

Special Issue

# Optimized Intelligent Design for Smart Systems Hybrid Beamforming and Power Adaptation Algorithms for Sensor Networks Decision-Making Approach

Ali Kamil Khisarullah<sup>1,\*</sup>, Ufuk Tureli<sup>1</sup>, Didem Kivanc<sup>2</sup><sup>1</sup> *Electronics And Communications Engineering, Yildiz Technical University, Istanbul, Turkey-34220*<sup>2</sup> *Electrical And Electronics Engineering, Okan University, Istanbul, Turkey-34959*

## ARTICLE INFO

### Article History

Received 01 Sep 2019

Accepted 10 Nov 2019

### Keywords

Optimized intelligent design for smart systems

Wireless network beamforming

Power adaption

Interference

Strategic decision-making

Game theory

Reinforcement learning

## ABSTRACT

During last two decades, power adaptation and beamforming solutions have been proposed for multiple input multiple output (MIMO) Ad Hoc networks. Game theory based methods such as cooperative and non-cooperative joint beamforming and power control for the MIMO ad hoc systems consider the interference and overhead reduction, but have failed to achieve the trade-off between communication overhead and power minimization. Cooperative method using game theory achieves the power minimization, but introduced the overhead. The non-cooperative solution using game theory reduced the overhead, but it takes more power and iterations for convergence. In this paper, a novel game theory based algorithms proposed to achieve the trade-off between power control and communication overhead for multiple antennas enabled wireless ad-hoc networks operating in multiple-users interference environment. The optimized joint iterative power adaption and beamforming method designed to minimize the mutual interference at every wireless node with constant received signal to interference noise ratio (SINR) at every receiver node. First cooperative potential game theory based algorithm designed for the power and interference minimization in which users cluster and binary weight books along used to reduce the overhead. Then the non-cooperative based approach using the reinforcement learning (RL) method is proposed to reduce the number of iterations and power consumption in networks, the proposed RL procedure is fully distributed as every transmit node require only an observation of its instantaneous beamformer label which can be obtained from its receive node. The simulation results of both methods prove the efficient power adaption and beamforming for small and large networks with minimum overhead and interference compared to state-of-art methods.

© 2019 The Authors. Published by Atlantis Press SARL.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION

Ad Hoc multiple input multiple output (MIMO) communication system is base for the future generation wireless communication system Massive MIMO. Multiuser-MIMO (MU-MIMO) boosts the design and development of Massive MIMO technologies [1]. MIMO communication system efficiency mainly depends on adaptability of transmit parameters such as beamformer selection, transmit power, transmission rate, modulation technique, and so on. The design of beamformer selection plays the important role for the adaptive wireless communication systems. For MIMO, the beamforming methods designed for different types of networks such as point-to-point, cellular, and ad hoc [2]. Number of beamforming methods designed to address the challenges related to each of such networks. The conventional method of beamforming shows the power control, throughput maximization, and capacity enhancement especially for point-to-point and cellular networks. For MIMO ad hoc networks, the distributed beamforming solutions show the system throughput enhancement and energy minimization but suffer from the challenges of interference and

communication overhead. Several distributed beamforming methods were designed to minimize the interference and communication overhead in literature.

As discussed earlier, there are three categories in which the beamforming solutions presented for MIMO. The point-to-point communication based MIMO beamforming solutions described in [3–6] and the beamformers and linear precoders (eigencoders) proposed for point-to-point MIMO communications in [7,8]. The beamforming algorithms designed for the cellular networks in [9–11] reduces the power and improve the capacity of single antenna mobile transmitters and array-equipped base stations. The beamforming and power control for point-to-point and cellular MIMO networks achieved with minimum overhead and interference challenges. However, the problem is for ad hoc MIMO networks where operate without centralized controller and moving nodes. For MIMO ad hoc networks, the distributed beamforming techniques enhance network throughput and minimize energy consumption [12,13]. The main challenge using optimization based solutions for ad hoc networks is that they needs the systematic and proper study as the environment of such networks is interference limited as well as overhead introduced by beamforming algorithms

\* Corresponding author. Email: [kamil.ali8219@gmail.com](mailto:kamil.ali8219@gmail.com)

affects the performance of MIMO ad hoc networks. Earlier were distributed spatial beamforming techniques introduced in [14,15] for multi-user ad hoc MIMO networks under channel reciprocity conditions. However, in such methods the transmission overhead is introduced during the power control and at each iteration. To minimize the communication overhead. Further, many iterative algorithms were introduced for ad hoc MIMO networks. However, the trade-off is not achieved with such techniques as well as convergence of iterative algorithms has not been investigated.

On the other hand, there are a lot of formats can view the information structure that allowed a full understanding and showed a good view of the problems that we are interesting to, when the predictive are enabling such its capability can be determined in order to enhance the view of information such that (modify, reinforce, or maybe reject). Mathematical modeling such that game theory and information theory can capture important reality features and meaningful data structure in a manageable manner with efficient information processing to achieve a smart and trust decision-maker strategies, a qualified intuition is strongly needed. A good decision aid methodology implies understanding reality in a manageable way [16]. Game theory is useful mathematical tools for Studying environments where multiple players are interacting and making decisions, it often talk about rational players (they try to optimize their own interests) [17].

The Game theory based solutions produced the efficiency along with convergence investigation for important problems in wireless communications like joint code-division multiple access [18], distributed power control algorithm [19], and optimum transmission signaling strategies [20,21]. Thus, game theory based beamforming and power adaptation methods designed for multi-user ad hoc MIMO network communications to achieve the power efficiency with minimum overhead. In [22], the initial cooperative and non-cooperative game theory based methods designed for joint channel and power control for wireless mesh networks, and channel allocation and beamforming for ad hoc networks was studied in [23], however the solutions are suboptimal for wireless communications. For ad hoc networks the reduction of power using the distributed algorithms and transmit beamformer selection is difficult tasks and hence its challenging to formulate the beamforming games for multi-user wireless communications. In [24,25], the first attempt over joint discrete transmits beamforming and power adaptation introduced using the cooperative and non-cooperative game theory based solutions. The decentralized approach designed for optimizing the transmit beamformer and power adaptation by exploiting the local information with the acceptable computational burden under constraint of constant received target signal to interference noise ratio (SINR). However, designing the optimized game theory based approach for the power adaptation and beamforming in ad hoc MIMO systems is still the complex research problem the both solutions failed to achieve the trade-off between power minimization and overhead. In [25], the first solution cooperative power minimization algorithm (COPMA) achieved the power minimization but introduced the overhead and second non-cooperative solution regret-matching-based joint transmit beamformer and power selection game (RMSG) minimizes the communication overhead but it takes more power and iterations to converge. In this paper, we further optimize COPMA and RMSG techniques to minimize the overhead and power consumption respectively. In [26] the

reinforcement learning (RL) based algorithm used to determine the suitable policy for selection between two different techniques, first one is beamforming while the second is power control for sensor array networks, by taking the benefit of the individual features of each method according to SINR threshold. The key issue of the RL is to select the action (BF or PC) according to the learning state in this case (SINR), that is satisfied with the system requirement. Q-learning with an epsilon greedy ( $\epsilon$ -greedy) policy was used in the RL algorithm that is trained by using the offline method. Artificial intelligence (AI) techniques have been described and applied in problems for beamforming, power control, and MIMO wireless communication systems [27–30].

The key contributions of this paper summarized as:

1. Enhanced co-operative power minimization algorithm (**ECOPMA**) for MIMO Ad hoc networks is proposed to overcome the challenges of COPMA. We proposed cooperative method using the potential game approach in which first compute allocated power and beamformer for each user until the convergence to the steady state according to Nash equilibrium (NE) technique, then divide users in clusters [31] and enable them to converge at same time in order to minimize the overhead. In short, in this case the users cooperate each other to reduce the power and interference using the potential game theory and NE.
2. To reduce the communication overhead in ECOPMA, the binary weight books are used rather than using the complex Grassmannian weight books. The binary weight book reduces the overhead incurred by cooperative solution.
3. Proposed reinforcement learning based power allocation and beamformer algorithm (**RLPBA**) for multi-user ad hoc MIMO systems. In RLPBA, optimized non-cooperative solution designed using the RL game based approach in which the local information used for the beamforming and power adaptation decisions. The Reinforcement approach propose to reduce the number of iterations and hence the transmit power for convergence.
4. Extensive experimental results and comparative evaluations presented for ECOPMA and RLPBA with centralized and state-of-art cooperative and non-cooperative algorithms.

In Section 2, the proposed system model, algorithms for ECOPMA and RLPBA presented. In Section 3, the simulation results are discussed. In Section 4, conclusion and future work described.

## 2. METHODOLOGY

This paper proposes the two technologies ECOPMA and RLPBA proposed to address the challenges of power minimization with minimum interference and communication overhead with guaranteed QoS for multi-user MIMO ad hoc systems with constant SINR constraint. ECOPMA is based on potential cooperative games using binary weight books rather than using the complex Grassmannian weight book to reduce the communication overhead. ECOPMA is similar to the COPMA [25] with inclusion of binary weight books. The RLPBA is non-cooperative games based power allocation and

beamforming technique similar to RMSG [25] with use of reinforced learning to reduce the power consumption and number of iterations to converge. In Section 2.1, we present the MIMO ad-hoc communication model and game theory problem. In Section 2.2, we present the design of centralized beamforming and power allocation technique. In Section 2.3, design of ECOPMA presented, in Section 2.4, the design of RLPBA described.

## 2.1. System Model

Figure 1 demonstrates the system model referred from [25] which consist of wireless ad hoc systems with multiple antennas nodes pair under the same channel. The interference generated by other nodes pairs operating on similar channel. Total  $N$  number of node pairs with each pair  $q$  consist on single transmit-receive wireless node. Every transmitting and receiving node is equipped with  $A$  antennas. Every node having the beamformer pair  $(w_q, t_q)$ . The complex transmitted symbol stream is  $s_q \in C$  and received symbol stream is  $\hat{s}_q \in C$ . Thus the received signal vector  $R_q \in C$  for  $q^{th}$  receiving node is computed as:

$$R_q = \sqrt{E_q} U_{q,q} t_q s_q + \sum_{j \neq q} \sqrt{E_j} U_{q,j} t_j s_j + n_q \quad (1)$$

where  $U_{q,j}$  denotes the  $A \times A$  MIMO channel between the  $j^{th}$  transmitting node and the  $q^{th}$  receiving node and is assumed to be quasi-static and  $E_q$  is the power of the  $q^{th}$  transmitting node. The additive white Gaussian noise terms  $n_q$  have identical covariance matrices.

The worst case is considered in this work in which all the pairs always having packets to transmit through the wireless channel. The network assumed to be synchronous. The available codebook beamformers set for the  $q^{th}$  node pair is represented by  $\partial_q = \{t_q^1, t_q^2, \dots, t_q^Y\}$  with the cardinality  $Y$ . The transmit beamformer from codebook selected by receiving nodes and feeds back with selected beamformer index. Every node can select among  $Y$  transmit beamformers codebook. Consider that  $t_q$  i.e.  $\partial_q$  selected transmit beamformer for the  $q^{th}$  node pair.  $\emptyset = [t_1, t_2, \dots, t_N]^A$  and  $E = [E_1, E_2, \dots, E_N]^A$  are denoted as transmit beamformer selection and transmission power vectors for  $N$  number of nodes respectively. The  $A \times A$  matrix of the interference and noise covariance at  $q^{th}$  receiving node is:

$$R_q(\emptyset_{-q}, E_{-q}) = \sum_{j \neq q} E_j U_{q,j} t_j^U U_{q,j}^U + \sigma^2 I \quad (2)$$

where  $\emptyset_{-q}$  and  $E_{-q}$  are transmit beamformer and powers of the nodes other than  $q$ . Further, normalized receive beamformer at  $q^{th}$  receiving node computed as:

$$w_q = \frac{\widehat{w}_q}{\|\widehat{w}_q\|} \quad (3)$$

where  $\widehat{w}_q = R_q^{-1} U_{q,q} t_q$ . The outcome of SINR at  $q^{th}$  receiving node due to desired transmitter of  $q^{th}$  node pair is:

$$\tau_q = \frac{E_q |w_q^U H_{q,q} t_q|^2}{\sum_{j \neq q} E_j |w_q^U H_{q,j} t_j|^2 + \sigma^2} \quad (4)$$

Thus the proposed methods in this work attempt to achieve the target SINR through the transmit powers adjustment. According to this, the optimization problem of this work is described as below.

The objective is to reduce the transmit energy of all the nodes pairs  $q \in \{1, 2, \dots, N\}$  in network under the constraint of constant  $SINR_{\gamma_0}$ . We define the optimization problem as:

$$\text{Minimize } (\emptyset, E) \sum_{q=1}^N E_q \quad (5)$$

subject to  $\tau_q \geq \gamma_0, \|w_q\| = \|t_q\| = 1$

$E_{min} < E_q \leq E_{max}$ , where the  $E_{min}$  and  $E_{max}$  are the minimum and maximum transmit powers respectively.

This problem is represented in game theory approach as normal form of game as:

$$\pi = (N, C, \{F_q\}_{q=1}^N) \quad (6)$$

where,  $N$  is the set of players,  $C$  is set of available actions of all  $N$  players, and  $\{F_q\}_{q=1}^N$  is set of utility functions that the players associate with their strategies. The actions  $c_q \in C_q$  for a player  $q$  are the selection of transmit powers  $E_q \in [C_{min}, C_{max}]$  and the transmit beamformer  $t_q$  i.e.  $\partial_q$ . In this case, the players in game selects the appropriate action to enhance their utility functions. The convergence point in proposed case is set of strategies, set of beamforming selections  $\emptyset = [t_1, t_2, \dots, t_N]^A$  and  $E = [E_1, E_2, \dots, E_N]^A$  from which no player would deviate. Such set of strategies is known as NE [30]. The NE is a set of strategy profiles  $c$  from which no player can increase his utility by unilateral deviations. NE used to decide and change the strategy profile while keeping the actions of other players same. In this paper we designed two scenarios of node pairing such as cooperative and non-cooperative to obtain the best outcomes to satisfy the objective function with guaranteed convergence.

## 2.2. Centralized Method

Prior discussing the decentralized solutions, this section we formulate the centralized solution for MIMO ad hoc communication systems. Using the centralized agent [32], joint transmit beamformers and corresponding transmit powers selected to reduce the overall transmit power of all the antennas as:

$$(\emptyset^*, E^*) = \arg \min_{\emptyset, E} \sum_{q=1}^N E_m(\emptyset, E_{-q}) \quad (7)$$

where,  $\emptyset^*$  and  $E^*$  are the optimal transmit beamformer and power solutions respectively. The transmit power (Eq)  $q^{th}$  nodes pair is computed as:

$$E_q(\emptyset, E_{-q}) = \frac{\gamma_0}{t_q^U U_{q,q} R_q^{-1} U_{m,m} t_m} \quad (8)$$

In this case, the centralized agent computers the total network power for  $Y^N$  possible beamforming vector combinations which leads the complex tasks in case large scale wireless ad hoc networks. The sever complexity invalidates the centralized method for multi-user MIMO ad hoc systems. To solve such problems we introduced the decentralized solutions in this paper.

## 2.3. ECOPMA

As the name indicates, ECOPMA is cooperative game based solution (Where nodes can cooperative each other to achieve the optimum solution) for beamforming and adaptive power allocation for multi-user MIMO ad hoc systems. As discussed earlier, the ECOPMA is similar to COPMA [25] in which we used the two users paired in clustering and binary codebook to reduce the computational overhead. In this work, we generated the binary codebook of size 16 (represented as  $\Upsilon$ ) code length and 4 (represented as  $A$ ) is dimension of code [33].

The clustering mode, in which we assume that the node pairs with similar transmit power properties are grouped into one cluster. The mobile users clustered in order to enable them to converge at same time to minimize the energy and overhead. In [34], it is proved that for multi-user scenario, two users paired into cluster leads to common transmit beamforming vector sharing and hence improves the QoS performance with minimum overhead and power consumption. In ECOPMA, finding the joint optimal transmit power allocation and transmit beamformer so that total power consumption in network is minimized. Let the each user in  $q^{th}$  paired represented into  $M$  clusters in such manner that they share common transmit beamforming vector. The  $q^{th}$  nodes pair assumed that it paired to  $m^{th}$  cluster,  $m \in [1, 2, \dots, M]$ . For the clustering, we assume the channel conditions as for transmit and receive beamforming vectors as: (1) zero-forcing (ZF) precoding at each receiving node to remove the inter-cluster interference and (2) signal alignment is conducted at the receiver between users in the same cluster. According to that, let  $E_q^{m,m}$  and  $\varnothing_q^{m,m}$  are denoted as transmission power vectors and transmit beamformer selection vectors for  $q^{th}$  nodes pair across the  $m^{th}$  transmit and receive clusters respectively. Thus, the objective function of ECOPMA is:

$$F_{\text{net}}(\varnothing, E) = - \sum_{m=1, q=1}^{M, N} E_q^{m,m}(\varnothing, E_{-q}^{m,m}) \quad (9)$$

Eq. (9) will assume as the every users utility function that is,

$$\begin{aligned} F_i(\varnothing, E) &= F_{\text{net}}(\varnothing, E) \\ &= - \sum_{m=1, q=1}^{M, N} E_q^{m,m}(\varnothing, E_{-q}^{m,m}) \end{aligned} \quad (10)$$

The special case of potential games is modelled in this paper called identical interest game [25,35]. Therefore it becomes easier to verify that at least one pure NE produced by all the identical interest games which represent any action profile that enhances  $F_i(\varnothing, E)$ . The working of ECOPMA technique described as: (See Algorithm 1 on next column)

The functionality of smoothing factor is referred from [25] with the core contribution of this research paper. The above Algorithm 1 shows the difference between COPMA [25] and ECOPMA using the clustering. The  $ID_q^m$  is the single user unique number that belongs to  $q^{th}$  nodes pair and  $m^{th}$  cluster.

## 2.4. RLPBA

The second solution proposed in this paper is based on non-cooperative game theory based approach for beamforming and adaptive power allocation in multi-user ad hoc MIMO systems using the RL scheme. The main aim is to design the distributed learning technique for joint transmits beamformer and power

---

### Algorithm 1: ECOPMA

---

#### 1. Inputs:

$k$ : predefined number of iterations

$M$  = number of clusters

2. Apply the users clustering on all the mobile users regardless of users pairing

3. **FOR** each cluster  $m$  in  $M$ , do

4. **Initialization:** For each pair  $q$ , initialize transmit beamformers and transmit powers:

4.1.  $\|w_q^{m,m}\| = \|w_q^{m,m}\| = 1, q \in N, m \in M$

4.2.  $E_q^{m,m} = E_{\text{max}}, q \in N$

5. Repeat: random selection of  $q^{th}$  nodes pair with probability of  $1/N$ .

5.1. Set  $t_q^{m,m}(n) = t_q^{m,m}(n-1)$  (current transmit beamformer for  $q^{th}$  nodes pair)

5.2. Calculate  $E_q^{\text{current}}$  using Eq. (8)

5.3. Random selection of transmit beamformer  $t_q^{\text{updated}}$  and compute corresponding transmit power  $E_q^{\text{updated}}$  using Eq. (8) required while using updated transmit beamformer.

5.4. Form a vector  $[ID_q^m, E_q^{\text{current}}, E_q^{\text{updated}}]$  and broadcast all other nodes pairs  $j, j \in N$ .

5.5. On receiving data vector, for each  $j$   
IF ( $E_j$  changes due to interference at  $j^{th}$  receiver)

Every other node pairs sets  $E_j^{\text{current}} = E_j^{\text{updated}}$ ,

Computes new transmit power and set it as  $E_j^{\text{updated}}$

ELSE

Unchanged  $E_j^{\text{current}}$  and  $E_j^{\text{updated}}$

END

Send back vector  $[ID_j^m, E_j^{\text{current}}, E_j^{\text{updated}}]$  to node pair  $q$

5.6. node pair  $q$  computes the total network transmit power as  $E_{\text{current}} = \sum_{q=1}^N E_q^{\text{current}}$  and updated power as  $E_{\text{updated}} = \sum_{q=1}^N E_q^{\text{updated}}$  with  $t_q^{\text{updated}}$

5.7. Updating node pair  $q$  selects the  $t_q^{\text{updated}}$  using the probability based on smoothing factor ( $\tau > 0$ ) computed by  $1 / (1 + \exp((E_{\text{updated}} - E_{\text{current}}) / \tau))$ .

5.8.  $q^{th}$  nodes pair broadcasts the decision signal whether new transmitter beamform is kept.

5.9. If not kept, then every other node pair  $j, j \in m/N$  keeps  $E_j^{\text{updated}} = E_j^{\text{current}}$

6. Until predefined number of iterations steps  $k$

7. **END FOR**

---

selection scheme which needs only local information for updates for ad hoc MIMO systems. The utility function for non-cooperative users used in this case. As compared to cooperative solution, the  $q^{th}$  nodes pair focused only on their own power minimizations rather than complete network power. Every player's utility function based on the selection of transmits beamformers and its power, and other player's selections for beamformers and transmit power through perceived interference. The novel solution of RLPBA proposed that exhibits "rewards" of selecting the action or strategy [26,36]. We formulate non-cooperative game theory approach using the Actor-Critic Learning algorithm known as Continuous Actor-Critic Learning Automaton in this paper to predict and select the beamformer and accordingly calculate the new transmit power.

One of the important methods in Temporal-Difference (TD) methods is Actor-Critic which based on separated memory structure in order to describe the independent of policy as compared with the value function (as contrasted with the reward  $R_e$  instead of having a short-term return, the value function anticipated of long-term along with decreasing factor). Where policy structure (the action which the agent uses to evaluate the next action strategy depends on the current state) is recognized as an actor, for the reason that it is employed to choose actions, and determined value of a function can be known as the critic, due to it criticizes actions that made by the actor. The critic has to observe and justify if the policy is being followed by way of the actor or not [37].

As like in Figure 1, let  $S$  denotes the number of *states* required that contains the  $\gamma$  number of transmit beamformer codebooks which selected by every node. Let vector  $\hat{\delta}_q$  denotes all *action set* for user  $q$ , that is,  $\hat{\delta}_q = \{t_q^1, t_q^2, \dots, t_q^\gamma\}$ , where  $t_q(i)$  denotes the transmit beamformer vector selected by  $q^{th}$  user in iteration  $i$ . Define the reward function  $Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k)$  of  $q^{th}$  user for an action  $\hat{\delta}_q$  at  $k^{th}$  iteration as:

$$Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k) = \sum_{i=1}^{k-1} \mu_q (\hat{\delta}_q, t_{-q}(i)) - \mu_q (t_q(i)) \quad (11)$$

Every user  $q$  computes the  $Re_{\hat{\delta}_q}^{\hat{\delta}_q}$  for each action for all the past steps when all other players action remains unchanged. Every user  $q$  update its reward function value for every set of action  $\hat{\delta}_q$  as:

$$Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k+1) = Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k) + \mu_q (\hat{\delta}_q, t_{-q}(i)) - \mu_q (t_q(i)) \quad (12)$$

After update, we perform the selection the transmit beamformer  $t_q(k)$  by computing the probability as:

$$P_{\hat{\delta}_q}^{\hat{\delta}_q}(k) = \frac{Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k)}{\sum_{\hat{\delta}_q \in q} Re_{\hat{\delta}_q}^{\hat{\delta}_q}(k)} \quad (13)$$

Every user  $q$  chooses an action or strategy according to the outcome of  $P_{\hat{\delta}_q}^{\hat{\delta}_q}(k)$  while checking the decision condition as:

$$P_{\hat{\delta}_q}^{\hat{\delta}_q}(k+1) > P_{\hat{\delta}_q}^{\hat{\delta}_q}(k) \quad (14)$$

According to the selection of transmit beamformer  $t_q(k)$ , the new transmit power  $E_q$  is computed using Eq. (8) on selected  $t_q(k)$ . Algorithm 2 demonstrates the above steps.

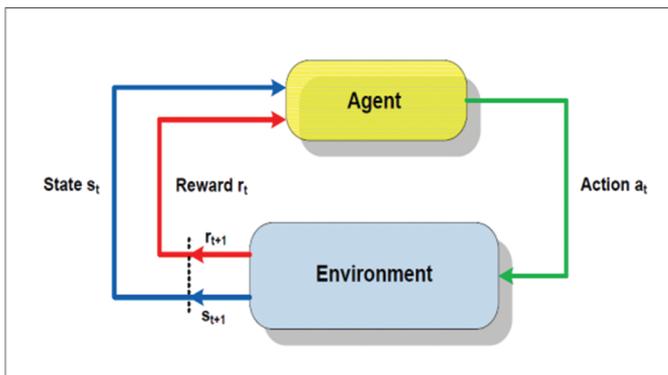


Figure 1 | Environment and agent system interaction method [26].

---

#### Algorithm 2: RLPBA

---

Inputs

$k$ : predefined number of iterations

1. FOR  $i = 1, 2, \dots, k$

2. FOR  $q = 1, 2, \dots, N$

3. Compute the reward value using Eq. (11)

4. Update Reward table using Eq. (12)

5. Compute the probability values using Eq. (13)

6. Take the decision on  $t_q(k)$  selection using Eq. (14)

7. Compute new transmit power for  $t_q(k)$  using Eq. (8)

8. END FOR

9. END FOR

---

### 3. SIMULATION RESULTS AND DISCUSSIONS

This section presents the performance investigation of proposed ECOPMA and RLPBA methods with state-of-art solutions such as centralized optimization, RMSG [25], and COPMA [25]. While evaluating such methods we designed ad hoc networks with assumption of 5 (small) and 10 (large) homogenous pairs with each having one transmitter and receiver node. The complete sets of parameters are described for 5 pairs and 10 pairs in Tables 1 and 2 respectively. For COPMA [25], the Grassmannian codebook of size  $Y = 16$  with  $A = 3$  antennas for all users, however as the Grassmannian codebook complex, for ECOPMA we used binary codebook of size  $Y = 16$  with  $A = 4$  antennas for all users in this work.

#### 3.1. Evaluations of 5-Pairs Wireless Ad hoc System

Table 1 shows the simulation parameters to use to investigate said methods.

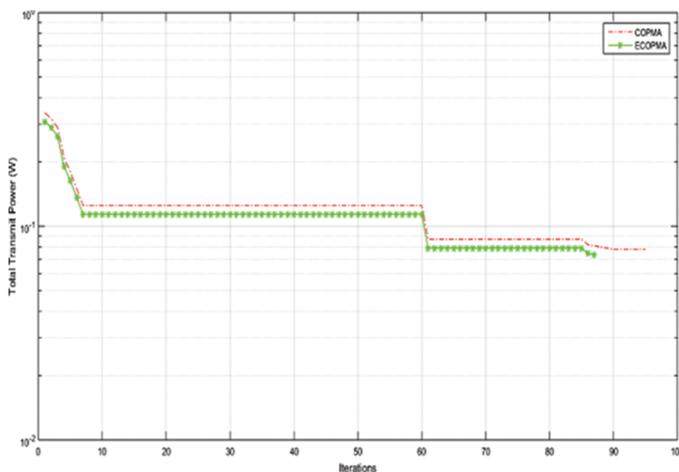
As per the above table which is designed for small ad hoc networks evaluations, we measured the performance of total power consumption. Figure 2 demonstrates the comparative analysis of total power consumed by network for cooperative methods COPMA and ECOPMA. The purpose of ECOPMA is to reduce the overhead by reducing the number of iterations compared to COPMA method which is achieved in results (Figure 2). We observe that ECOPMA's performance settles at the global optimum combination after 87 iterations as compared to COPMA (93 iterations). The reduction in iterations reduces the communication overhead as well as total power consumption of network. The performance of ECOPMA is improved due to the clustering and simple binary codebook. Similarly we evaluated the non-cooperative distributed learning based techniques in Figure 3. The RLPBA designed to reduce the power consumption with the reduction of iterations for optimum allocation solutions for network pairs. We exploited the advantages of RL technique over the regret based learning approach in RLPBA. The outcomes of RLPBA show that it reduced the total transmission power for small wireless ad hoc MIMO networks compared to RMSG technique. It takes 110 iterations to converge total network transmission power compared to RMSG (120 iterations).

Figure 4 demonstrates the comparative study of all the methods. It shows that total power in network varies using the non-cooperative methods over the 120 iterations. The cooperative

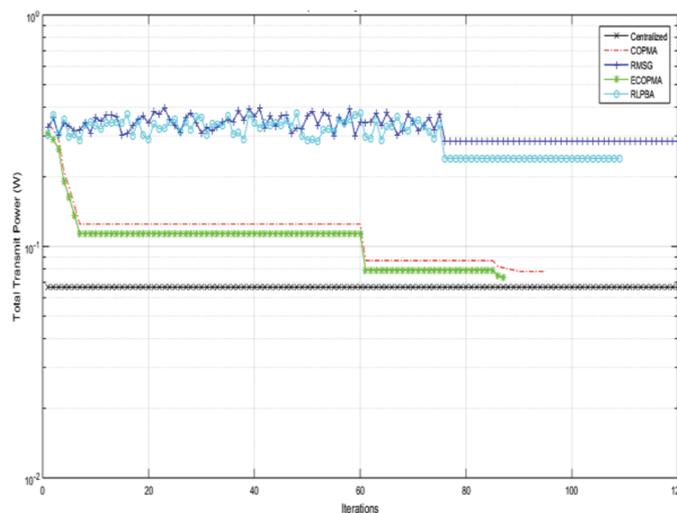
**Table 1** | Simulation parameters for 5-Pairs wireless ad hoc MIMO system.

Parameter	Value
Number of wireless ad hoc pairs	5
Size of network	30 m × 30 m
Constant Received SINR	10 dB
Grassmannian codebook	( $\Upsilon = 16, A = 3$ )
Binary codebook	( $\Upsilon = 16, A = 4$ )
$E_{max}$	100 mW
$E_{min}$	1 mW
Propagation channel	Radio propagation channel (path loss component is 4)
Predefined number of iterations ( $k$ )	120
Smoothing factor (For COPMA and ECOPMA)	$0.01/k^2$
Clusters (ECOPMA)	2
Methods Investigated	Centralized, COPMA, RMSG, ECOPMA, and RLPBA.

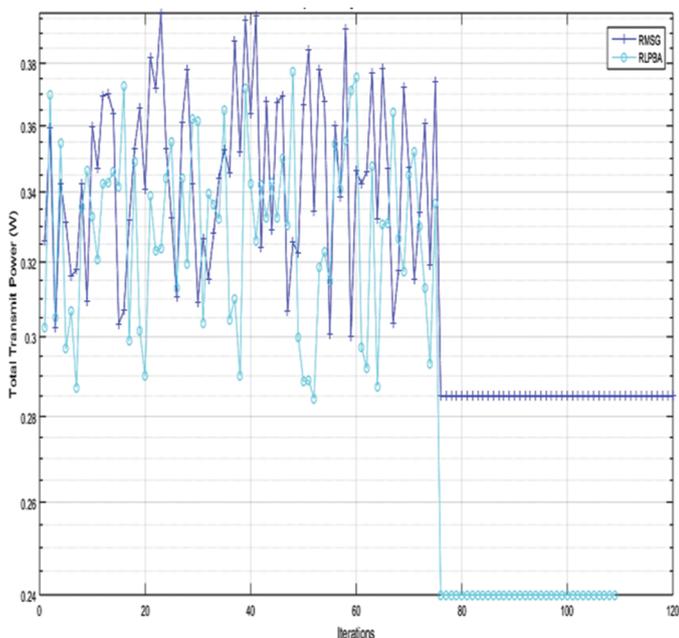
Notes. COPMA = cooperative power minimization algorithm; ECOPMA = enhanced co-operative power minimization algorithm; MIMO = multi-input multiple output; RLPBA = reinforcement learning based power allocation and beamformer algorithm; RMSG = regretratching-based joint transmit beamformer and power selection game; SINR = signal to interference noise ratio.



**Figure 2** | Total transmit power versus iterations (N = 5) for cooperative methods evaluations.



**Figure 4** | Total transmit power versus iterations (N = 5).



**Figure 3** | Total transmit power versus iterations (N = 5) for non-cooperative methods evaluations.

methods (COPMA and ECOPMA) achieved the better total transmit power reduction performance compared to non-cooperative methods (RMSG and RLPBA). But the cooperative methods introduced the significant overhead compared to non-cooperative methods. The updating task needs less overhead for non-cooperative methods. The proposed cooperative and non-cooperative achieved the optimum performances compared to existing solutions in this paper.

Figure 5 demonstrate the power trajectories in ECOPMA for every user pair in network. As observed in figure, every user starts with maximum power levels initially (i.e., 100 mW), and then power is updated iteratively as per the behaviour of ECOPMA algorithm till the NE achieved.

Figure 6 demonstrates the variations in probability mass function (P.M.F) of RLPBA method computed by Eq. (13) after iterations 1, 12, 50, and 100 for single user. At the start, the user selects the transmit beamformers with equal probability, further it changes according to working RL.

### 3.2. Evaluations of 10-Pairs Wireless Ad hoc System

Table 2 shows the simulation parameters for large wireless ad hoc MIMO systems performance evaluations using the different

methods. As observed in Table 2, the network area, smoothing factor, and number of clusters changed as per the network topology compared to Table 1 parameters.

Similar to scenario of N = 5, we computed the performance of cooperative, non-cooperative, and all methods in Figures 7–9 respectively for total transmission power of network.

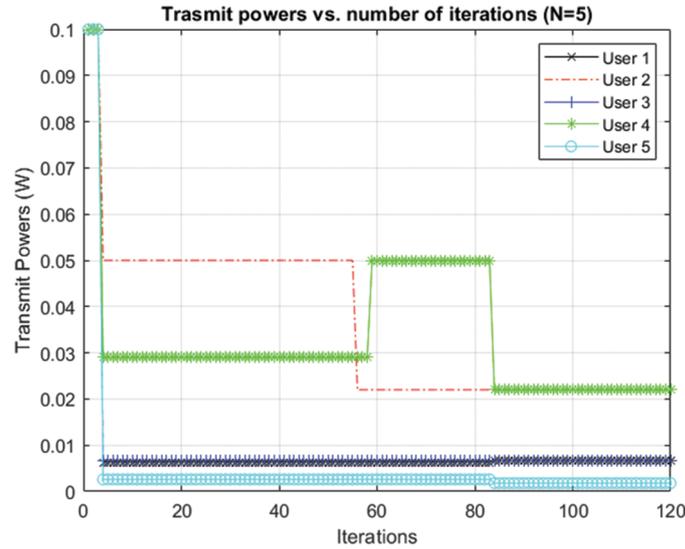


Figure 5 | Transmit power versus iterations with N = 5.

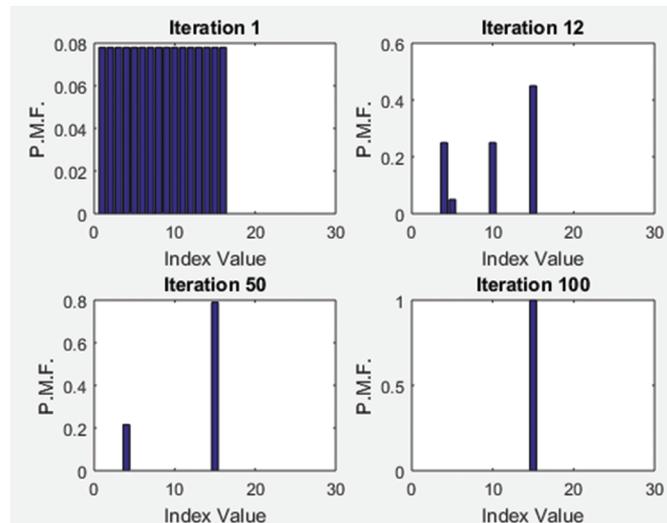
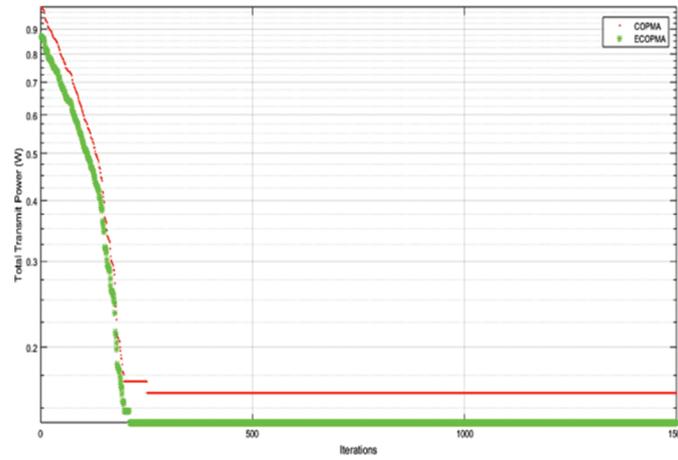


Figure 6 | Probability mass function (P.M.F.) of based power allocation and beamformer algorithm (RLPBA) method for single user (N = 5).

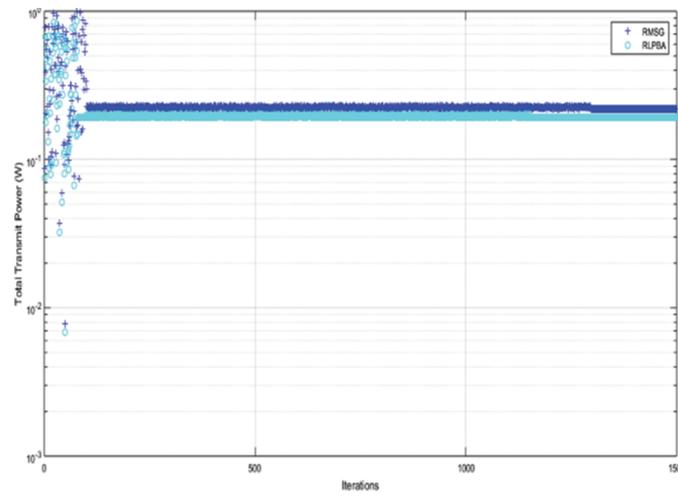
Table 2 | Simulation parameters for 10-Pairs wireless ad hoc MIMO system.

Parameter	Value
Number of wireless ad hoc pairs	10
Size of network	100 m × 100 m
Constant Received SINR	10 dB
Grassmannian codebook	( $\Upsilon = 16, A = 3$ )
Binary codebook	( $\Upsilon = 16, A = 4$ )
$E_{max}$	100 mW
$E_{min}$	1 mW
Propagation channel	Radio propagation channel (path loss component is 4)
Predefined number of iterations ( $k$ )	1500
Smoothing factor (For COPMA and ECOPMA)	$200/k^2$
Clusters (ECOPMA)	4
Methods Investigated	Centralized, COPMA, RMSG, ECOPMA, and RLPBA.

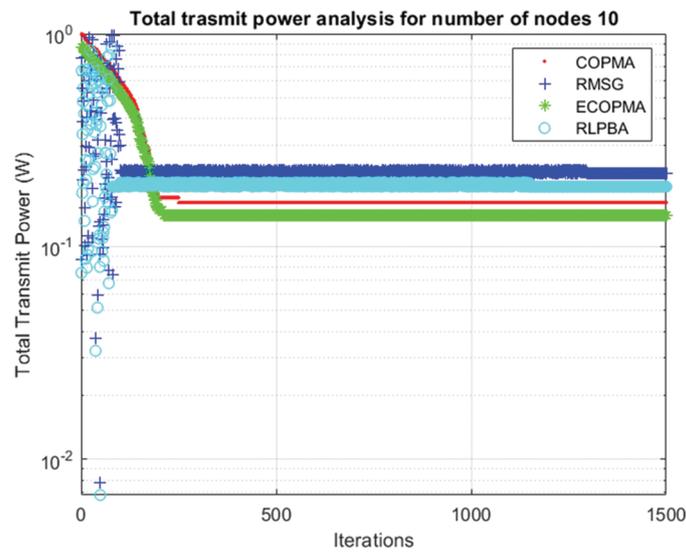
Notes. COPMA = cooperative power minimization algorithm; ECOPMA = enhanced co-operative power minimization algorithm; MIMO = multiple input multiple output; RLPBA = reinforcement learning based power allocation and beamformer algorithm; RMSG = regretmatching-based joint transmit beamformer and power selection game; SINR = signal to interference noise ratio.



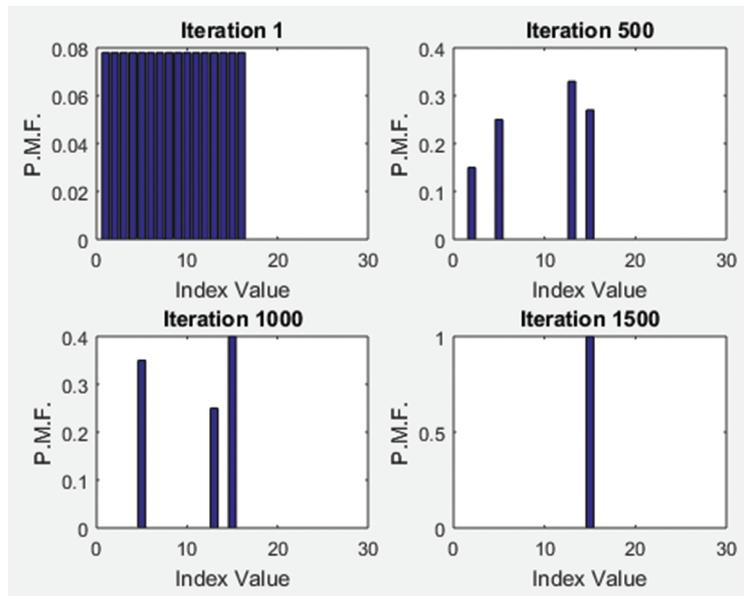
**Figure 7** | Total transmit power versus iterations (N = 10) for cooperative methods evaluations.



**Figure 8** | Total transmit power versus iterations (N = 10) for non-cooperative methods evaluations.



**Figure 9** | Total transmit power versus iterations (N = 10).



**Figure 10** | Probability mass function (P.M.F.) of based power allocation and beamformer algorithm (RLPBA) method for single user (N = 10).

The results demonstrate that using the proposed cooperative and non-cooperative methods, we achieved the total transmission power minimization and overhead reduction compared to existing methods cooperative and non-cooperative methods. A proposed solution reduces the number of iterations to achieve the NE and hence delivers the highest throughput compared to existing techniques. The centralized approach is no longer feasible in this scenario due to the enormous strategy space required [25]. As seen in Figure 9, the cooperative methods achieved the reduction in total transmit power as compared to non-cooperative as they converge early. The non-cooperative methods takes too much iterations to achieve the NE step, hence failed to reduce the transmit power, but it take the minimum overhead compared to co-cooperative solutions. For N = 10, the proposed solutions further optimize the cooperative methods. The RLPBA reduced the number of iterations for convergence and hence total transmit power reduction compared to RMSG. Figure 10 demonstrates the variations in P.M.F. of RLPBA method computed by Eq. (13) after iterations 1, 500, 1000, and 1500 for single user in this scenario as well.

The proposed solutions achieved the reduction in communication overhead compared to previous methods. The clustering with binary codebook for cooperative method helps to reduce the communication overhead as compared to COPMA and the RL helps to reduce communication overhead compared RMSG techniques.

#### 4. CONCLUSION AND FUTURE WORK

In this paper we initiated our study on design of beamformer selection and it's the important role for the adaptive wireless communication systems. We proposed two game theory based algorithms to optimize the performance of ad hoc wireless MIMO systems for joint transmit power and transmit beamforming with minimum power consumption and overhead. We designed cooperative (ECOPMA) solution using clustering and binary codebook and

non-cooperative (RLPBA) solution using the distributed reinforce learning in this paper. The simulation outcomes demonstrate the convergence properties of the proposed techniques and their performance in terms of overall power minimization, convergence rate, and communication overhead in network. The proposed solutions achieved the significant performance improved compared to state-of-art methods. For future work, we further suggest to investigate performance under different propagation channels and Signal to Interference Noise Ratio (SINR) constraints.

#### REFERENCES

- [1] A.J. Paulraj, D.A. Gore, R.U. Nabar, H. Bolcskei, An overview of MIMO communications—a key to gigabit wireless, *Proc. IEEE*. 92 (2004), 198–218.
- [2] D. Gesbert, S. Hanly, H. Huang, S.S. Shitz, O. Simeone, W. Yu, Multi-cell MIMO cooperative networks: a new look at interference, *IEEE J. Selected Areas Commun.* 28 (2010), 1380–1408.
- [3] H. Sampath, P. Stoica, A. Paulraj, Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion, *IEEE Trans. Commun.* 49 (2001), 2198–2206.
- [4] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, Uniform power allocation in MIMO channels: a game theoretic approach, *IEEE Trans. Inf. Theory*. 49 (2003), 1707–1727.
- [5] D.J. Love, R.W. Heath, T. Strohmer, Grassmannian beamforming for multiple-input multiple-output wireless systems, *IEEE Trans. Inf. Theory*. 49 (2003), 2735–2747.
- [6] F. Zhao, H. Wu, H. Chen, W. Wang, Game-theoretic beamforming and power allocation in MIMO cognitive radio systems with transmitter antenna correlation, *Mob. Inf. Syst.* 2015 (2015), 1–7.
- [7] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization, *IEEE Trans. Signal Process.* 51 (2003), 2381–2401.

- [8] D.P. Palomar, M.A. Lagunas, J.M. Cioffi, Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints, *IEEE Trans. Signal Process.* 52 (2004), 1179–1197.
- [9] M. Schubert, H. Boche, Solution of the multiuser downlink beamforming problem with individual SINR constraints, *IEEE Trans. Veh. Technol.* 53 (2004), 18–28.
- [10] K. Wong, R.D. Murch, K.B. Letaief, Performance enhancement of a multiuser MIMO wireless communication system, *IEEE Trans. Commun.* 50 (2002), 1960–1970.
- [11] B.-J. Lee, S.-L. Ju, Kim, N.-i. K.-S. Kim, Enhanced transmit-antenna selection schemes for multiuser massive MIMO systems, *Wirel. Commun. Mobile Comput.* 2017 (2017), 1–6.
- [12] M.C. Bromberg, B.G. Agee, Optimization of spatially adaptive reciprocal multipoint communication networks, *IEEE Trans. Commun.* 51 (2003), 1254–1257.
- [13] F. Rashid-Farrokhi, K.J.R. Liu, L. Tassiulas, Transmit beamforming and power control for cellular wireless systems, *IEEE J. Select. Areas Commun.* 16 (1998), 1437–1450.
- [14] R. Iltis, S. Kim, D. Hoang, Non-cooperative iterative MMSE beamforming algorithms for ad-hoc networks, *IEEE Trans. Commun.* 54 (2006), 748–759.
- [15] M.C. Bromberg, Optimizing MIMO multipoint wireless networks assuming Gaussian other-user interference, *IEEE Trans. Inf. Theory.* 49 (2003), 2352–2362.
- [16] M. Espinilla, J. Montero, J. Tinguaro Rodríguez, Computational intelligence in decision making, *Int. J. Comput. Intell. Syst.* 7 (2014), 1–5.
- [17] W. Saad, M. Bennis, *Game Theory for Future Wireless Networks: Challenges and Opportunities*, ICC, London, UK, 2015–Tutorial.
- [18] S. Buzzi, H.V. Poor, Joint receiver and transmitter optimization for energy efficient CDMA communications, *IEEE J. Sel. Areas Commun.* 26 (2008), 459–472.
- [19] C.U. Saraydar, N.B. Mandayam, D.J. Goodman, Efficient power control via pricing in wireless data networks, *IEEE Trans. Commun.* 50 (2002), 291–303.
- [20] G. Scutari, D.P. Palomar, S. Barbarossa, Optimal linear precoding strategies for wideband non-cooperative systems based on game theory part I: Nash equilibrium, *IEEE Trans. Signal Process.* 56 (2008), 1230–1249.
- [21] G. Scutari, D.P. Palomar, S. Barbarossa, Optimal linear precoding strategies for wideband non-cooperative systems based on game theory part II: algorithms, *IEEE Trans. Signal Process.* 56 (2008), 1250–1267.
- [22] Y. Song, C. Zheng, Y. Fang, Joint channel and power allocation in wireless mesh networks: a game theoretical perspective, *IEEE J. Sel. Areas Commun.* 26 (2008), 1149–1159.
- [23] E. Zeydan, D. Kivanc, U. Tureli, Joint iterative Channel Allocation and beamforming algorithm for interference avoidance in MIMO ad hoc networks, in *MILCOM 2007-IEEE Military Communications Conference*, Orlando, 2007, pp. 1–7.
- [24] E. Zeydan, D. Kivanc, U. Tureli, Joint iterative beamforming and power control for MIMO ad hoc networks, in *IEEE Global Telecommunications Conference GLOBECOM*, Miami, FL, 2010, pp. 1–5.
- [25] E. Zeydan, D. Kivanc, U. Tureli, C. Comaniciu, Joint iterative beamforming and power adaptation for MIMO ad hoc networks, *EURASIP J. Wirel. Commun. Netw.* 2011 (2011), 79.
- [26] N.C. Almeida, M.A.C. Fernandes, A.D.D. Neto, Beamforming and power control in sensor arrays using reinforcement learning, *Sensors (Basel)*. 15 (2015), 6668–6687.
- [27] N. Noori, S.S. Razavizadeh, A. Attar, Joint beamforming and power control in MIMO cognitive radio networks. *IEICE Elect. Express.* 7 (2010), 203–208.
- [28] J. Yoshida, A. Hirose, Beamforming for impulse-radio UWB communication systems based on complex-valued spatio-temporal neural networks, in *Proceedings of the International Symposium on Electromagnetic Theory*, Hiroshima, Japan, 2013, pp. 848–851.
- [29] Z.D. Zaharis, C. Skeberies, T.D. Xenos, P.I. Lazaridis, J. Cosmas, Design of a novel antenna array beamformer using neural networks trained by modified adaptive dispersion invasive weed optimization based data, *IEEE Trans. Broadcast.* 59 (2013), 455–460.
- [30] K. Terabayashi, A. Hirose, Ultra-short-pulse acoustic imaging using complex-valued spatio-temporal neural-network for null-steering: experimental results, in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Beijing, 2014, pp. 3410–3413.
- [31] A.F.M. Shahen Shah, H. Ilhan, U. Tureli, CB-MAC: a novel cluster-based MAC protocol for VANETs, *IET Intell. Trans. Syst.* 13 (2019), 587–595.
- [32] S. Abraham, D.C. Popescu, O.A. Dobre, Joint beamforming and power control in downlink multiuser multiple-input multiple-output systems, *Wirel. Commun. Mob. Comput.* 15 (2013), 552–560.
- [33] S.K. Varma V, G.V.S. Padma Rao, Design of Grassmannian weightbooks and binary weightbooks for MIMO beamforming systems, *Int. J. Adv. Res. Comput. Sci.* 4 (2013), 83–87.
- [34] M. Zeng, A. Yadav, O.A. Dobre, H. Vincent Poor, A fair individual rate comparison between MIMO-NOMA and MIMO-OMA, in *IEEE Globecom Workshops (GC Wkshps)*, Singapore, 2017.
- [35] J.R. Marden, G. Arslan, J.S. Shamma, Regret based dynamics: convergence in weakly acyclic games, in *AAMAS '07: Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, ACM, New York, 2007, pp. 1–8.
- [36] L. Wang, S. Fortunati, M.S. Greco, F. Gin, Reinforcement learning-based waveform optimization for MIMO multi-target detection, in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, 2018, pp. 1329–1333.
- [37] R.S. Sutton, A.G. Barto, *Reinforcement Learning: An Introduction*, A Bradford Book The MIT Press Cambridge, Massachusetts, London, England, 2014.