IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 11 March 2023, accepted 31 March 2023, date of publication 28 April 2023, date of current version 4 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3271423

RESEARCH ARTICLE

Deep Reinforcement Learning Based Algorithm for Symbiotic Radio IoT Throughput Optimization in 6G Network

GERGES MANSOUR SALAMA[1], **SAMAR SHAKER METWLY**[2], **E. G. SHEHATA**[1], **AND AHMED M. ABD EL-HALEEM**[2], **(Member, IEEE)**

[1]Electronic and Communication Engineering Department, Minia University, Minya 61511, Egypt
[2]Electronic and Communication Engineering Department, Helwan University, Helwan, Cairo 11792, Egypt

Corresponding author: Samar Shaker Metwly (samarshaker2904@heng.helwan.edu.eg)

ABSTRACT Internet of Things (IoT) -based 6G is expected to revolutionize our world. Various candidate technologies have been proposed to meet IoT system requirements based on 6G, symbiotic radio (SR) is one of these technologies. This paper aims to use symbiotic radio technology to support the passive Internet of things and enhance uplink transmission performance. The IoT tag information is sent to the cloud for analysis through a macro base station (MBS) or a wireless access point (WAP), where the smartphones are used as a relay to transmit this information to the MBS or WAP. In this paper, an optimization problem was formulated into two phases to maximize the total throughput of the system. The first phase is, the problem of achieving the optimum mode selection of the LTE or Wi-Fi Network, aiming to maximize the system throughput. A matching game algorithm is used to solve this problem. Second phase, the problem of achieving optimum clustering of tags, where the tags are divided into virtual clusters, and finding which smartphones' LTE/Wi-Fi downlink signals all cluster members can ride to maximize the system throughput. A double deep Q-network (DDQN) model was proposed to solve this problem. Simulation results show that our proposed algorithms increase the total system data rate by an average of 90% above the system using the LTE network first without DDQL algorithm. Furthermore, it enhances the capacity of the system on average by 100% above LTE network first system without the DDQL algorithm.

INDEX TERMS IoT, DDQL, LTE, Wi-Fi, matching game, backscattering, symbiotic radio.

I. INTRODUCTION

The Internet of Things (IoT) is considered as one of the major applications of future wireless networks that enables devices to communicate with each other [1]. IoT devices have grown significantly, so substantial amounts of spectrum resources and energy are required [2]. It is a challenge for 5G wireless systems to satisfy the IoT requirements. 5G has introduced some technologies for IoT, but cannot fully satisfy IoT requirements. Therefore, 6G is a promising generation of wireless communications that is expected to overcome 5G in enhancing the performance of these systems. Many 6G promising technologies have been proposed, such as artificial intelligence (AI), and symbiotic radio(SR). These technologies are expected to meet the requirements of the IoT system such as enhancing the system rate, capacity, and energy efficiency where the IoT system has a low downlink data rate of 250 kb/s and an uplink data rate of 15 kb/s [3], [4], [5]. AI is a critical and highly valued technology in the future's 6G where it provides intelligence to wireless networks. AI is used as a tool for system data analysis which helps in making correct decisions automatically [6], [7].

Backscattering (BackCom) is a technology used to improve wireless system performance. Backscattering provides spectrum resources and energy for IoT tags, thereby improving the system spectrum and energy efficiency. Backscattering is presented in the LTE (LScatter) and Wi-Fi (BackFi) frequency bands, where the tag's information is

The associate editor coordinating the review of this manuscript and approving it for publication was Hassan Tariq Chattha.

ridden and modulated over continuous wave signals generated by a carrier emitter [8]. Recently, ambient backscatter communication (AmBC) system have been used for passive Internet of Things that do not require a dedicated carrier emitter. The ambient system enables backscattering tags to transmit their information over the existing ambient Radio Frequency (RF) signal (e.g., Cellular or Wi-Fi signal) that shares the transmitter and spectrum of the primary signal; however, it has some decoding problems at the reader [9]. Therefore, symbiotic radio SR is used for passive Internet of things communication, which is one a 6G promising technologies. SR provides highly reliable backscattering communications, which compensate for the disadvantage of AmBC. In SR, IoT tags not only share the transmitter and spectrum of the primary signal but also share the primary receiver. The primary receiver decodes the signals from the primary transmitter and backscattering tags [10]. Therefore, using SR can significantly enhance the performance of IoT systems, where mutualism spectrum sharing increases the system spectrum efficiency and backscattering communication increases its energy efficiency. Decoding at the primary receiver by utilizing successive interference cancellation (SIC) can solve the decoding problems and provide a reliable system. SIC was used to estimate the primary signal first. Then, the receiver subtracts this estimated signal from the received signal and finally, detects the tags' backscattered signal. Recently, mode selection communication scheme is studied with objective of enhancing the IoT system performance by selecting between active (battery- assisted communication) and passive mode (battery- free communication) based on the network topology and the location of the IoT device related to the receiver. The passive mode is more required for IoT network where it solves the problem of longer communication range and support "green communication" [11], [12], [13]. Furthermore, machine learning is used based on backscatter for providing IoT system performance enhancement as the bit error rate(BER), throughput and capacity of the system. A trained model can be obtained by using machine learning based on the input data of the system which improve the system performance [14].

A. LITERATURE REVIEW

In this section, many studies are have attempted to improve IoT performance by maximizing throughput, capacity, and energy efficiency using different 6G techniques. In [15] backscattering (BackCom) is an emerging technique in communication systems. backscatter technique and tag-power harvesting are coupled, which enhances the spectrum and energy efficiency of the system. In [16] and [17] exposed the ambient backscatter communication (AMBC) technique that enables smart tags to communicate using ambient radio frequency (RF) signals without requiring active RF transmission. IoT systems can use this technique to enhance their performance. In [18] AMBC was integrated into RF-powered underlay Cognitive Radio Non-Orthogonal Multiple Access (CR-NOMA) networks which enabled secondary transmitters to utilize the harvested energy for simultaneous data transmission depending on power-domain NOMA. RF-powered underlay CR-NOMA networks can increase secondary system throughput. Although, the promising aspect of ambient backscattering is successful in enhancing the system performance, it has a detection problem at the reader, where the ambient backscatter system detects the signal of other systems' that are strangers to the reader, and the channel parameters cannot be estimated by the reader.

Other studies have supported the symbiotic radio (SR) technique for communication systems. Therefore, in [10] and [19] SR is introduced and detection is performed at the primary receiver using successive interference cancellation (SIC), which detects the primary signal first and then subtracts this signal from the received signal and decodes the backscattered signal. SR can improve IoT performance more than BackCom and AMBC. In [20], an analysis of the downlink of a NOMA-enabled cellular system batched with a symbiotic radio, derived the rate for the symbiotic radio system and formulated an expression for the outage probability of the signal-to-interference and noise ratio. In [21], three practical transmission schemes were proposed with different symbiotic relationships between backscatter and primary systems. The achieved transmissions rates of the primary and backscatter systems can be derived for the fading state and an optimal power allocation is acquired for the transmission scheme. In [11] communication mode selection scheme is proposed for IoT network where the IoT devices select between active transmission mode and backscattering to nearby smart phone mode according to the network topology and transmit power budget. A mode selection problem is formulated to maximizing the connection density of the system where at passive mode the IoT tags backscatter the nearest smart phone signal which act as a relay to deliver the tag signal to the MBS. In [12] an adaptive radio mode selection is proposed for IoT network where the IoT device can transmit its information through active or passive mode. Active mode supports the traditional battery-assisted and passive mode supports backscatter- assisted communication. An adaptive communication problem was formulated to maximize the power and minimize the outage performance of the system. In [22] hybrid active and passive antenna selection for signal transmit/ backscatter system is proposed using multi input single output (MISO) antennas where two working mode is allowed for each antenna (active and passive modes) which is determined by choosing the maximum value of the transmit power in the active mode and the backscatter power in the passive mode. This scheme aims to enhance the system reliability at active mode and reduce the power consumption at passive mode. In [23] AMBC-assisted decode-and-forward (DF) relaying scheme is proposed so, the relay can decode and forward the backscattered signal of the tag where received power at the relay is split between the traditional DF scheme and the AMBC scheme which enhance the performance of the system. In [24] an unsupervised machine learning

(UML) algorithm is proposed for signal detection based on AMBC communication system where the tags received signal features are exploited to groups them into clusters which improve the system detection performance.

B. CONTRIBUTION

Our work aims to enhance the uplink IoT system data rate to contribute in solving the growth of IoT devices issue (Massive IoT), which requires better system performance. The problem is proposed to be solved into two optimization problem phases that are solved in series. In the first phase solution, a many-to-one matching game is used to select the network (LTE or Wi-Fi) for information transmission where, the smartphone act as a relay to deliver decoded IoT tag information to the MBS or WAP, which maximizes the overall system rate. In the second phase solution, the SR technique is used for enabling IoT tags to reflect their information to the smartphones by backscatter the downlink signal between the MBS and smart phones then this information is decoded at the smartphones. IoT tags were virtually divided into optimum clusters using the DDQL algorithm. The information of Each cluster tag is backscattered on a smartphone through (LTE or Wi-Fi) network. The contributions of our paper are as follows:

1) A symbiotic radio communication technique was proposed to enhance the throughput of IoT systems.
2) Two optimization problem phases are considered to maximize the total uplink rate of IoT system.
3) A Many-to-one matching game algorithm is used to solve the first problem phase by selecting the network (LTE or Wi-Fi) for the transmission of the smartphone decoded information that maximize the system rate.
4) The DDQL algorithm is utilized to solve the second problem phase by achieving optimum tags clusters that backscatter on each smartphone signals.

The reminder of this paper is organized as follows. In section II, we present the proposed system model. In section III, we formulate the throughput maximization problem of the system. In section IV, we present the matching game algorithm and explain the details of the proposed algorithm. In section V, we provide the theoretical description of the double deep reinforcement learning algorithm and the details of the proposed algorithm. In section, VI, the network parameters are presented. Section VII introduces network parameters. Finally, we conclude the paper in section VIII.

II. SYSTEM MODEL

Consider a symbiotic radio communication system with K IoT tags, which is represented by $\mathcal{K} = \{1, 2, \ldots\ldots, K\}$, M smart phones are represented by $\mathcal{M} = \{1, 2, \ldots\ldots, M\}$ and N Wi-Fi access point $\mathcal{N} = \{1, 2, \ldots\ldots, N\}$ under the coverage of a microcell base station (MBS) as shown in Figure (1). The IoT tags are backscattered on the downlink primary signal from the MBS or WAP and reflect their information to the smartphones by using the SR technique to enhance the system performance by increasing the spectrum and energy efficiency of the system, where it shares the primary signal transmitter, receiver, and spectrum resource. Then, the smartphone is used as a relay that decodes the tag information using the SIC technique and transmits it to the MBS through the LTE network or WAP through the Wi-Fi network to reach the data cloud. The data cloud is an important part of the IoT network, where it is responsible for tag information analysis and computing to control system tags. The IoT tags are distributed virtually into clusters, and each cluster tags are allowed to backscatter their information to one of the neighboring smartphones. The optimum cluster is one in which its members can achieve the maximum total rate. Then, the tags information is decoded at the smartphone and transmitted through a network of (LTE or Wi-Fi). The WAPs operate in a frequency range different from that of the LTE MBS. Moreover, Orthogonal Frequency Division Multiple Access (OFDMA) is presumed to be utilized by the LTE network to access the channel, whereas WAPs use the IEEE 802.11 distributed coordination function (DCF) mechanism which is dependent on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol for channel access. The following sections describe the received data rate formulas of the IoT tags information at the smartphone, and the data rate of the smartphone received information at the MBS or WAP based on the OFDMA technique for transmission using a LTE network and orthogonal frequency division multiplexing (OFDM) technique for transmission through a Wi-Fi network [25].

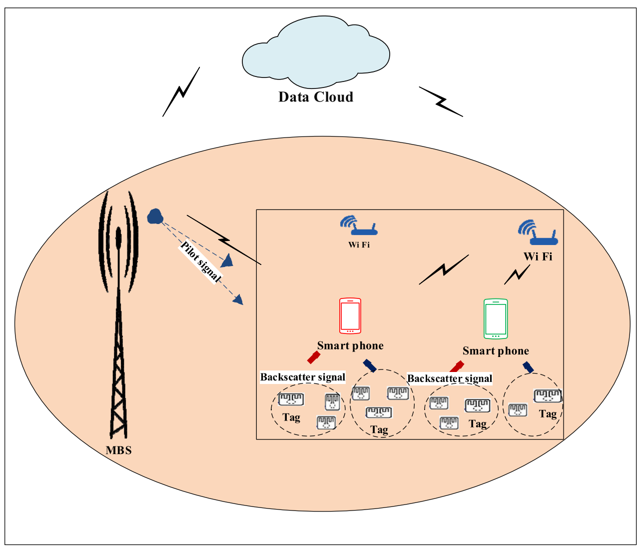

FIGURE 1. SR for IoT system model.

A. RATE CALCULATION FOR THE SMARTPHONE

In the first problem solution, the expected detected tag information at the smartphone is associated with the MBS through the LTE network or with the WAP through the Wi-Fi network based on the maximum data rate achievement. Therefore, the data rate of the tags' information signal is calculated network selection for the tags' information transmission.

The SNR of the tags' information signal transmitted from the smartphone m to the MBS can be expressed as,

$$\gamma_{ml}^{LTE} = \frac{p_{\text{ml}} h_{\text{ml}}}{N_0} \tag{1}$$

where, p_{ml} is the power transmitted from the smartphone to the MBS, h_{ml} is the channel gain between smartphone and the MBS as shown in Figure (2). Subsequently, the physical data rate of the signal can be determined instead of the simple and theoretical form of Shannon by mapping with the SNR (γ_{LTE}) as:

$$R_{ml}^{LTE} = \frac{N_{fB}.N_{sub}.N_n.N_{bx}.ECR_x}{T_{Sf}} \times \frac{1}{R_{epx}} \tag{2}$$

where N_{fB} is the number of frequency blocks in bandwidth B, N_{sub} is the number of subcarriers in one Block, N_n is the number of each subcarrier's symbols, T_{Sf} is the subframe duration, R_{epx} is the repetition factor. Furthermore, N_{bx} is the number of symbol bits, and ECR_x is the code rate [26], [27].

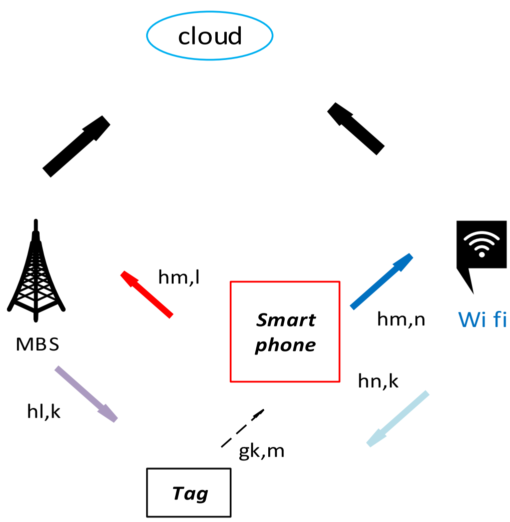

FIGURE 2. Channels model of SR for IoT system.

For WAP association, the SNR of the tags information signal that is transmitted from smartphone m to WAPn through the Wi-Fi band (2.4 GHZ) can be expressed as follows:

$$\gamma_{mn}^{Wi-fi} = \frac{p_{\text{mn}} h_{\text{mn}}}{N_0} \tag{3}$$

We denote p_{mn} as the transmitting power of the smartphone m to WAP n. h_{mn} is the channel between the smartphone m and WAP n. A contention extent may be on the Wi-Fi channel, which affects system throughput. This contention depends on the number of users associated with the WAP, in which a collision between them may occur. Thus, each WAP has a maximum capacity; based on its characteristics and the technology used; to avoid collision. The data rate of the tag information signal at WAP is calculated based on the enhanced 802.11 DCF back-off model to obtain the channel contention considering the CW resetting scheme and the packet collision effect [28], [29] as follows:

$$R_{mn}^{Wi-fi} = \frac{\tau(1-\tau)^{N_w} D}{T + \left(\frac{D\tau(1-\tau)^{N_w}}{R_{mn}^{ph}}\right)} \tag{4}$$

where τ represents the probability of channel contention, N_w denotes the total number of users associated with the WAP, D is the maximum allowed size of the user packets. T can be denoted as,

$$T = (1-\tau)^{N_w+1} e + (1 - \left(1-\tau)^{N_w+1}\right)(T_{cTS} + T_{DIFS}) + (N_w+1)\tau(1-\tau)^{N_w}(T_{RTS} + T_{ACK} + 3T_{sIFS}) \tag{5}$$

where e is the duration of an empty slot time; T_{RTS}, T_{cTS}, T_{DIFS}, T_{sIFS} and T_{ACK} denote the duration of the request to send (RTS) short frame, clear to send (CTS) short frame, DIFS, SIFS, and ACK, respectively. R_{mn}^{ph} is the physical rate of the transmitted signal from the smartphone to the WAP which can be calculated as follows:

$$R_{mn}^{ph} = \frac{N_{spatial}.N_{sub}.N_b.C_w}{T_S} \tag{6}$$

where, $N_{spatial}$ is the number of streams, N_{sub} is the total number of data sub-carriers, T_S represents the OFDM symbol duration, N_b is the number of bits per symbol, C_w is the coding rate, and N_b C_w are determined by mapping the measured SNR (γ_{mn}^{Wi-fi}) value at the WAP [27].

B. RATE CALCULATION FOR THE IoT TAGS' BACKSCATTERED INFORMATION

As mentioned, IoT tags are backscattered on the downlink primary signal from the MBS to a smartphone through the LTE band. The received backscattered signal at the smartphone can be represented as,

$$y_{sp}^{L}(t) = h_{lm}s(t) + \eta g_{km}\, h_{lk}x(t)\, s(t) + N_0 \tag{7}$$

where, s(t) is the downlink primary signal of the smartphone and x(t) is the tag information signal. As shown in Figure 2, h_{lm} is the channel gain between the RF source (MBS) and the smartphone m. h_{lk} is the channel gain between the RF source (MBS) and the tag k. g_{km} is the channel gain between the smartphone m and the tag k. N_0 denotes the additive white Gaussian noise AWGN. η is the reflection coefficient [20], [30]. The SINRγ_{km}^{L} of the tag k signal that backscatters the down link primary signal of smartphone m can be computed as follows:,

$$\gamma_{km}^{L} = \frac{\eta p_l g_{km} h_{l,k}}{\sum_{i=1, i\neq k}^{k_{lm}-1} \eta g_{im} p_{l,} h_{l,i} + N_0} \tag{8}$$

where, p_l denotes the power of the downlink RF signal. $\sum_{i=1,i\neq k}^{k_{lm}-1} \eta g_{im} p_{l,} h_{l,i}$ represent the interference caused by other $k_{lm} - 1$ tags which backscatter the same downlink primary signal between the MBS and smartphone m. The smartphone detects these backscattered signals using SIC. SIC detects the tag's signal after subtracting the primary signal, which provides a reliable system.

In addition, IoT tags use Wi-Fi transmissions from the WAP to a Wi-Fi client as the excitation signal to ride it. The tag receives the Wi-Fi signal and reflects its information on the smartphone. The received signal at the smartphone through the Wi-Fi frequency band can be expressed as,

$$y_{sp}^{w}(t) = h_{mn}w(t) + \eta g_{km}h_{nk}x\,(t)\,w(t) + N_0 \tag{9}$$

where, w(t) is the primary Wi-Fi transmitted signal, h_{mn} is the channel gain between smartphone m and WAP n, h_{nk} denotes the channel gain between WAP n and IoT tag k. SINR of the received tag's k backscattered information signal at smartphone m through the Wi-Fi frequency band (2.4 GHz) can be calculated as,

$$\gamma_{km}^{w} = \frac{\eta p_w g_{km} h_{nk}}{\sum_{i=1,i\neq k}^{k_{nm}-1} \eta g_{im} p_{w,} h_{ni} + N_0} \tag{10}$$

where, p_w denotes the power of the WAP-transmitted signal. $\sum_{i=1,i\neq k}^{k_{nm}-1} \eta g_{im} p_{w,} h_{ni}$ represent the interference caused by other $k_{nm} - 1$ tags which backscatter the same primary Wi-Fi signal between WAP n and the smartphone m [31], [32].

The rate of each tag k $\in \mathcal{K}$ at the smartphone m $\in \mathcal{M}$ can be calculated by using the Shannon Model where the tag rides the primary signal without any control as,

$$R_{mk} = Blog_2(1+\gamma_{mk}) \tag{11}$$

where, γ_{km} is the SINR of tag k which may be γ_{km}^{L} or γ_{km}^{w}.

III. PROBLEM FORMULATION

In this section, the system problem is formulated to maximize the overall IoT system throughput, which contributes to massive IoT devices communication. This is represented by two optimization problem phases. The first problem phase, p1 (12), can be solved by matching the game algorithm to determine the optimum mode selection of the network (LTE or Wi-Fi). The second problem phase, P2 (13), is solved by finding the optimal cluster of tags using the DDQL algorithm, which satisfies the maximum system throughput.

$$\begin{aligned} &P1 : \max \sum\nolimits_{m\in \mathrm{M}} \sum\nolimits_{n\in N} (R_{ml}^{LTE} + R_{mn}^{Wi-fi}) \\ &P2 : \max \sum\nolimits_{m\in \mathrm{M}} \sum\nolimits_{k\in K} X_{km}R_{km} \end{aligned} \tag{12}$$

$$\begin{aligned} &\text{Subject to} \\ &\quad C1 : p_{ml} > p_{lth} \text{ and } p_{mn} > p_{wth}, \\ &\quad C2 : \gamma_{lm}^{L} > \gamma_{km}^{L} \text{ and } \gamma_{mn}^{w} > \gamma_{km}^{w}, \\ &\quad\quad \forall k \in \mathcal{K}, m \in \mathcal{M} \text{ and n} \in \mathcal{N} \\ &\quad C3 : X_{km} = \{0, 1\} \end{aligned} \tag{13}$$

Constraint C1 confirms that the smartphone is covered by the MBS and the WAP where, p_{lth} and p_{wth} denote the threshold transmitting power of MBS and WAP respectively that guarantee the coverage. Constraint C2 was used to obtain the a condition of perfect SIC. Constraint C3 is a backscattering indicator where, X_{km} is a binary value that be 0 or 1. In the following sections, DDQL and matching game algorithms are descibed.

IV. MATCHING GAME ALGORITHM

*The matching game is*a low-complexity algorithm; it is one of the algorithms used to solve maximization problems to enhance the performance of the system. Matching game has many types: one- to- one matching game in which one agent from one side of matching is matched with one agent on the other side, many-to-one matching game in which many agents from one side can be matched with one agent on the other side, and many-to-many matching game in which many agents from one side can be matched with many other agents on their side [33], [34], [35], [36].

A. THEORETICAL DESCRIPTION

In this work, we applied a many-to-one matching game to enable each smartphone to determine the optimum selection of the network LTE or Wi-Fi band which transmits the tags information through it to maximize its signal rate. Thus, the smart mobile device constructs a preference list according to the utility function U_{tm} which can be defined as,

$$U_{tm} = R_m, \tag{14}$$

where, R_m is the smartphone mobile rate which may be R_{ml}^{LTE} or R_{mn}^{Wi-fi}. The utility function that the MBS or WAP accepts the request for the smartphone to maximize the overall system rate, is calculated as:

$$U_{tb} = R_T, \tag{15}$$

where, R_T is the total system rate. The smart mobile request may not be accepted at the MBS or WAP, in this case, the smartphone tries to select the next choice in its preference list.

B. THE PROPOSED ALGORITHM 1

In the proposed algorithm, the smartphone calculates its utility function; based on (14), and constructs its preference list. Then, it tries to select the frequency band (LTE or Wi-Fi) for tag information transmission by sending a request to the MBS or WAP according to the first choice in its preference list. The request is accepted or rejected based on (15) and repeated until stable matching Φ is achieved that all the smart phones are connected with MBS or WAP.

V. DOUBLE DEEP REINFORCEMENT LEARNING ALGORITHM

The reinforcement learning algorithm is a learning process in which agents take action and wait for the reaction of the environment. A reward is provided by the environment to the agent for every action and the Reinforcement learning (RL) agent selects the correct action that maximizes the reward of the new state. Deep reinforcement learning (DRL) algorithms can be classified into three types:

(I) Value-based methods that determine a policy by learning only the value function Deep Q -learning(DQL).

(II) Policy-based methods that find policy directly based on the gradient related to the policy.

Algorithm 1 Many –to-One Matching Game

- Discover the tags backscattered information signals at each smart phone and calculate its channel gain through MBS and WAP.
- Construct the preference list of each smart mobile based on its utility function U_{tm} in (14).
- Requests to select the frequency band for information signal transmission of each smart mobile according to its preference list.
- Construct the MBS and WAP preference list based on (15) and decide wether to accept or reject each mobile request according to it.
- repeat
  The rejected request re-apply to the next choice in its preference list.
- Until
  All smart mobiles connect to (MBS or WAP) for signal transmission. A stable matching Φ was achieved.
- Calculate the system rate.

(III) Actor-critic methods that are a hybrid of the previous two types of value-based and policy-based methods [37].

A. THEORETICAL DESCRIPTION

The DQL algorithm was proposed to solve problem P2. DQL is a multi-layered neural network that for a given state $\mathrm{s_t}$ outputs an action $\mathrm{a_t}$ where, θ^{Q} are the parameters of the network at time t. Firstly, it takes an action $\mathrm{a_t}$ at state $\mathrm{s_t}$ and then observes the environment's immediate reward $\mathrm{r_{t+1}}$ with the resulting state $\mathrm{s_{t+1}}$, The DQL update equation is:

$$\mathrm{Q}\left(\mathrm{s}, a \mid \theta^{Q}_{t+1}\right) = \mathrm{Q}\left(\mathrm{s}, a \mid \theta^{Q}_{t}\right) + v\left[\mathrm{r_{t+1}} + \xi \max_{a}^{\prime}\left(\mathrm{s_{t+1}}, a' \mid \theta^{Q}_{t}\right) - \mathrm{Q}\left(\mathrm{s_t}, a_t \mid \theta^{Q}_{t}\right)\right] \tag{16}$$

$$= \mathrm{Q}\left(\mathrm{s}, a \mid \theta^{Q}_{t}\right) + v\,[\mathrm{r_{t+1}} + \xi \max_{a'} \mathrm{Q}\left(\mathrm{s_{t+1}}, \arg\max_{a'} \mathrm{Q}\left(\mathrm{s_{t+1}}, a' \mid \theta^{Q}_{t}\right) \mid \theta^{Q}_{t}\right) - \mathrm{Q}\left(\mathrm{s_t}, a_t \mid \theta^{Q}_{t}\right)] \tag{17}$$

where, ν denotes the learning rate. The learning term $\mathrm{Q}(\mathrm{s}, a|\ \theta^{Q}_{t+1})$ is the updated Q network which is calculated depending on the current Q network $\mathrm{Q}(\mathrm{s}, a\ |\theta^{Q}_{t})$ at action , $a'.max_{a'}\mathrm{Q}(\mathrm{s_{t+1}}, argmax_{a'}\mathrm{Q}(\mathrm{s_{t+1}}, a'|\theta^{Q}_{t})|\theta^{Q}_{t})$ represent a systematic overestimation of the Q-value. Double DQL is an algorithm used to reduce the overestimation. In DDQL two networks Q and Q' are learned by randomly assigning each experience to update one of the two networks, where, there are two parameters θ^{Q} and $\theta^{Q'}$. In each update, one of these parameters is used to obtain the greedy policy and the other to obtain its target value. DDQL reduces overestimations by decomposing the max operation into selection and evaluation action a', a that evaluates the greedy policy according to the target network [38], [39]. The DDQL update equation(Bellman) can be represented as,

$$\mathrm{Q}\left(\mathrm{s}, a \mid \theta^{Q}_{t+1}\right) = \mathrm{Q}\left(\mathrm{s}, a \mid \theta^{Q}_{t}\right) + v\,[\mathrm{r_{t+1}} + \xi \max_{a'} \mathrm{Q}\left(\mathrm{s_{t+1}}, arg\max_{a'} \mathrm{Q}\left(\mathrm{s_{t+1}}, a' \mid \theta^{Q'}_{t}\right) \mid \theta^{Q}_{t}\right) - \mathrm{Q}\left(\mathrm{s_t}, a_t \mid \theta^{Q}_{t}\right)] \tag{18}$$

The parameter $\theta^{Q'}$ of the Q' network can be calculated periodically using the Polyak averaging method based on parameter τ to represent the parameter θ^{Q} of Q network after time step t_0 as [40], and [41].

$$\theta^{Q'}_{t+t_0} = (1-\tau)\theta^{Q'}_{t} + \tau\theta^{Q}_{t} \tag{19}$$

Our goal is to solve the maximization problem of the system throughput by designing a DRL system for virtual tags clusters. We optimized these virtual clusters using the DDQL algorithm. Figure (3) shows a block diagram of the DDQL Scheme, which is used to solve the problem in (12).

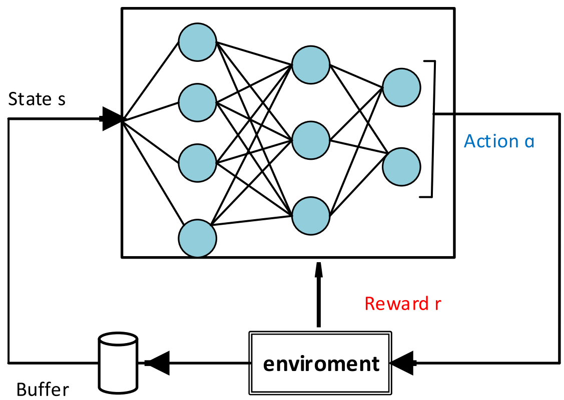

FIGURE 3. DDQL model.

The design includes the tag virtual clusters state s and, the reward r, which is given by the environment when taking an action a through the DDQL algorithm training and reply buffer observation of different actions and then randomly sample from the buffer experience. The state s of IoT tags K is taken as a Vector of SINR value for each tag in its virtual cluster and the reward is the throughput of the system.

B. THE PROPOSED ALGORITHM 2

We begin by initializing the environment with one MBS, K IoT tags, M smartphones, and N WAP. At the initial state $\mathrm{s_0}$ the input of the neural network is a vector of the SINR of the IoT tags based on (8), and (10), regardless of whether it backscatters the smartphone signal through the LTE or Wi-Fi band. The DDQL agent takes action a_t at every time step, t, of the episode. The environment sends an immediate reward to the agent and the state transitions to a new state $\mathrm{s_{t+1}}$. The neural network updates the parameter θ^{Q}_{t} to θ^{Q}_{t+1} (19) to calculate $\mathrm{Q}(\mathrm{s}, a\ |\theta^{Q}_{t+1})$ using the Bellman equation (18).

Algorithm 2 DDQL Algorithm for IoT Tags Clustering

• Initialization: Discover each IoT tags, smart phones, WAP location coordinates.
• Randomly initialize the Q network Q(s, a |θ_t^Q).
• Initialize the target network Q'
• $\theta^{Q'} \Leftarrow \theta^{Q}$
• Initialize replay buffer R
• for episode = 1……; E do
• initializing the Environment state s_0
• calculate the SINR for each tag as an input vector of the neural network.
• for t = 1…..; T do
  Select the clustering action a_t ➧ $arg\ \max_a$ Q(s_t, a_t)
  Execute action a and observe reward r_t and observe the new state s_{t+1}
  Store transition (s_t, a_t, r_t, s_{t+1}) in Buffer R
  Sample a random mini-batch of L transitions from the buffer R.
• Training of DDQL
  Select a' ➧ $argmax Q'(s_{t+1}, a'|\theta_t^{Q'})$
  Update the Q based on bellman equation in (17).
  Update the DDQL target network if mod (t, p) = 0
  end for
  end for

VI. NETWORK PARAMETERS

The proposed algorithm results were evaluated based on parameters listed in Table 1. Moreover, the indoor path loss model between tags and MBS and, WAP at the LTE and Wi-Fi bands, respectively, are represented by (20) and, (21) as,

$$p^{\mathrm{LTE}} = 128.1 + 37.6 \log(d_L) \tag{20}$$

$$p^{\mathrm{Wi\,Fi}} = 20 \log f_w + \eta_w \log d_w + p_f(M_w) - 28 \tag{21}$$

where d_L is the distance between the MBS and the IoT tags, d_w is the distance between the WAP and the IoT tags, f_w is the frequency band of the Wi-Fi, η_w is the coefficient of distance power loss and is assumed to be 30, $p_f(M_w)$ represents the penetration loss factor and is equal to $p_f\ (M_w) = M_w + 13$; M_w denotes the number of walls and is assumed to be 3 [27].

Tables 1 and 2 list the parameters used for the simulation and calculation of the matching game algorithm and the DDQL algorithm, respectively [29], [40]. The smartphones can be in movements so, our proposed algorithm runs periodically to update its calculations based on the recent locations of the smartphones.

VII. PERFORMANCE EVALUATION

To evaluate our proposed algorithms, the performance of the proposed matching game algorithm which is used as a solution to the first problem of the system, is compared with the LTE and Wi-Fi first association techniques. Then, the DDQL algorithm, which is proposed to solve the second problem of the system, is compared with the system that does not use the nearest smartphone algorithm. Furthermore, the performance of the system in terms of data rate and capacity was studied using different numbers of smartphones to explain the optimum number of smartphones needed for this system.

TABLE 1. Simulation parameters of matching game algorithm [27].

Simulation parameters	Value
The cell radius	500 m
IoT bandwidth (LTE)	180 kHz
Number of subcarriers for 3.75 kHz/ subcarrier	48 subcarriers
Additive white Gaussian noise	-174 dBm/Hz
Max transmit power (P_{max}) at LTE	23 dBm
The number of frequency blocks in bandwidth B	1 block
Wi-Fi technology	802.11g
Max transmit power (P_{max}) at Wi-Fi	200 mW
Minimum contention window(W)	32
Slot time	9 µs
DIFS, SIFS	50,10 µs
RTS,CTS	208,160 bits
ACK	160 bits
Data Payload size (D)	1500 bytes

TABLE 2. Simulation parameters of DDQL algorithm [38].

Simulation parameters	Value
State s = {s_1, …. s_k }	$\{\gamma_1, \ldots. \gamma_k\}$
Reward r	R_{mk}=$B\ log_2(1+\gamma_{mk})$
Discount factor ξ	0.99
Learning rate ν	$5*10^{-5}$
Buffer size	20000
Episode, Time step	2000, 500

Figure (4) shows our proposed matching game algorithm which is designed to follow the smartphone utility function to choose the most suitable network with a specific transmission mode (LTE or Wi-Fi), through which an initial smartphone load is proposed to be transmitted. Matching game algorithm succeeded in maximizing the total data rate of smartphones by 45% compared with the system using LTE first (all operation of LTE protocol), and by 12% compared with Wi-Fi first (all operation of Wi-Fi Protocol). The figure shows that the system data rate using the matching game is increased by increasing the number of smartphones that have more information, and then it will be saturated as a result of exhausting the spectrum resources.

Thus, the total data rate of the system is saturated, regardless of the increase in smartphones and their information.

Figure (5) shows that the proposed DDQL algorithm outperforms the backscattering of IoT tags at nearest smartphone by 16% in terms of the total rate of the system second phase

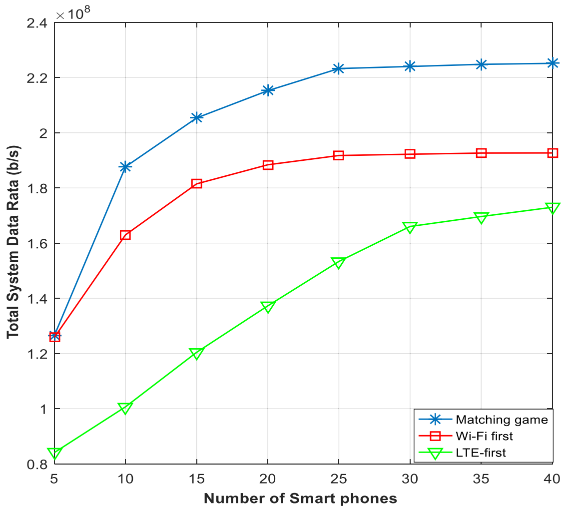

FIGURE 4. Total data rate vs. number of smartphones.

versus different number of tags at M = 2. This is because the DDQL algorithm can find the optimum clusters of backscatter tags for each smartphone which maximize the total data rate of the tags when using only two smart phones at the second phase of the system. As we can see from the figure the total data rate firstly increases by increasing the number of tags and then, it decreases gradually as a result of increasing mutual interference between tags in the same cluster.

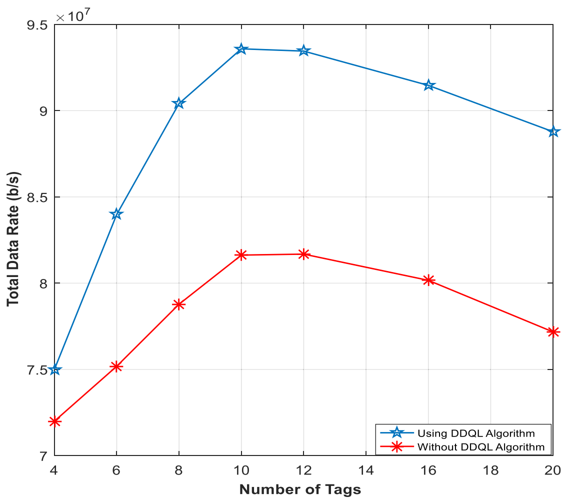

FIGURE 5. The total data rate vs. number of tags.

Figure (6) illustrates that our proposed algorithm which use the matching game and DDQL algorithms successfully maximizes the overall system data rate compared to the system performance by using the matching game but without the DDQL algorithm, using the LTE network first with the DDQL algorithm, and using the LTE network first without DDQL. As shown in the figure, our proposed scheme can enhance the total system data rate by an average of 90% above the system without any algorithms, 20% above the system without the DDQL algorithm, and 90% above the system with the LTE network first. This figure explains the performance of the system at a different number of tags $\mathcal{K}$ and number of smartphones $\mathcal{M} = 4$ and Wi-Fi access point $\mathcal{N} = 2$ where the total system data rate increase by increasing the number of tags and then, it saturated according to the achieved data rate of the system with matching game by using a number of smart phones $\mathcal{M} = 4$. So, we studied the performance of different numbers of smartphones, as shown in Figure (7).

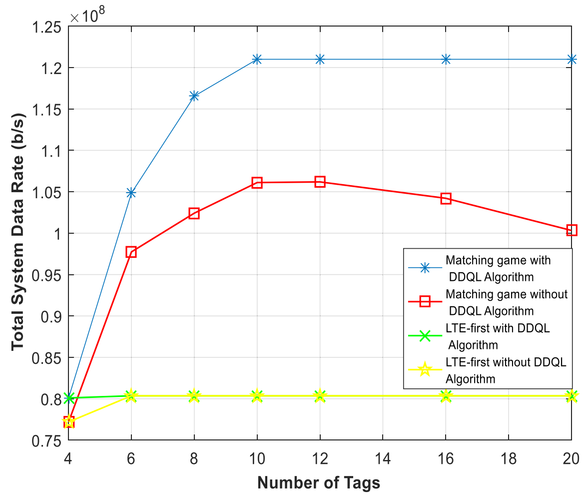

FIGURE 6. Total system data rate vs. the number of tags.

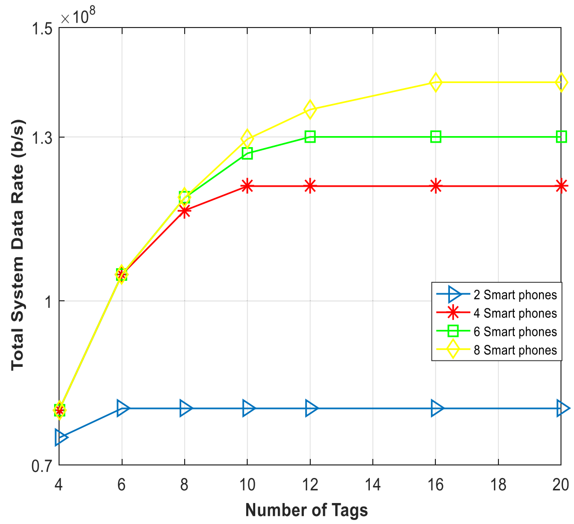

FIGURE 7. The total system data rate vs. the number of tags at different numbers of smartphones.

The total system data rate of our proposed algorithms (DDQL and Matching game) is illustrated in Figure (7) for different numbers of smartphones to obtain the optimum number of smartphones required for our system. We observe that for a number of smartphones, $\mathcal{M} = 4$, the total s ystem data rate is increased by an average of 90% compared to the system-based number of smartphones $\mathcal{M} = 2$. This is because when smartphones increase the number of tags per smartphone cluster, the achieved rate of each cluster increase as a result of less mutual interference between the cluster tags However, at $\mathcal{M} = 6$ and 8, the total system rate increases slightly by 5% and, 8% over the total data rate of the system-based number of smartphones, $\mathcal{M} = 4$, because each smartphone cluster has approximately one or two tags. Based on the aforementioned, the optimum number of smartphones for our system is $\mathcal{M} = 4$ because after four smartphones any increase in the number of smartphones has a low effect on the total system data rate.

Figure (8) shows the effect of increasing the number of smartphones on the system capacity of IoT tags at the required data rate per IoT tag of 10 Mb/s assuming that the number of tags is K=20. This figure shows that with an increase in the number of smartphones from $\mathcal{M} = 2$ to $\mathcal{M} = 4$, the capacity of the system greatly increases as a result of fewer cluster tags and subsequently a high cluster data rate.

Subsequently, the increase in the number of smartphones achieved a slight system capacity enhancement. Table 3 presents the system capacity for different numbers of smartphones M. Furthermore, the figure illustrates that the system capacity of our proposed algorithms using a matching game with DDQL algorithms outperforms the capacity of the system with LTE first with DDQL algorithms, where it succeeded in enhancing the system capacity on average by 100% above the system using the LTE network first with the DDQL algorithm at a number of smartphones $\mathcal{M} = 4$.

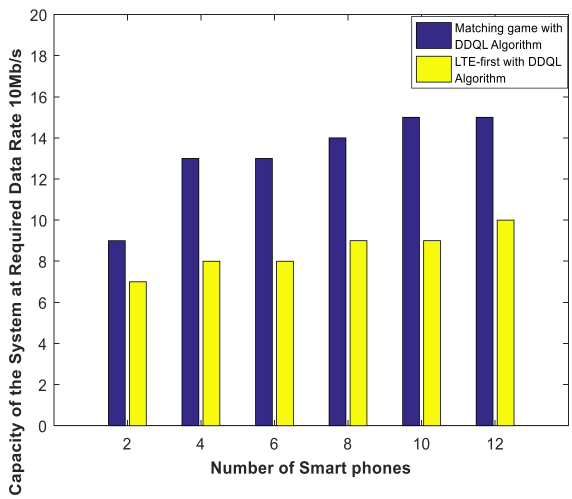

FIGURE 8. The capacity of the system vs. the number of smart phones.

However, for a high number of smartphones $\mathcal{M} = 10, 12$ the capacity of the system becomes the same. This is because each smartphone cluster has one tag; therefore, the DDQL algorithm has no effect in this case.

Figure (9) represent a comparison between the total data rate of the system second phase by using the DDQL algorithm and using pairing and scheduling algorithm versus different number of tags at $\mathcal{M} = 4, 6$. The figure shows that the DDQL algorithm overcome the pairing and scheduling algorithm in

TABLE 3. Capacity of the system vs. number of smart phones.

Number of Smart phones ($\mathcal{M}$)	2	4	6	8	10	12
Capacity of the System	9	13	14	15	16	16

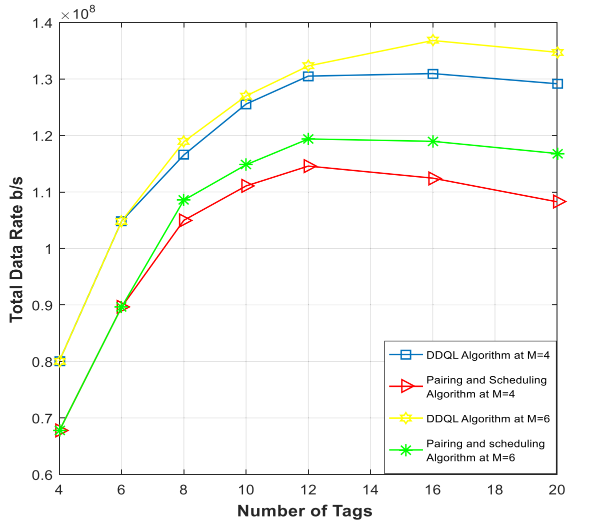

FIGURE 9. The total system data rate vs. number of tags at $\mathcal{M} = 4, 6$.

increasing the total data rate by 10% at $\mathcal{M} = 4$ and 11% at $\mathcal{M} = 6$ where, DDQL algorithm by finding the optimum clusters of IoT tags success in increasing the total data rate than scheduling and pairing algorithm which allows the tags that backscatter the nearest smart phone and uses it as relay to deliver its information to the MBS. Furthermore, the figure shows that firstly, the system data rate increases by increasing the number of tags then it decreases as a result of increasing interference between tags.

Figure (10) illustrates that our proposed algorithm outperforms the pairing and scheduling algorithm by 50% in term of the capacity of system at required rate per IoT device equal 10Mb/s versus different number of smart phones at $K = 20$. We observe that the capacity of the system increases by increase the number of smart phones in the two algorithm. However, the system capacity increases slightly after $\mathcal{M} = 8$ by using pairing and scheduling algorithm because the number of tags that are backscattered on each smart phone decrease after $\mathcal{M} = 8$. AT our proposed scheme the system capacity increases slightly after $\mathcal{M} = 4$ as a result of fewer cluster (LTE or Wi-Fi) tags where each smart phone has two clusters LTE and Wi-Fi. So, the matching with DDQL algorithm achieves better performance than pairing and scheduling algorithm with lower number of smart phones.

In Figure (11) the DDQL algorithm can outperform both of the UML (clustering) algorithm and the paring and scheduling algorithm in enhancing the system second phase data rate by average 7% and 10% respectively. The DDQL algorithm can find the optimum clusters which maximize the system

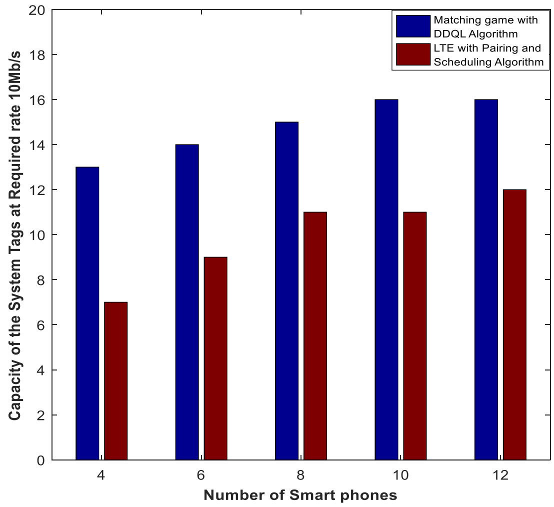

FIGURE 10. **The capacity of the system vs. the number of smart phones.**

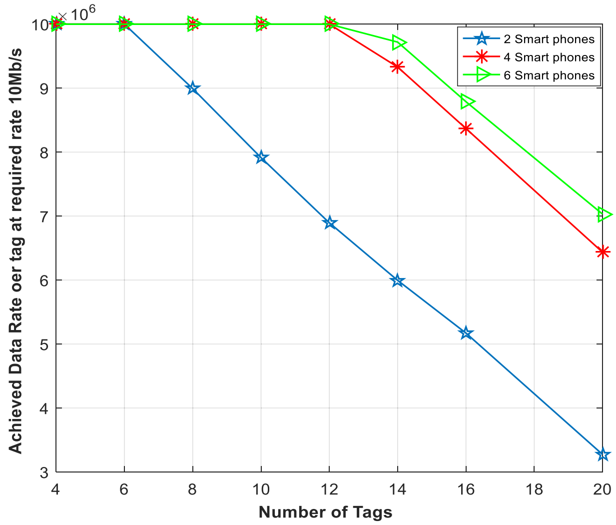

FIGURE 12. **The achieved data rate per tag vs. number of tags.**

total data rate better than the unsupervised clustering algorithm which only groups the tags into clusters according to their features to backscatter the smart phone signal [24] and better than the pairing and Scheduling algorithm which allows only the IoT tags that achieve the threshold SINR to be backscattered on the primary smart phone signals.

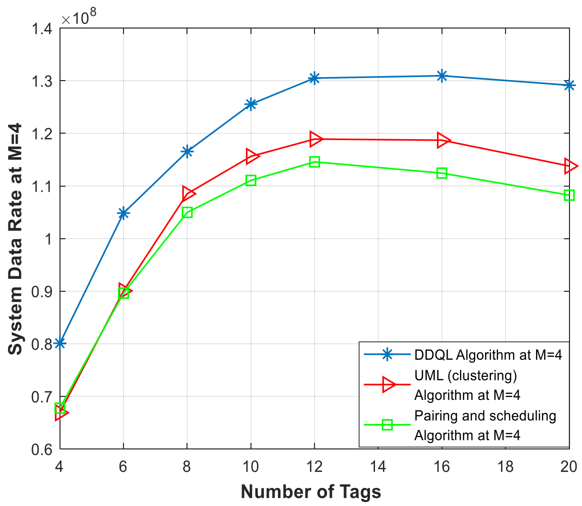

FIGURE 11. **The total system data rate vs. number of tags at $\mathcal{M} = 4$.**

The achieved data rate per tag at the required data rate of the IoT tag (10 Mb/s) is represented by different numbers of tags in Figure (11). We observe that the achieved rate is decreased by increasing the number of tags as a result of more mutual interference in each cluster which reduces the rate per tag in the cluster. Moreover, the achieved data rate for the number of smartphones $\mathcal{M} = 4$ succeeded in achieving the required rate for more tags than the system with smartphone number $\mathcal{M} = 2$ and it increased slightly by increasing the number of smartphones M.

VIII. CONCLUSION

In this paper, a framework is proposed to jointly optimize the smartphone mode selection of (LTE or Wi-Fi) network that acts as a relay for the backscattered information received from IoT tags. The mode selection (LTE or Wi-Fi) problem has been formulated as an optimization problem to maximize system throughput, where a many-to-one matching game is used to solve the first optimization problem phase. In addition, IoT tags clustering is optimized using the SR technique to backscatter on the neighboring smartphone. The IoT tags clustering problem was formulated as a second optimization problem phase to maximize the system throughput to solve this problem, and a DDQL algorithm was proposed to obtain the optimum cluster members. The simulation results show that the proposed scheme outperforms than the system without using DDQL and system using pairing and scheduling algorithm. Our proposed algorithm outperforms the system using the LTE network first without the DDQL algorithm on average by 90% and succeeded in increasing the system capacity on average by 100% above the system using the LTE network first without the DDQL algorithm at number of smartphones M = 4. Furthermore, the system performance is studied to find the optimum number of smartphones that is needed for the system where we find that the optimum number of smartphones required for this system is four.

REFERENCES

[1] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the Internet of Things: Communication technologies and challenges," *IEEE Access*, vol. 6, pp. 3619–3647, 2017.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[3] M. Chafii, F. Bader, and J. Palicot, "Enhancing coverage in narrow band-IoT using machine learning," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.

[4] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band Internet of Things," *IEEE Access*, vol. 5, pp. 20557–20577, 2017.

[5] H. Malik, H. Pervaiz, M. M. Alam, M. Y. Le, A. Kuusik, and M. A. Imran, ''Radio resource management scheme in NB-IoT systems,'' *IEEE Access*, vol. 6, pp. 15051–15064, 2018.

[6] Y. Zhao, W. Zhai, J. Zhao, T. Zhang, S. Sun, D. Niyato, and K.-Y. Lam, ''A comprehensive survey of 6G wireless communications,'' 2020, *arXiv:2101.03889*.

[7] B. Barakat, A. Taha, R. Samson, A. Steponenaite, S. Ansari, P. M. Langdon, I. J. Wassell, Q. H. Abbasi, M. A. Imran, and S. Keates, ''6G opportunities arising from Internet of Things use cases: A review paper,'' *Future Internet*, vol. 13, no. 6, p. 159, Jun. 2021.

[8] M. A. ElMossallamy, Z. Han, M. Pan, R. Jantti, K. G. Seddik, and G. Y. Li, ''Backscatter communications over ambient OFDM signals using null subcarriers,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[9] C. Boyer and S. Roy, ''Invited paper—Backscatter communication and RFID: Coding, energy, and MIMO analysis,'' *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 770–785, Mar. 2014.

[10] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, ''Symbiotic radio: Cognitive backscattering communications for future wireless networks,'' *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 4, pp. 1242–1255, Dec. 2020.

[11] A. E. Mostafa and V. W. S. Wong, ''Transmit or backscatter: Communication mode selection for narrowband IoT systems,'' *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 5477–5491, May 2022.

[12] D. Li, H. Zhang, and L. Fan, ''Adaptive mode selection for backscatter-assisted communication systems with opportunistic SIC,'' *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2327–2331, Feb. 2020.

[13] G. Li, X. Lu, and D. Niyato, ''A bandit approach for mode selection in ambient backscatter-assisted wireless-powered relaying,'' *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9190–9195, Aug. 2020.

[14] Y. Hu, P. Wang, Z. Lin, M. Ding, and Y.-C. Liang, ''Machine learning based signal detection for ambient backscatter communications,'' in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[15] C. Yang, X. Wang, and K.-W. Chin, ''On max–min throughput in backscatter-assisted wirelessly powered IoT,'' *IEEE Internet Things J.*, vol. 7, no. 1, pp. 137–147, Jan. 2020.

[16] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, ''Ambient backscatter communications: A contemporary survey,'' *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2889–2922, 4th Quart., 2018.

[17] G. Wang, F. Gao, R. Fan, and C. Tellambura, ''Ambient backscatter communication systems: Detection and performance analysis,'' *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4836–4846, Nov. 2016.

[18] Y. Zhuang, X. Li, H. Ji, H. Zhang, and V. C. M. Leung, ''Optimal resource allocation for RF-powered underlay cognitive radio networks with ambient backscatter communication,'' *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15216–15228, Dec. 2020.

[19] Y. Liu, P. Ren, and Q. Du, ''Symbiotic communication: Concurrent transmission for multi-users based on backscatter communication,'' in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2020, pp. 835–839.

[20] A. Raza, S. J. Nawaz, A. Ahmed, S. Wyne, B. Muhammad, M. N. Patwary, and R. Prasad, ''A NOMA-enabled cellular symbiotic radio for mMTC,'' *Wireless Pers. Commun.*, vol. 122, no. 4, pp. 3545–3571, Feb. 2022.

[21] H. Guo, Y.-C. Liang, R. Long, and Q. Zhang, ''Cooperative ambient backscatter system: A symbiotic radio paradigm for passive IoT,'' *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1191–1194, Aug. 2019.

[22] D. Li, ''Hybrid active and passive antenna selection for backscatter-assisted MISO systems,'' *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 7258–7269, Nov. 2020.

[23] D. Li, ''Two birds with one stone: Exploiting decode-and-forward relaying for opportunistic ambient backscattering,'' *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1405–1416, Mar. 2020.

[24] Q. Zhang and Y.-C. Liang, ''Signal detection for ambient backscatter communications using unsupervised learning,'' in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–6.

[25] A. B. Narasimhamurthy, M. K. Banavar, and C. Tepedelenlioglu, ''OFDM systems for wireless communications,'' *Synth. Lect. Algorithms Softw. Eng.*, vol. 2, no. 1, pp. 1–78, 2010.

[26] P. Poornima, G. Laxminarayana, and D. S. Rao, ''Performance analysis of channel capacity and throughput of LTE downlink system,'' *Int. J. Comput. Netw. Commun.*, vol. 9, no. 5, pp. 55–69, Sep. 2017.

[27] M. Anany, M. M. Elmesalawy, and A. M. A. El-Haleem, ''Matching game-based cell association in multi-RAT HetNet considering device requirements,'' *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9774–9782, Dec. 2019.

[28] H. Wu, S. Cheng, Y. Peng, K. Long, and J. Ma, ''IEEE 802.11 distributed coordination function (DCF): Analysis and enhancement,'' in *Proc. IEEE Int. Conf. Commun.*, vol. 1, Apr. 2002, pp. 605–609.

[29] N. A. Elmosilhy, M. M. Elmesalawy, and A. M. A. Elhaleem, ''User association with mode selection in LWA-based multi-RAT HetNet,'' *IEEE Access*, vol. 7, pp. 158623–158633, 2019.

[30] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, ''Symbiotic radio: A new communication paradigm for passive Internet of Things,'' *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1350–1363, Nov. 2020.

[31] D. Bharadia, K. R. Joshi, M. Kotaru, and S. Katti, ''BackFi: High throughput WiFi backscatter,'' *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, pp. 283–296, Oct. 2015.

[32] B. Kellogg, A. Parks, S. Gollakota, J. R. Smith, and D. Wetherall, ''Wi-Fi backscatter: Internet connectivity for RF-powered devices,'' in *Proc. ACM Conf. SIGCOMM*, Aug. 2014, pp. 607–618.

[33] S. Bayat, Y. Li, L. Song, and Z. Han, ''Matching theory: Applications in wireless communications,'' *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 103–122, Nov. 2016.

[34] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, ''Matching theory for future wireless networks: Fundamentals and applications,'' *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 52–59, May 2015.

[35] C. An and Y. Liu, ''A matching game algorithm for spectrum allocation based on POMDP model,'' in *Proc. 7th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2011, pp. 1–3.

[36] S. Metwaly, A. A. El-Haleem, and O. El-Ghandour, ''NOMA based matching game algorithm for narrowband Internet of Things (NB-IoT) system,'' *Ingénierie des Systèmes d Inf.*, vol. 25, no. 3, pp. 345–350, Jun. 2020.

[37] F. Tan, P. Yan, and X. Guan, ''Deep reinforcement learning: From *Q*-learning to deep *Q*-learning,'' in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2017, pp. 475–483.

[38] H. Hasselt, ''Double *Q*-learning,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 23, 2010, pp. 1–9.

[39] H. Van Hasselt, A. Guez, and D. Silver, ''Deep reinforcement learning with double *Q*-learning,'' in *Proc. AAAI Conf. Artif. Intell.*, 2016, vol. 30, no. 1, pp. 1–7.

[40] Y. Al-Eryani, M. Akrout, and E. Hossain, ''Multiple access in cell-free networks: Outage performance, dynamic clustering, and deep reinforcement learning-based design,'' *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1028–1042, Apr. 2021.

[41] Z. Qadir, K. N. Le, N. Saeed, and H. S. Munawar, ''Towards 6G Internet of Things: Recent advances, use cases, and open challenges,'' *ICT Exp.*, vol. 30, pp. 1–17, Jun. 2022.

GERGES MANSOUR SALAMA received the B.Sc. degree in electrical engineering and the M.Sc. degree in electronics and communications engineering from Minia University, Minya, Egypt, in 1999 and 2006, respectively, and the Ph.D. degree from the Faculty of Telecommunication Networks, Switching Systems, and Computer Technology (FTN, SS, and CT), St. Petersburg State University of Telecommunications (named after Prof. M. A. Bonch-Bruevich). He was with the Russian Federation Federal Communications Agency, Ministry of Communications and Mass Media, in 2012. He has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Minia University, where he is currently an Associate Professor with the Faculty of Engineering. His current research interests include image enhancement, image restoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications.

SAMAR SHAKER METWLY received the B.Sc. and M.Sc. degrees in electronics and communications engineering from Helwan University, Cairo, Egypt, in 2003 and 2021, respectively. She is currently pursuing the Ph.D. degree in electronics and communication engineering. Since 2021, she has been a Teaching Assistant with the Electronics and Communication Department, Canadian International College. Her current research interests include mobile communication, the Internet of Things (IoT), AI application in wireless communication, LTE–WLAN communication, unmanned aerial vehicle (UAV) communications, and 6G technologies.

E. G. SHEHATA received the B.Sc.Eng., M.Sc., and Ph.D. degrees in electrical engineering from the Faculty of Engineering, Minia University, Egypt, in 2002, 2007, and 2013, respectively. He is a member of the Academic Staff with the Department of Electrical Engineering, Minia University. Since 2018, he has been an Associate Professor with the Department of Electrical Engineering, Minia University, where he was an Assistant Professor, from 2013 to 2017. His research interests include control and design of electric machines for electric vehicles, renewable energy applications, dc micrigrid control, and wireless power transfer for charging electric vehicles.

AHMED M. ABD EL-HALEEM (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electronics and communications engineering from Helwan University, Cairo, Egypt, in 2001, 2006, and 2012, respectively. He is currently an Associate Professor with the Electronics and Communications Engineering Department, Faculty of Engineering, Helwan University. He is a member of a research team that awards several applied research projects funded by national and international funding agencies in the field of wireless communication, the IoT and its applications, and smart education systems. His current research interests include mobile/vehicular ad-hoc communication networks, 5G and 6G radio access networks, cognitive radio networking, device-to-device communication, the Internet of Things (IoT), reconfigurable intelligent surface (RIS), and AI application in wireless communication. This includes mobility management techniques and routing schemes for mobile ad-hoc networks (MANET), security and secure routing algorithms for ad-hoc wireless networks, and routing protocols for cognitive radio communication networks. This also includes AI-based applications for education and wireless networks, the Internet of Things (IoT) communications, unmanned aerial vehicle (UAV) communications, reconfigurable intelligent surface (RIS), and smart grid communication systems.