User-Centric Learning for Multiple Access Selections

Sumayyah Dzulkifly, Wahidah Hashim, Ahmad Fadzil Ismail, Mischa Dohler

Abstract: We are in the age where business growth is based on how user-centric your services or goods is. Current research on wireless system is more focused on ensuring that user could achieve optimal throughput with minimal delay, disregarding what user actually wants from the services. Looking from connectivity point of view, especially in urban areas these days, there are multiple mobile and wireless access that user could choose to get connected to. As people are looking toward machine automation, we understand that the same could be done for allowing users to choose services based on their own requirement. This paper looks into unconventional, non-disruptive approach to provide mobile services based on user requirements. The first stage of this study is to look for user association from three new perspectives. The second stage involved utilizing a reinforcement learning algorithm known as q-learning, to learn from feedbacks to identify optimal decision in reaching user-centric requirement goal. The outcome from the proposed deployment has shown significant improvement in user association with learning aware solution.

Keywords:q-learning, user-centric, heterogeneous wireless networks, user association, access selection

I. INTRODUCTION

Managing user association in heterogenous wireless network (HWN) is complicated due to distinctnode's transmit powerandaccess mechanism operating adjacent to each other. Wireless standards look at enabling separate association for uplink and downlink to delegate the access node burden[1], [2]. The idea is to enable downlink association with the access node operating at higher transmission power while uplink association with the low powered nodes. While this is a good solution, delegating users while maintaining control at the nodes will put more strain on network management[1].

User association from its perspective requires intelligent autonomous system, capable to selectan optimal access. Optimal access from user perspective could vary. Hence, there is a need to understand the features that could influence the association behavior from user perspective. Among influencing factors that can be considered are:

Revised Manuscript Received on September22, 2019.

SumayyahDzulkifly, Institute of Informatics and Computing Energy, Universiti TenagaNasional, Malaysia

Wahidah Hashim, Institute of Informatics and Computing Energy, Universiti TenagaNasional, Malaysia

Ahmad Fadzil Ismail, IIUM Strategic Technologies and Engineering Research Unit, Research Management Center, International Islamic University Malaysia

Mischa Dohler, Department of Informatics, 30 Aldwych, King's College London, WC2B 4BG London, United Kingdom

- Network congestion – in wireless communication, resources can be limited due to the need to cater for a large volume of services. Hence, it is crucial to identify whether the BS has enough resources before initiating connectivity.
- User energy consumption - unlike access nodes, user's battery life has limited capacity for every single charge. The battery life can be prolonged by monitoring user's energy consumptions.
- Subscribed mobile data quota - user within HWN coverage is exposed to both free (i.e. WiFi) and paid access (i.e. mobile network). Access selection should be performed based on this factor

Machine learning algorithm is needed to enable user to have an intelligent mechanismin finding an optimal solution. There are extensive studies on machine learning application in HWN, depending on specified targets [3].Among popular machine learning approach is a subset of reinforcement learning (RL), known as q-learning algorithm. Q-learning (QL) enables agent (i.e. the entities who sense and execute final decision) to learn about the environment from trial and error to reach a conclusion. It is fundamentally from Markov Decision Process and is model free[4]. This means that the transition of the states is not dependent upon the transition probability.

This paper looksintoinitiating user centric association using q-learning algorithm to analyze optimal decision based on discussed possible factors affecting user association. The contributions of this paper are as follow:

1. Proposes general framework to observe user distribution in HWN, comprising of access nodes with varied access mechanism (i.e. mobile networks and WiFi)
2. Apply QL algorithm to enable user-centric access selection for user-centric attributes proposed, namely user energy consumption

The rest of the paper is organized as follow. Section 2 reviews the literature behind this study. Section 3 presents designed HWN framework utilized in this study. Section 4 discussed general idea behind RL and its evolvement to QL. Section 5 looks into varied QL applications for the three different user-centric attributes discussed. Section 6 discussed and analyzed the outputs. Finally, section 7 concludes the study.

II. RELATED WORKS

RL is widely deployed in heterogeneous cellular network (HCN) to manage different aspect in HCN such as resource allocation, interference mitigation, mobile handover, energy, network selection, traffic management and others[3].

In [5], fuzzy logic and q-learning is applied to identify user association based on the network backhaul throughput. The justification behind the proposed method is that user might experience service interruption if backhaul network condition is ignored. Utilizing fuzzy logic to create association rule based on backhaul throughput is convenient.However, backhaul throughput can be extremely dynamic, which makes it unsuitable to rely solely on simple fuzzy logic decision.

Meanwhile [6] utilizes q-learning to assist in channel selection in improvised cognitive radio scenario. QL is also employed in [4] to assist resource management for separate association in HCN. While managing channel selection and radio resources could mitigate user congestion issues, the tasks increase the network processing burden. With the introduction of HCN, user association solely based on highest signal strength is no longer relevant. Various improvement in user association rule is proposed. In [7] user association rule based on content cache assignment is employed to effectively reduce delay from user perspective. In [8], user dynamics (i.e. movement and mobility) within HCN clusters are utilized as part of the association rule. Study by [9] proposes an alternative to conventional user association, namely linking device-to-device (D2D) content cachingwith base station (BS) to enable user with similar content requests for better data rates.

Authors in [10] propose combining conventional user association technique based on the highest signal strength with time bases user scheduling to balance BS loads in HCN. In [11], user association based on energy optimization is proposed for heavily congested network. The focus of the study is to improve energy consumption of the BS while catering a huge volume of users. Though, the authors failed to notice that energy consumption could also severely vary based on the type access (i.e. WiFi or mobile) user is connected to at the time.

Meanwhile, [12] proposes thefusion of user association based on service's quality with resource allocation in licensed-assisted access-based long term evolution (LAA-LTE) system. LAA-LTE system take advantage on WiFi ability to operate in an unlicensed band to improve overall user throughput. Managing user association and resource allocation in such system can be complex due to different mechanism involved.Despite the proposed approach, LAA-LTE system is still not widely deployed in practice due to the complexity of the system and high installation costs.

III. USER ASSOCIATION FRAMEWORK

Designing HWN architecture in a simulation form can be simple or complex, depending on the specific requirement. In this study, the focus is more towards user association mechanism. Hence, the association attributes within the designed framework are identified as follows:

There are 3 different BS with varied transmit power and access mechanism within a specified area.

Large number of users are distributed uniformly within the specified area.

Linkmodellingwith close replication to HWN propagation scenario.

These are elaborated further in the subsequent subsection. Other HWN attributes are ignored to simplify the study and concentrates on user association scenario.

Simulated Architecture

Heterogeneous wireless network (HWN) comprises of licensed and unlicensed BS transmitting at varied power. The highest transmission power could be up to 40 dBm while the lowest could be approximately 20 dBm. To reflect HWN scenario as precise as possible, this study proposes 3 different BS with varied transmission power and operating bands in a specified simulation area. The demography of the simulation is specified within an area of 600 meters by 500 meters.

Users are prompted to continually change their position at every iteration to emulate user's dynamic movement. The simulation scenario is as illustrated in figure 1. In the figure, the blue diamond represents MBS while the red squares represent SBS, green triangles signify WiFi access point (AP) and the black dots represent UE. Note that large number of UEs are distributed randomly in the area.

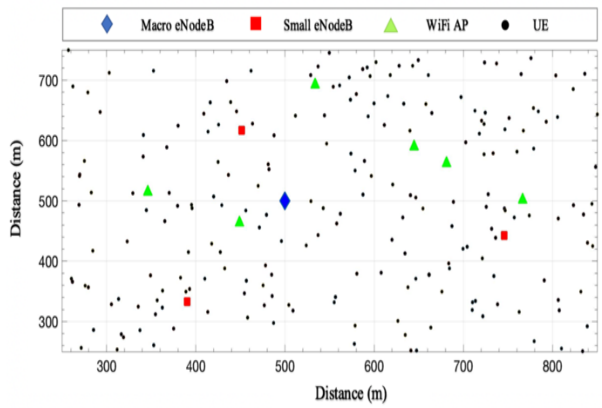

Fig. 1Simulated scenario illustrations

Link Modelling

In physical layer of wireless communication, the received power is measured during transmission while it experiences radio propagation and interference. Note that this study employs 3 BS with different type of transmission power and operating bands. Hence, modelling the link involves 3 distinct path loss formulation based on specific conditions. The path loss and shadowing factor is determined based on the operating carrier frequency. This is in reference to the propagation model as proposed in [13]

As the study took into account all users within the generated area, the distance between users and BS as well as the line-of-sight is taken into considerations. The shadowing factors experienced by users are multiplied with computational random value. This is to replicate signal fluctuations during the transmissions. Fundamental path loss formulation in reference to table A1-2 in [13]is as follow:

$$h_{s,cf} = c_{cf,d} * \log_{10} d + \chi_{cf,d} \qquad (1)$$

Where s is the link status, whether user is within BS line-of-sight (LOS) or not. Meanwhile C and X represent a constant value determined by selected operating carrier frequency, cf and distance, d. The line-of-sight status is determined from table A1-3 in [13]. The signal-to-interference-plus-noise-ratio (SINR) is calculated using the following equation:

$$\gamma_i = \frac{P_i h_i}{\sum_{i\neq j}^{n} P_j h_j + N} \tag{2}$$

Given the SINR, the channel capacity of each user, ξ_i is calculated using (3). From (3), the achievable rate is formulated as an average over the number of instances, m_i, as shown in (4).

$$\xi_i = \log_2(1 + \gamma_i) \tag{3}$$

$$\partial_i = \frac{\sum_{m_i} \xi_i}{m_i} \tag{4}$$

Network Congestion

Among thecommon issue in wireless network is network congestion, which occur due to limitation of resources despite the number of users requesting the services. Hence, it is crucial to identify whether the BS selected has enough resources to cater user's need. Hence, a sum of user resource requests, $\mathfrak{Y}_i$, cannot exceed an identified BS resource limit, $\mathfrak{W}_{max}$, as shown in (5).

$$\sum_{i=1}^{K} \mathfrak{Y}_i \leq \mathfrak{W}_{max} \tag{5}$$

The resources are analysed from the type of services selected by user. In this study, 3 different user services are identified, namely web browsing, video and voice-over-IP (VoIP). The specifications for the services are as presented in Table 1.

Table. 1 Bandwidth requirement for each service type

Type of service(s)	Bandwidth requirement(s) in kbps
Video streaming	1000 - 4000
VoIP	128
Web browsing	1

Video bandwidth requirement is specified in reference typical video streaming these days from the lowest to a higher definition. The selection of the bandwidth will be based on the channel condition experienced by user at the time. Meanwhile, for VoIP, the value is set to the minimum bandwidth requirement for video calling, using common VoIP codec, known as SiLK codec. For web browsing, there is no specific minimum bandwidth requirement, so long as users could get access to the network. Hence, the value is set to 1 kbps to ensure that user could still access the network successfully.

$$\max_i \mathcal{U}_i\,(\mathring{A}_{ij} + \partial_{ij}) \tag{6}$$

$$s.t. \qquad \mathring{A}_{ij} = \begin{cases} 1; & if\ \bar{\bar{\mathbb{R}}}_i \leq \xi_{ij} \\ 0; & if\ \bar{\bar{\mathbb{R}}}_i > \xi_{ij} \end{cases}$$

$$\Big(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\Big) \leq \breve{\forall}_{max}$$

To reflect BS congestion status, an optimisation formula is derived as shown by (6). The aim of the optimization is to maximize the association and also throughput. Hence, the optimization is specified in terms of utility function.

User Energy Consumption

As this study looks into specific user behaviour, namely mobile phone users, the energy consumption is formulated based on specific attributes, namely the type of BS selected as well as the type of service utilise at the time. Hence, energy consumption formulation as shown in (6), in reference to [14].

$$\ddot{e}_i = t_i * (r_m + \partial_{ij} * r_n) \tag{7}$$

Where $t_i \in \{t_{i,min}, t_{i,max}\}$ is the time window when users is active using selected time of services, r_m is consumption power of user mobile device, ∂_{ij} is the computed achievable rate and r_n is the data energy consumption rate.Interpreting energy consumption in an optimisation formula, we have:

$$\min_{i \in I, j \in J} \ddot{e}_{ij} \tag{8}$$

$$s.t. \qquad \mathring{A}_{ij} = \begin{cases} 1; & if\ \bar{\bar{\mathbb{R}}}_i \leq \xi_{ij} \\ 0; & if\ \bar{\bar{\mathbb{R}}}_i > \xi_{ij} \end{cases}$$

$$\Big(\sum_{i=1,j=1} \mathring{A}_{ij} * \ddot{e}_{ij}\Big) \leq \ddot{e}_{th}$$

Where the aim of the optimisation is to minimise the energy consumption while selecting the best BS to connect to and to ensure that the energy consumption does not reach critical level.

User Data Quota

Mobile phone users are usually subscribed to a certain data plan. The subscribed data plan enables users to get connected to paid mobile networks. Depending on user's socio-economic status, the paid data quota can be less than 4 MB and as high as unlimited data quota (usually with bandwidth restrictions if there is heavy internet usage) monthly.Also, with user's dynamic movement, there will be instances where users are within the proximity of free WiFi access.

This especially true for users who has almost similar daily pattern such as office workers who work fixed hour from 9 a.m. to 5 p.m. Hence, transforming this into an optimisation formula:

$$\max_i \partial_{ij} \tag{9}$$

$$s.t. \qquad \mathfrak{q}_i = \begin{cases} 1; & if\ unlimited\ plan \\ 0; & if\ limited\ plan \end{cases}$$

$$\left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\right) \leq \breve{\forall}_{max}$$

Where the aim is to maximize the achievable throughput, at the same time ensuring that BS does not run out of resource at the time the connection is requested.

IV. REINFORCEMENT LEARNING

The fundamental of QL is RL. RL basis is Markov Decision Process (MDP), where the agent (in this case, user) interacts with the environment through a defined set of state and action. RL is fundamentally a machine learning approach where the agent is prompted to learn from trial and error. Optimal approach is defined through an accumulated reward function. QL is widely favoured due to is character, which is model-free. In other word, agent is forced to explore and exploit the state and action space to find the best decision.

QL originated from the combination of dynamic programming and temporal difference technique in RL. In RL, state represent the possible conditions agent could experience in the environment. While action is a set of possible choices that agent can choose. From selected pair of state and action, a reward is calculated. The rewards are accumulated over time once each pair of state and action is tested, yield:

$$R = \sum_{t=0}^{\infty} \gamma^t \mathcal{R}_{t+1} \quad (10)$$

Given that the policy is defined, and the rewardaccumulated over time, the state value function is evaluated as shown in (11):

$$\vartheta_\pi(\mathfrak{x}) = \mathbb{E}_\pi(\sum_{t=0}^{\infty} \gamma^t \mathcal{R}_{t+1} \mid \mathfrak{x}_t = \mathfrak{X}) \quad (11)$$

The valuation in (10) can be improved to understand the state and action relationship while accumulating the rewards. This is presented by (12):

$$q_\pi(\mathfrak{x}) = \mathbb{E}_\pi(\sum_{t=0}^{\infty} \gamma^t \mathcal{R}_{t+1} \mid \mathfrak{x}_t = \mathfrak{X},\ a_t = \mathcal{A}) \quad (12)$$

To find the optimal decision, (12) is utilised

$$q^*(\mathfrak{x}, a) = \max_\pi q_\pi(\mathfrak{x}, a) \quad (13)$$

Putting (12) into (13), we have (14) in a simplified version

$$q_\pi(\mathfrak{x}, a) = \mathcal{R}_{t+1} + \gamma^t \vartheta^*(\mathfrak{x}_{t+1}) \quad (14)$$

Finally, the equation is further expanded for model free QL as shown in (14):

$$q_\pi(\mathfrak{x}, a) = \mathcal{R}_{t+1} + \gamma^t \vartheta^*(\mathfrak{x}_{t+1}) \quad (15)$$

V. DECENTRALIZED USER-CENTRIC LEARNING

This section elaborates on the methodology proposed in this study to enable user-centric decision based on the attributes as discussed in the section 1 (introduction). Details are discussed in the subsequent sub sections,

Learning based on Network Congestion Status

In this section, user association decision is based on available resource of each BS and user's required throughput. BS is assumed congested once its resource is fully utilised. As this study deals with BS that has different access mechanism, q-learning algorithm is deployed to enable user-centric decision. To synchronise the congestion status during decision making, it is assumed that LTE resources are allocated using the PRB and WiFi resources are distributed based on the average number of users allowed for association per period.

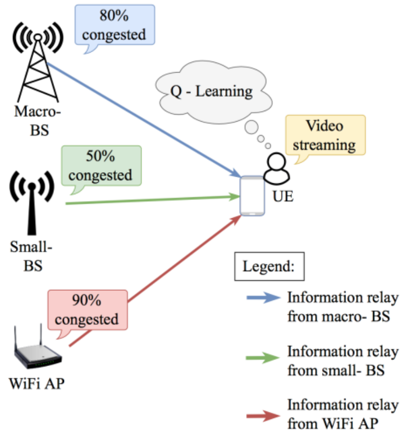

Fig. 2Illustrations on BS congested scenario

In practice, network congestion occurs due to packet queue and packet drop caused by signal loss. However, for the ease of this study, it is assumed that no packet drop once association is established. The application of the proposed approach is as portrayed in Fig. 2. In the figure, user need to access video streaming while each AN relay congestion status towards UE to enable UE to decide which AN is better to connect based on its own required throughput at the time.

The association rule is designed to fulfil two main requirements, namely user's required throughput and congestion status. (6) isutilised and translated into q-learning algorithm based on (15). Details in the parameters used are as presented in table 2.

Table. 2 Parameters description for learning based on network congestion status

Parameter(s)	Value
state, $\mathfrak{x}_t$	$\mathfrak{x}_t = \{\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3, \mathfrak{x}_4\}$
action, a_t	$a_t = \{ \eth_m, \eth_s, \eth_w\}$
Explorationstrategy, ϵ	$\epsilon = 0.5$

Where the state is expressed based on the possible congestion status. This is further elaborated in (16). The action consists of the number of BS with available resources as the possible choices. Meanwhile, the exploration strategy is based on the percentage of exploration deemed good enough to arrive to the optimal solution.

$$\mathfrak{X} = \begin{cases} \mathfrak{x}_1; & if\ \bar{\bar{\mathbb{R}}}_i \le \xi_{ij}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\right) \le \breve{\forall}_{max} \\ \mathfrak{x}_2; & if\ \bar{\bar{\mathbb{R}}}_i \le \xi_{ij}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\right) > \breve{\forall}_{max} \\ \mathfrak{x}_3; & if\ \bar{\bar{\mathbb{R}}}_i > \xi_{ij}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\right) \le \breve{\forall}_{max} \\ \mathfrak{x}_4; & if\ \bar{\bar{\mathbb{R}}}_i > \xi_{ij}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij}\right) > \breve{\forall}_{max} \end{cases} \quad (16)$$

The reward is formulated such that when the corrected pair of state and action is selected, the agent receive a reward value. Otherwise, the agent received receive no value, as shown in (17).

$$rw_t = \begin{cases} K * e^{-(\partial_{ij})}; & f\ \bar{\bar{\mathbb{R}}}_i \le \xi_{ij}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ 0; & otherwise \end{cases} \quad (17)$$

Learning based on User Energy Consumption

Similar to section 5.1, the learning is formulated such that user choose the BS with that prompt user to consume less energy. Based on optimisation equation in (8), the learning algorithm is designed with the set of state, action, exploration strategy and reward similar to parameters displayed in table 2. The only difference is that the exploration value is set to 0.1. This is due to poor performance is too much exploration is carried out.

$$\mathfrak{X} = \begin{cases} \mathfrak{x}_1; & if\ \ddot{e}_{ij} \le .\ddot{e}_{th}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ \mathfrak{x}_2; & if\ \ddot{e}_{ij} \le .\ddot{e}_{th}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) > \breve{\forall}_{max} \\ \mathfrak{x}_3; & if\ \ddot{e}_{ij} > .\ddot{e}_{th}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ \mathfrak{x}_4; & if\ \ddot{e}_{ij} > .\ddot{e}_{th}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) > \breve{\forall}_{max} \end{cases} \quad (18)$$

The set of state for this section is as defined in (18). Noted that similar to (17), (18) has similar state sets. The state set is defined in reflection to the optimisation model in (8). In this section, the reward function is described as follow:

$$rw_t = \begin{cases} e^{-(\tilde{\gimel})}; & if\ \ddot{e}_{ij} \le .\ddot{e}_{th}\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ 0; & otherwise \end{cases} \quad (19)$$

Learning based on User Data Quota

For this learning, it is assumed that user has an option to select WiFi access to preserve the mobile data quota. However, once mobile network/BS is selected, the learning process is executed. The learning parameters are defined similar to table 2.

The distinctions are the exploration strategy is set to 0.9. Due to the nature of the selection, the exploration strategy is set to a higher value to enable the agent to identify the best state-action pair. The state set is as shown in (20), in reference to (9).

$$\mathfrak{X} = \begin{cases} \mathfrak{x}_1; & if\ \widetilde{\mathbb{q}}_i = 1\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ \mathfrak{x}_2; & if\ \widetilde{\mathbb{q}}_i = 1\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) > \breve{\forall}_{max} \\ \mathfrak{x}_3; & if\ \widetilde{\mathbb{q}}_i = 0\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ \mathfrak{x}_4; & if\ \widetilde{\mathbb{q}}_i = 0\ and \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) > \breve{\forall}_{max} \end{cases} \quad (20)$$

Similar to the previous sections, the reward function is defined based on whether the agent select the best set of state and action. This is as shown in (21).

$$rw_t = \begin{cases} e^{-(\tilde{\gimel})}; & if\ \widetilde{\mathbb{q}}_i = 1\ or \left(\sum_{i=1,j=1} \mathring{A}_{ij} * \breve{\forall}_{ij} \right) \le \breve{\forall}_{max} \\ 0; & otherwise \end{cases} \quad (21)$$

VI. RESULT AND ANALYSIS

This section presents the outputs generated from proposed method discussed in section 5.1 to 5.3. Analysis from the outcomes are summarised in this section.

Outcomes: Learning based on Network Congestion Status

The main aim of the proposed approaches is to optimise user association with higher achievable throughput. Hence, the graph comparing average achievable rate and the number or user is presented in Fig. 3. Note that the proposed approach, which utilises association rule based on user's required throughput and congestion status (indicated by the orange bar) has shown significant improvement as compared to the conventional approach (indicated by purple bar).

Table 3 presents the differences between the proposed approach and the conventional maximum received signal strength indicator (RSSI) for user association. From the values presented in the table, it is evident that there is an increase of approximately 15 percent in achievable rate when the proposed approach is applied, despite a large number of UE.

Table.3 Summary of the difference between the proposed approach and max. RSSI

	50 users	**100 users**	**150 users**
Difference in achievable rate (Mbps)	0.2	0.12	0.2
Difference in percentage (%)	20	10	15

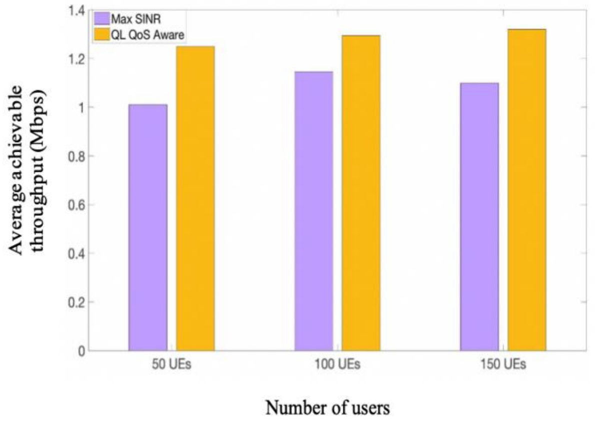

Fig. 3Average achievable rate per number of users

Outcomes: Learning based on User Energy Consumption

The aim of the proposed method in this section is to minimize energy consumption while selecting the best association.This study compares the proposed approach based on energy consumption against the common approach based on maximum RSSI. Fig. 4 portrays the comparison between the association based on RSSI and the proposed approach. As shown in the figure, there is significant reduction in energy consumption in 15th and 50th percentage cumulative distribution function (CDF).

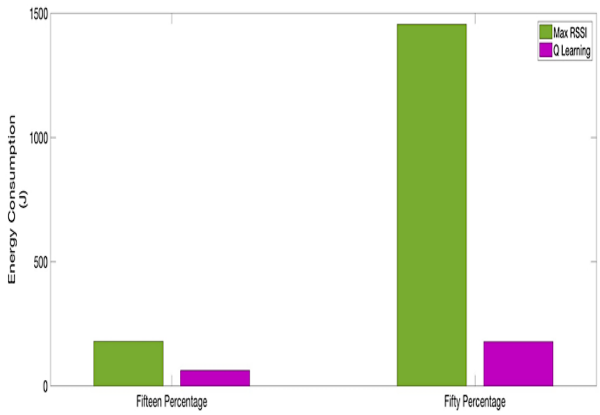

Fig. 415th and 50th percentile CDF (energy consumption)

Table 4 summarizes the amount of conserved energy when the proposed approach is applied. From the table, 65 and 88 percent improvement is shown when the proposed approach is utilised. Note that high energy value is possibly due to the simplified system model and energy formula. Hence, the study on energy consumption can be extended to utilise a more complex energy model in HWN to thoroughly observe the amount energy conserved.

Table. 4 Summary on the conserved energy consumption for 15th and 50th percentile

	15^{th} percentile	**50^{th} percentile**
Conserved energy consumption (J)	116	1277
Difference in percentage (%)	65	88

Outcomes: Learning based on User Data Quota

The key analysis of this study is to observe the achievable rate when the selection is made based on the proposed association rule. Optimal achievable rate reflects that user is able to select the network that conserve its mobile data quota. Fig. 5 portrays the achievable rate between the proposed approach and maximum RSSI.

Note that due to underlying rule in updating the congestion status, it is easier for each individual UE to reach convergence, as shown in Fig. 6. In the figure, UE is able to reach convergence lower iteration value. This is possibly due to basic association rule and system model applied.

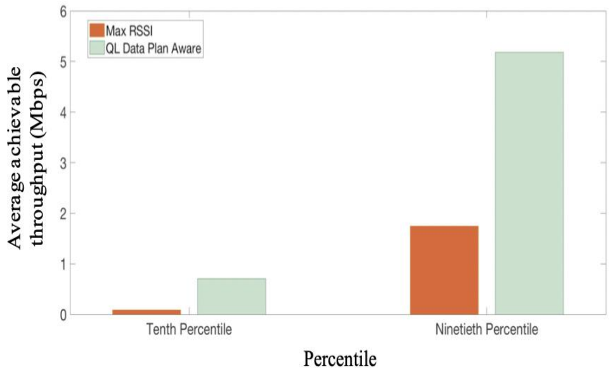

Fig. 510th and 90th percentile cdf (energy consumption)

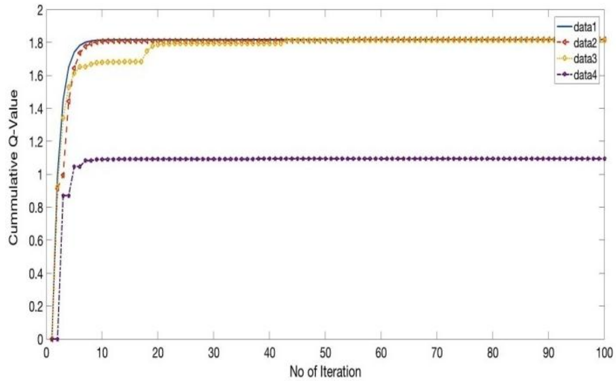

Fig. 6Convergence graph for q-learning implementation based on UE subscribed data plan

VII. CONCLUSIONS

This study proposes user-centric learning mechanism to enable user-centric learning and association. Looking at association from the perspective of user, a number of variables are considered, namely congestion status, energy consumption and mobile data quota. Enabling user-centric association requires machine automation and intelligence, which is proposed in this study. From the analysis presented, significant improvements can be seen when the proposed solutions are deployed, as seen in section 6. For example, 10 to 20 percent improvement is seen when learning association based on network congestion is applied. Similar enhancement is seen for energy consumption bases association. To conclude, the proposed approaches have shown massive improvement to user-centric association with machine learning assistance.

REFERENCES

1. F. Boccardi et al., "Why to Decouple the Uplink and Downlink in Cellular Networks and How to Do It," IEEE Commun. Mag., vol. 54, no. 3, pp. 110–117, 2016.
2. R. Li, X. Liu, K. Luo, T. Jiang, and S. Jin, "Decoupled Access in Het-Nets with Backhaul Constrained Small Base Stations," IEEE Access, vol. 6, pp. 27028–27038, 2018.
3. Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues," pp. 1–34, 2018.
4. Y. Hu, R. Mackenzie, and M. Hao, "Expected Q-learning for Self-Organizing Resource Allocation in LTE-U with Downlink-Uplink De-coupling," Eur. Wirel. 2017; 23th Eur. Wirel. Conf. Proc., pp. 185–190, 2017.
5. F. Pervez, M. Jaber, J. Qadir, S. Younis, and M. A. Imran, "Fuzzy Q-Learning-Based User-Centric Backhaul-Aware User Cell Association Scheme," IEEE Access, vol. 6, pp. 39595–39605, 2018.
6. A. Carie, M. Li, C. Liu, P. Reddy, and W. Jamal, "Hybrid Directional CR-MAC based on Q-Learning with Directional Power Control," Futur. Gener. Comput. Syst., vol. 81, pp. 340–347, 2018.
7. W. Jing, X. Wen, Z. Lu, and H. Zhang, "User-Centric Delay-Aware Joint Caching and User Association Optimization in Cache-Enabled Wireless Networks," IEEE Access, vol. 7, pp. 74961–74972, 2019.
8. Y. Sun, G. Feng, S. Qin, and S. Sun, "Cell Association With User Behavior Awareness in Heterogeneous Cellular Networks," IEEE Trans. Veh. Technol., vol. 67, no. 5, pp. 4589–4601, 2018.
9. M. I. Ashraf, M. Bennis, W. Saad, M. Katz, and C. S. Hong, "Dynamic Clustering and User Association in Wireless Small-Cell Networks with Social Considerations," IEEE Trans. Veh. Technol., vol. 66, no. 7, pp. 6553–6568, 2017.
10. X. Ge, X. Li, H. Jin, J. Cheng, and V. C. M. Leung, "Joint User Association and User Scheduling for Load Balancing in Heterogeneous Networks," IEEE Trans. Wirel. Commun., vol. 17, no. 5, pp. 3211–3225, 2018.
11. C. Ma, F. Liu, Z. Zeng, and S. Zhao, "An Energy-Efficient User Association Scheme Based on Robust Optimization in Ultra-Dense Networks," 2018 IEEE/CIC Int. Conf. Commun. China, ICCC Work. 2018, pp. 222–226, 2019.
12. J. Tan, S. Xiao, S. Han, Y. C. Liang, and V. C. M. Leung, "QoS-Aware user association and resource allocation in LAA-LTE/WiFi coexistence systems," IEEE Trans. Wirel. Commun., vol. 18, no. 4, pp. 2415–2430, 2019.
13. M. Series, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," 2009.
14. R. Andreucetti, S. Chen, Z. Yuan, and G. M. Muntean, "Smartphone energy consumption of multimedia services in heterogeneous wireless networks," IWCMC 2014 - 10th Int. Wirel. Commun. Mob. Comput. Conf., pp. 1147–1151, 2014.