WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2011; **11**:1646–1656
DOI: 10.1002/wcm.1225

SPECIAL ISSUE PAPER

A game-theoretic approach for relay assignment over distributed wireless networks

Xuedong Liang[1]*, Min Chen[2] and Victor C. M. Leung[1]

[1] Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver V6T 1Z4, Canada
[2] School of Computer Science and Engineering, Seoul National University, Seoul, 151-742, Korea

ABSTRACT

For full cooperative diversity gains to be achieved while still maintaining spectral and energy efficiency, relay assignment schemes for cooperative communications have been extensively studied in recent research. These schemes select only the best relay from multiple relaying candidates to cooperate with a communication link. However, it is challenging to find the optimal relay in distributed wireless networks because of the dynamic nature of such networks. In this paper, we first formulate the problem of relay assignment as a noncooperative, mixed-strategy, repeated game, where relaying candidates are modeled as rational players. We then propose a game-theory-based relay assignment scheme *GTRA*, in which each player plays against all the other players and determines whether to cooperate with a communication link on a packet-by-packet basis in a distributed manner. To adapt to dynamic environments, players utilized an adaptive learning algorithm, that is, modified-regret-matching algorithm, to learn optimal strategies of relay assignment, as well as to orient the game to converge to a set of correlated equilibriums, which is often more system efficient than a Nash equilibrium. To evaluate the performance of *GTRA*, we compare it with *BR*, a fictitious two-player game-based approach. Simulation results have shown that *GTRA* outperforms *BR* in terms of network throughput, especially in environments where the channel fading becomes severe. It is also shown that *GTRA* can converge to a correlated equilibrium in a short period that enables the *GTRA* to work well in dynamic environments.

KEYWORDS
cooperative communications; relay assignment; game theory; wireless networks

***Correspondence**
Xuedong Liang, Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver V6T 1Z4, Canada.
E-mail: xuedongl@ece.ubc.ca

1. INTRODUCTION

Relay assignment [1] for cooperative communications [2,3], that is, dynamically choosing the best relay from multiple relaying candidates to cooperate with a communication link and thus forming multiple transmission paths or virtual multiple-input multiple-output systems to the destination, has emerged as an effective technique to improve the performance of wireless networks by exploiting the spatial diversity of the wireless medium. Optimal relay assignment can enable a cooperative communication system to achieve full diversity gains while still obtaining high spectral and energy efficiencies [4,5]. However, finding the optimal relay in distributed wireless networks is challenging, as the link qualities vary over time and the network topology may change.

A number of adaptive relay assignment schemes have been proposed recently. In the literature, most of the schemes perform relay assignment in a centralized manner and choose relays based on measured link metrics. For instance, the metrics, such as *distance* towards either the source or the destination [6–9], *signal-to-noise ratio (SNR)* [10,11], and *channel gain* [12–16] are used for relay selection, and these metrics are assumed to be available at the source, destination, and all the relaying candidates. However, there is often no centralized control in distributed wireless networks, and the measured metrics are often inaccurate and tend to vary in dynamic environments. Moreover, significant communication overhead will be incurred in acquiring and disseminating such information to all of the cooperative participants, especially for the cooperative protocols, as in [12,13], that instantaneous channel state information is required at all the candidates for performing relay selection.

For adaptive and distributed relay assignment schemes for wireless networks to be developed, game theory [17] based approaches [18] have received much research attention. In [19], game theory is used to model a wireless network consisting of selfish nodes, wherein a credit-based model involving exchanges of virtual currencies is proposed to manage node interactions. Whether a node cooperates with a communication link depends on the credit that can be earned and the resource needed for relaying a packet. The authors in [20] proposes a relay assignment scheme based on a two-level Stackelberg game, in which the source node and relay nodes are modeled as buyer and sellers, respectively. The proposed scheme jointly considers the benefit of both the buyer and the sellers and can achieve the best system performance with minimum power consumption. Chen and Kishore [21] present a game-theoretic analysis of a decode-and-forward [16] cooperative communication system over additive white Gaussian noise [22] and Rayleigh fading channels. The analysis shows that a mutually Nash equilibrium exists if a proper power control method is utilized and users care about their long-term performance. The game-theory-based schemes in the literature often assume that players have complete information of the game, that is, all players' identities, strategies, payoffs, and/or utility functions [20]. Furthermore, the game's history, for example, all actions have been taken by players in previous stages, and the the actions' corresponding outcomes are also assumed to be known to all players in a multistage game [19]. However, these assumption do not always hold in realistic scenarios, as nodes in wireless networks usually only have locally observed information and limited knowledge of others nodes' behavior. Therefore, adaptive learning, for example, estimating payoffs that may obtain by taking certain actions, predicting the other players' possible behavior should be involved in game designs [23].

In this paper, we propose a game-theory-based relay assignment scheme *GTRA*, in which the process of relay assignment is modeled as a noncooperative, mixed-strategy, repeated game and the relaying candidates are modeled as players. Packet transmissions between a source and a destination are modeled as repeated stages in the multistage game. At every stage of the proposed game, each player plays against all the other players and determines whether to cooperates with the communication link to assist the communication between the source and the destination in a distributed manner. To adapt to dynamic environments, players utilized a modified-regret-matching (MRM) algorithm [24] to learn optimal strategies and to orient the game to converge to a set of correlated equilibriums (CEs) [23]. To evaluate the performance of *GTRA*, we compare it with another game-theory-based approach *BR* [19], which models the process of relay assignment as a fictitious two-player game. Simulation results have shown that *GTRA* outperforms *BR* in terms of network throughput, especially in environments where the channel fading becomes more severe. It is also shown that *GTRA* can converge to a set of CE(s) in a short period, which enables the *GTRA* to fit well in dynamic environments.

The rest of the paper is organized as follows. We present the system model in Section 2. Section 3 formulates the problem of relay assignment as a mixed-strategy game and shows how to orient the game to converge to a set of CE(s) by using adaptive learning algorithms. Section 4 analyzes the *GTRA* algorithm's overhead and computational complexity. The performance analysis is presented in Section 5. Finally, Section 6 concludes the paper and discusses the future research directions.

2. SYSTEM MODEL

2.1. Cooperative communication model

As shown in Figure 1, we consider a wireless network consisting of uniformly and randomly distributed nodes, which are functionally equivalent in terms of radio communications, signal processing, and power supply. Multiple pairs of source and destination nodes are randomly selected for data packet transmissions. For a pair of source s and destination d, we assume that there exit a number of common-neighboring nodes N that are connected with both of s and d, as nodes are usually densely deployed. Therefore, a node, for example, node $n_i \in N$, may overhear the packet transmission between s and d because of the broadcast nature of the wireless medium. Node n_i may cooperate with the communication link between s and d by retransmitting the packet overheard from s.

We consider the use of decode-and-forward as the cooperative protocol, which operates in two phases, namely, direct transmission and relay transmission. In the direct transmission phase, the source transmits a packet to the destination and all the relaying candidates. In the relay transmission phase, a relay is chosen from the relaying candidates to retransmit the packet that overheard in the direct transmission phase to the destination. Then the destination combines the signals received from both the source and the relay and applies maximal ratio combining [25] for optimal packet decoding. For network operations and power consumption to be reduced, the relay transmission will be invoked only when a packet transmission

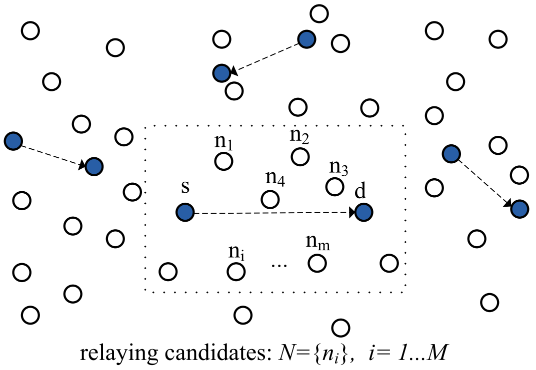

Figure 1. Cooperative communications with relay assignment.

fails in the direct transmission. We assume that the destination can feedback the improvement on received signals' SNR to the assigned relay by using acknowledgment (ACK) or negative-acknowledgment packets with extended data fields.

2.2. Wireless channel model

The wireless channel is modeled as a log-distance shadowing path loss channel that is defined as $PL(d) = PL(d_0) + 10 \cdot \eta \cdot \log(\mathrm{d}/\mathrm{d}_0) + X_\sigma$, where $PL(d)$ is the path loss at distance d, $PL(d_0)$ is the known path loss at a reference distance d_0, η is the path loss exponent, and X_σ is a zero-mean Gaussian random variable with standard deviation σ. For the links between two nodes, for example, i and j, the channel coefficient h_{ij} that captures the effects of path loss, shadowing, and fading are assumed to be reciprocal, that is, $h_{ij} = h_{ji}$. The channel coefficients are constant for a given transmitted block, or a codeword, but are independent and identically distributed for different blocks [11]. For different links, the channel fading coefficients are statistically independent and identically distributed, which is a reasonable assumption as nodes are usually spatially deployed [26].

3. ANALYSIS OF THE PROPOSED GAME

3.1. Description of the proposed game

The process of relay assignment is modeled as a noncooperative, mixed-strategy, repeated game G, which is denoted as $G = (N, A, U)$, where

- $N = \{n_1, \ldots, n_m\}$ denotes the set of players.
- $A = \{A_1, \ldots, A_m\}$ denotes the set of actions.
- $U = \{u_1, \ldots, u_m\}$ denotes the set of utility functions.

In the proposed game, the nodes are modeled as rational players, which means that the nodes are expected to follow a common set of strategies and choose actions from the strategies to maximize their utilities. In a distributed system, the nodes behaves selfishly, that is, a node always chooses actions to maximize its own utility, without considering the utilities of other nodes. For a player, for example, n_i, the strategy A_i consists of two actions, that is, $A_i = \{cc, ncc\}$, where cc denotes that the player n_i decides to cooperates with the communication link between s and d by retransmitting the overheard packet and ncc denotes that n_i chooses to remain silent. As a mixed-strategy game, each player takes its actions in accordance with a probability distribution. That is, the player n_i chooses the actions of cc and ncc with the probabilities of $p_i(cc)$ and $p_i(ncc)$, respectively, such that $p_i(cc) + p_i(ncc) = 1$.

Because the number of packets transmitted between the source and the destination is often assumed to be large, we model the process of relay assignment as a multistage game, and the process of relay assignment iterates in each stage of the game. Each candidate simultaneously plays the game against the other candidates in a distributed manner. Whether a candidate cooperates with a communication link or not depends on the payoff it may obtain. As the environment is assumed to be dynamic, each player, for example, n_i, updates its strategy by adjusting the probability distribution over the actions of cc and ncc, according to the link qualities between $s \rightarrow n_i$ (source–relay) and $n_i \rightarrow d$ (relay–destination).

Each packet transmission is modeled as a stage in the multistage game and consists of the following steps:

1. The source transmits a packet to the destination and all the players.
2. Each player chooses an action of cc or ncc autonomously and simultaneously, on the basis of the probability distribution of $p_i(cc)$ and $p_i(ncc)$, respectively.
3. Each player evaluates the quality of the selected action by computing the obtained payoff.
4. Each relay updates its strategy by adjusting the probability distribution of $p_i(cc)$ and $p_i(ncc)$.

3.2. Payoff calculation

In this game-theory-based approach, payoff is used to describe the difference between the benefit, that is, improvement on link quality achieved by relay retransmission, and the cost associated, that is, the channel occupancy time of a relay, and the energy consumed by the relay. The utility function designed to compute the payoff is defined as

$$u_i = \omega_1 \frac{SNR_{s,n_i,d} - SNR_{s,d}}{SNR_{s,d}} - (\omega_2 \frac{T_{n_i,t_x} - T_{n_i,r_x}}{T_{\mathrm{avr}}} + \omega_3 \frac{P_{n_i}}{P_m}) \tag{1}$$

where $SNR_{s,d}$ denotes the SNR of the received signal at the destination d that is transmitted by the source s in the direct transmission phase and $SNR_{s,n_i,d}$ denotes the SNR of the signal that combines the signals received from the source s and the selected relay n_i, respectively. T_{n_i,t_x} and T_{n_i,r_x} represent the packet retransmitting time and the packet receiving time at the relay n_i, respectively. The value difference between T_{n_i,t_x} and T_{n_i,r_x} reflects the packet processing, queuing, and channel access contention delays at node n_i. T_{avr} is the average amount of time needed for preparing a packet retransmission without considering processing, queuing, and channel access contention delays. T_{avr} is calculated as

$$T_{\mathrm{avr}} = T_{\mathrm{TA}} + T_{\mathrm{BO}} \tag{2}$$

where T_{TA} denotes the transceiver's receiving to transmitting turnover time, which is a constant value for a specific radio hardware, and T_{BO} is the average backoff time at n_i without any channel contention, and the value is determined by the underlying medium access control layer protocol.

P_{n_i} is the transmission power level of player n_i. P_m is the medium power level between $P_{\min}$ and $P_{\max}$, where $P_{\min}$ and $P_{\max}$ are the minimum and maximum available transmission power levels of player n_i, respectively.

ω_1, ω_2, and ω_3 are the weighting factors for the metrics of SNR, delay, and energy consumption, respectively. The values of the weighting factors can be adjusted to adapt to the quality of service requirements of the communication link.

In Equation (1), the first term represents the improvement (in percentage) on the link quality in terms of SNR by employing the relay transmission. The second term denotes the relative delay for preparing a packet retransmission, including processing, queuing, and channel access contention delays. The third term represents the relative energy efficiency, compared with using a fixed, medium transmission power level. If a player takes the action of cc, the achieved payoff computed by using Equation (1) reflects both the benefit and cost; otherwise, if a player chooses the action of ncc, the payoff is zero, as there is neither benefit nor cost caused by the action.

3.3. Correlated equilibrium

In a game, if players can receive a signal containing strategy recommendations from a central coordinator and all the players follow the recommendations on how to play the game, the outcome of the game can converge to a set of CE(s), which is often the most system efficient state of the game [27]. In a CE, it is the best interest of a player to follow the recommendations, that is, a player will not have a higher payoff by taking any other actions, provided all other players follow the recommendations. In the context of relay assignment, a CE can be interpreted as a *steady* state, as none of the nodes has incentive to unilaterally deviate from the recommendation profile to increase its payoff.

For a game, a probability distribution p is a CE of the game, if and only if $\forall\, n_i \in N, a_i \in A_i, \forall a_i^{'} \in A_i$; it holds that

$$\sum_{a_{-i} \in A_{-i}} p(a_i, a_{-i}) \left[u_i(a_i{}', a_{-i}) - u_i(a_i, a_{-i})\right] \le 0 \tag{3}$$

or equivalently,

$$\sum_{a_{-i} \in A_{-i}} p(a_{-i}|a_i) \left[u_i(a_i^{'}, a_{-i}) - u_i(a_i, a_{-i})\right] \le 0 \tag{4}$$

where A_i denotes n_i's action space and A_{-i} denotes the action space of n_i's opponents, that is, all player except n_i. a_i is the action chosen by n_i from A_i, and a_{-i} is the action combination of n_i's opponents. $u_i(a_i, a_{-i})$ is n_i's obtained payoff by taking the action of a_i, and its opponents taking the action combination of a_{-i}. $p(a_i, a_{-i})$ is the joint probability distribution over actions for all players. Equation (4) can be interpreted as that when the player n_i is recommended to choose action a_i; then choosing $a_i'(a_i' \in A_i, a_i' \neq a_i)$ will not lead to a higher payoff.

3.4. Convergence to correlated equilibrium

Orienting the relay assignment game to converge to a set of CE(s) is not trivial, as a central coordinator broadcasting recommendations on how to play the game is often not available in distributed wireless networks. A feasible approach [28] of orienting a game to a set of CE(s) is using the common history of the game as a coordinator. That is, a game can converge to a set of CE(s) if each player adjusts its strategy by tracking a series of *regret* values, which are quantitative measures for not taking certain actions in previous stages.

3.4.1. Calculation of regret values.

Regret matching [28], also called no-regret learning, can be used by players to calculate the regret values.

Assuming player n_i has taken action a_i at each of the past M stages, the difference of average payoff $D_i^M(a_i', a_i)$ between the player that has actually obtained and the player that would have obtained if it had taken the action a_i' instead of action a_i is defined as

$$D_i^M(a_i', a_i) = \frac{1}{M} \sum_{m \le M} (u_i^m(a_i', a_{-i}) - u_i^m(a_i, a_{-i})) \tag{5}$$

where $u_i^m(a_i', a_{-i})$ is the payoff the player would have obtained if the player had taken action a_i' at stage m and $u_i^m(a_i, a_{-i})$ is the payoff the player has actually obtained by taking action a_i at stage m.

For any two actions $a_i^{'}$ and a_i, the regret value for not taking action $a_i^{'}$ at the previous M stages is defined as

$$R_i^M(a_i', a_i) = \max\{D_i^M(a_i', a_{-i}), 0\} \tag{6}$$

The regret value is proportional to the difference of the average payoffs and is lower bounded by zero to ensure that the probabilities of taking any actions are positive. If the regret value is zero, it means that the player has obtained a higher payoff by taking action a_i than taking $a_i^{'}$, and thus, there is no regret. Otherwise, if the regret value is greater than zero, it means that the player would have obtained a higher payoff if the player had taken action $a_i^{'}$.

As observed in Equations (5) and (6), the implementation of regret matching requires that the player should know the payoffs it would have obtained if its actions in previous stages had been different from the actions that the player has actually taken. However, it is difficult for a player to compute the payoff that the corresponding action has not been taken. A common approach [27] in the literature is that using a coordinator broadcasting references to all the players at each stage of the game, indicating the potential payoffs that players would have obtained if the players had taken certain actions in the previous stage. By observing the references, the players can compute the payoff differences and thus calculate the regret values. However, this approach is often not feasible in distributed wireless networks wherein central coordinating is often not available.

Modified regret matching has been proposed for players to estimate the payoffs that the corresponding actions have not been actually taken, by only using the available historical information. That is, a player only need to know the probability distribution over its actions, and the payoffs it has obtained in the previous stages, to determine the probabilities of actions from the actually realizations only. MRM can be interpreted as a *reinforcement* or *stimulus-response* mechanism, as in the procedure of MRM, a relative high payoff at stage m will tend to increase the probability of playing the same action at stage $m+1$.

In MRM, the difference of the average payoffs of player n_i would have obtained if it had taken the action $a_i^{'}$ every time in the previous stages instead of taking a_i is defined as

$$C_i^M(a_i', a_i) = \frac{1}{M} \sum_{m \le M} \frac{p_i^m(a_i')}{p_i^m(a_i)} u_i^m(a_i, a_{-i}) - \frac{1}{M} \sum_{m \le M} u_i^m(a_i, a_{-i}) \quad (7)$$

where $p_i^m(a_i^{'})$ and $p_i^m(a_i)$ denote the probabilities of taking $a_i^{'}$ and a_i at stage m, respectively. The first term in Equation (7) is an estimation of the average payoff the player would have obtained if player n_i had taken the action of $a_i^{'}$ in the previous M stages; the second term is the average payoff that player n_i has actually obtained by taking action a_i at every stage in the past M stages.

The modified regret value of not playing $a_i^{'}$ is defined as

$$Q_i^M(a_i^{'}, a_i) = \max(C_i^M(a_i^{'}, a_i), 0) \quad (8)$$

Equations (7) and (8) show that a player can estimate its regret value of not taking a certain action by only using locally available information, that is, the probabilities of actions, and the payoffs actually obtained in the previous stages.

3.4.2. Strategy update: adjusting probability distribution of actions.

To maximize payoffs in a multistage game, a player updates its strategy by adjusting the probability distribution over different actions throughout the game. In MRM, a player, for example, n_i, updates its probabilities of actions at stage $m+1$ as

$$\begin{aligned} p_i^{m+1}(a_i^{'}) &= (1 - \frac{\delta}{m^\gamma}) \min(\frac{1}{\mu} Q_i^m(a_i^{'}, a_i), \frac{1}{K_i - 1}) \\ &\quad + \frac{\delta}{m^\gamma}\frac{1}{K_i}; \\ p_i^{m+1}(a_i) &= 1 - \sum_{a_i^{'} \neq a_i} p_i^{m+1}(a_i^{'}) \end{aligned} \quad (9)$$

where $a_i^{'}$ and $a_i \in A_i$ and $a_i^{'} \neq a_i$, K_i is the number of actions available for player n_i, and μ is a sufficiently large number that controls the probabilities of action switching and convergence speed. For the game converge to a limit CE to be ensured, the term δ/m^γ decreases to zero as m increases, where $0 < \delta < 1$ and $0 < \gamma < 0.25$. More details on parameter settings can be found in [24].

Equation (9) can be interpreted as follows. The first term, with weight $(1-(\delta/m^\gamma))$, denotes the modified regret value of not taking $a_i^{'}$ and shows how strongly the player intends to switch its action from a_i to $a_i^{'}$ at the next stage $m+1$. The second term, with weight δ/m^γ, denotes the uniform distribution over the available actions of n_i. This uniform distribution guarantees that all possible actions at stage $m+1$ can be taken with the probabilities of δ/m^γ at least. The first term is upper bounded to ensure that the sum of the probabilities does not exceed one.

The mathematic property of Equation (9) shows that if an action can have a relative high payoff at stages m, then the belief of taking the same action at $m+1$ is *reinforced*. For an action, a higher payoff will generate a greater reinforcement. All the effects, for example, belief reinforcement, action switching, decrease with the evolution of the game, as m increases over time.

3.5. *GTRA* algorithm based on modified-regret-matching learning

In the relay selection game, for player n_i, the probabilities of taking actions cc and ncc at stage m are denoted as $p_i^m(cc)$ and $p_i^m(ncc)$, respectively. The pseudocode of the GTRA algorithm implemented at n_i is listed in Algorithm 3.5.

3.6. Convergence of the GTRA algorithm

For the number of stages M, the relative frequency of players' action a played until M stages is defined as

$$z_M(a) = \frac{1}{M} |\{m \le M : a_m = a\}| \quad (10)$$

where a_m denotes all the users' action at stage m. It has been proved in [24] that z_M is guaranteed to converge almost surly at $m \to \infty$ to the set of CE(s) of the game G, if each player plays according to the adaptive MRM learning procedure, and adjusts its probability distribution

Algorithm 1 The GTRA algorithm based on MRM learning

```
begin
initialization
    Generate an arbitrarily probability distribution of $p_i^0(cc)$ and $p_i^0(ncc)$, and $p_i^0(cc) + p_i^0(ncc)$ =1.
for m=1,2,3, $\cdots$
    1. Compute the difference of average payoff $C_i^m(a_i^{'}, a_i)$ using (7).
    2. Compute the regret value $Q_i^m(a_i^{'}, a_i)$ using (8).
    3. Update the probability distribution $p_i^{m+1}(cc)$ and $p_i^{m+1}(ncc)$ at stage $m+1$ using (9).
    if $cc$ is the action chosen at each of the $m$ stages,
    then adjust the probabilities of taken actions of $ncc$ and $cc$ as:
        $p_i^{m+1}(ncc) = (1 - \frac{\delta}{m^\gamma}) \min(\frac{1}{u} Q_i^m(ncc, cc), 1) + \frac{\delta}{2m^\gamma}$;
        $p_i^{m+1}(cc) = 1 - p_i^{m+1}(ncc)$.
    else adjust the probabilities of taken actions of $cc$ and $ncc$ as:
        $p_i^{m+1}(cc) = (1 - \frac{\delta}{m^\gamma}) \min(\frac{1}{u} Q_i^m(cc, ncc), 1) + \frac{\delta}{2m^\gamma}$;
        $p_i^{m+1}(ncc) = 1 - p_i^{m+1}(cc)$.
end for
```

over actions as defined in Equation (9). As the number of packets transmitted from the source to the destination is often assumed to be large, that is, the number of stages is sufficiently large, the relay assignment game will converge to a set of CE(s).

4. ANALYSIS OF THE *GTRA* ALGORITHM'S OVERHEAD AND COMPLEXITY

4.1. Communication overhead

The information exchange in the *GTRA* algorithm, for example, the SNR's value of the packet received at the destination node needs to feedback to the node that retransmitted the packet, increases the number of information bits transmitted in the network. To reduce the communication overhead, we have extended the ACK packet, which already exists in the network, with new data fields to include the SNR information. The original ACK packet, with packet size of 5 bytes, is originally designed to indicate the successful packet transmission in the IEEE 802.15.4 standard. Adding a new data field with 1 byte to the ACK packet will not introduce much communication overhead, compared with the size of data packet with the maximum payload length of 127 bytes.

The SNR information feedback is the only communication overhead introduced by the *GTRA* algorithm, as the delay caused by packet processing, queuing, and channel access contention, is locally calculated at each node and does not require information exchange.

4.2. Computational complexity analysis

As a game-theoretic algorithm, computational overhead is inevitably introduced. To evaluate the algorithm's scalability, that is, how well the *GTRA* algorithm performs when the number of nodes in the network increases, we now evaluate the algorithm's computational complexity.

There are mainly three kinds of computation in the *GTRA* algorithm, that is, payoff calculation, regret value calculation, and the update of actions' probability distribution. For the payoff calculation, as shown in Equation (1), the computational complexity is a constant value, that is, O(c). As the payoff calculation is a simple linear function and is locally calculated at the node that takes the action of packet retransmitting, the nodes that choose to remain silent just set the value to zero. The most intensive computation is involved in the regret value calculation, as nodes need to store all the the payoff values that were actually obtained in the previous stages, and uses the history information to calculate the average payoff differences. For instance, for the payoff difference $Q_i^M(a_i^{'}, a_i)$, as shown in Equations (7) and (8), to be calculated, a node needs to compute M times of the simple calculation of payoff difference. Therefore, the computational complexity is a linear function of the number of previous stages, that is, O(n). This means that the more number of packets transmitted in a communication link, the more memory is needed to store the payoff values that were obtained in the previous stages and the more time is needed to compute the average payoff differences. Therefore, it may happen that when the number of packets transmitted in the communication link increases over a threshold, a node's memory may overflow and/or the node cannot afford the heavy computation. A simple solution to solve the problem is that a node may maintain a *sliding window*, that is, a node only stores the most recently obtained payoff values within the *sliding window* and removes the history information outside the window. By doing so, nodes do not require a large memory and can effectively reduce the heavy computational burden. However, the *GTRA* algorithm may need more time to reach convergence because of loosing some history information. Therefore, tradeoff should be considered when choosing the size of the *sliding window*. For the update of the action's probability distribution, as shown in Equation (9), the computation complexity is a constant value, that is, O(c), as the computation is related to neither the number of packets transmitted in the communication link nor the number of nodes in the network.

5. PERFORMANCE EVALUATION

To study the performance of *GTRA*, we compare it with *BR*, which models the process of relay assignment as a fictitious game. In *BR*, a player deems all the other players in the game as a fictitious opponent and plays against the opponent. At every stage of the game, if a player estimates that the retransmission probability of the fictitious opponent at the stage is less than the retransmission probability at the mixed Nash equilibrium, the player will retransmit the packet; otherwise, the player will remain silent.

5.1. Simulation environment

We simulate a wireless sensor network where 20 sensor nodes are randomly distributed in a 50 m × 50 m area. A constant bit rate traffic with five packets per second is used as the communication pattern, and the source and destination nodes are chosen randomly in each simulation run. We define the network lifetime as the time when the first node exhausts its battery's energy. Table I lists the detailed simulation parameters.

The Castalia [29,30] wireless sensor network simulator, which is based on the OMNeT++ [31,32] discrete event simulation platform, is used as the simulation environment. The data link layer in Castalia is modified to facilitate maximal ratio combining and decoding, and we also extend the ACK packet with a new data field to feedback the received signals' SNR information.

Table I. Simulation parameters.

Parameters	Value
Wireless channel model	Log shadowing path loss channel
Path loss exponent	2.4
Channel deviation (in dB)	1
Collision model	Additive interference model
Physical and MAC layer	IEEE 802.15.4 standard
Packet length	50 bytes
Transmitting power level	[−25, −15, −10, −7, −5, −3, −1, 0] dBm
Node's initial energy	12 J
Data transmission rate	250 kbps
Simulation time	100 s
ω_1	0.5
ω_2	0.3
ω_3	0.2

MAC, medium access control.

5.2. Performance evaluation

Figure 2 illustrates the evolution of the *GTRA* game in which five players cooperating with a source-destination link, and δ and γ are set to 0.5 and 0.1, respectively. In this game, the total number of joint action space is 32, as each player chooses either cc or ncc in a distributed manner.

The result shows that a player takes its actions of cc and ncc with arbitrary probabilities in the beginning of the game. Then each player adjusts the probabilities of different actions by computing a series of regret values. After about 60 iterations, the *GTRA* game converges, that is, the joint probability distribution of players converge to a set of CEs. We can observe that the two joint actions, that is, (ncc, cc, ncc, ncc, ncc) and (ncc, ncc, ncc, ncc, cc) are chosen with probabilities of about 0.74 and 0.24, respectively. The other joint probability distributions are all of small values, that is, less than 0.01 (we only plot the curves of six joint probabilities because of limited space). The result can be interpreted as follows. A strategy recommendation signal, generated at each player by using the MRM algorithm based on historical information, recommends the players on how to play the relay assignment game. That is, with the probabilities of about 0.74 and 0.24, the signal recommends the player n_2 and n_5 to act as relays, respectively. At the same time, the signal recommends the other players to remain silent. The joint probabilities except $p(ncc, cc, ncc, ncc, ncc)$ and $p(ncc, ncc, ncc, ncc, cc)$ are all of very small values, which means that the possibilities of joint actions except the two joint actions (ncc, cc, ncc, ncc, ncc) and (ncc, ncc, ncc, ncc, cc) can be neglected. It can be also observed in Figure 2 that there are some deviations from the recommended strategies even after the game converges. The are two reasons for the strategy deviation. First, as a mixed-strategy game, a player takes its actions with a probability distribution, which means that the player's strategy is of a *probabilistic* nature throughout the game. The other reason is the result of applying the MRM-based algorithm. That is, a player takes each of its actions with the probability of $\delta/2m^{\gamma}$ at least at every stage, as shown in Equation (9), to ensure that all actions have chances to be evaluated. By doing so, a player explores the dynamic environment continuously.

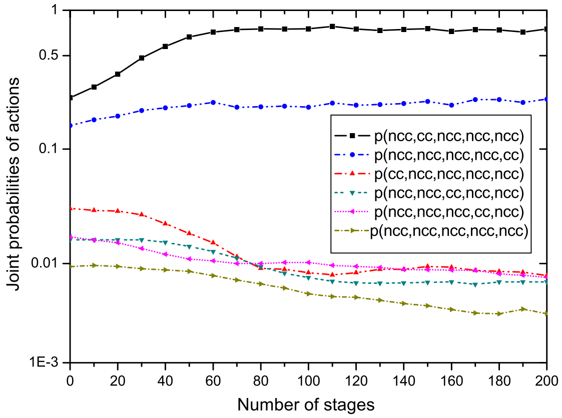

Figure 2. Evolution of the *GTRA* game.

For the performance of *GTRA* in a wireless channel with different fading to be investigated, the average network throughput versus the channel deviation is shown in Figure 3.

The simulation results show that *GTRA* outperforms *BR*, especially when the channel deviation σ becomes higher. We explain this as follows. The parameter of X_σ with standard deviation σ reflects the signal attenuation caused by the channel fading. That is, the higher the channel deviation σ, the more variation of the instantaneous strength of the received signals. In a wireless channel with higher variations, packets transmitted between a source and a destination are more likely to be corrupted. Therefore, it is more critical to choose the best relay to cooperate with the communication link to help the packet delivery. In *BR*, a player, for example, n_i, assumes that its opponent n_{-i}'s strategy is stationary and estimates n_{-i}'s possible behavior by simply tracking the frequency of actions that has been taken by n_{-i} in previous stages. This approach works well in static environments but does not fit in dynamic

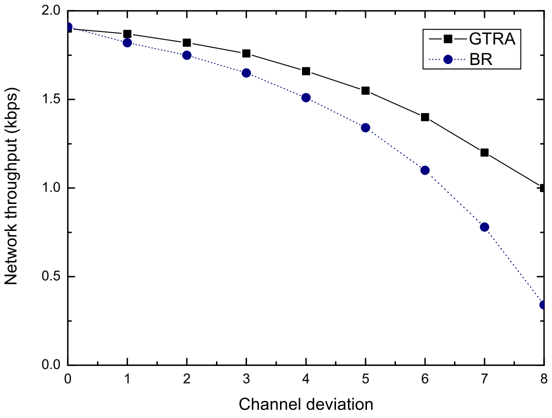

Figure 3. Network throughput versus channel deviation σ.

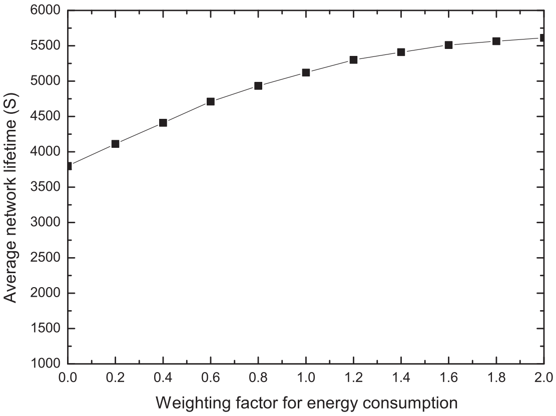

Figure 5. Network lifetime versus weighting factor for energy consumption.

environments. In contrast, *GTRA* is more adaptive in relay assignment, as players continuously evaluate the qualities of the actions that have been taken in previous stages, as well as evaluating the actions that have not been taken by using the adaptive learning algorithm MRM. Then players update their strategies by adjusting the probability distribution of the actions based on the payoff differences. Therefore, the flexible nature of the learning algorithm allows *GTRA* to adapt to dynamic environments, especially in networks with varying link qualities.

Figures 4 and 5 show the impacts of the weighting factor ω_3 for energy consumption on the performance of network throughput and lifetime, respectively.

Figure 4 shows that the network throughput decreases with the increment of ω_3. The reason is that when ω_3 is small, the benefit of acting as a relay (SNR improvement) outweighs the cost (energy consumption); thus, players tend to retransmit packets to obtain higher payoffs that leads to a better performance on network throughput. However, when the ω_3 increases over a certain value, the cost becomes a dominating factor in computing the payoffs. For higher payoffs to be obtained, players tend to remain silent instead of retransmitting packets, which results in a lower network throughput.

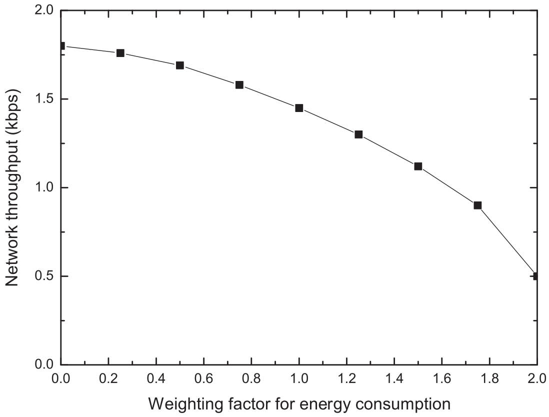

Figure 4. Network throughput versus weighting factor for energy consumption.

Figure 5 illustrates that the network lifetime always increases with the increment of ω_3. The reason is that when ω_3 becomes sufficient large, the cost is an important factor in payoff computing. Thus, all players tend to remain silent instead of retransmitting packets, which leads to a longer network lifetime. However, as observed in Figures 4 and 5, the longer lifetime is achieved by sacrificing the performance on network throughput. Therefore, a tradeoff must be considered when choosing the value of the weighting factor ω_3.

6. CONCLUSIONS AND FUTURE RESEARCH

In this paper, we have studied the problem of relay assignment for cooperative communications and have formulated the problem as a noncooperative, mixed-strategy, and repeated game, in which each player plays against all the other players and determines whether to cooperate with a communication link on a packet-by-packet basis in a distributed manner. For optimal cooperating strategies in dynamic environments to be learned, the MRM adaptive learning algorithm has been implemented at each player to adjust the probability distribution over actions, as well as orienting the game to converge to a set of CE(s). Simulation results have shown that *GTRA* outperforms *BR* in terms of network throughput and can converge in a short period that enables it to work well in dynamic environments.

In future research, we will examine the issue of system fairness to ensure that each node achieves an effort balance and to receive a fair share of the channel access in both retransmitting packets for other nodes and sending its own packets. Furthermore, we will also consider employing a power allocation scheme to prolong the network lifetime, as well as reducing concurrent transmission interferences to improve the network performance.

ACKNOWLEDGEMENTS

This work was supported in part by the Canadian Natural Sciences and Engineering Research Council under grant RGPIN 44286-09. This work was also supported in part by The NAP of Korea Research Council of Fundamental Science & Technology and The MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0004)).

REFERENCES

1. Sadek AK, Su W, Liu KJR. Multinode Cooperative Communications in wireless networks. *IEEE Transactions on Signal Processing* 2007; **55**: 341–355.
2. Nosratinia A, Hunter TE, Hedayat A. Cooperative communication in wireless networks. *IEEE Communications Magazine* 2004; **42**: 74–80.
3. Hong Y-W, Huang W-J, Chiu F-H, Kuo C-CJ. Cooperative communications in resource-constrained wireless networks. *IEEE Signal Processing Magazine* 2007; **42**: 47–57.
4. Feng X, Lau FCM, Zhou QF, Dian-Wu Y. Outage performance of cooperative communication systems using opportunistic relaying and selection combining receiver. *IEEE Signal Processing Letters* 2009; **16**(2): 113–136.
5. Xuedong L, Balasingham I, Leung VCM. Cooperative communications with adaptive relay selection for qos provisioning in wireless sensor networks, In *The 52nd IEEE Global Communications Conference (GLOBECOM'09)*, Honolulu, HI, 2009; 1–8.
6. Jakllari G, Krishnamurthy SV, Faloutsos M, Krishnamurthy PV, Ercetin O. A cross-layer framework for exploiting virtual MISO links in mobile ad hoc networks. *IEEE Transactions on Mobile Computing* 2007; **6**: 579–594.
7. Chen M, Liang X, Leung VCM, Balasingham I. Multi-hop mesh cooperative structure based data dissemination for wireless sensor networks, In *The 11th International Conference on Advanced Communication Technology (ICACT'09)*, Phoenix Park, 2009; 102–106.
8. Chen M, Kwon T, Mao S, Yuan Y, Leung VCM. Reliable and energy-efficient routing protocol in dense wireless sensor networks. *International Journal on Sensor Networks* 2008; **4**(1/2): 104–117.
9. Liang X, Chen M, Xiao Y, Balasingham I, Leung, VCM. A novel cooperative communication protocol for QoS provisioning in wireless sensor networks, In *The 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops (TridentCom'09)*, Washington, DC, 2009; 1–6.
10. Lin Y, Song J-H, Wong VWS. Cooperative protocols design for wireless ad-hoc networks with multihop routing. *Mobile Networks and Applications* 2009; **14**(2): 143–153.
11. Hunter TE, Nosratinia A. Distributed protocols for user cooperation in multi-user wireless networks, In *The IEEE 47th Annual Global Telecommunications Conference (GLOBECOM'04)*, Dallas, Texas, USA, 2004; 3788–3792.
12. Ibrahim AS, Sadek AK, Su W, Liu KJR. Cooperative communications with relay-selection: when to cooperate and whom to cooperate with? *IEEE Transactions on Wireless Communications* 2008; **7**(7): 2814–2827.
13. Bletsas A, Khisti A, Reed DP, Lippman A. A simple cooperative diversity method based on network path selection. *IEEE Journal on Selected Areas in Communications* 2006; **24**(3): 659–672.
14. Zhou Z, Zhou S, Cui J-H, Cui S. Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks. *IEEE Transactions on Wireless Communications* 2008; **7**(8): 3066–3078.
15. Bletsas A, Shin H, Win MZ. Cooperative communications with outage-optimal opportunistic relaying. *IEEE Transactions on Wireless Communications* 2007; **6**(9): 3450–3460.
16. Laneman JN, Tse DNC, Wornell GW. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Transactions on Information Theory* 2004; **50**(12): 3062–3080.
17. Felegyhazi M, Hubaux JP. Game theory in wireless networks: a tutorial. *Technical Report: LCA-REPORT-2006-002, EPFL*, 2006.
18. Maskery M, Krishnamurthy V, Zhao Q. Decentralized Dynamic spectrum access for cognitive radios: cooperative design of a non-cooperative game. *IEEE Transactions on Communications* 2009; **57**(2): 459–469.
19. Sergiilia S, Pancaldi F, Vitetta GM. A game theory approach to selection diversity in wireless adhoc networks, In *The 8th 2009 IEEE international conference on Communications (ICC'09)*, Dresden, Germany, 2009; 4184–4189.
20. Wang B, Han Z, Liu KJR. Distributed relay selection and power control for multiuser cooperative communication networks using Stackelberg game. *IEEE Transactions on Mobile Computing* 2009; **8**(7): 975–990.
21. Chen Y, Kishore S. A game-theoretic analysis of decode-and-forward user cooperation. *IEEE*

Transactions on Wireless Communications 2008; **7**(5): 1941–1951.

22. Simon MK, Alouini M-S. *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*. John Wiley & Sons, Inc: Hoboken, NJ, USA, 2000.
23. Wang B, Wu Y, Liu KJR. Game theory for cognitive radio networks: an overview. *Computer Networks* 2010; **54**(14): 2537–2561.
24. Hart S, Mas-Colell A. A reinforcement procedure leading to correlated equilibrium. In *Economic Essays: A Festschrift for Werner Hildenbrand*, Debreu G, Neuefeind W, Trockel W (eds). Springer: Berlin, 2001; 181–200.
25. Brennan DG. Linear diversity combining techniques. *Proceedings of the IEEE* 2003; **91**(2): 331–356.
26. Sadek AK, Wei Y, Liu KJR. When does cooperation have better performance in sensor networks? In *The 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON'06)*, Reston, VA, 2006; 188–197.
27. Chen L. A distributed access point selection algorithm based on no-regret learning for wireless access networks, In *The 71st IEEE Vehicular Technology Conference (VTC2010-Spring)*, Taipei, 2010; 1–5.
28. Hart S, Mas-Colell A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 2000; **68**(5): 1127–1150.
29. Pham HN, Pediaditakis D, Boulis A. From simulation to real deployments in WSN and back, In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'07)*, Espoo, Finland, 2007; 1–6.
30. The Castalia Website. http://castalia.npc.nicta.com.au/, 2011.
31. Varga A. The OMNeT++ discrete event simulation system, In *Proceedings of The 15th European Simulation Multiconference (ESM'01)*, Prague, Czech, 2001; 319–324.
32. The OMNeT++ website. http://www.omnet++.org/, 2011.

theory and its applications, resource allocation and optimization, quality of service provisioning, formal modeling, and validation of wireless networks.