An application of Learning Automata Based ARL to Subchannel Allocation in Cellular OFDMA System

Hesam Montazeri and Mohmmad Reza Meybodi
Computer Engineering and IT Department
Amirkabir University of Technology
Tehran, Iran
{hmontazeri, mmeybodi }@aut.ac.ir

***Abstract*— In this paper, a new subchannel allocation schemes for cellular OFDMA networks employing an adaptive frequency reuse factor (FRF) strategy is considered. The allocation algorithm is semi-distributed solution comprising two phases. In the first phase, the Radio Network Controller (RNC) adaptively determines the FRF of each subchannel in a centralized manner. In the second phase, each base station autonomously allocates subchannels to the users using a simple algorithm (i.e. MaxC/I). To solve the first phase, we introduce a hybrid associative reinforcement learning (ARL) model combining self organizing map (SOM) and Learning Automata (LA) to deal with large size and continuous nature of the problem space. The simulation results illustrate that the proposed model achieves a better throughput gain in comparison with other allocation algorithms. It is noteworthy that the proposed algorithm has a low computational cost and achieving this throughput gain is only due to proper assignment of FRF to subchannels.**

Resource Allocation, Frequency Reuse Factor, Associative Reinforcement Learning, Learning Automata

I. INTRODUCTION

Orthogonal frequency division multiple access(OFDMA) system have been proposed to provide high data rate transmission in wireless communication [1]. Since the total bandwidth given to an OFDMA system is limited, allocation schemes play a key role in effective use of radio channels. With respect to the fact that multi-cell OFDMA networks are more applicable for practical use, resource allocation in these networks is considered more intensely.

Several algorithms have been proposed to deal with resource allocation problem. These algorithms can be divided into two centralized and distributed categories. The centralized algorithms assign the resources considering the state of all cells, while in the distributed ones, the allocation is performed autonomously in each cell without a need to consider the state of the other cells. The heuristic algorithms proposed in [2], [3], [4], [5], and [6] allocate the available resources to minimize total transmission power at the base stations. The algorithms presented in [2], [3], and [6] are centralized, whereas the algorithm in [4] is a distributed one. A distributed algorithm presented in [7] utilizes a game theoretic approach. It first defines a utility function that can represent the system performance while taking into account the co-channel interference among cells. Then it models resource allocation problem as a non-cooperative game that leads to a solution which maximize total system throughput. In [8], a throughput maximization algorithm is proposed, in which users experiencing high co-channel interference (CCI) receive less resources. The distributed algorithm proposed in [9] attempts to minimize the maximum value of the QoS violation for users in different cells. The algorithm works effectively in the conditions upon which the network resources are barely enough for the user demands. In [10], a linear programming formulation with the constraints of QoS and limited bandwidth is made to achieve optimal system throughput in a centralized manner. Since finding optimal solution for linear programming has high computational complexity, a heuristic algorithm yielding a semi-optimal solution is proposed.

This paper specifically addresses the problem of resource allocation in OFDMA cellular networks using a semi distributed approach. In the proposed method, assignment of subchannels to frequency reuse factor is performed in a centralized manner, while each cell allocates the subchannels to the users independently. We utilize hybrid model based on learning automata and self organizing map to determine FRF of each subchannel. The SOM quantizes the state space of the allocation problem via distributing some neurons in the state space. A learning automaton is associated with each neuron to adaptively select the best FRF for that state.

The remainder of the paper is organized as follows. Section II describes the system model and formulates the allocation problem. Section III presents the proposed solution. In this section, LA, SOM, and the Associated Reinforcement Learning (ARL) based on SOM and LA are described, and proposed subchannel allocation algorithm based on ARL is described. Section IV verifies the performance of the proposed solution by simulations. Finally, conclusion remarks are given in section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a downlink multi-cell OFDMA network which utilizes adaptive modulation and coding (AMC). In this network, transmission units are subchannels, comprising a group of subcarriers. Each subchannel consists of some data subcarriers and one pilot subcarrier. In addition, different base stations use different code signatures when broadcasting pilot signals. Then each user can measure the received powers of the pilot signals coming from multiple base stations and calculate the signal to interference and noise ratio SINR. After that the user feeds back his channel condition (SINR) to the associated base station. We assume that subchannels have identical transmission power, while varying the transmission rate by using adaptive modulation. The parameter such as guard interval and frequency offset are assumed to be designed so that the inter-channel interference can be disregarded. Also, on the ground of the fact that the subcarriers in a subchannel are randomly selected over the total bandwidth, frequency diversity and interference average can be obtained.

Our objective is to maximize the total system throughput. The goal is to find the subchannel allocation with a proper set of FRF and the optimal number of subchannels for each user. The formal optimization problem can be formulated as follows:

$$Maximize \sum_{b\in B}\sum_{i\in U^b}\sum_{f\epsilon F} N_f^{ib} R_f^{ib} \tag{1}$$

subject to:

1) $\sum_{i\in U^b} N_f^{ib} = A_f$ $\quad\forall b, f$
2) $\sum_{i\in U^b} f A_f = K$ $\quad\forall b$
3) $\sum_{f\in F} N_f^{ib} R_f^{ib} \geq \dot{R}^{ib}$ $\quad\forall i, b$

where R_f^{ib} is the maximum achievable date rates for user *i* in cell *b* with FRF *f*; N_f^{ib} is the number of subchannels the algorithm should assign to user *i* in cell *b* with FRF *f*; U^b is the set of users in cell *b*; *B* is the set of all cells; *F* is the set of all FRFs; *K* is the total number of subchannels; $\dot{R}^{ib}$ is the required data rate of user *i* in cell *b*; and A_f is the number of channels with FRF *f* used in each cell.

Constraint 1 implies that the number of subchannels with each FRF should be same for all cells. Constraint 2 dictates that the total number of assigned subchannels to the users in a cell should be exactly as same as the number of total subchannels. Constraint 3 is the one upon which the required data rate of all users will be guaranteed.

III. PROPOSED SOLUTION

A. Learning Automata

Learning automaton can be defined as an abstract model which randomly selects one action out of its finite set of actions and performs it on a random environment. Environment then evaluates the selected action and responses to the automaton with a reinforcement signal. Based on selected action, and received signal, the automaton updates its internal state and selects its next action.

Environment can be defined by the triple $E = \{\alpha, \beta, c\}$ where $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ represents a finite input set, $\beta = \{\beta_1, \beta_2, \ldots, \beta_r\}$ represents the output set, and $c = \{c_1, c_2, \ldots, c_r\}$ is a set of penalty probabilities, where each element c_i of c corresponds to one input of action α_i. An environment in which β can take only binary values 0 or 1 is referred to as P-model environment. A further generalization of the environment, known as Q-model, allows finite output set with more than two elements that take values in the interval [0, 1]. A further step in this direction is the S-model whose responses can take continuous values over the unit interval [0, 1].

Learning automata can be classified into fixed-structure LA, and variable-structure LA. Fixed-structure automata are characterized by state transition probabilities that are fixed. A variable-structure automaton is defined by the quadruple $\{\alpha, \beta, p, T\}$ in which $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ represents the action set of the automata, $\beta = \{\beta_1, \beta_2, \ldots, \beta_r\}$ represents the input set, $p = \{p_1, p_2, \ldots, p_r\}$ represents the action probability set, and finally learning algorithm is defined as $p(n+1) = T[\alpha(n), \beta(n), p(n)]$. This automaton operates as follows. Based on the action probability set p, automaton randomly selects an action α_i, and performs it on the environment. After receiving the environment's reinforcement signal, automaton updates its action probability set based on following reinforcement scheme, equations (2) for favorable response, and equations (3) for unfavorable one. It is noteworthy that this reinforcement scheme is for multi-action learning automata acting in the P-model environment.

$$\begin{aligned} p_i(n+1) &= p_i(n) + a\big(1 - p_i(n)\big) \\ p_j(n+1) &= p_j(n) - a p_j(n) \qquad \forall\, j \neq i \end{aligned} \tag{2}$$

$$\begin{aligned} p_i(n+1) &= (1-b) p_i(n) \\ p_j(n+1) &= \frac{b}{r-1} + (1-b) p_j(n) \quad \forall\, j \neq i \end{aligned} \tag{3}$$

In these two equations, a and b are reward and penalty parameters respectively. For a=b, learning algorithm is called L_{R-P}, for b << a, it is called $L_{R\varepsilon P}$, and for b = 0, it is called L_{R-I} [11].

B. Self Organizing Maps

The self organizing map (SOM) usually consists of an array of units arranged in a grid. Associated with unit *t* is weight vector $\vec{w}^t = [w_1^t, w_2^t, \ldots, w_D^t]$ where *D* is the dimensionality of the input data. The aim is to find a suitable set of weights for each unit so that the network estimates the distribution of the input data.

The learning rule responsible for finding a suitable set of weights is simple: given input vector $\vec{x} = [x_1, x_2, \ldots, x_D]$, the distance between unit *t* of the SOM and the input vector is calculated by Euclidean metric. The unit with the smallest distance is the one that most closely represents the current input, and thus considered the winner for that input. The weights of the winning unit are now updated towards the input. In addition to the winning unit being updated towards the current input vector, the winning unit's neighbors are also moved in this direction but by an amount that decays with the distance of those neighbors from the winning unit.

The weights of the map are initialized to random values and then the above process is iterated for each input vector in the data set, effectively resulting in a competition between different regions of the input space for units of the map. Dense regions of the input space will tend to attract more units than sparse ones, with the distribution of units in the weight space ultimately reflecting the distribution of the input data in the input space. Neighborhood learning also encourages topology preservation with units close in the topology of the map ending up close in the weight space too.

C. *Automata Based ARL*

In this section, a hybrid model for associative reinforcement learning (ARL) obtained by combining self organizing map and learning automata is described [12]. The ARL agent's interaction with its environment is similar to the Interaction of learning automata with its environment except that in addition to the environment response it also gets a context vector from the environment.

The hybrid model is defined by $\{S, \zeta, N, \phi, T\}$ in which S is the self organizing map used to represent state space; $\zeta = \{L_i | 1 \le i \le M\}$ *is* a team of M learning automata, where M is less than or equal to the number of SOM neurons; N is the topology of learning automaton team, which is assumed to be the same as the topology of SOM in this paper; ϕ is a function which specifies associations between each neuron of SOM and each learning automaton of LA team (ζ); and finally T is the learning algorithm [12].

Figure 1 depicts the model structure and its interaction with the environment. The hybrid model contains two layers. First layer incorporates a SOM which is used to quantize the state (context) space. The second layer comprised of a learning automata team which is used to select the optimal action. First layer is mapped to the second layer via ϕ function. In other words, each learning automaton is in correspondence with only one neuron of the SOM. The interaction of the model with the environment is as follows: environment provides the ARL agent with a context vector; the winner neuron which most closely represents this context vector is determined; learning automaton associated to the winner neuron is determined using ϕ function; the selected learning automaton chooses an action and performs it in the environment; the environment evaluates the action and provides the LA with a reinforcement signal; using this signal, LA and Its neighbors update their action selection strategies [12].

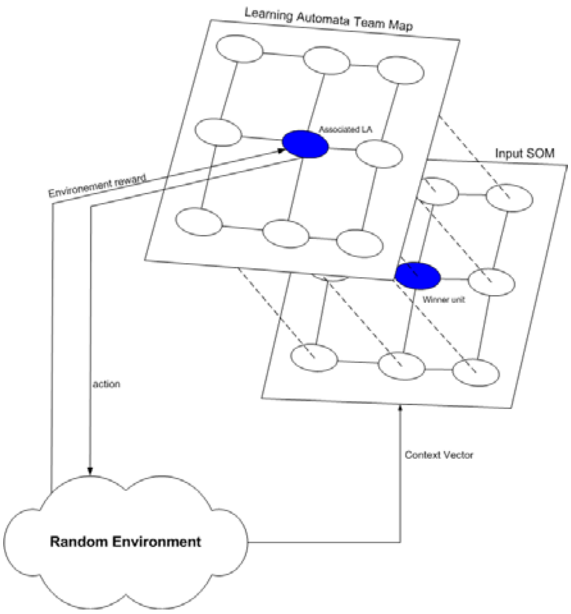

Figure 1. The Proposed model

D. *Proposed Subchannel Allocation algorithm*

In this section, the proposed subchannel allocation is presented. The algorithm has following steps:

1. MS report: Each mobile station calculates its SINR and reports it to the associated BS .
2. BS report: Each BS calculates the achievable data rate and reports it to a RNC that communicates with BSs.
3. FRF assignment: RNC assigns FRF to each subchannel using associative reinforcement learning.
4. Subchannel allocation to MS: In a BS, a simple algorithm such as MaxC/I is used to allocates the subchannels to the users.

The A_f is calculated with below function:

$$A_7 = \alpha K/7$$
$$A_3 = \beta(K - 7A_7)/3$$
$$A_1 = K - 3A_3 - 7A_7$$

The different values for α and β result in different algorithms which is shown in table 1. Algorithms with constant values of these parameters do not yield optimal selection of FRF in each state of the cells. It is obvious that for different state of the cells, different values for α and β should be used.

In the proposed method, the values for α and β are calculated using an associative learning method with feedback signal from environment. It's noteworthy that our algorithm has learning phase which will be used to train the algorithm. When the learning is finished, the learned parameters are used to select the number of FRF. The steps of the algorithm in the learning phase is as follow:

1. Calculate (N_1^b, N_3^b, N_7^b) for all *b*. Where N_f^b is the number of required subchannels of each SRF [10].
2. Maximum, minimum, and average of the N_f^b for all cells is calculated as input state to the ARL
3. ARL first determines α and β, then calculates A_f.
4. After subchannel allocation is done, throughput is calculated and gives to the ARL as reward of the selected action.

TABLE I. α AND β VALUES IN ALLOCATION ALGORITHMS

Algorithm	α , β
FRF1	$\alpha = 0, \beta = 0$
FRF3	$\alpha = 0, \beta = 1$
FRF7	$\alpha = 1, \beta = 0$
FRF 1,3	$\alpha = 0, \beta = 0.75$
SSAS	Using a simple heuristic[10]
Proposed Model	Using associative RL method

When the learning is finished, both of the BS and RNC have computational cost. Each BS allocates the subchannels to the users with a simple algorithm (i.e. MaxC/I) which has very low computational cost. The RNC selects the closest state in the SOM to the current cell state and calculates the A_f according to learned parameters for that state. As it can be seen when the learning is finished, the task of RNC is trivial and has low computational cost. Also in the learning phase, the proposed model does not have very high computational cost. The used associative RL is comprised of LA and SOM. Both of these models have low computational cost.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed solution, our simulator emulates an OFDMA multi-cell network based on the WiBro system [13]. The measured SINR of a subchannel in cell b by a user is expressed by

$$SINR = 10\log_{10}\frac{s_b d_b^{-\alpha}}{\sum_{i(\neq b)\in N^b} s_i d_i^{-\alpha} + \vartheta} \qquad (19)$$

where s_i is the transmit power of the base station in cell i, d_i is the distance between the user and cell i, α is the path loss exponent, ϑ is the background noise, and N^b is the neighboring cells of cell b. For the sake of simplicity, we neglect the ϑ noise and assume that the transmission powers of subchannels are the same. We set the path loss exponent 3.5 and use eight SINR levels in Table 2 which shows a possible modulation and coding scheme (MCS) in the WiBro system [13].

The other assumptions are as follows. There are 49 cells, each of which has the radius of 1Km and the total number of subchannels is 512 for each cell.

The performance of our algorithm is examined in comparison with some fixed channel assignment algorithms with various FRFs such as 3, and 7 and SSAS algorithm [10] which is a dynamic channel assignment algorithm. A cell with FRF 1 is not practical because the user near the cell boundary can not be served. Furthermore, a fixed channel assignment with hybrid FRF 1-3 is examined, 128 subchannels are allocated for FRF 1 and FRF 3, respectively. Note that all algorithms use MaxC/I algorithm [14] to allocate subchannels to the users after the values of A_f have been determined.

Figure 2 depicts the performance of the algorithms for various users' required data rates.

TABLE II. MCS LEVELS [13]

level	modulation	coding rate	required SINR	data rate (kbps)
1	QPSK	1/12	-1.8	9.78
2	QPSK	1/6	-0.3	19.55
3	QPSK	1/3	2.6	39.11
4	QPSK	2/5	4.2	46.93
5	16QAM	1/4	5.2	58.67
6	16QAM	1/3	6.8	78.21
7	16QAM	2/5	8.3	93.78
8	16QAM	1/2	11.3	117.33

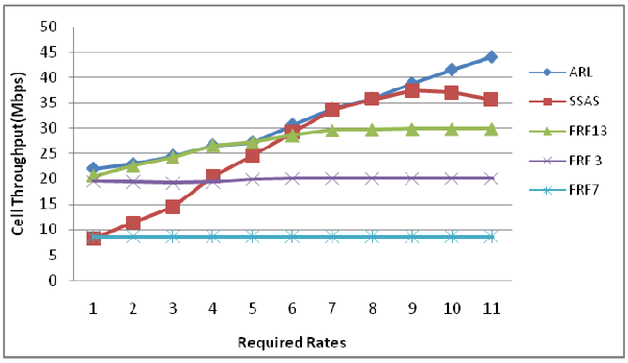

Figure 2. Cell throughput vs. users' required rates

As it can be seen from the figure 2, the proposed algorithm outperforms the others in almost all situations. Among the other methods, when the required data rate is lower than 1000, the hybrid FRF 1-3 method is better than other methods and its performance is near to the proposed method. In other cases the SSAS has greater throughput than others. It can be inferred that the fixed algorithms only do well for some parts of the problem space. But, dynamic allocation algorithms such as proposed method and SSAS are robust and perform well in almost all situations.

Figure 3 illustrates the performance of the algorithms for different number of users. Again the proposed algorithm outperforms the others in scenario considered. Figure 1 and 2 indicate that the algorithm is examined in various situations and it has reasonable performance in almost all situations.

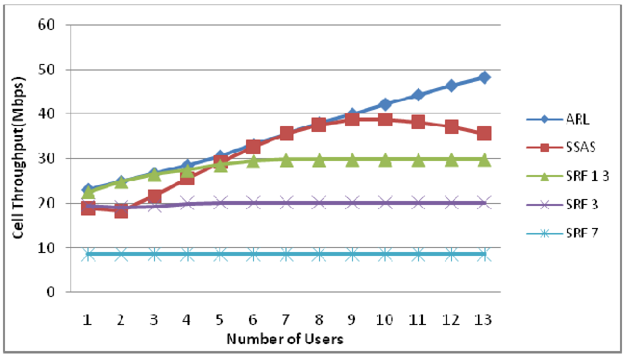

Figure 3. Cell throughput vs. number of users

V. CONCLUSION

In this paper, we have presented a novel resource allocation algorithm for cellular OFDMA networks through an ARL approach. Our semi-distributed allocation solution comprises of two phases. In the first phase, the RNC adaptively determines the FRF of each subchannel in a centralized manner. In the second phase, each BS autonomously allocates subchannels to the users. Our main contribution in this paper was solving the first phase problem using a hybrid ARL model combining SOM and LA models. This ARL approach has low computational cost and could elegantly deal with huge size and continuous nature of the problem. Simulation results indicate that proposed allocation algorithm provides the system throughput increase by properly selecting frequency reuse factor for each subchannel. Also, it is experimentally demonstrated that proposed model achieves a better throughput gain in comparison with SSAS and some fixed channel assignment algorithms.

REFERENCES

[1] I. Koffman, V. Roman, and R. Technol, "Broadband wireless access solutions based on OFDM access in IEEE802. 16," Communications Magazine, IEEE, vol. 40, pp. 96-103, 2002.

[2] S. Pietrzyk, G. J. M. Janssen, A. N. Unit, P. T. C. Sp, and P. Warsaw, "Radio resource allocation for cellular networks based on OFDMA with QoS guarantees," in Global Telecommunications Conference, 2004, GLOBECOM'04. IEEE. vol. 4, 2004, pp. 2694-2699.

[3] L. I. Jun-qiang, K. I. M. Hojin, and L. E. E. Yongsuk, "A novel broadband wireless OFDMA scheme for downlink in cellular communications," in Proceedings of IEEE Wireless Communications and Networking Conference. vol. 3, 2003, pp. 1907-1911.

[4] L. Yan, Z. Wenan, and S. Junde, "An adaptive subcarrier, bit and power allocation algorithm for multicell OFDM systems," in IEEE CCECE 2003. Canadian Conference on Electrical and Computer Engineering. vol. 3, 2003, pp. 1531-1534.

[5] I. Koutsopoulos and L. Tassiulas, "Cross-layer adaptive techniques for throughput enhancement in wireless OFDM-based networks," IEEE/ACM Transactions on Networking (TON), vol. 14, pp. 1056-1066, 2006.

[6] S. Hamouda, P. Godlewski, and S. Tabbane, "Enhanced Capacity for Multi-Cell OFDMA Systems with Efficient Power Control and Reuse Partitioning," in 10th IEEE Singapore International Conference on Communication systems, 2006, pp. 1-5.

[7] H. Kwon and B. G. Lee, "Distributed Resource Allocation through Noncooperative Game Approach in Multi-cell OFDMA Systems," in ICC'06: IEEE International Conference on Communications. vol. 9, 2006, pp. 4345-4350.

[8] G. Li and H. Liu, "Downlink dynamic resource allocation for multi-cell OFDMA system," in Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th. vol. 3, 2003, pp. 1698-1702.

[9] H. Kwon, W. I. Lee, and B. G. Lee, "Low-Overhead Resource Allocation with Load Balancing in Multi-cell OFDMA Systems," IEEE 61st Vehicular Technology Conference, vol. 5, 2005.

[10] H. Kim, Y. Han, and J. Koo, "Optimal subchannel allocation scheme in multicell OFDMA systems," in IEEE 59th Vehicular Technology Conference vol. 3, 2004, pp. 1821-1825.

[11] K. S. Narendra and M. A. L. Thathachar, Learning automata: an introduction: Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1989.

[12] H. Montazeri and M. Meybodi, "A new hybrid associative reinforcement learning" accepted in third International conference in Information and Knowledge Technology, Mashhad, Iran, 2007.

[13] Y. J. Choi, C. S. Kim, and S. Bahk, "Flexible Design of Frequency Reuse Factor in OFDMA Cellular Networks," in ICC'06: IEEE International Conference on Communications. vol. 4, 2006, pp. 1784-1788.

[14] Y. M. Ki, E. S. Kim, S. I. Woo, and D. K. Kim, "Downlink Packet Scheduling with Minimum Throughput Guarantee in TDD-OFDMA Cellular Network," Lecture notes in computer science, pp. 623-633, 2005.