed Q-learning algorithm will run on each gateway to find the appropriate slot allocation.

IV. Results and Discussions

Fig. 1. Capillary Networks

In this section, performance of the proposed algorithm is verified via simulation. We compare the convergence time with respect to varying learning rate (α) and reward (R) given in (6).

The simulation setup is as follows: We consider four capillary networks ($K = 4$), having three devices ($u = 3$) in each, as shown in Figure 1. For simplicity of analysis, we assume that each network has multiple frequencies for internal network data exchange. The frame has four time slots ($T = 4$), one slot for each gateway [1]. Each gateway will transmit data coming from eNB to its associated devices/networks in one of the selected time slots using the proposed Q-learning algorithm. In the first scenario, gateways will randomly select a slot and transmit data to its associated devices. It will result in a collision if two or more gateways choose to send in the same slot. After collision, controllers will retransmit again by randomly selecting slot. For the channel-based allocation, controllers are allowed to pickup the channel with the best channel gain, without considering the selection made by neighbouring gateways. This may result in collision if the same slot is selected by others for the same channel. This transmission is considered unsuccessful and retransmission takes place following the same procedure.

Figure 2 shows the slot selection by all four gateways using Q-learning. Initially when more than one gateway selected the same slot, with a number of iterations, collision is observed

[1]Note: By increasing number of devices, only frame size has to be adjusted and algorithm will successfully converge. No other adjustment is required.

and later with experience, gateways learn to send in the best slots. In the figure shown, convergence occurs after 20 iterations, at which time all the gateways learn to send their data in unique slots. That is , the gateways distributively self-organize their transmission slots to avoid inter-gateway interference.

Fig. 2. Convergence for Unique Slot Selection

A. *Effect of Learning Rate (α)*

We assign different values of α and variable positive and negative reward values for R. In Figure 3, convergence time is shown with respect to learning rate for various rewards. We define it as the time taken by the gateways to learn with repeated data transmission and resulting Q values, for the selection of unique time slot for their data transmission. Convergence is declared when there is no change in slot assignment. We calculate convergence time obtained over 1000 iterations and average the results. Different values of α varying from 0.1 to 0.9 in a step-size of 0.1 are used to study the rise and fall of convergence time. Convergence time (in terms of slots) varies from 4 to 10. We define convergence rate as, how fast or slow gateways learn to transmit in unique slots without collision. It is noted that as α is increasing, the convergence rate is increasing and it takes less time to converge. This is an expected behavior as with an increase in α, more weight is given to the reward than to the current Q value. This can be seen by rewriting (6) as follows:

$$Q_{t+1}(j,T,w) = Q_t(j,T,w)(1-\alpha) + \alpha R \qquad (9)$$

We can see that by increasing α, the ratio of weight assigned to R compared to $Q_t(j,T,w)$ is given as $\frac{\alpha}{1-\alpha}$. Increasing value of α gives more weight to R, as a result quicker convergence is achieved due to more emphasis being given to current values and actions. Hence, as the ratio increases, the convergence time decreases. Similar behaviour prevails for three sets of reward values at ±1, ±1.75 and ±2.

Fig. 3. Convergence Time at Various Rewards (raw data and curve-fitted values)

Fig. 4. Convergence Probability at Various Rewards (raw data and curve-fitted values)

Figure 4 shows convergence probability versus α for various rewards. Convergence probability is defined as ratio of number of times simulation converge and the total number of simulations, i.e. 1000. α is varied from 0.1 to 0.9 in a step size of 0.1. A decreasing trend in convergence probability can be observed between 0.1 and 0.8, except for a small increase between 0.8-0.9. From Figures 3 and 4, we can observe that a higher

α decreases both convergence time and probability. Thus a lower α is preferable for increased convergence probability at the expense of increased convergence time. Also, similar behaviour is observed with different sets of reward values as long as there is equal weight given to positive reward and negative penalty. Note that cubic curve fitting is employed to smoothen the plot.

B. Cumulative Success Probability

Cumulative success probability is defined to be the probability of observing less than or equal to a given number of successes for slot allocation. It also refers to accumulated value of convergence or non-convergence of proposed algorithm. In Figure 5, cumulative success probability for different values of α is plotted with increasing time. We can see that for increasing values of α, the cumulative success probability decreases that in turns means that success rate decreases. This decrease further compliments the results in the Figure 4, in which convergence probability decreases with the increase of learning rate α. This convergence probability only accounts for success or convergence of the proposed algorithm. For Q-learning, the probability increases almost linearly until 20 slots and becomes constant after that. This constant value indicates the convergence point for the Q-learning algorithm. Once converged, i.e., once each gateway learns to successfully transmit in a specific time slot without any collision, it will continue to transmit into the same slot. As with Q-learning, each gateway learns with experience which slot is better in terms of data transmission and reduced collisions, resulting in better success rates.

Fig. 5. Cumulative Success Probability at Various Alpha

C. Channel Based Selection of Gateways

In Figure 6, Q-learning is applied in the channel-based access (CBA) and random channel access (RCA) for comparison.

Fig. 6. Q-Learning with Channel Selection

Q-learning is applied with $\alpha = 0.1$ and $R = \pm 1$. Average SNR is plotted with the increasing number of MTCDs per network. For simplicity of analysis, equal transmission power of 1W and a path-loss attenuation factor of 3 is assumed for our simulation. We can observe that when Q-learning is applied in channel-based access, it performs better than random channel access as each gateway has to pick up the best channel for transmission to its associated devices/networks, which results in increasing average SNR. The SNR difference between CBA and RA-based allocation increases due to channel diversity as expected with the increase in number of MTCDs per network. However, we also see increase in average SNR for RA-based allocation with increasing number of devices per network, this is due to user diversity. Therefore, average SNR increases with the number of devices in the network, with the fixed number of MTCDs being selected for transmission. If it had been consideration for average SIR than there will be possibly a less increase due interference among devices being transmitted in the same slot.

V. CONCLUSION

A capillary network provides local connectivity to devices using short-range radio access technologies while it connects to the backhaul cellular network through a node called a capillary gateway. We have used Q-learning for the resource allocation of these gatways and analyzed the relationship between learning rate and assigned reward. We also showed that when the learning rate is increasing, convergence rate is also increasing, i.e., it takes less time to converge. In addition, it is also observed that convergence probability decreases with the increase in learning rate. In future work, we aim to use Q-learning for versatile capillary networks using different short-range communication technology such as Bluetooth, Zigbee etc. This combination of various networks will change the required data rate characteristics and hence the slot allocation decisions. Also, in our present work we assumed that all the networks have same number of MTCDs in them and we intend to work with variable size networks in the future.

REFERENCES

[1] F. Hussain, A. Anpalagan, and R. Vannithamby, "Medium Access Control Techniques in M2M Communication: Survey and Critical Review," *Wiley Transactions on Emerging Telecommunications Technologies*, September 2014.

[2] H. Shariatmadari, P. Osti, S. Iraji, and R. Jantti, "Data Aggregation in Capillary Networks for Machine-to-Machine Communications," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 2277– 2282, September 2015.

[3] S. Lien and K. Chen, "Toward Ubiquitous Massive Accesses in 3GPP Machine-to-Machine Communications," *IEEE Communications Magazine*, vol. 49, pp. 66 – 74, April 2011.

[4] G. Kumar and H. Mehra, "An Hybrid Clustering Algorithm for Optimal Clusters in Wireless Sensor Networks," *IEEE Student Conference on Electrical, Electronics and Computer Science*, pp. 1–6, March 2014.

[5] H. Harb and A. Makhoul, "K-Means Based Clustering Approach for Data Aggregation in Periodic Sensor Networks," *IEEE Conference on on Wireless and Mobile Computing, Networking and Communications*, pp. 435– 441, December 2014.

[6] O. Novo, N. Beijar, and M. Ocak, "Capillary Networks - Bridging the Cellular and loT Worlds ," *IEEE World Forum on Internet of Things (WF-IoT)*, vol. 1, pp. 571–578, December 2015.

[7] J. Matamoros and C. Anton-Haro, "Data Aggregation Schemes for Machine-to-Machine Gateways: Interplay with MAC Protocols," *Future Network and Mobile Summit (FutureNetw)*, pp. 1–8, July 2012.

[8] M. A. Alsheikh and S. Lin, "Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications," *IEEE Communication Surveys and Tutorials*, vol. 16, pp. 1996– 2018, November 2014.

[9] G. Serrano and L. Giupponi, "Distributed Q-learning for aggregated interference control in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 1823–1834, May 2010.

[10] M. Hasan and E. Hossain, "Random Access for Machine-to-Machine Communication in LTE-Advanced Networks: Issues and Approaches," *ieeecm*, pp. 90– 97, June 2013.

[11] B. L. Mohammed and M. Paul, "Application of Q-Learning for RACH Access to Support M2M Traffic over a Cellular Network," *IEEE Proceedings of European Wireless Conference*, pp. 1–6, May 2014.

[12] A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for Interference Control in OFDMA-based Femtocell Networks," *IEEE VTC*, May 2010.

[13] P. Mitchell and D. Grace, "ALOHA and Q-Learning Based Medium Access Control for Wireless Sensor Networks," *International Symposium on Wireless Communication Systems*, pp. 511–515, August 2012.

[14] R. M. Huq, K. P. Moreno, and H. Zhu, "On the Benefits of Clustered Capillary Networks for Congestion Control in Machine Type Communications over LTE ," *International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–7, August 2015.

[15] S. Singh and K.-L. Huang, "A Robust M2M Gateway for Effective Integration of Capillary and 3GPP Networks ," *IEEE International Conference on Advanced Telecommunication Systems and Networks (ANTS)*, pp. 1– 3, December 2011.

[16] J. J. Nicklas Beijar, Oscar Novo and J. Melen, "Gateway Selection in Capillary Networks," *IEEE International Conference on Internet of Things (IOT)*, pp. 90– 97, October 2015.

[17] R.Sutton and A.Barto, *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press, 1998.