

Modeling and Stimulating Node Cooperation in Wireless Ad Hoc Networks

Abbas Arghavani, Mahdi Arghavani, Abolfazl Sargazi, and Mahmood Ahmadi

In wireless networks, cooperation is necessary for many protocols, such as routing, clock synchronization, and security. It is known that cooperator nodes suffer greatly from problems such as increasing energy consumption. Therefore, rational nodes have no incentive to cooperatively forward traffic for others. A rational node is different from a malicious node. It is a node that makes the best decision in each state (cooperate or non-cooperate). In this paper, game theory is used to analyze the cooperation between nodes. An evolutionary game has been investigated using two nodes, and their strategies have been compared to find the best one. Subsequently, two approaches, one based on a genetic algorithm (GA) and the other on learning automata (LA), are presented to incite nodes for cooperating in a noisy environment. As you will see later, the GA strategy is able to disable the effect of noise by using a big enough chromosome; however, it cannot persuade nodes to cooperate in a noise-free environment. Unlike the GA strategy, the LA strategy shows good results in a noise-free environment because it has good agreement in cooperation-based strategies in both types of environment (noise-free and noisy).

Keywords: Cooperation, game theory, genetic algorithm, learning automata, prisoner's dilemma model, wireless network.

Manuscript received Dec. 18, 2013; revised May 8, 2014; accepted May 28, 2014.

Abbas Arghavani (corresponding author, a.arghavani@yahoo.com), Mahdi Arghavani (ma.arghavani@kssriau.ac.ir) are with the Department of Information Technology, Applied Science and Technology University, Ravansar Branch, Iran.

Abolfazl Sargazi (a.sargazi@qiau.ac.ir) is with the Department of Computer Engineering and Information Technology, QIAU, Qazvin, Iran.

Mahmood Ahmadi (m.ahmadi@razi.ac.ir) is with the Faculty of Computer Engineering Department, Razi University, Kermanshah, Iran.

I. Introduction

Wireless networks suffer from environmental destructive phenomena, such as fading, blocking, and shadowing. To reduce the effects arising from these phenomena, methods under the title of *diversity* have been propounded and studied. In computer networks, diversity is a method for improving the quality of a signal and increasing the reliability of packet transfer. Different types of diversity have been propounded, such as space-time coding diversity and frequency diversity [1]–[6]. One of the profound methods in reducing the effect of the aforementioned phenomena is cooperative diversity (see also user cooperation diversity [7]–[8]), which has been studied in detail in [6]–[10]. In cooperative diversity, to achieve reliable communication and higher diversity gain, nodes share their antenna with each other. In fact, they create a system with virtual or distributed antenna diversity.

Cooperative communication can provide increased capacity and power savings in ad hoc networks. Cooperative communications in wireless ad hoc networks have gained a lot of interest recently in the research community. Cooperation is used in routing protocols [11], clustering [12], localization [13], and other protocols.

In this paper, the basic assumption is that the considered network includes a set of nodes that are assumed to be rational, autonomous, and selfish. A selfish node is a node that acts selfishly. A selfish node performs a function that leads to achieving a higher payoff [14] and prefers its advantages as opposed to those of the network. In essence, a selfish node does not feel inclined to forward the packets of other nodes because this leads to a use of its own resources (energy). Therefore, the main purpose of this paper is to provide a method that will stimulate selfish nodes to cooperate with each other so that they

may help to profoundly increase network throughput.

There are few researches that focus on the behavior of selfish nodes in cooperative networks. For example, [15] has considered the behavior of selfish nodes in ad hoc networks and has discussed a trust and reputation mechanism that would stimulate cooperation between such nodes.

In [16]–[23], by using virtual currency schemes and reputation-based schemes, rational nodes have been forced to cooperate with each other.

To date, a lot of problems in wireless networks have been analyzed through game theory. Due to the similarities in the behavior of the nodes and their decision-making trends, [24] has analyzed the cooperation of nodes through game theory and the prisoner's dilemma model. In [25], a prisoner's dilemma game has been used to model the behavior of nodes in multi-hop wireless networks. The basic assumption in [14] and [25] is carried over in this paper; that is, we assume all nodes to be selfish and rational.

In a wireless network in which cooperation is the founding principle behind many of its protocols, selfish nodes can decrease the overall performance of the network. Accordingly, the purpose of the proposed approach is summarized as follows:

- To model the behavior of nodes and their interactions with each other (cooperation and noncooperation) using game theory and an iterative prisoner's dilemma (IPD) game [24].
- To use a genetic algorithm (GA) and L_{RP} learning automata (LA) to encourage the nodes to cooperate with each other.

The remainder of this paper is organized as follows. The main problem is stated in Section II. Section III is devoted to describing the IPD game and evolutionary games, as well as evaluating several well-known solutions to the IPD game in the field of social and economics science. Section IV will present the GA- and LA-based approaches, whereby we can encourage the nodes to cooperate with each other. In Section V, we consider the domino effect, which is a chain reaction; that is, when a small change instigates a continuing chain of similarly small changes.

II. Problem Definition

As we mentioned in the previous section, environmental destructive phenomena have harmful effects on the performance of wireless networks, and cooperative diversity can reduce the impact of such effects. In this method, if a direct communication between two nodes is not possible, then these nodes can communicate with each other through their neighbors. For further explanation, let us now consider Fig. 1. Suppose that node i wishes to send a data packet to node D_i . Because of the obstacle between the two nodes, it is not

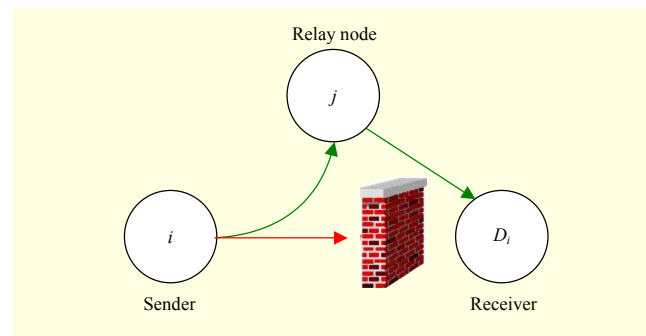


Fig. 1. Node i needs node j 's cooperation.

possible to send it directly. Cooperative diversity suggests that node j can act as a relay node, which means that it receives the data packet sent from i and then forwards it toward D_i .

As we know, in a traditional network, node j has to cooperate with node i . The modeling and monitoring of the behavior of autonomous nodes in a nontraditional network is our main goal in this paper. A nontraditional network comprises autonomous nodes that behave rationally. An autonomous node is a node that independently decides about its own actions, no matter its status. It should also be noted that a rational node is a type of selfish node, since it always chooses the action that is most beneficial to it from those that are available (that is, cooperate or non-cooperate). In other words, a rational node prefers its own profit over that of the network.

The main challenge in such networks is the cost of cooperation. It is obvious that helping node i incurs costs to cooperative node j . This cost is a function of the power consumed in forwarding the packet of node i . Although it will be low, node j — which is assumed to be autonomous and rational — will be discouraged from cooperating. In such a nontraditional network, nodes do not cooperate with each other because it is inconsistent with the assumption that all nodes are autonomous and rational.

On the other hand, as a consequence of the high volume of interactions between nodes in a network, a situation could arise whereby node j requires node i 's assistance. However, once again, because of the assumption that all nodes are autonomous and rational, node i will be discouraged from cooperating (Consider Fig. 2). In this scenario, for similar reasons, forwarding data packets from node j to D_j is not possible directly. Now, it has been expected that node i compensates node j 's behavior in the previous scenario, but the decision whether to cooperate is not that simple for node i , because the scenario explained in Fig. 1 may happen again (that is, it may require node j 's assistance once again).

As we mentioned above, the basic assumption in this paper is that the considered network includes a set of rational and autonomous nodes that always perform an action that benefits

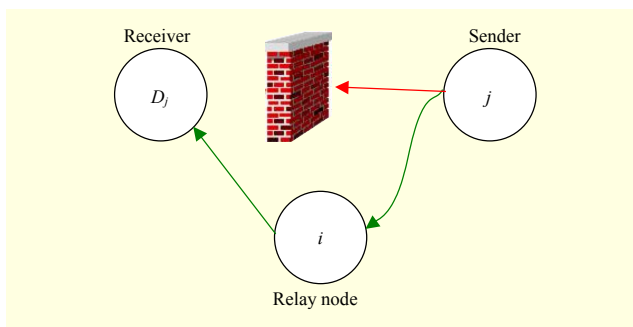


Fig. 2. Node j needs node i 's cooperation.

them. In a wireless network in which cooperation is the foundation of many protocols, this behavior will decrease the network's performance.

III. IPD Game, Evolutionary Games, and Strategies

Game theory is a bag of analytical tools designed to help us understand the phenomena that we observe when decision-makers interact. The basic assumptions that underlie the theory are that decision-makers pursue well-defined exogenous objectives (they are rational) and that they take into account their knowledge or expectations of other decision-makers' behavior (they reason strategically) [24].

The main characteristic of making a "decision" in a game is that before players select and make a choice, they should analyze the reactions to the potential choices of the other game players, and then having done so, the players making the decisions should choose the ones best suited to them based upon the potential reactions [26]. An environment in which there is such a reciprocal effect and reaction among individuals' decisions, is called a strategic environment.

1. IPD Game

The IPD game, which has been studied in [27]–[32], is an example of a non-zero sum game. In non-zero sum games, the profit of one player does not necessarily incur a loss to the other, and in some situations, it may be the case that both of them may profit. It has been considered that cooperation can be ascertained only in non-zero sum games, because only in such games can we expect more points for all players who participate in a cooperation. In other words, all players can gain more points through cooperation. The best way to illustrate the cooperation of two players is to use the IPD game. In the IPD game, there are two players. Each has two choices; namely, cooperate (C) or defect (D). Each player should make their choice without knowing what the other will do. No matter what the other one does, defection yields a higher payoff than

Table 1. Payoff matrix.

	D	C
D	P, P	T, S
C	S, T	R, R

Table 2. Normalized payoff matrix.

	D	C
D	$0, 0$	$\alpha, -\beta$
C	$-\beta, \alpha$	$\alpha - \beta, \alpha - \beta$

cooperation. This results in a dilemma, for if both players defect, then they will both lose out. This simple model game will provide the basis for the entire analysis used in this paper.

A game is shown in Table 1. One player chooses a row, either cooperating or defecting. The other player simultaneously chooses a column, either cooperating or defecting. Together these choices result in one of the four possible outcomes shown in this table. If both players cooperate, then both do fairly well and get R points, which is the reward for mutual cooperation. If one player cooperates, but the other defects, then the defecting player gets T points, while the other gets S points. If both defect, then both get P points.

The payoff matrix of for the IPD game must satisfy the following two conditions related to the payoff values for different actions:

- $T > R > P > S$.
- $2 \times R > (T + P)$.

It is obvious that irrespective of what their partner chooses, a player always has an incentive to choose D over C. Thus, this game has a unique steady-state equilibrium solution or Nash equilibrium, [33], which is (D, D). In game theory, the Nash equilibrium is a solution concept of a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing their own strategy unilaterally. If each player has chosen a strategy and no player can benefit by changing his or her strategy while the other players keep theirs unchanged, then the current set of strategy choices and the corresponding payoffs constitute a Nash equilibrium.

In wireless ad hoc networks, each node has to decide whether to forward or drop packets belonging to other nodes. Since forwarding packets consumes a node's energy, this cost is considered as β . Furthermore, when a node's packet is relayed by its neighbor, it receives α points as the reward (see

Table 2). As in [14], here it is assumed that any two neighbors have uniform network traffic and that any two nodes can send packets to each other while simultaneously deciding whether to forward or drop any that they may be about to receive. This decision-making process is repeated multiple times by all nodes in the system.

Now, we move on to consider a simplified situation. Consider a packet-relaying game that is divided into n rounds. This packet-relaying game is the same as the repeated IPD game. It is played several times, and in each round, actions and outcomes of previous rounds are observable. In these games, by considering the history of the opponent's actions, players are encouraged to cooperate with each other.

2. Evolutionary Games

Evolutionary games have been applied most widely in the area of evolutionary biology — the domain in which the idea was first articulated by J.M. Smith and G.R. Price. Evolutionary games are based on the idea that the strategy that best fits a given scenario will be the one to produce more offspring.

Suppose there are M strategies in an environment. At generation n , each strategy is represented by a certain number of players; that is, as $W^n(S_i)$, where S_i is one of the M strategies and $W^n(S_i)$ represents the number of players (nodes) playing strategy S_i in the n th generation.

The score of strategy S_i when it plays with strategy S_j is represented by $V(S_i|S_j)$. This score is calculated from Table 1. Equation (1), below, is used for computing each strategy's score. The score of strategy S_i at the end of generation n is denoted by $g^n(S_i)$.

$$g^n(S_i) = \sum_{j=1}^M (W^n(S_j)V(S_i|S_j)) - V(S_i|S_i). \quad (1)$$

The size of strategy S_i in the $(n + 1)$ th generation, $W^{n+1}(S_i)$, is determined by its score in the n th generation.

$$W^{n+1}(S_i) = \frac{W^n(S_i)g^n(S_i)\sum_{j=1}^M W^n(S_j)}{\sum_{j=1}^M W^n(S_j)g^n(S_j)}. \quad (2)$$

3. Strategies

A strategy is an algorithm that determines what should be done by each player in any game. The objective of each strategy is to get more points for its player. A player's strategy in game theory refers to one of the actions he can choose in a setting where the outcome depends not only on his own actions but on the actions of others, and it will determine the action the player will take at any stage of the game.

A strategy profile (sometimes called a strategy combination)

is a set of strategies for each player, which fully specifies all actions in a game. A strategy profile should include one, and only one, strategy for every player. The strategy concept is sometimes wrongly confused with the concept of a move. A move is an action taken by a player at some point during the play of a game; (for example, in chess, moving White's Bishop from a2 to b3). On the other hand, a strategy is a complete algorithm for playing the game and guides a player toward what to do in every possible situation throughout the game.

A. Tit for Tat

Tit for tat (TFT) is probably the most studied strategy in game theory [14]. A player using this strategy will initially cooperate and then respond in kind to an opponent's previous action. If the opponent was previously cooperative, then the player will cooperate; otherwise, he will not. TFT was submitted by Anatol Rapoport in Axelrod's Tournament, in 1984. If a player continually uses a TFT strategy, then the player cannot score higher or lesser than its playing partner. It means TFT doesn't win or lose in any repeated game.

B. Gradual

In *gradual* strategy, a player uses cooperation on the first move and then continues to do so as long as its opponent cooperates. Then, after the first defection of the opponent, it defects once and cooperates twice; after the second defection of the opponent, it defects twice and cooperates twice, and so on and so forth, until after the N th defection it reacts with N consecutive defections and then calms its opponent down by cooperating twice [14].

C. ALLC Strategy

ALLC strategy is a strategy that cooperates in all games. ALLD strategy is a strategy that defects in all games, and in every game, *random* strategy selects one of the actions (cooperation or defection) with equal possibility.

D. Optimal Strategy

It is much less obvious as to what constitutes an *optimal* strategy or even how an optimal strategy should be defined [34]. The question of what makes a strategy optimal in the IPD game is very difficult to answer. As many researchers have noted, the performance of a strategy is highly dependent on which other strategies it interacts with. This has led to several conflicting definitions of the term "optimality," with resulting differences in which strategies (if any) are considered optimal. According to this, optimality has been defined relative to a given set of opponents: the optimal strategy is the one that achieves the highest score (with respect to some measure)

against that set of opponents. One typical measure of performance is the average score in a round-robin tournament interaction. In reality, there is no fixed strategy that performs best against every given set of opponents in this type of interaction, because achieving optimality requires the forecasting of opponents' actions before they act, and this is impossible [34]. So, the problem that we consider in this paper is to find a suboptimal strategy that will ensure victory in a packet-relaying tournament.

A good strategy is one that allows a player to remain alive in the population for the longest possible time and in the biggest possible proportion [35]. In evolutionary games, such a strategy is called an evolutionary stable strategy (ESS). A strategy is an ESS when the whole of the population is using this strategy. Any small group of invaders using a different strategy will eventually die off over multiple generations [24].

A strategy can be suboptimal provided that it has four characteristics [14]. These characteristics are as follows:

- It should begin by cooperating; that is, it is good.
- It should be able to retaliate; that is, it can return an opponent's defection.
- It should be generous; that is, it forgets the past if the defecting opponent cooperates again.
- It should not be memory-less; that is, it utilizes its interaction history.

IV. Proposed Strategies

To be optimal, a strategy should have the ability to predict. In other words, before playing with a strategy, a player should know the method of every strategy. For example, before playing with a TFT strategy, a player should know how to interact with this strategy to have the best play, which in this case is cooperation. As we have already said, planning such a strategy is impossible, but instead of predicting, a method can be suggested that tries to identify the behavior of the opponent's strategy by playing several games. In other words, our proposed strategies, which are based on GA and LA, try to perform several games and study the consequences arising from each game. Then, they analyze the behavior of the opponent's strategy and gradually incline toward the best performance in each game.

1. GA Strategy

In [36] and [37], a GA has been investigated completely. A GA is a heuristic search that mimics the process of natural evolution. This heuristic search is routinely used to generate useful solutions for optimization and search problems.

GAs belong to a larger class of algorithms called

evolutionary algorithms which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, selection, cross over, and so on.

A. GA and IPD

The important thing in solving IPD by GA is how to formulate the problem. Two genes have been supposed to exist in a game environment. One is cooperation, and the other is defection. Cooperation and defection are respectively indicated by a value of "0" and "1." An initial population consists of two chromosomes. We believe that the behavior of any particular given strategy should finally be inclined to either absolute cooperation or absolute defection. As a result, two chromosomes with a length of 16 genes are considered. One of them comprises all 0s, and the other comprises all 1s; that is, we have 0000000000000000 and 1111111111111111. The length of the chromosomes dictates the number of games that are spent analyzing the behavior of the opponent's strategy. Then, the proposed strategy plays with the above two chromosomes, which means that it continuously cooperates for 16 plays (the first chromosome) and then continuously does not cooperate for 16 plays (the second chromosome). The obtained score by each chromosome will be computed based on a fitness function. The fitness function that has been used here is the same as that in the game table of the prisoner's dilemma. Now, intercourses should be done. To choose two parents for intercourse, the elitist selection and roulette wheel methods have been used. The elitist method selects the best chromosome as one of the parents. The best chromosome is the one that has obtained the higher score. In the roulette wheel method, a chromosome's score dictates the probability that it is chosen. Finally, by these two methods, two parents will be chosen for generating future chromosomes.

A simple method, known as the one point crossover, has been used for generating new children. In this method, the

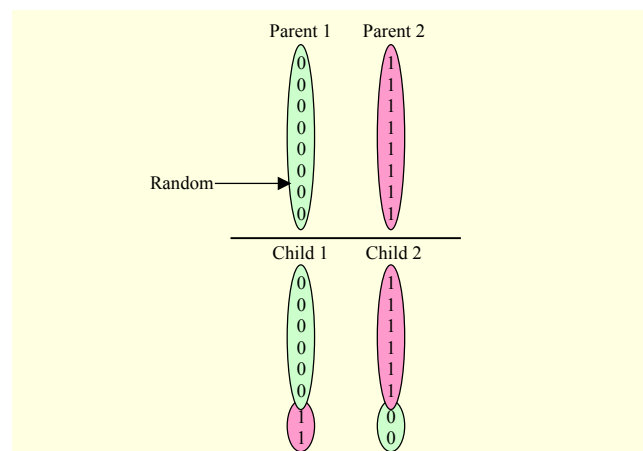


Fig. 3. Example of crossover between two chromosomes.

genes within either parent's chromosome are subject to exchange with one another (swap) at a random point within the given sequences. Figure 3 shows an example of a possible intercourse between two parents' chromosomes comprising eight genes. The proposed algorithm continues playing with the children's chromosomes. Then, the score of each chromosome is computed again, and the algorithm enters into a new phase of selecting new parents and generating new children. This process then continues repeatedly.

B. Simulation Results

One way of finding a superior strategy is to hold competitive interactions among strategies. At the end of every repetition, based on the evolutionary games law [(1) and (2)], the populations of superior strategies will increase and the populations of weak strategies will be decreased; thus, the upshot of this will be that the weak strategies will eventually leave the game as their generation becomes extinct. If a weak strategy is not able to gain a high payoff, then the population to which it belongs will decrease.

In the simulation phase, a fully connected network with 600 rational nodes is considered. IPD involves two players, and in a fully connected network with 600 nodes, $(600 \times 599) / 2 = 17,970$ games per each generation will be held to find the best strategy.

The network topology, packet size, energy, and so on do not influence a node's behavior when playing an IPD game. The following six strategies are to be used in the simulation: TFT, ALLC, ALLD, Random, Gradual, and GA. It is obvious that superior genes that obtain more scores will be in the later-generation chromosomes.

At first, no node knows the superior strategy; therefore, each node selects one of the strategies randomly and plays accordingly. So, it is assumed that an equal number of nodes will use each of the aforementioned six strategies in the first phase. When two nodes are playing the packet-relaying game, after each round, they should be aware of each other's actions.

It is supposed that each node becomes aware of the opposite node's behavior by a central or distributed mechanism. So, the correctness of the information provided by this mechanism is important because strategies like TFT and Gradual adjust their behavior based on this information.

An important feature of interactions in the real world is that choices cannot be implemented without error because the other player does not necessarily know whether a given action is an error or a deliberate choice. Noise, in the form of random errors, in implementing a choice is common in real-world interactions.

To date, different methods have been suggested for playing in noise-free environments, and these methods have often been successful, but playing in a noisy environment poses many

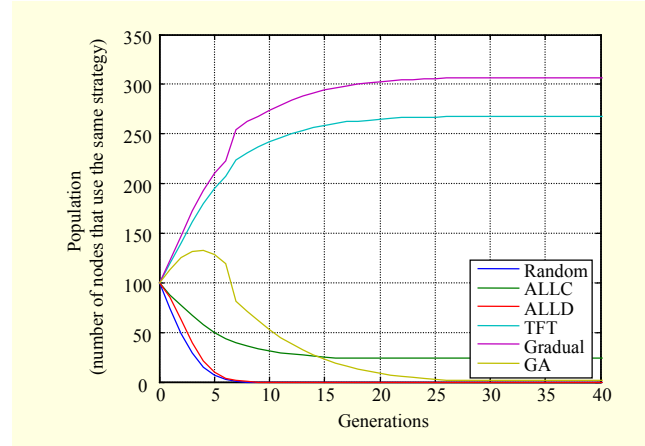


Fig. 4. Simulation with noise ratio $\beta = 0\%$.

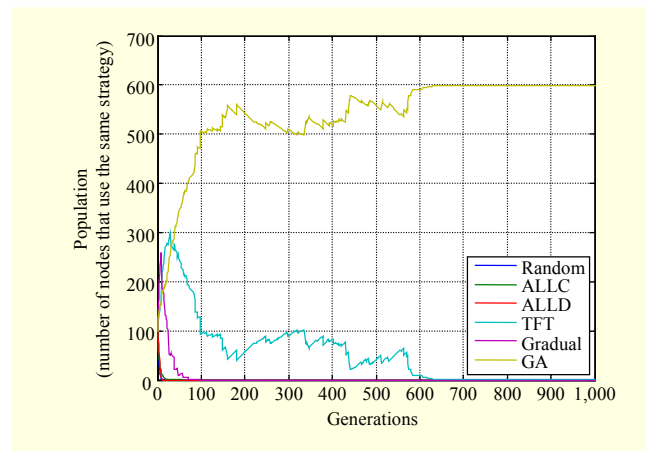


Fig. 5. Simulation with noise ratio $\beta = 5\%$.

challenges, and as of yet, no comprehensive solution has been suggested for such a case. In [37] and [38], some approaches related to noisy environments have been investigated.

A noisy environment is one that makes use of an informing mechanism that contains some error. For example, consider two nodes, i and j , which are playing the packet-relaying game. If the environment has a noise rate, say Φ , then it means that the action of each node with possibility of Φ is inversely informed to its opposite node. Figs. 4, 5, 6, and 7 show the results of holding the above tournament in both a noise-free environment and a noisy environment, with noise rates of 0%, 5%, 10%, and 20%, respectively. An interesting point here is that the GA strategy in an environment with a noise rate of 20% will act better than in an environment with less noise ratio, and unfortunately, it doesn't work well in a noise-free environment at all.

The reason for this is because as the noise ratio of an environment increases, the circumstances under which a strategy can survive become harsher, and so some strategies will ultimately be forced to leave the game environment. Thus,

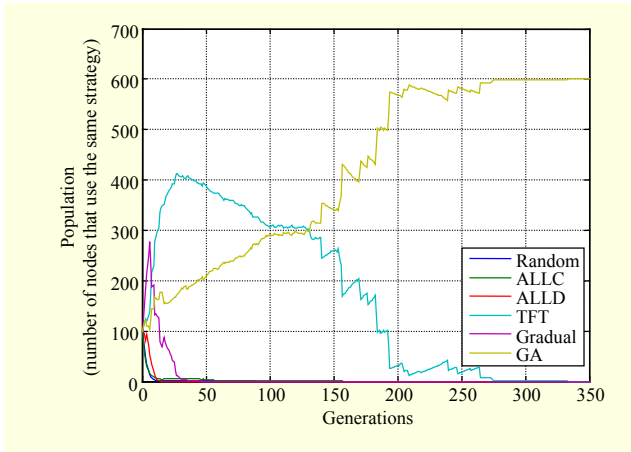


Fig. 6. Simulation with noise ratio $\beta = 10\%$.

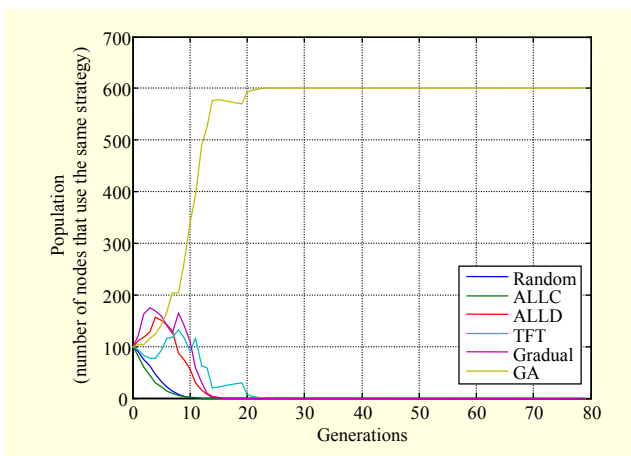


Fig. 7. Simulation with noise ratio $\beta = 20\%$.

when there is uncertainty in a node's behavior because of noise existence, then the GA strategy helps its user to take more benefits and coexist together. This may cause a faster growth in the population of GA strategy. The main reason for the success of the GA strategy in a noisy environment relates to the length of a chromosome, because it is considered to be comparatively longer than those chromosomes belonging to other strategies; thus, the strategy is exerted on a lot of games. Therefore, the appearance of noise has no great effect on the sum of a chromosome's scores.

2. LA Strategy

A. LA and IPD

As mentioned before, it is impossible to achieve an optimal strategy, but the expectation of finding a way to estimate the behavior of an opponent's strategy is not illogical. Learning is a process, and living beings need it for making changes in their behavior and being compatible with the environment. Stochastic LA is a decision-making algorithm that acts in a

stochastic environment and updates its strategy for the next action based on the response that it gets from an interaction with the environment.

LA doesn't know the environment at the beginning of the game; therefore, it tries to learn of it through a process of trial and error. So, at first, it performs actions randomly. The environment of the game (here is the opposite strategy) reacts in front of each action, and LA reduces or increases the possibility of performing this action based on the amount of desirability or undesirability of it before the next game begins [39].

The environment of the game is defined by (α, c, β) , where α is the sum of the actions that the automata can perform. It's evident that, in the mentioned environment, α equals $\{C, D\}$, where "C" means cooperation and "D" means defection. Set c is the set of all possible penalties, provided that an action is performed from the set of actions. The set β is determined by taking into consideration the proportions for penalty and reward outlined in Table 1.

Based on Fig. 8, the reactions of the environment to the behavior of the automata comprise the set $\beta = \{S, P, R, T\}$. The members of this set allocate amounts from the interval $[0, 1]$ to themselves. For this reason, the reactions of environment are divided by X , where $X = S + P + R + T$. So, the set of reactions of the environment are defined as $\beta = \left\{ \frac{S}{X}, \frac{P}{X}, \frac{R}{X}, \frac{T}{X} \right\}$.

In continuation, an indicator is needed to determine the desirability of the reactions of the environment. It is known that if two nodes make an effort to delete each other's packets, then they have neither profit nor loss (obtaining score P). A node will profit when its opposite node directs its packet (obtaining score R or T). So, a desirable action is one that offers a potential score (the reaction of the environment) that is greater than P , and an undesirable action is one that offers a potential

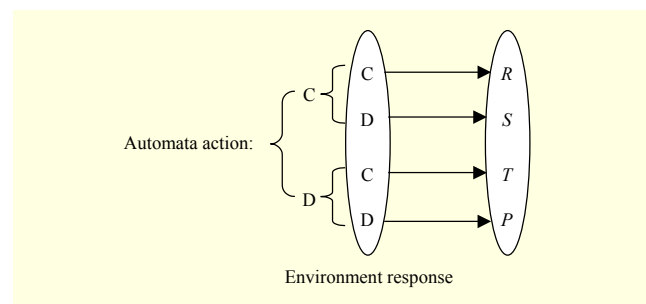


Fig. 8. Environment response to behavior of automata.

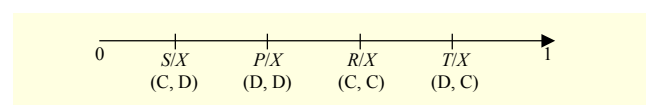


Fig. 9. Evaluating desirability of responses of environment.

```

if ( $P_i(n)$  had a good response) then
     $P_i(n+1) = P_i(n) + \text{Reward} \times (1 - P_i(n))$ 
     $P_{-i}(n+1) = (1 - \text{Reward}) \times P_{-i}(n)$ 
else
     $P_i(n+1) = (1 - \text{Penalty}) \times P_i(n)$ 
     $P_{-i}(n+1) = \text{Penalty} + (1 - \text{Penalty}) \times P_{-i}(n)$ 

```

Fig. 10. Learning automata pseudo code (L_{RP}).

Table 3. Determining amount of penalty and reward.

Action	Bad response	Good response
C	$(S-P)/X$	$(R-P)/X$
D	$(P-P)/X$	$(T-P)/X$

score that is less than P . The amount of desirability and undesirability is dependent on the difference between an action's score and that of the value P (see Fig. 9).

Considering the above, the number of members in set β is limited and kept in the interval $[0, 1]$. The model of the game's environment is that of the Q model. We have used L_{RP} automata for updating the possibility of performing the actions of the set of automata's actions. This is because we believe that an automata can be victorious when it rewards or fines its action with every reaction that it receives from the environment, because only in such a situation can it be possible to confront malevolent strategies, such as ALLD. If $P_C(n)$ is considered as the possibility of performing function C in the n th game and $P_D(n)$ as the possibility of performing function D in the n th game, then the following pseudocode (see Fig. 10) shows the trend of updating each of the above possibilities, where i is the action that the automata has performed in the n th game ($i \in \alpha$) and $-i$ represents the other actions of the members of set α .

The important problem is determining the amounts of reward and penalty. As mentioned before, the amount of desirability of the received reactions determines the amount of reward or penalty, which is allocated to the action that is done. Table 3 shows the amount of reward or penalty for each action that is performed.

B. Simulation Results

The game will be held using six strategies — TFT, ALLC, ALLD, Random, Gradual, and LA. As previously mentioned, a fully connected network is considered, where for every 100 nodes, each node must choose one, and only one, of the six aforementioned strategies. Each node plays the IPD game with all of the other nodes, and after an iterative game, each strategy

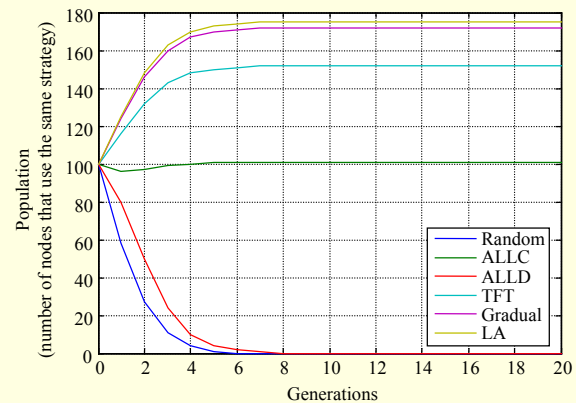


Fig. 11. Simulation with noise ratio $\beta = 0\%$.

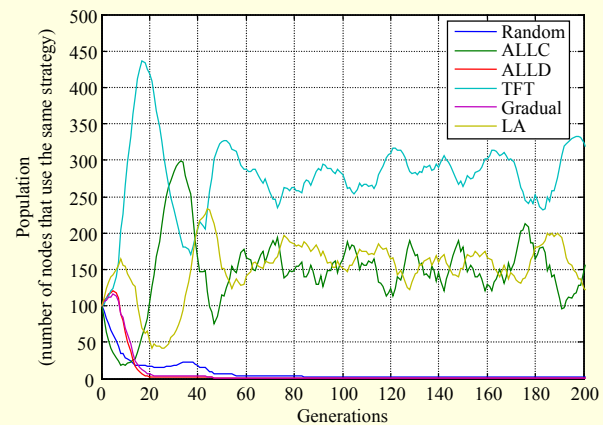


Fig. 12. Simulation with noise ratio $\beta = 5\%$.

calculates its payoff. Then, to find the best strategy, the population of each strategy will be computed based on (2). We remind the reader here that network topology, packet size, energy, and so on do not influence a node's behavior during the course of the game.

Figures 11, 12, 13, and 14 show the results of holding the tournament in the aforementioned environment. As can be seen from Fig. 11, the generation of unintelligent strategies (that is, ALLD, ALLC, and Random) has disappeared and left the competition early. The LA strategy has the highest population, followed respectively by the Gradual and TFT strategies. This means that the LA strategy has prevailed over all of the other strategies. The reason for the failure of the other strategies, like Gradual and TFT, is that, at the time of making a decision, they make use of only a single criterion (that is, they make use of only the event that happened in the last game), while LA considers multiple criterion — that being, recent games. In a noise-free environment, the LA strategy incites rational nodes to cooperate; thus, network throughput will be increased.

Figures 12, 13, and 14 show the simulation results for noise

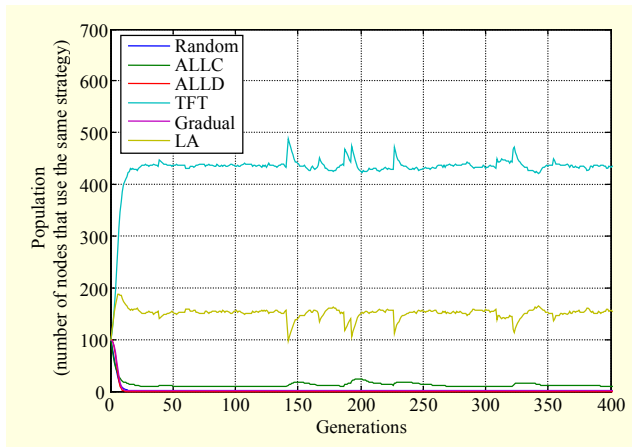


Fig. 13. Simulation with noise ratio $\beta = 10\%$.

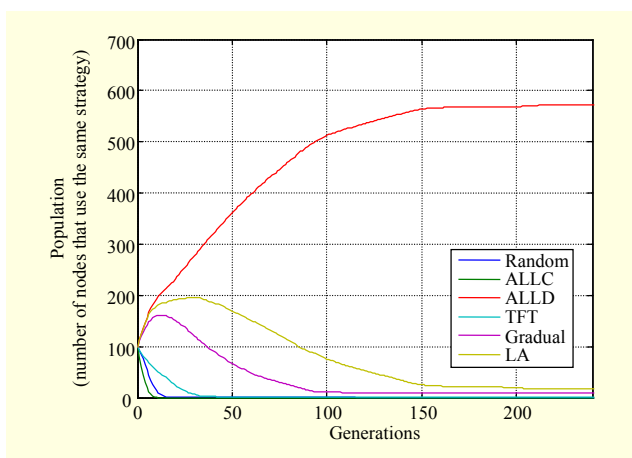


Fig. 14. Simulation with noise ratio $\beta = 20\%$.

ratios of 5%, 10%, and 20%, respectively. As mentioned in the previous section, the only difference between a noise-free environment and a noisy environment is the existence of some error within the notification mechanism.

It can be seen that there is no resistance strategy that works well in noisy environments, and this encourages nodes to cooperate with each other. Simulations were also performed under other noise ratios. With increasing noise, cooperative strategies are affected more than other strategies; thus, for high noise ratios, all rational nodes are inclined toward adopting the ALLD strategy. This means that when nodes do not have enough correct information about the behavior of other nodes, they prefer noncooperation.

V. Conclusion

In this paper, we first presented and surveyed a model for cooperation among nodes. The model originates from game theory and attempts to solve one of the classical puzzles from this domain; that is, it shows the intentions of rational nodes

when they are in cooperation with each other. A rational node is a node that performs a function to increase its profit. Following this, repetitive and evolutionary games were explained, and we concluded that one way to find a superior strategy is to hold a tournament between nodes. Then, we studied the profound strategies within this domain and concluded that a comprehensive method for playing a game in a noisy environment has not yet been proposed. Finally, by presenting a strategy based on a genetic algorithm, it has been shown that the effect of noise in a game can be prevented.

References

- [1] J.N. Laneman and G.W. Wornell, "Distributed Space-Time-Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, Oct. 2003, pp. 2415–2425.
- [2] R.U. Nabar, H. Bolcskei, and F.W. Kneubuhler, "Fading Relay Channels: Performance Limits and Space-Time Signal Design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, Aug. 2004, pp. 1099–1109.
- [3] I.E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Trans. Telecommun.*, vol. 10, no. 6, Nov.–Dec. 1999, pp. 585–595.
- [4] M. Godavarti, "Multiple Antennas in Wireless Communications: Array Signal Processing and Channel Capacity," Ph.D. dissertation, Department of Electrical Engineering and Computer Science, University of Michigan, MI, USA, 2001.
- [5] A. Narula, M.D. Trott, and G.W. Wornell, "Performance Limits of Coded Diversity Methods for Transmitter Antenna Arrays," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, Nov. 1999, pp. 2418–2433.
- [6] J.G. Proakis, "Digital Communications," New York, NY, USA: McGraw-Hill, 1995.
- [7] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity—Part I: System Description," *IEEE Trans. Commun.*, vol. 51, no. 11, Nov. 2003, pp. 1927–1938.
- [8] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity—Part II: Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, Nov. 2003, pp. 1939–1948.
- [9] A. Nosratinia, T.E. Hunter, and A. Hedayat, "Cooperative Communication in Wireless Networks," *IEEE Commun. Mag.*, vol. 42, no. 10, Oct. 2004, pp. 74–80.
- [10] V. Srinivasan et al., "Cooperation in Wireless Ad Hoc Networks," *Proc. IEEE INFOCOM*, San Francisco, CA, USA, vol. 2, Mar. 30–Apr. 3, 2003, pp. 808–817.
- [11] X. Huang, H. Zhai, and Y. Fang, "Robust Cooperative Routing Protocol in Mobile Wireless Sensor Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, Dec. 2008, pp. 5278–5285.

- [12] J.W. Yoo and K.H. Park, "A Cooperative Clustering Protocol for Energy Saving of Mobile Devices with WLAN and Bluetooth Interfaces," *IEEE Trans. Mobile Comput.*, vol. 10, no. 4, Apr. 2011, pp. 491–504.
- [13] F. Chiti et al., "Cooperative Localization Protocols for Wireless Sensor Networks," *Proc. IEEE GLOBECOM*, Washington, DC, USA, Nov. 26–30, 2007, pp. 1048–1052.
- [14] L. Yan and S. Hailes, "Cooperative Packet Relaying Model for Wireless Ad-hoc Networks," *Proc. ACM. FOWANC*, Hong Kong, China, May 27–30, 2008, pp. 93–100.
- [15] J. Hu, "Cooperation in Mobile Ad-Hoc Networks," Tech. report, Computer Science Department, Florida State University, FL, USA, 2005.
- [16] L. Buttyan and J.-P. Hubaux, "Enforce Service Availability in Mobile Ad-hoc WANs," *Proc. MobiHoc*, Boston, MA, USA, Aug. 6–11, 2000, pp. 87–96.
- [17] J.-P. hubaux et al., "The Terminode Project: Towards Mobile Ad-hoc WANs" *IEEE Int. Workshop Mobile Multimedia Commun.*, San Diego, CA, USA, Nov. 15–17, 1999, pp. 124–128.
- [18] S. Buchegger and J. Boudec, "Performance Analysis of the CONFIDANT Protocol: Cooperation of Nodes Fairness in Dynamic Ad-hoc Networks," *Proc. ACM. Int. Symp. MobiHoc*, Lausanne, Switzerland, June 9–11, 2002, pp. 226–236.
- [19] S. Zhong, J. Chen, and Y.R. Yang, "Sprite: A Simple, Cheat Proof, Credit-Based System for Mobile Ad-hoc Networks," *Proc. IEEE INFOCOM*, San Francisco, CA, USA, vol. 3, Mar. 30–Apr. 3, 2003, pp. 1987–1997.
- [20] P. Michiardi and R.K. Molva, "Core: A Collaborative Reputation Mechanism to Enforce Node Cooperation in Mobile Ad-hoc Networks," *Int. Federation Inf. Process.*, Portoroz, Slovenia, vol. 100, Sept. 26–27, 2002, pp. 107–121.
- [21] S. Bansal and M. Baker, "Observation-Based Cooperation Enforcement in Ad-hoc Networks," Technical Paper, July 2003.
- [22] L. Buttyan and J.-P. Hubaux, "Stimulating Cooperation in Self-Organizing Mobile Ad-hoc Networks," *J. Mobile Netw. Appl.*, vol. 8, no. 5, Oct. 2003, pp. 579–592.
- [23] L. Buttyan and J.-P. Hubaux, "Nuglets: A Virtual Currency to Stimulate Cooperation in Self-Organized Mobile Ad-hoc Networks," Swiss Federal Institute Technol., Lausanne, Switzerland, DSC/2001/001, Jan. 2001.
- [24] R. Axelrod and W.D. Hamilton, "*The Evolution of Cooperation*," New York, NY, USA: Basic Books, 1984, pp. 390–396.
- [25] L. Yan, "Cooperative Packet Relaying in Wireless Multi-hop Networks," *Proc. Int. Conf. WAINA*, Bradford, UK, May 26–29, 2009, pp. 345–350.
- [26] J.N. Webb, "*Game Theory: Decisions Interaction and Evolution*," 1st ed., New York, NY, USA: Springer, 2006, pp. 118–132.
- [27] Z. Fang and B. Bensaou, "Fair Bandwidth Sharing Algorithms Based on Game Theory Frameworks for Wireless Ad-hoc Networks," *IEEE INFOCOM*, Hong Kong, China, vol. 2, Mar. 7–11, 2004, pp. 1284–1295.
- [28] V. Srivastava et al., "Using Game Theory to Analyze Wireless Ad-hoc Networks," *IEEE Commun. Surveys Tutorials*, vol. 7, no. 4, Dec. 2005, pp. 46–56.
- [29] L.A. DaSilva and V. Srivastava, "Node Participation in Ad-hoc and Peer-to-Peer Networks: A Game-Theoretic Formulation," *Workshop Games Emergent Behavior Distrib. Comput. Environments*, Birmingham, UK, Sept. 18, 2004.
- [30] L. Anderegg and S. Eidenbenz, "Ad-hoc-VCG: A Truthful and Cost-Efficient Routing Protocol for Mobile Ad-hoc Networks with Selfish Agents," *Proc. MobiCom*, San Diego, CA, USA, Sept. 14–19, 2003, pp. 245–259.
- [31] S. Zhong et al., "On Designing Incentive-Compatible Routing and Forwarding Protocols in Wireless Ad-hoc Networks: An Integrated Approach Using Game Theoretical and Cryptographic Techniques," *Proc. ACM Int. Conf. MobiCom*, Cologne, Germany, Aug. 28–Sept. 2, 2005, pp. 117–131.
- [32] W. Wang et al., "OURS: Optimal Unicast Routing Systems in Non-cooperative Wireless Networks," *Proc. Annu. Int. Conf. MobiCom*, Los Angeles, CA, USA, Sept. 24–29, 2006, pp. 402–413.
- [33] J.F. Nash, "Equilibrium Points in N-person Games," *J. PNAS*, vol. 36, no. 1, 1950, pp. 48–49.
- [34] D.B. Neill, "Optimality under Noise: Higher Memory Strategies for the Alternating Prisoner's Dilemma," *J. Theoretical Biol.*, vol. 211, no. 2, July 2001, pp. 159–180.
- [35] X. Yao and P.J. Darwen, "An Experimental Study of N-person Iterated Prisoner's Dilemma Games," *Workshop Evolutionary Comput.*, Melbourne, Australia, vol. 956, no. 1, June 1995, pp. 90–108.
- [36] S.N. Sivanandam and S.N. Deepa, "*Introduction to Genetic Algorithms*," Berlin, Germany: Springer, 2007, pp. 145–201.
- [37] M. Mitchell, "*An Introduction to Genetic Algorithms (Complex Adaptive Systems)*," 3rd ed., MA, USA: Bradford Book, MIT Press, 1998, pp. 134–167.
- [38] J. Wu and R. Axelrod, "How to Cope with Noise in the Iterated Prisoner's Dilemma," *J. Conflict Resolution*, vol. 39, no. 1, Mar. 1995, pp. 183–189.
- [39] K.S. Narendra and M.A.L. Thathachar, "*Learning Automata: An Introduction*," NJ, USA: Prentice-Hall, 1989.



Abbas Arghavani is a teacher of wireless communication and information technology at the Applied Science and Technology University, Ravansar, Iran. He received his BS degree in software engineering from the University of Payam-e-Noor, Kermanshah, Iran, in 2008 and his MS degree in computer networks engineering from QIAU University, Qazvin, Iran, in 2012. His research interests include game theory and the lifetime, coverage, and placement of nodes in wireless networks.



Mahdi Arghavani received his BS degree in computer engineering from Shamsipour Technical and Professional University, Tehran, Iran, in 2007. Currently, he is an MS student in information technology at the Science and Research branch of Islamic Azad University, Kermanshah, Iran. In addition, he is teacher of wireless communication and information technology at the University of Applied Science and Technology, Ravansar, Iran and is with the Wireless Communications Business, Chinicord Company, Iran. His research interests include SDN, wireless communication, and game theory.



Abolfazl Sargazi received his BS degree from Sistan-va-Baluchestan University, Zahedan, Iran, in 2009 and his MS degree from QIAU University, Qazvin, Iran, in 2014, both in information technology. Currently, he is a network manager. His current research interests are in wireless networks and cognitive radio.



Mahmood Ahmadi received his BS and MS degrees in computer engineering from Isfahan University, Isfahan, Iran, in 1995 and Amirkabir University of technology of Tehran, Iran, in 1998, respectively. He joined the Razi University of Kermanshah, Iran, as a lecturer in 1998. In October 2005, he joined the Department of Computer Engineering, Delft University of Technology, Delft, Netherlands. Currently, he works as an assistant professor at the Department of Computer Engineering, Razi University of Kermanshah. His research interests include network processing, bloom filters, high-performance computing, performance modeling, reconfigurable architectures, and speech signal processing.