A Novel Method for Assigning Joint Power Spectrum and Power Selection in Device to Device Networks to Improve Performance

Anahita Jabbari[1], S. Mahmood Daneshvar Farzanegan[2]
1-Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: mail.jabbari@gmail.com
2-Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: smdaneshvar@gmail.com (Corresponding author)

Received: May 2019 Revised: June 2019 Accepted: July 2019

ABSTRACT:
Optimal utilization of frequency spectrum in wireless networks particularly in device to device communication is of significant importance owing to the growing demand. Traditional methods to optimal spectrum utilization of spectrum are not sufficiently efficient and result in loss of spectrum. Recently, application of Cognetive radio is suggested to solve this problem. Cognetive radio is a smart wireless system which is aware of the spectral traffic condition of its environment in an instantaneous way and through these spectral conditions, changes the power of transmitter and the type of modulation and it adapts to the environment. The main purpose of this paper is to investigate the problem of spectral sharing. Today, communication systems suffer from main problems including limited bandwidth, download speed increase, rate increase and saving in transmitted power. To solve such problems, new methods based on machine learning in spectrum sharing are necessary to overcome such challenges. In this work, using cellular learner automata, a method is proposed for simultaneous assigning of spectrum and resource. The aim of each pair of transmission is to transmit in an appropriate channel and power level so that it can maximize its compensation in cellular learner automata. In these scenarios, compensation is taken as the difference between operational (collective) and consumed power. The cost of the consumed power is the signal to interference noise ration. Proposed method is simulated on a LTE-A network as well as an NS2. Proposed algorithm is of rapid convergence and semi-optimal efficiency in low repetitions.

KEYWORDS: Frequency Spectrum, Cognitive Radio, Machine Learning, Device To Device Pair, Cellular Automata.

1. INTRODUCTION

In current wireless networks, frequency spectrum is assigned in a constant manner. I other words, overall frequency band is managed through a certain organization and each part is considered for a specific service. Despite of minimization of the interference between communication systems with such method of assignment, due to growing demand for utilization of frequency bands, it seems that the policy of constantly assigning frequency spectrum is no longer efficient to meet growing demands. Cognitive radio is in fact a smart wireless network which is instantaneously aware of the spectral conditions of its environment and is capable of adapting to its environment by means of changing its internal parameters such as transmitter power and the type of modulation through its environmental conditions. based on the definition of the Cognitive radio, in those networks, data transfer is generally performed based on spectrum sensing, spectrum management, spectrum sharing, and spectrum mobility. In recent years, each of the above items are evaluated from different aspects and numerous papers are published. However, no final conclusion is made and researches are ongoing. Overall contribution of this research is that D2D pairs work automatically in homogeneous cellular networks which are composed of many BSs in which no data exchange is performed. Therefore, no knowledge is available about the quality of channels. Moreover, it is supposed that spectra of BS include orthogonal channels which can overlap with each other and hence, it can increase the inter–cell interference among users during transmission in a channel.

D2D communication is a direct link between two users without any communication with BS of network center and hub. One of the benefits of D2D access is improved quality of services (QoS). However, in D2D communication, a twofold challenge is present:

1. Reduced overlap between users working in the same frequency band

2. Optimal management of the spectrum so that the assigned spectrum is not lost and there is no spectral loss.

Appropriate interference management and resource assignment schema can improve the efficiency of the cellular networks. This mater is defined as the efficiency of the spectrum, cellular coverage, network throughput power and user experience. Most of the methods presented in this context [1] consider the network as an integrated system in which resources are assigned by BSs. BSs assigning the resources have sufficient knowledge about the environment and the state of the CSI network though regular basic signals in cellular communication cannot be used for estimation of the D2D channels [2]. As a result, many of researchers discussed the strategies of assigning resources in which D2D users utilize secondary networks and they are allowed to use empty bands as well. Notably, methods proposed in [3] and [4], collaboration of cellular users, BSs as well as D2D pairs with respect to the data exchange including D2D links and number of users are for efficient sharing of the spectrum.

In [10], a new method based on the learning theory is presented. The issue of reinforced learning is not presented in the context of access to various spectra whose purpose is to decrease the interference between overlapping networks. The problem of reduced interference is addressed as a precise potential graphical game with pure NE strategy which seems that is cannot be implemented in real conditions.

In this research, two scenarios of D2D function are taken into account. In first one, D2D users are out of a certain resource. Therefore, there is no interference between cellular and D2D users in various frequencies. I the second one, D2D pairs exchange data in shared network of cellular users. The main purpose of each pair is to select the appropriate wireless channel as well as power level so that the compensation can be optimized. Compensation is the difference between the obtained throughput as well as the consumption of the power which is limited by the minimum value of SINR in each network application. This problem is taken as the purely random non-collaborative game. Moreover, it is considered as the learner cellular automata in a purely random space. In addition, in such space, each D2D pair functions completely independent and it has no information about its environment. In this space, each D2D user acts as a learning cell whose aim is to learn the best state leading to NE. this strategy is based on this idea that each learner cell in LCA must at so that it can be compensated from its environment and function and make updates based on its compensation.

Selection of channel is based on learning process as well as cellular automata. Selection of the communication channel is done using received compensations as well as lack of compensation. Performance of each cell which includes a D2D pair is done independent from other cells or D2D pairs. This action is taken for selection of the transmission power level as well and as a result, complexity of the method decreases.

In the performance of cellular automata, each time, D2D pair selects an action which can be random or based on the previous observations. Adjacent cells are either the other D2D users or the cellular users. Based on the action selected by other cells or the rules governing this automata, selected action is either compensated or fined. Based on the compensation, cellular leaner automata modifies its behavior and update is done on the structure. After updating, each automata cell which is the D2D user selects another action. This selection continues until spectrum sharing achieves a steady state.

Rest of this paper is organized as follows: in section 2, cellular automata is described. In section 3, model of network is explained. In section 4, proposed method of spectrum sharing based on learning in D2D is presented. Section 5 evaluates the proposed method and finally, in section 6, overall conclusion of the paper is provided

2. CELLULAR AUTOMATA

Cellular automata are mathematical methods which can be used for calculation and simulation of systems. Cellular automata are simple discrete systems which can have complex behaviors and calculations trough simple and local rules. Locality means that adjacent cells contribute to in determination of the new value for cell and farther cells have no contribution. Each cell has its own set of states and in each time, it decides about its own and neighboring cells' state. Rules of state change for cellular automata is constant during the process and remains unchanged. Cellular network can take various dimensions and can have one or more dimensions. According to the values which can be taken by cells, cellular automata are divided into binary and multiple-value ones. Understanding the behavior of cellular automata by their rules is very difficult and requires simulation. One of the problems in using cellular automata is to design rules for our intended action. There are different rules for updating the cells leading to different types of cellular automata [5]. For instance, rules can be expressed as definite or contingent bringing about definite or contingent cellular automata, respectively.

Learner automata is an item designed for contingent and uncertain environment. This machine can perform finite actions. Each of the learner automata has a vector of these probabilities. This vector shows the probability of each action. Sum of the vector arrays are equal. each action taken by automata and each selected action is evaluated and the result is given to the automata in the form of a positive or negative signal and it is affected by this result for subsequent selection [6].

The goal is to select the best action among a set of actions. The best action is the one which maximize the compensation from the environment. Performance of the learner automata in interaction with environment is shown in Fig. 1 [7].

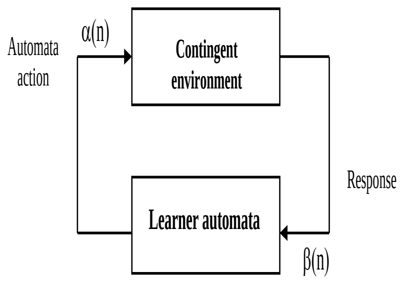

Fig. 1. Relationship between learner automata and environment [7].

Environment can be shown as the triplet $E \equiv \{\alpha, \beta, c\}$ in which:

- $\alpha \equiv \{\alpha_1, \alpha_2, ..., \alpha_r\}$ is set of inputs
- $\beta \equiv \{\beta_1, \beta_2, ..., \beta_m\}$ is the set of outputs
- $c \equiv \{c_1, c_2, ..., c_r\}$ is the set of fines.

If β has two members, environment is of type P. in such conditions, $\beta_1=1$ is taken as compensation and $\beta_1=0$ is taken as fine. In type Q environment. B can take any value in [0,1] interval and in type S environment, it takes a random value in this interval. The parameter c_1 is the probability that action α has unsuitable consequence. In static environment, the value of c_1 remains unchanged while in non-static environment, these values change with time.

Learner automata with constant structure: is represented by quintet $\{\alpha, \beta, F, G, \phi\}$ in which $\alpha \equiv \{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ is the set of automata actions, , $\beta \equiv \{\beta_1, \beta_2, \cdots, \beta_m\}$ is the set of inputs, is the set of internal conditions, $\phi \equiv \{\phi_1, \phi_2, \cdots, \phi_s\}$ is the function of the new state of automata $F: \phi \times \beta \rightarrow \phi$ and $G: \phi \rightarrow \alpha$ is the output function which maps the current state to the new one.

Learner automata with changing structure: can be represented by quartet $\{\alpha, \beta, p, T\}$ in which $\alpha = \{\alpha_1, ..., \alpha_r\}$ is the set of automata actions, $\beta = \{\beta_1, ..., \beta_m\}$ is the set of automata inputs and $p = \{p_1, .., p_r\}$ is the vector of the probability of selecting each of the actions. Learning algorithm is given by $p(n+1) = T[\alpha(n), \beta(n), p(n)]$. Following algorithm is a linear learning algorithm. Suppose action α_i is selected in n^{th} stage.

- Suitable response

$$\begin{aligned} p_i(n+1) &= p_i(n) + a[1 - p_i(n)] \\ p_j(n+1) &= (1-a)p_j(n) \qquad \forall j \quad j \neq i \end{aligned} \tag{1}$$

- Unsuitable response

$$\begin{aligned} p_i(n+1) &= (1-b)p_i(n) \\ p_j(n+1) &= (b/r - 1) + (1-b)p_j(n) \quad \forall j \quad j \neq i \end{aligned} \tag{2}$$

In Eq. 1 and 2, compensation parameter is a and penalty parameter is b. according to these parameters, three cases can be considered:

- When a and b are equal, algorithm is L_{RP}.
- When a is smaller than b, algorithm is $L_{R\varepsilon P}$.
- When b is equal to zero, algorithm is L_{RI}.

This probability updating is so that sum of probabilities is equal to unity [8].

A solely learner automata is not so efficient. If a lot of learner automata come in the vicinity and collaboration, they can solve difficult problems. As stated earlier, it is difficult to design constant rules for cellular automata and without simulation, imagination of the behavior of cellular automata is very difficult. Combination of the cellular automata and learner one can somehow solve the problem.

Tanking into account this problem and shortcomings of the cellular automata, a new model is created through combination of two models and is named as the cellular learner automata. In what follows, definition of the cellular learner automata is provided.

- Definition 1: d-dimensional cellular learner automata is a polynomial as $CLA = (Z^d, \phi, A, N, F)$ so that:
- Z^d is a network of d integers. This network can be a finite, semi-infinite or infinite network.
- φ is a finite set of states.
- A is the set of learner automata each of which is assigned to each cell of the cellular automata.
- $N = \{\bar{x}_1, ..., \bar{x}_{\bar{m}}\}$ is a finite set of Z^d which is called neighborhood vector.
- $F: \underline{\phi}^{\bar{m}} \rightarrow \underline{\beta}$ is the local rule of CLA so that β is the set of values which can be accepted as the reinforcing signal.

3. PROPOSED METHOD

In this paper, the main concentration is on the performance of down link in a homogeneous cellular network. Proposed method is capable of simulation in up-link as well. Fig. 2 illustrates the homogeneous cellular network of this work. In this figure, there are N BS bases represented by BS_1 through BS_N. Therefore, N={1…N}. each BS of the network can have high power level of macro-cell service providing or with low power level for microcell. Furthermore, it is assumed that they are permitted to transmit in the spectrum. Hence, they are assumed as the primary user of the spectrum which can overlap with each other. BSs provide services to M D2D pairs as well as L cellular users. For application of mathematical elements, even D2D users are enumerated: PU_{M+1} through PU_{M+N}. the prime in their indices refers to the opposing user pair. Cellular users are shown by U as $L = \{M + N + 1, \dots, M + N + L\}$. Cellular users are shown by L.

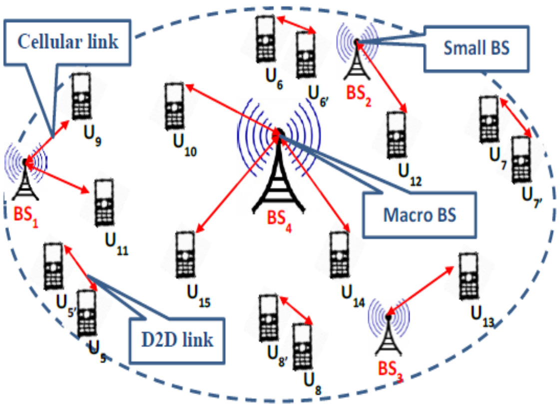

Fig. 2. Homogeneous cellular network with N=4 BS bases, number of pairs: M=4 and L=7 cellular users.

Described system is simulated by non-overlapping time divisions. There are t time divisions which are as much as T_s apart. It is assumed that in this system, communication is made in each time division and D2D users are synchronized using transmitted temporal signals of D2D and or GPS data. Wireless networks are assigned to cellular users based on BSD and this assignment is based on the predetermed program. However, contrary to the cellular users, D2D ones act completely automatically and they have no data exchange with BS and themselves. As a result, D2D users have no precise data about the performance of their environment and hence, they select their transmission power independently and based on their local observations. Moreover, in this network, it is assumed that users select their power level and their channel in the onset of the time division. In addition, there is another assumption in the network based on which cellular and D2D users remain independent from each other and network leaving is done independently. There is no certainty in the time interval of their presence in network and channel.

In this paper, two scenarios are considered for D2D communications:

- In the first scenario, cellular users and D2D pairs are active in various spectra and frequency bands.
- In the second scenario, cellular users and D2D pairs are active in the same channel.

There are K orthogonal channels which are reperesnted by C_1-C_k. these channels are accessible for D2D users. That is, K={1,.. k} is the index of channels corresponding to the D2D pairs.

In the first scenario, D2D pairs exchange data solely in these K channels and cellular users have their specific channels shown by K^c. therefore, $K \cap K^c = \emptyset$ and overall number of bandwidths of all channels is composed of $K \cup K^c$.

In the second scenario, cellular users and D2D pairs exchange data in K channels which is the overall number of available spectra of the network. For each user, PU_m and U_m a binary assignment vector is defined as $C^m t = (C^m 1(t) \dots C^m k(t)$ in which $m \in M \cup L$ whose elements are either zero or one. If a cellular user or a pair user send data in channel K, C_k=1 and otherwise, it is equal to zero. For each cellular user, a binary vector of BS assignment is considered as $b_t^l = \{b_1^l(t), \dots, b_N^l(t)\}$ and $\in L$. if BS is linked to user U_t, $b_n^l(t) = 1$, otherwise, it is equal to zero.

In the network studied in this paper, simultaneous function of D2D pairs is not limited in similar channels though in time division t, each D2D user can select at most one channel in the bandwidth. That is:

$$\sum_{k=K} C_k^m(t) \leq 1 \quad \forall m \in M \tag{3}$$

In the first scenario, all channels in set K are considered for providing services to D2D pairs. Therefore, we have

$$C_k^l(t) = 0 \quad \forall l \in L, \forall k \in K \tag{4}$$

In each time division, a finite set of channels for each PU_m is defined by Eq. (5)

$$C^m = \{C_t^m \mid \sum_{k=K} C_k^m(t) \leq 1 \ \}, \ \forall m \in M \tag{5}$$

3.1. Channel Model

If $k \in K, m \in M, j \in M \cup L$ $G_{m,j}^k(t)$ is the gain of the channel between U_m and U_j is in channel C_k and time division t and $G_{m,j}^k(t)$ is the gain of the link between base BS_n and U_j in the same channel and time division, instantaneous value of $G_{m,j}^k(t)$ can be measured by all cellular users and BSs for all $n \in N, l \in L, k \in K$ during base signal.

On the other hand, D2D users have no information about channel quality. Therefore, the value of $G_{m,j}^k(t)$ and $G_{n,j}^k(t)$ is unknown for D2D pairs.

In the first scenario, interference between each D2D pair is made by other D2Ds during operation. As a result, SINR value for each S2D user, that is, PU_m in channel C_k is given by[12]

$$SINR_k^m(t) = \frac{S_k^m(t)}{I_{m.k}^D(t)+\sigma^2}, \forall m \in M, \forall k \in K \quad (6)$$

In Eq. 6, σ^2 is the variance of the accumulative white noise and $S_k^m(t)$ is the useful power of PU_m in channel C_k and time division t which is calculated as follows[12]

$$S_k^m(t) = G_{m,m'}^k(t)C_k^m(t)P^m(t) \quad (7)$$

In above equation, $P^m(t) \leq P_{max}$ is the transmitted power. It must be noted that this value must not exceed a maximum value which is yielded by PU_m in time division. $I_{m,k}^D(t)$ is the interference between D2D users in channel C_k which is defined by Eq. 8[12]

$$I_{m,k}^D(t) = \sum_{j\in M\setminus\{m\}} G_{j,m}^k(t)c_k^j(t)P^j(t) \quad (8)$$

On the other hand, in second scenario, interference is made between D2D pairs and other D2Ds as well as cellular users. The value of SINR for D2D users, that is, PU_m in wireless channel C_k is given by

$$SINR_k^m(t) = \frac{S_k^m(t)}{I_{m.k}^D(t) + I_{m.k}^C(t) + \sigma^2}, \forall m \in M, \forall k \in K \quad (9)$$

In each time division with $I_{m.k}^C(t)$ value, interference is between PU_m and cellular users in channel C_k which is given by Eq. 10

$$I_{m.k}^D(t) = \sum_{i\in N}\sum_{j\in L} G_{i,m'}^k(t)b_i^j(t)c_k^j(t)\,P^j(t) \quad (10)$$

where, $P^j(t)$ is the instantaneous transmission power in DL state of the channel between cellular user U_j and base BS. For compatibility with the proposed method based on cellular automata, the level of transmitted power is quantized for each D2D pair. For each user pair, there are J power levels which are represented by $P_1 \dots P_j$. This level of quantized power is give by binary vector as $P_t^m = \left(p_1^m(t) \dots p_j^m(t)\right)$. $p_j^m(t) = 1$ is selected for transmitted power P_j in time division t and it will be zero for other cases. Because only one power level can be selected, Eq. 11 is defined[13]

$$\sum_{j=J} P_j^m(t) \leq l. \forall m \in M \quad (11)$$

In above relationship, instantaneous transmission power in each D2D user; PU_m is given as:

$$P^m(t) = \frac{P_{max}}{J} \sum_{j=J} P_j^m(t), \forall m \in \quad (12)$$

Moreover, the set of power levels selected by D2D pairs is given by

$$P^m = \left\{P^m \middle| \sum_{j=J} P_j^m(t) \leq 1\right\}, \forall m \in M \quad (13)$$

3.2. Problem conditions

He aim of the present research is to propose an automatic schema for simultaneous selection of power level and channel for D2D users so that each SINR of the intended pair is less than threshold SINR.

$$SINR^m(t) = \sum_{k\in K} SINR_k^m(t) \geq SINR_{min}^m, \forall m \in M \quad (14)$$

In Eq. 14, $SINR_{min}^m$ is the minimum value of the acceptable level in each PU_m. if in tie division t, Eq. 14 holds, this pair is compensated as much as u_t^m. Compensation is the difference between the group and the power cost. If Eq. 14 doesn't hold, pair will not be compensated and will be fined. Since each cellular and D2D user are in the same automata cell, the process of cellular automata will be updated. The compensation u_t^m received by the cell (D2D user) is calculated as follows

$$u_t^m = R^m(t) - v_0^m P^m(t) if SINR^m(t) \geq SINR_{min}^m \quad (15)$$

Otherwise:

In Eq. 15, cost of each unit (watt); $R^m(t)$, will be given by Eq. 16

$$R^m(t) = \omega log\left(1 + SINR^m(t)\right) = \omega log(1 + \sum_{k\in K} SINR_k^m(t)) \quad (16)$$

In Eq. 16, ω is the bandwidth of the channel and the transmitted power is given by Eq. 5b. it must be noticed that in all time divisions, the value of instantaneous compensation; PU_m in Eq. 8a depends upon the following conditions:

The value of power and the channel selected in pair PU_m, that is, $C^m t$ and $P^m t$ as well as the link gain for channel; $G_{m,m'}^k(t)$. These figures can be calculated for D2D pair or is predetermined.

Channels as well as the power level selected by other users and the link gain of the intended channel which are not observed by other D2Ds.

In te network studied in this paper, in each time division, each D2D user; PU_m, selects its channel and power level in a way that the compensation of the learner cellular automata is maximized. Moreover, the sum of received compensations and penalties must be a finite value.

The main assumption is that all D2D users try to maximize their compensation and minimize their penalties.

For simplicity, it is assumed that the value of discretization rate is along with each compensation which limits the overall compensation.

$$U_t^m = \sum_{\tau=t}^{+\infty} \gamma^{t-\tau} u_\tau^m \quad (17)$$

In order to use the model of network in cellular learner automata, A^m matrix is defined which includes the set of possible actions selected by each D2D user in each cell as $A^m = C^m \times P^m$. The aim of each PU_m user pair is to select the $\bar{C}_t^m, \bar{P}_t^m$ belonging to A^m. in each time division t, in long term, overall compensation is maximized according to Eq. 11.

$$(\bar{C}_t^m, \bar{P}_t^m) = argmax_{(C_t^m, P_t^m) \in A^m} U_t^m \quad (18)$$

That is, $\boldsymbol{\bar{C}_t^m}.\boldsymbol{\bar{P}_t^m}$ must be selected so that Eq. 18 is maximized in a limited way.

4. RESULTS

Now, the problem is the optimal and simultaneous selection of power level and the intended channel with m cells in each of which there is a D2D and each cell is not aware of the conditions of neighboring cells. conditions of all cells include the quality of selected channel as well as the intended power level. It is assumed that all cells operate independently and select their action in each time division. Each cell having D2D user selects its power level and channel independently and tries to maximize its compensation according to Eq. 17. Now, the problem space for each cell can be a set of decisions in A^m. Each user pair performs its selected action in corresponding automata; that is, $a^m\mathrm{t} = (C^m t, P^m\mathrm{t}) \in A^m$ which is the selection of channel as well as power level in time division t. for all cells including all users, it time division t, we have

$$a^{-m}\mathrm{t} = (a^1 t, a^{m-1}\mathrm{t}, a^{m+1}\mathrm{t}, \dots, a^m\mathrm{t})$$

Where, $A^{-m} = X_{i \in M \backslash \{m\}} A^i$. In intended network, the value of throughput, $R^m(t)$ for each cell and PU_m users depends upon the value of SINR in intended channel which determines the selected action of the user in learner automata. That is, $a^m\mathrm{t}, a^{-m}\mathrm{t}$ provides the instantaneous values of the link gain for matrix G^mt which is given in Eq. 13.[15]

$$G^m t = \begin{bmatrix} G_{1,m'}^1(t) & \dots & G_{1,m'}^1(t) \\ \vdots & \vdots & \vdots \\ G_{N+M,m'}^1(t) & \dots & G_{N+M,m'}^1(t) \end{bmatrix} \quad (19)$$

Indeed, instantaneous value of SINR for each D2D pair; $SINR^m(t)$ in first and second scenario for all m values is defined as follows

$$SINR^m(a_t^m, a_t^{-m}, G_t^m) = \sum_{k \in K} \frac{S_k^m(a_t^m, G_t^m)}{I_{m,k}^D(a_t^{-m}, G_t^m) + \sigma^2} \quad (20)$$

and in the second scenario

(21)

$$SINR^m(a_t^m, a_t^{-m}, G_t^m) = \sum_{k \in K} \frac{S_k^m(a_t^m, G_t^m)}{I_{m,k}^D(a_t^{-m}, G_t^m) + I_{m,k}^c + \sigma^2} v$$

In above relations, values of $S_k^m(t)$ and $I_{m,k}^D$ are explicitly a function of a_t^m ، a_t^{-m} and G_t^m in each time division. It must be remembered that in each time division, each cell can mesaure its own SINR.
The value of instantaneous compensation of each D2D user pair will be zero after the operation if the measured is less that the threshold value of Eq. 16.
Accordingly, in each time division, following state can be defined for each user in the cell[16]

(22)

$$s_t^m = \begin{cases} 1, & if SINR^m\left(a_t^m, a_t^{-m}, G_t^m\right) \geq SINR_{min}^m, \forall m \in M \\ 0 & otherwise \end{cases}$$

This state is completely observable since it depends upon the $SINR^m(t)$ values and is measured by each time division. Using Eq. 22, instantaneous compensation received by each cell in each time division will be given by

(23)

$$u_t^m = u^m(a_t^m, a_t^{-m}, S_t^m) = S_t^m[R^m(a_t^m, a_t^{-m}, G_t^m) - v^m P^m(a_t^m)]$$

For all m values and

(24)

$$R^m(a_t^m, a_t^{-m}, G_t^m) = \omega \log(1 + SINR^m(a_t^m, a_t^{-m}, G_t^m)$$

And according to the intended scenario, the value of $SINR^m(a_t^m, a_t^{-m}, G_t^m)$ is selected. After receiving the compensation based on selection of the cell, it is time for updating. That is,each cell in which the intended user is present must perform the S_t^m update so that the final compensation od Eq. 17 is maximized. Learning algorithm used in this research is a learner algorithm with changing structure. Practically, updating process takes place in following time divisions.

In this state of network, D2D user; that is, the cell of cellular automata selects another pair from a_{t+1}^m, C_{t+1}^m and hence, the state s_{t+1}^m is created randomly from selected A_m. in this case, condition S_{t+1} is completely dependent on the state a_{t+1}^m, a_{t+1}^{-m}. In fact, state change and performed action are unknown to the user. Consequently, problem space for cellular learning automata is completely random and the problem can be formulated as follows

- There are a limited number of cases; that is, S which is the state space is limited and for all $m \in M$, S=$X_{m \in M} S^m, with\ S^m = \{0,1\}$.
- There are a limited number of cellular automata.
- There are a limited number of actions taken by each D2D user.

- Probability transfer function $T(S^m, a, S'^m)$ includes all probabilistic functions of the following states. That is, $s'^m = m_{t+1}^m \in S^m$ for all common activities $a = (a_t^1, \dots, a_t^m) \in A$ in s^m states.
- Compensation vector for cell in performed action for the rest of the process as well as the training of learner automata and proportionality to the NE problem holds the Eq. 25 for all $s^m \in S^m$ cells.

$$u^m(\bar{a}^m, \bar{a}^{-m}, S^m) \geq u^m(a^m, \bar{a}^{-m}, S^m), \forall a^m \in A^M \tag{25}$$

That is, in Ne case, action of each cell which includes D2D user is the best response compared to the other cells.

Precise and correct estimation of the efficiency of a network is a vital issue. For this reason, selection and design of the efficiency parameter and evaluation functions is an important issue in D2D networks. Due to the dynamic nature of the networks, designers of the smart networks try to get a better understanding about the internal relationship between goals, efficiency parameters, evaluation functions, efficiency of networks, links and operational environments. Management of level and spectrum efficiency is along with the optimization of the methods so that the network can increase its capacity, decrease the delay and improve the reliability of the network regardless of the accessible bandwidth and the failure in transmission.

Simulation of the proposed method is performed in NS2 programming system under Windows 10 OS. Hardware used for simulation is an Intel® Core™ i5-8500 processor with 16GB ram reserved from SSD memory.

The aim of the proposed method is simultaneous assignment of the optimal spectrum and power to D2D users in a homogeneous cellular network using cellular automata. In simulated model, N=3 (number of Bss) in which BS1 is in Pico, BS2 is in macro and BS3 is in micro scale. the pattern of the used antenna is explained in [14]. In the first scenario, overall bandwidth of the network is composed of the unit channels $K \cup k^c = \{1 - 50\}$ including five blocks of LTE source. Among available BSs, k={1-10} and k={26-40} are used for communication of D2D and $k = \{11 - 25\}$ and $k = \{41 - 50\}$ are used for cellular users. In the second scenario, all Ks can be used for communication of the data through D2D or cellular users. These spectra are summarized in table 1 and 2 for first and second scenario, respectively. In this table, resources blocks are proportional to the cellular users of C_k and K_n channels which are the authorized channels and can be used for D2Ds. K_n^c represents the channels belonging to BS_n which are reserved for cellular users. Intended bandwidth is as much as 180kHz.

BSs provide services to a set of D2Ds and cellular users in a random way. Moreover, it is supposed that assignment to cellular users of BSs is based on a precise value of CSI. In all simulations, number of cellular users is L=100. All devices have external operation and simulation environment is a common one. Device of each user has its own traffic and is able to make any type of traffic. For simplicity, it is assumed that traffic of each user is modeled as a complete buffer and each user transmits 10 pockets per second. furthermore, each pocket is 1500 bytes. Minimum SINR for both cellular and D2D users; $m \in MUL$, $SINR_{min}^m = SINR_{min} = 0$ is the number of power levels; J=10 while maximum allowable power level is $P_{max} = 23dB_m$. In the second scenario, maximum allowable power level is based on Eq. 36 and $P_{max} = 20dB_m - 10log$. In addition, main simulated parameters are given in table 3.

Table 1. Frequency bands in BSs of the first scenario.

$\boldsymbol{C_k, k \in K_n^c}$ Cellular chennals	$\boldsymbol{C_k, k \in K}$ ۃ D2D chnnels	$\boldsymbol{\lvert K_n^c \cup K_n \rvert}$ bandwidth	BS No.
$K_1^c = \{11, \dots, 25\}$	$K_1 = \{1, \dots, 10\}$	$25RB_s(5\ MHz)$	BS_1
$K_2^c = \{11, \dots, 25, 41, \dots, 50\}$	$K_2 = \{1, \dots, 10, 26, \dots, 40\}$	$50RB_s(10\ MHz)$	BS_2
$K_3^c = \{41, \dots, 50\}$	$K_3 = \{26, \dots, 40\}$	$25RB_s(5\ MHz)$	BS_3

Table 2. Frequency bands in BSs of the second scenario.

$\boldsymbol{C_k, k \in K}$ ۃ **D2D Common**	$\boldsymbol{K_n}$ **bandwitdth**	**BS$_n$ No.**
$K_1 = \{1, \dots, 25\}$	$25RB_s(5\ MHz)$	BS_1
$K_2 = \{1, \dots 50\}$	$50RB_s(10\ MHz)$	BS_2
$K_3 = \{26, \dots, 50\}$	$25RB_s(5\ MHz)$	BS_3

Table 3. Simulated parameters in LTE-A model.

Value	Parameter
TDD	Frame structure
1 msec	T_s Time divisions
0	TDD Partitions
46dBm	Trnasmission power of eNode
23 dBm	Maximum trnamission power
-74 dBm	noise power
$128 + 37.6log\|d\|$ 7m	Loss of cellular link
$40log\|d\| + 30log\|f\| + 4g$	Loss of D2D link

Proposed method is compared to the methods presented in [15]. The method of that reference paper is based on the reinforced learning theory. Action – value updating in the method is performed using reinforced learning while updating in the proposed method is based on cellular learner automata.

For further comparison, proposed method is compared to URS [16] as well. In tis method, the value of intended pairs is selected in uniform interval and time division.

Fig. 3-6 illustrate the results of simulation in first scenario with $\gamma = 0.5$ and $v^m = v = 1$ for all $m \in M$. Moreover, updating model of cellular automata is LRI with $a = b = 0{,}09$. Average compensation for D2D pairs is $v_t = \frac{1}{M}\sum_{m\in M} u_t^m$.

Based on the number of repetitions for constant number of D2D users M=100, Fig. 4 represents the difference between maximum instantaneous compensation and the average compensation value. Output value is based on constant number of users in 50 time divisions.

Fig. 5 shows the average convergence time for algorithms. URS failed to converge. Output is based on D2D pairs. As can be seen, proposed method converges faster than the control method.

It is notable that average instantaneous compensation of proposed method is higher that u_t^{max} and proportional to the maximum throughput power as well as the transmission cost. In other words

$$u_t^{max} = R_t^{max} - \frac{vP_{max}}{J} \approx 120 - 2 = 118 \qquad (26)$$

Hence, the optimal value in figures 1-4 through 3-4 has low convergence speed.

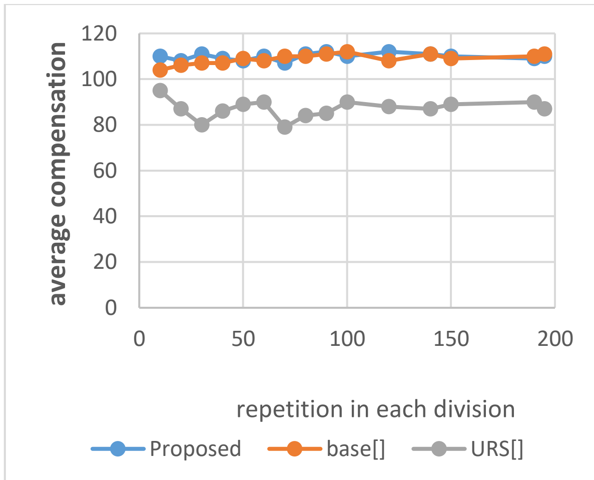

Fig. 3. Average instantaneous compensation U_t versus algorithm repetition for M=100.

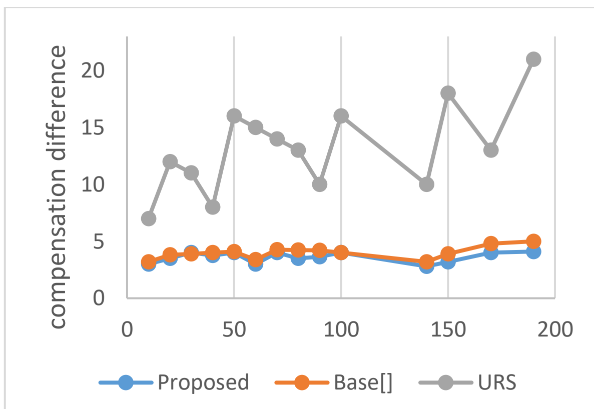

Fig. 4. Compensation difference as a function of the number of D2D users after 50 repetitions, M=100.

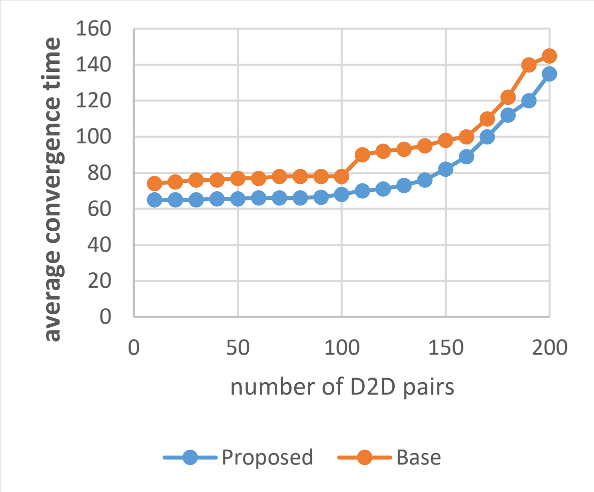

Fig. 5. Average number of repetitions required for convergence of algorithm, M=100.

Figures 6 and 7 show the efficiency of the network in second scenario by level of minimum SINR. In this relationship, ${R_t}^D$ and R_t^C are given in $SINR_{min}$ function. There are 50 algorithm repetitions. Number of D2D users is M=100.

As can be seen from diagrams, throughput power for cellular and D2D users is explained by concave functions of $SINR_{min}$. Moreover, maximum level depends upon the selected channel and the way these channels are selected. Minimum value is obtained using URS method. These results suggest that slight adjustments for $SINR_{min}$ leads to decreased throughput of users which is expected because channel conditions are not suitable. when $SINR_{min}$ is very high, throughput value will decrease as a result of lack of spectrum and bandwidth as well as the appropriate channels since they fail to meet all requirements.

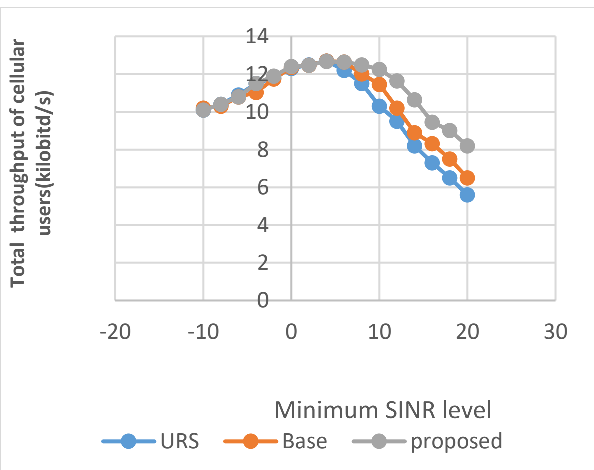

Fig. 6. Throughput of cellular users in second scenario as a function of $SINR_{min}$ for 50v repetitions of algorithm in M=100.

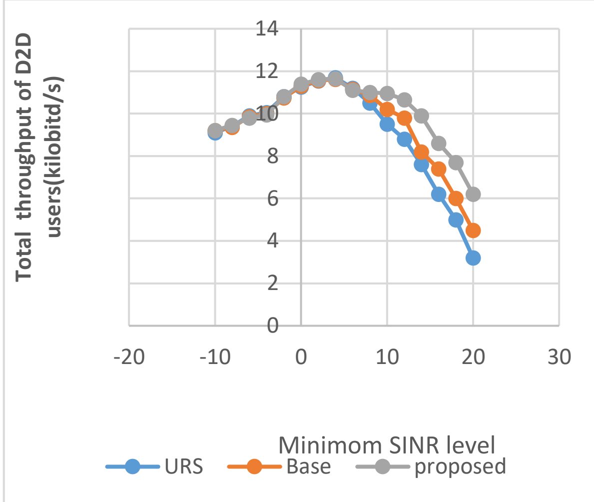

Fig. 7. Throughput of D2D users in second scenario as a function of $SINR_{min}$ for 50v repetitions of algorithm in M=100.

5. CONCLUSION

In this paper, a new method based on cellular learning automata in random space is presented for selection of the best power level and channel by D2D users in cellular networks. As can be seen from results of two scenarios, proposed method which is based on the cellular learner automata in which ICI is taken into account yields better results. Further, in this research, the problem of distributed resources assignment for D2D pairs is simulated as cellular learner automata in random and independent space while no cell is aware of the function of the adjacent one and power level and channel selection is done independently and based on the relationships given for compensation and penalty. Therefore, this method can be semi-optimal owing to fast convergence. However, throughput of the second scenario decreases as the number of users increases. But, if the number of users is constant (M=100), $SINR_{min}$ function acts as a concave function and hence, according to the results, it is possible to implement and use the method in real situations

REFERENCES

[1] A. Asheralieva and Y. Miyanaga, "**Dynamic buffer status-based control for LTE-A network with underlay D2D communication**," *IEEE transactions on communications*, Vol. 64, No. 3, pp. 1342-1355, 2016.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "**NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey**," *Computer networks*, Vol. 50, No. 13, pp. 2127-2159, 2006.

[3] W. Y. Lee, "**Spectrum management in cognitive radio wireless networks**," Georgia Institute of Technology, 2009.

[4] A. Vosoughi, J. R. Cavallaro, and A. Marshall, "R**obust consensus-based cooperative spectrum sensing under insistent spectrum sensing data falsification attacks**," in *2015 IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1-6: IEEE**.**

[5] W. Zhang, Z. Wang, Y. Guo, H. Liu, Y. Chen, and J. Mitola III, "**Distributed cooperative spectrum sensing based on weighted average consensus,**" in *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011*, 2011, pp. 1-6: IEEE**.**

[6] S. Tanwar, S. Tyagi **,**N. Kumar, and M. S. Obaidat, "**LA-MHR: Learning Automata Based Multilevel Heterogeneous Routing for Opportunistic Shared Spectrum Access to Enhance Lifetime of WSN**," *IEEE Systems Journal*, no. 99, pp. 1-11, 2018**.**

[7] N. Vucevic, I. F. Akyildiz, and J. Pérez-Romero, "**Cooperation reliability based on reinforcement learning for cognitive radio networks,**" in *2010 Fifth IEEE Workshop on Networking Technologies for Software Defined Radio Networks (SDR)*, 2010, pp. 1-6: IEEE**.**

[8] Y. Li, D. Jin, J. Yuan, and Z. Han" ,**Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks**," *IEEE Transactions on wireless communications,* Vol. 13, No. 7, pp. 3965-3977, 2014**.**

[9] Y. Xiao, K.-C. Chen, C. Yuen, Z. Han, and L. A. DaSilva, "A **Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks**," *IEEE Transactions on Wireless Communications,* Vol. 14, No. 7, pp. 4034-4051, 2015**.**

[10] L. Rose, S. Lasaulce, S. M. Perlaza, and M. Debbah, "**Learning equilibria with partial information in decentralized wireless networks,**" *IEEE communications Magazine,* Vol. 49, No. 8, pp. 136-142, 2011**.**

[11] S. Rasaneh and M. Jahanshahi, "**A QoS aware learning automata based channel assignment method in cognitive network**," *Wireless Personal Communications,* Vol. 97, No. 1, pp. 495-519, 2017**.**

[12] S. Gheisari and M. R. Meybodi, "**LA-CWSN: A learning automata-based cognitive wireless sensor networks,**" *Computer Communications,* Vol. 94, pp. 46-56, 2016**.**

[13] B.-Y. Huang, S**-.**T. Su, C.-Y. Wang, C.-W. Yeh, and H.-Y. Wei, "R**esource allocation in D2D communication-A game theoretic approach,**" in *2014 IEEE International Conference on Communications Workshops (ICC)*, 2014, pp. 483-488: IEEE**.**

[14] A. Larmo, M. Lindström, M. Meyer, G **.**Pelletier, J. Torsner, and H. Wiemann, "**The LTE link-layer design,"** *IEEE Communications magazine,* Vol. 47, No. 4, pp. 52-59, 2009**.**

[15] A. Asheralieva and Y. **Miyanaga, "An autonomous learning-based algorithm for joint channel and power level selection by D2D pairs in heterogeneous cellular networks**," *IEEE transactions on communications,* Vol. 64, No. 9, pp. 3996-4012, 2016**.**

[16] T. Alpcan, H. Boche, M. L. Honig, and H. V. Poor, ***Mechanisms and games for dynamic spectrum allocation*****.** Cambridge University Press, 2013**.**