A New Game Theory-Based Algorithm for Target Coverage in Directional Sensor Networks

Elham Golrasan*
Faculty of Electrical and Computer Engineering, Malek Ashtar University of Technology, Iran
egolrasan@yahoo.com
Marzieh Varposhti
Department of Computer Engineering, Shahrekord University, Shahrekord, Iran
varposhti@aut.ac.ir

Received: 19/Oct/2020 Revised: 22/Mar/2021 Accepted: 15/Apr/2021

Abstract

One of the challenging problems in directional sensor networks is maximizing target coverage while minimizing the amount of energy consumption. Considering the high redundancy in dense directional sensor networks, it is possible to preserve energy and enhance coverage quality by turning off redundant sensors and adjusting the direction of the active sensor nodes. In this paper, we address the problem of maximizing network lifetime with adjustable ranges (MNLAR) and propose a new game theory-based algorithm in which sensor nodes try to adjust their working direction and sensing range in a distributed manner to achieve the desired coverage. For this purpose, we formulate this problem as a multiplayer repeated game in which each sensor as a player tries to maximize its utility function which is designed to capture the tradeoff between target coverage and energy consumption. To achieve an efficient action profile, we present a distributed payoff-based learning algorithm. The performance of the proposed algorithm is evaluated via simulations and compared to some existing methods. The simulation results demonstrate the performance of the proposed algorithm and its superiority over previous approaches in terms of network lifetime.

Keywords: Directional Sensor Networks; Target Coverage; Network Lifetime; Game Theory; Payoff-Based Learning Algorithm.

1- Introduction

Directional sensor networks (DSNs) contain several directional sensors deployed densely and randomly to cover a set of targets. Applications of such networks have been grown and widely applied in the field of industry and our daily life. In comparison with omni-directional sensor nodes, directional sensors have their unique characteristics, such as angle of view, working direction, and line of sight, therefore DSN applications require specific solutions for enhancing target coverage. Motility capability of a directional sensor node has a noticeable impact on the coverage enhancement in randomly deployed DSNs. Directional sensor nodes exploit motility to adjust their working direction. So motility can be used to minimize the overlapped regions. Because of limited energy resources in these networks, providing desired target coverage is a challenging problem [1-3].

In many applications, a large number of directional sensor nodes are randomly deployed in an area of interest. The availability of redundant sensors enhances the fault tolerance capability of the network. However, keeping all the sensor nodes active is not efficient because it leads to higher energy consumption. Therefore, one of the goal of this paper is to enhance the target coverage in a network of self-orienting sensor nodes. The other goal is decreasing the power consumption and increasing the network lifetime.

Since we are interested in the automated self-orienting of the nodes, we cast this problem as a non-cooperative game as it is a well-established tool for modeling coordination problems. There are many reasons to choose game theory as a method to solve the coverage problem in sensor networks. First, the principle of game theory allows sensor nodes to operate independently and calculate their proper orientation in a distributed manner. A well-designed gain function that takes into account all the limitations of the problem including the energy consumption, makes it possible to establish acceptable coverage in the sensor network. So, we can provide scalable network coverage using game theory. Finally, the game theory method is resistant to node failures and environmental disturbances [4].

In this paper, we model the coverage problem as a finite strategic game and propose a game theory-based algorithm (GT-based algorithm), in which the utility function is designed to capture the tradeoff between the worth of the

* Corresponding Author

covered area and the energy consumption due to sensing. An important issue is to devise distributed learning algorithms, using local information and processing abilities, to reach a Nash equilibrium (NE) of the game. To this end, we use a distributed payoff-based learning algorithm [5]. It has been proved that if all sensors (players) adhere to this algorithm, then each sensor selects the action profile that maximizes the total payoff of the sensors.

In most game theory based algorithm, a challenging problem is achieving the Nash equilibrium using a distributed learning algorithm [4, 6-10]. In [8, 10, 11], distributed learning algorithms are proposed to achieve NE in coverage problems. The authors in [12, 13] have studied distributed systems and propose distributed learning algorithms to achieve NE in potential games. However, there are two main drawbacks in this context: First of all, most of the results in this area have focused on converging to the NE, while in many cases it is not the optimum solution. Detecting this inefficiency is an extremely active research area in algorithmic game theory [14]. Secondly, it is often impossible to frame the interaction of a given system as a potential game [5]. We measure the performance of an action profile using the sum of the sensor's utility functions. Therefore, by designing the appropriate utility function for each player (sensor node) and applying a distributed payoff-based learning algorithm, coverage in the sensor network converges to a Pareto optimal action profile. The utility function is defined based on the tradeoff between coverage and energy consumption. Then each sensor learns how to adjust its working direction to maximize its utility function which corresponds to find its best orientation based on local information.

In the following, the main contributions of this study are presented:

- We formulate the maximum network lifetime with adjustable ranges (MNLAR) problem as a multiplayer repeated game in which each sensor as a player tries to maximize its utility function. The utility function is designed to capture the tradeoff between the worth of covered targets and the energy consumption due to sensing.
- We propose a distributed payoff-based learning algorithm that converges to an efficient action profile.
- The performance of the proposed algorithm is evaluated via simulations and compared to previous approaches. The simulation results show that the proposed algorithm results in activating the minimum number of directional sensors. In addition, active sensors learn how to adjust their sensing ranges to maximize the coverage. These bring about less energy consumption along with network lifetime extension.

The paper is organized as follows: In section 2, we briefly review recent studies in the context of sensor coverage problems. In section 3, we introduce the MNLAR problem in DSNs. Section 4 presents the proposed method and its formulation based on game theory. In section 5, simulation results are presented through several experiments. Finally, we conclude the paper in section 6.

2- Related Work

In this section, we briefly review the research work on coverage in wireless sensor networks. The coverage problem is usually divided into three categories: area coverage, point coverage, and barrier coverage [2]. The purpose of area coverage is to cover the whole area. Point coverage is the problem of covering Points of Interest (PoI) in the area. The barrier coverage guarantees that every movement that crosses a barrier of sensors will be detected.

Habibi et al. [15] have proposed a distributed Voronoi-based strategy to maximize the sensing coverage in a mobile sensor network. In this algorithm, each sensor moves through a gradient-based nonlinear optimization approach and is placed inside its Voronoi cell.

Ai et al. [16] have studied the problem of covering targets with directional sensors. They have formulated the problem as an optimization problem to maximize the coverage with a minimum number of sensors and proved that it is NP-complete. They have proposed several greedy heuristic methods to solve the problem.

Mohamadi et al [17] have proposed two Greedy-based algorithms for target coverage in directional sensor networks with adjustable sensing ranges. To maximize the sensor lifetime, they have used both scheduling and adjusting sensing range techniques to form cover sets to cover all targets in the network.

In [18], the authors have provided a GA-based algorithm to find cover sets of directional sensors with appropriate sensor ranges in order to solve the MNLAR (Maximum Network Lifetime with Adjustable Ranges) problem.

Yu et al. [19] have addressed the problem of providing K-coverage along with prolonging the network lifetime in wireless sensor networks with both centralized and distributed protocols. They have introduced a new concept of Coverage Contribution Area (CCA). Based on this concept, a lower sensor spatial density is provided.

The authors in [20] have designed a probabilistic coverage preservation protocol (CPP) to achieve energy efficiency and ensure a certain coverage rate. The purpose of the proposed protocol is to select the minimum number of probabilistic sensors to reduce energy consumption.

A graph model named Cover Adjacent Net (CA-Net) is proposed by Weng et al in [21] to simplify the problem of k-barrier coverage while reducing the complexity of computation. Based on the developed CA-Net, two distributed algorithms, called BCA and TOBA, are presented for energy balance and maximum network lifetime.

Mostafaei et al. [11] have proposed a distributed boundary surveillance (DBS) algorithm to cover the boundary and reduce the energy consumption of sensors. DBS selects the minimum number of sensors to increase the network lifetime using learning automata.

Li et al. [22] have proposed the Voronoi-based distribution approximation (VDA) algorithm. In the proposed algorithm, in order to maximize the coverage of the desired area, the most Voronoi edges are covered. In [23], the authors have proposed the distributed Voronoi-based self-redeployment algorithm (DVSA), aiming to improve the overall field coverage of mobile directional sensor networks. This paper has utilized the geometrical features of Voronoi diagram and the advantages of a distributed algorithm.

Recently, game-theoretic approaches have been taken into consideration to solve coverage problems in WSNs [4, 6, 24, 25]. In [26], the authors have proposed an algorithm based on game theory for the problem of maximizing coverage and reducing energy consumption. They have shown that the desired solution in this model is an NE strategy profile.

In [27], the authors have proposed a distributed learning method to maximize the area coverage in mobile directional sensor networks. Each sensor in collaboration with its neighbors tries to determine its best position and orientation.

The authors in [28] have considered the problem of area coverage without location information in mobile sensor networks. They have modeled this problem as a potential game and proposed a distributed learning algorithm to achieve an NE.

In [29] coverage of an unknown environment is investigated by robots. The state-based potential game was designed to control the robots' actions. The reward of sensing the areas and the penalty of energy consumption due to the sensors' movement are considered in the utility function. The sensors update their action profile using the Binary Log-Linear Learning (BLLL) in which the sensors must know an estimation of the outcome of their future actions. Hence, an estimation algorithm is proposed to assist the sensors in predicting the probability of targets in unknown areas. An improved EM algorithm is introduced to estimate the number of targets and other probability distribution parameters. In this study, we propose a game theory-based algorithm to optimally cover targets and reduce energy consumption.

3- Problem Definition

We consider a two-dimensional mission space, where n directional sensors with motility capability and a set of m target points, T, are initially located within a given area. We have defined several power levels, p, and a set of working directions, D, for sensor nodes. So, each sensor s_i has two parameters, working direction and power level. Sensor nodes can rotate around their axis and adjust their power level to cover a sector area. So, a directional sensor, s_i, monitors all targets within its sensing range and field of view. Each sensor has a limited energy resource. The amount of energy consumption is a function of sensing range; the greater sensing range, the more energy is consumed [30]. All the sensors in the network have the same characteristics in terms of initial battery power and energy consumption function.

Since increasing the power level is equivalent to increasing the sensing range which results in covering more targets, for each sensor direction $d_{i.j}$ and each power level $p > 1$, we have $T_{(d_{i.j}.\ b)} \subseteq T_{(d_{i.j}.\ p)}\ \forall b \in \{1. \dots . p-1\}$, which $(d_{i,j}, p)$ is i^{th} sensor with j^{th} direction activated at power level of p. $T_{(d_{i.j}.\ b)}$ is the set of covered targets by sensor s_i. The power level p which is sufficient to cover target t_j by sensor $(d_{i.j}.\ p)$ is minimal if $t_j \in T_{(d_{i.j}.\ p)}$ and $t_j \notin T_{(d_{i.j}.\ b)}\ \forall b \in \{1. \dots . p-1\}$. It means that target t_j cannot be covered by power level less than p. The notations used in this paper are listed in Table 1.

We assume the sensors are non-rechargeable. According to [17], the energy consumption due to displacement between the directions of a sensor is negligible, so it is ignored. Here, a positive parameter Δ^p is defined at each power level p [30]. The parameter Δ^p indicates the ratio between battery consumption at level p and level 1. Level 1 is the lowest and cheapest level. For example, if Δ^p=2, then the battery consumption at level p is twice that of level 1 (Δ^p=1).

Table 1: Notations

Notation	Meaning
n	Number of sensors
m	Number of targets
w	Number of working directions, $w \geq 1$
p	Number of different power levels, $p \geq 1$
s_i	i^{th} sensor, for all $i \in \{1. \dots . n\}$
t_k	k^{th} target, for all $k \in \{1. \dots . m\}$
l_i	Lifetime of sensor s_i
$d_{i,j}$	Jth direction of ith sensor
D	Set of the working directions

S	Set of sensors, $\{s_1. \dots. s_n\}$
T	Set of targets, $\{t_1. \dots. t_m\}$
$(d_{i,j}, p)$	i^{th} sensor with j^{th} direction activated at power level of p
$T_{(di,j,\,p)}$	all the targets covered by i^{th} sensor with j^{th} direction at power level of p

Problem. How to divide the available sensors into cover sets so that each cover set covers all the targets in the area of interest with the goal of prolonging the network lifetime as much as possible. In other words, the main challenge of this process is assigning the appropriate working direction and sensing range to each sensor in a way that full target coverage and maximum network lifetime can be achieved.

For a better understanding, consider Fig. 1.A, which includes a scenario with three targets and four directional sensors to monitor the targets in the network. Fig. 1.B shows that each sensor has three directions $d_{i,j}$ ($1\leq j \leq 3$) and two sensing ranges. Consider the set of sensor nodes with their best parameters that cover each target t_j. $C(t_1) = \{(d_{3.1}. 1). (d_{1.2}. 2)\}$, $C(t_2) = \{(d_{2.3}. 1). (d_{3.2}. 2)$, $C(t_3) = \{(d_{3.3}. 2). (d_{4.1}. 1)\}$. the purpose is to form the best cover set. The possible cover sets include:

$C_1 = \{(d_{1.2}. 2). (d_{2.3}. 1). (d_{3.3}. 2)\}.$

$C_2 = \{(d_{1.2}. 2). (d_{2.3}. 1). (d_{4.1}. 1)\}.$

$C_3 = \{(d_{1.2}. 2). (d_{3.2}. 2). (d_{4.1}. 1)\}.$

$C_4 = \{(d_{3.1}. 1). (d_{2.3}. 1). (d_{4.1}. 1)\}.$

Therefore, in this example, the cover set C_4 is more desirable because of less energy consumption.

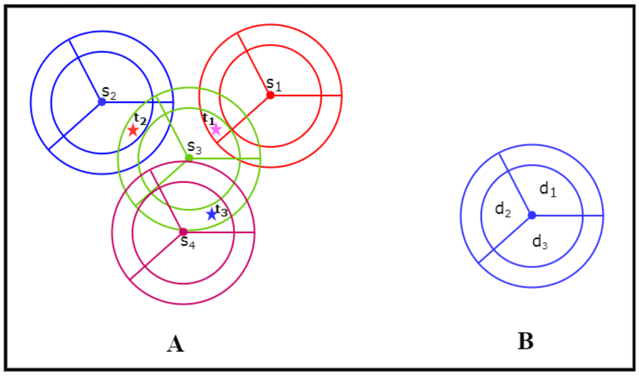

Fig. 1. (A) Example network with four directional sensors and three targets. (B) A directional sensor with three directions and two sensing ranges.

4- Proposed GT-Based Algorithm

In this section, we propose a game theory-based algorithm (GT-based algorithm) to target coverage. The new method is a solution to the MNLAR problem in DSNs. It converges to an efficient action profile using a distributed payoff-based learning algorithm. The output of the proposed algorithm is a cover set containing sensors with appropriate sensing ranges and working directions that can monitor all targets within the network. To calculate the activation time of the constructed cover set, we consider the sensor that minimizes $\frac{L_i}{\Delta p}$. Then the residual energy of the sensors in the cover set is calculated. The sensors that have no residual energy are removed from the list of available sensors. The GT-based algorithm continues its operation until all targets are fully covered. Finally, the sum of the activation times is returned as the network lifetime.

4-1- Background in Game Theory

In this section, we consider a brief review of the concepts in game theory. More information about game theory and learning in game theory is mentioned in [31, 32].

The strategic game $G \coloneqq \langle N. A. U \rangle$ has three components:

The finite set of N players (sensors) where $N \coloneqq \{1. \dots. n\}$.

An action set $A = A_1 \times \dots \times A_n$ where A_i is the finite action set of player i.

The set of utility (payoff) functions U. where the utility function $U_i\colon A \rightarrow R$ models the benefit of player i over action profiles.

For an action profile $a = (a_1. a_2. \dots. a_n) \in A$. $a_{-i} = \{a_1. \dots. a_{i-1}. a_{i+1}. \dots. a_n\}$ denotes the action profile of all players other than player i. Therefore, the action profile a can be represented as $(a_i. a_{-i})$.

The welfare of an action profile $a \in A$ is defined as follows:

$$W(a) = \sum_{i \in N} u_i(a) \tag{1}$$

if an action profile $a \in A$ maximizes the welfare, then the action profile a is efficient. In other words:

$$a \in \arg max_{a' \in A} W(a') \tag{2}$$

4-2- Game Formulation

Suppose *m* targets $T = \{t_1 . t_2 . \dots . t_m\}$ located in known locations in the area. A set of directional sensors $S = \{s_1 . s_2 . \dots . s_n\}$ with adjustable sensing ranges are deployed adjacent to the targets to completely monitor them. Sensors are static with variable sensing ranges between r_{min} and r_{max}. We assume that the communication range of each sensor i (R_i) is at least twice the r_{max} ($R_i \geq 2r_{max}$). Thus, each sensor can transmit state information to its neighbors and interact with its neighbors.

The worth of each target $t_k (1 \leq k \leq m)$ is denoted with $f_k \geq 0$. Each sensor s_i selects its mode from the set $Q_i = \begin{cases} 1 & if\ sensor\ i\ is\ on \\ 0 & if\ sensor\ i\ is\ off \end{cases}$.

The sensor's direction is determined by d_i and the set $D_i = \{d_{i.1} . d_{i.2} . \dots . d_{i.w}\}$ includes the defined working directions of sensor nodes. Each sensor s_i chooses its sensing range r_i from the discrete set $R_i = \{r_{i.1} . \dots . r_{i.p}\}$. The action of each sensor (player) s_i is shown by a vector a_i and defined as follows: $a_i = (c_i . d_i . r_i) \in A_i = Q_i \times D_i \times R_i$.

As mentioned before, $T(d_{i.j} . a)$ is a set of targets covered by sensor direction $d_{i.j}$ while its power level is a. For each target $t_k \in T(d_{i.j} . a)$, $n_k(s)$ represents the number of sensors that can observe the target point *k* and is defined as follows:

$$n_k(s) = \sum\nolimits_{i \in n} c_i \; ; \; t_k \in T(d_{i.j} . a) \tag{3}$$

The profit of observing the target point t_k, which is shown by f_k , is evenly divided by the sensors that observe t_k. So, the utility that sensor s_i obtains due to sensing is equal to

$$c_i (\sum_{t_k \in T(d_{i.j} . a)} \frac{f_k}{n_k(s)}) \tag{4}$$

Due to energy constraints, we consider the energy consumption parameter in the utility function. We assume that the energy consumption of sensor nodes is because of their sensing activity. So, the energy consumption of each sensor node depends on its sensing range and is defined as (5).

$$E_i^{sense}(a_i) = k_i c_i (r_i)^2 \tag{5}$$

in which $k_i > 0$is a weighting factor related to energy consumption. Therefore, the utility function of the sensor s_i represents its contribution to the coverage task and energy consumption due to sensing. We consider the utility function for sensor s_i as follows:

$$u_i(a) = c_i \left(\sum_{t_k \in T(d_{i.j} . a)} \frac{f_k}{n_k(s)} \right) - E_i^{sense}(a_i) \tag{6}$$

In the following, we present a new distributed learning algorithm that leads to Pareto optimal outcomes.

4-3- Payoff-Based Learning Algorithm

The game G is repeated each time $t \in \{0.1.2. \dots\}$. In time stamp *t*, the sensors simultaneously select their actions, so the action profile is $a(t) = (a_1(t) . \dots . a_n(t))$ and each sensor s_i receives utility $u_i(a(t))$. The sensor s_i will select the action $a_i(t)$ according to the probabilistic distribution $p_i(t) \in \Delta(A_i)$. $p_i(t)$ represents the strategy of sensor s_i at time *t*. $p_i^{a_i}(t)$ indicates the probability that the sensor s_i selects action $a_i(t)$ at time *t* according to the strategy $p_i(t)$. The sensor's strategy at time *t* depends on observations in previous times $\{0.1.2. \dots . t-1\}$.

The strategies of the sensors are updated by the information they have gathered. We know that the sensors here have limited observations. In this situation, sensors must learn to play an action profile that maximizes welfare. In this case, the sensors only have access to the actions they played and the utilities they received. Therefore, the strategy adjustment mechanism of sensor s_i is as follows:

$$p_i(t) = F_i(\{a_i(\tau) . u_i(a(\tau))\}_{\tau = 0.1. \dots . t-1}) \tag{7}$$

Such an algorithm is called payoff-based or completely uncoupled [33]. It is proved in [5, 34-37] that for finite strategic games, there are completely uncoupled learning rules that lead to Pareto optimal Nash equilibria. We use the learning rule presented in [5]. This distributed learning rule leads to the convergence of the game into a Pareto optimal action profile, which maximizes the welfare.

Each sensor has a baseline action and a baseline utility, that is expected to play and receive, respectively. Each sensor has an internal state variable called mood. Mood defines the sensor behavior as follows. There are two distinct types of moods: content and discontent. When a sensor is in the content mood, baseline action is selected with high probability. When a sensor is in a discontent mood, an action differs from the baseline action is selected with a high probability. Each sensor updates its mood after choosing an action and receiving a payoff by comparing the action played and the payoff received with its baseline action and baseline payoff. At each time step, the sensor's state is represented by a triple $[\bar{a}_i . \bar{u}_i . m_i]$, where

- $\bar{a}_i \in A_i$ is the baseline action.
- $\bar{u}_i$ is the baseline utility.
- m_i is the mood of sensor *i*, which can be content (C) or discontent (D).

The main steps of the distributed learning algorithm for Pareto optimality are described as follows:

Step 1- Initialization:

- At stage $t = 0$, each player randomly selects and plays any action, $a_i(0)$.
- This action will be initially set as the player's baseline action at stage 1, $\bar{a}_i(1) = a_i(0)$.
- Likewise, the player's baseline utility at stage 1 is initialized as $\bar{u}_i(1) = u_i(a(0))$.

- the player's mood at stage 1 is set as $m_i(1) = D$.

Step 2- Action Selection:
At each subsequent stage $t > 0$, each player selects his action according to the following rules.

- If the mood of sensor *i* is content, i.e., $m_i(t) = C$, the sensor chooses an action $a_i(t)$ according to the following probability distribution
$$p_i^{a_i}(t) = \begin{cases} \dfrac{\varepsilon^c}{|A_i| - 1} & for\ a_i \neq \bar{a}_i \\ 1 - \varepsilon^c & for\ a_i = \bar{a}_i \end{cases}$$
where $|A_i|$ represents the cardinality of the set A_i and *c* is a constant that satisfies $c > n$.
- If the mood of sensor *i* is discontent, i.e., $m_i(1) = D$, the sensor chooses an action a_i according to the following probability distribution
$$p_i^{a_i}(t) = \frac{1}{|A_i|} \quad for\ every\ a_i \in A_i$$
Note that the benchmark action and utility play no role in the sensor dynamics when the sensor is discontent.

Step 3- Baseline Action, Baseline Utility, and Mood Update:

Each sensor *i* sends its status information including $a_i(0).F(s_i)$ to its neighboring nodes or nodes that are less than or equal to twice the sensing range of the node *i*. Then each sensor *i* computes the payoff $u_i\big(a_i(t).a_{-i}(t)\big)$ based on the data collected from neighbors. the state is updated according to the following rules.

- First, the baseline action and baseline utility at stage *t+1* are set as
$\bar{a}_i(t+1) = a_i(t)$
$\bar{u}_i(t+1) = u_i\big(a_i(t).a_{-i}(t)\big)$
- The mood of sensor *i* is updated as follows.
$$if \begin{bmatrix} \bar{a}_i(t) \\ \bar{u}_i(a(t)) \\ m_i(t) \end{bmatrix} = \begin{bmatrix} a_i(t) \\ u_i(a(t)) \\ C \end{bmatrix} then\ m_i(t) = C$$
Else
$$m_i(t+1) = \begin{cases} C & with\ probability\ \varepsilon^{1-u_i(a(t))} \\ D & with\ probability\ 1 - \varepsilon^{1-u_i(a(t))} \end{cases}$$

Step 4- Return to Step 2 and repeat.

The learning algorithm produces a sequence of action profiles $a(1).....a(t)$, in which the behavior of a sensor *i* in each time $t = 1.2. ...$ depends on the baseline action $\bar{a}_i$, the baseline utility $\bar{u}_i$, and the mood $m_i(t) \in \{C.D\}$. To converge the game into an efficient action profile, the game's structure must be interdependent [35]. In the following, the definition of interdependence is fully described.

Definition1 (Interdependence): A finite game G is interdependent If for any action profile $a \in A$ and any appropriate subset of the sensors $J \subset N$, There is a sensor $i \notin J$ and a selection of actions $a'_j \in \prod_{j \in J} A_j$ so that $u_i(a'_J.a_{-J}) \neq u_i(a_J.a_{-J})$.

In general, the interdependence condition states that the sensors cannot be divided into two distinct subsets that do not interact with one another. For this reason, we assume that the sensors are deployed in the area so that they cannot be divided into two distinct subsets and the condition of interdependence is established in the game.

Theorem1: Consider a finite interdependent game with *n* players. Under the distributed learning algorithm for Pareto optimization defined above, a state $z = [\bar{a}.\bar{u}.m] \in Z$ is stable if and only if the following conditions are met.

- The action profile $\bar{a}$ optimizes the welfare $W(\bar{a}) = \sum_{i \in N} u_i(\bar{a})$.
- The baseline actions and payoffs are aligned, i.e. for each sensor *i*, $\bar{u}_i = u_i(\bar{a})$.
- All sensors are in the content mood i.e. for each sensor *i*, $m_i = C$.

Proof of the theorem depends on the resistance trees for Markov's decision process, and it has been proven in [38].

5- Simulation Results

In this section, we evaluate the performance of the proposed algorithm through several experiments. The algorithms are simulated on MATLABR2017b and implemented on a system with an Intel Core i7 processor, 3.4 GHz CPU, and 4 GB RAM. The most important criterion for evaluating the performance of the algorithm is the network lifetime. The network lifetime is defined as the time that the sensor nodes can cover all of the targets. Each experiment examines the impact of different parameters on the network lifetime. Here, to model a DNS, *m* targets and *n* directional sensors are deployed randomly and uniformly in a $500(m) \times 500(m)$ area. Each sensor has several working directions and sensing ranges. For each sensor node, only one working direction and one sensing range can be activated at each unit of time. By default, the number of sensors and targets are 100 and 10, respectively. Also, we have considered 3 working directions for each sensor. The sensing range of each sensor can be adjusted from 80-120(m) with incremental step 10(m). We assume that each sensor initially has one unit of energy.

To establish the interdependence condition, those targets that are not monitored by any sensor direction are ignored and all sensors that cannot cover any targets are removed from the list of sensors. The simulation parameter is chosen as $k_i = 3 \times 10^{-5}.f_q = 0 \cdot 1.c = 100$. According to [38], employing an annealing schedule $\varepsilon_t = \frac{1}{\sqrt{t}}$ in the learning algorithm guarantees the convergence to an efficient action profile.

Experiment 1. This experiment is performed to provide an intuitive example for the implementation of the proposed algorithm. Five targets (m=5) and 20 directional sensors (n=20) are randomly deployed in the area. Fig. 2 shows the initial configuration of the network, and Fig. 3 shows the final configuration of the network after 10^5 iterations. According to Fig. 3, in addition to full coverage of the targets, energy consumption in the area has decreased due to the adjustment of working direction and sensing range of all sensor nodes. In this figure, all targets are covered with the minimum number of sensors with the least overlapping area (each target is exactly in one sector). Consequently, it is clear that the result is a Pareto optimal action profile. Fig. 4 presents the evolution of the welfare for $\varepsilon_t = \frac{1}{\sqrt{t}}$. The result shows the convergence of the welfare function to its maximum value. The reason is that energy consumption due to sensing is considered in the utility function.

Experiment 2. This experiment is designed to evaluate the relationship between the number of directional sensors and the lifetime of the network. To this end, we have considered sensor networks with 60-100 sensor nodes. Fig. 5 shows that increasing the number of sensor nodes results in enhancing the network lifetime. The reason is that each target is covered with more sensors, so more cover sets are constructed and the network lifetime is increased. Simulation results demonstrate that as the number of sensors increases, the proposed GT-based algorithm performs better than the Genetic-based algorithm [18] and Greedy-based algorithm [17]. This is due to the iterative property of learning algorithms and the more efficient management of energy consumption in the proposed game.

Experiment 3. This experiment is performed to determine how the number of targets affects the network lifetime. For this purpose, we have considered an area of interest with 6-14 target points. According to Fig. 6, network lifetime decreases when the number of targets increases. The reason is that as the number of targets increases, more sensors are needed to monitor them. This will cause the sensors to run out of energy earlier, so results in reducing the network lifetime.

Fig. 7 shows the energy consumption of the cover sets in Greedy-based, Genetic-based, and GT-based algorithms based on the number of targets. As expected, an increase in the number of targets increases energy consumption. The reason is that more sensors are needed to cover more targets. The results confirm that the GT-based algorithm consumes less energy compared to the other two algorithms since the proposed algorithm activates fewer sensor nodes, so the energy consumption due to sensing in GT-based algorithm is less than Genetic-based algorithm and Greedy-based algorithm.

Experiment 4. This experiment is performed to investigate the effect of the sensing range on the network lifetime. The sensing range is fixed between 80 and 120 with incremental step 10. According to the results presented in Fig. 8, an increase in the sensing range leads to an improvement in the network lifetime. This is because of this fact that increasing the sensing range results in covering more targets, so fewer sensors are needed to cover all targets. In comparison with the other two algorithms, experiment results confirm that the proposed algorithm is more successful in terms of maximizing network.

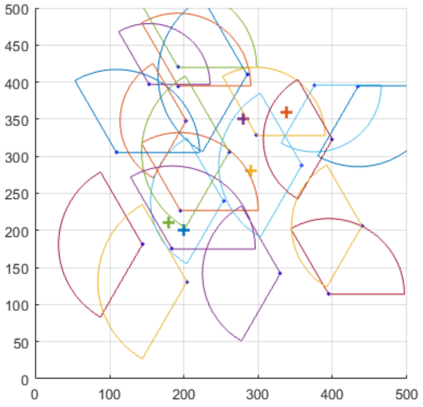

Fig. 2. The initial configuration of the network where "•" and "+" are sensor nodes and targets, respectively.

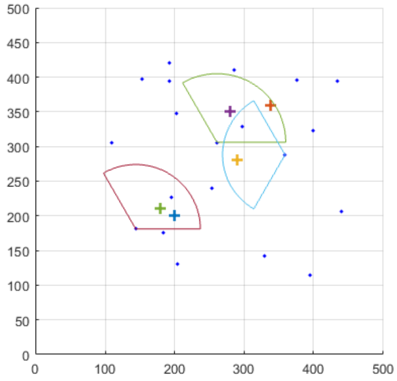

Fig. 3. The final configuration of the network after 10×10^4 iterations.

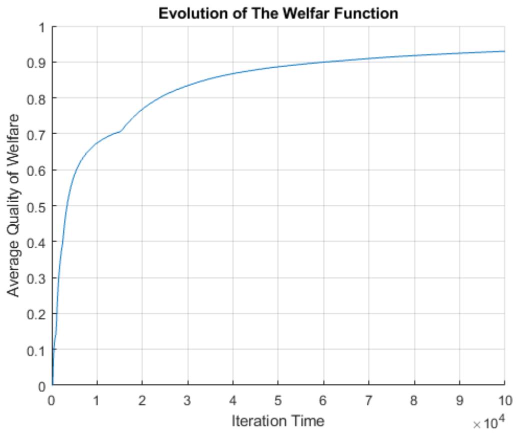

Fig. 4. The evaluation of the welfare function in $\varepsilon_t = \frac{1}{\sqrt{t}}$.

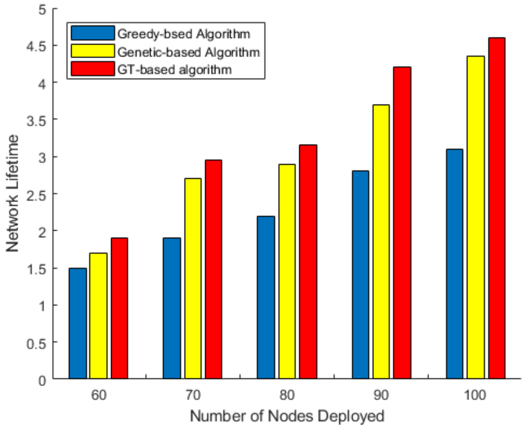

Fig. 5. Effect of the number of sensors on the network lifetime.

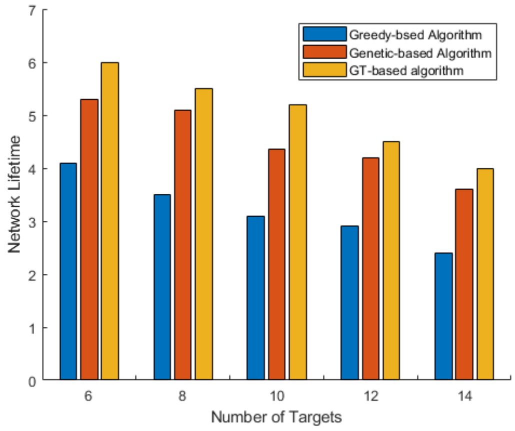

Fig. 6. Effect of the number of targets on the network lifetime.

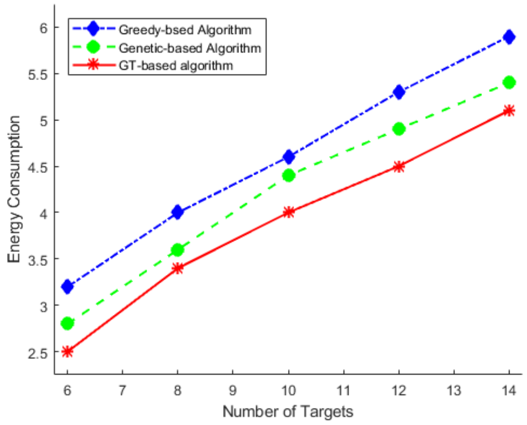

Fig. 7. Effect of the number of targets on energy consumption.

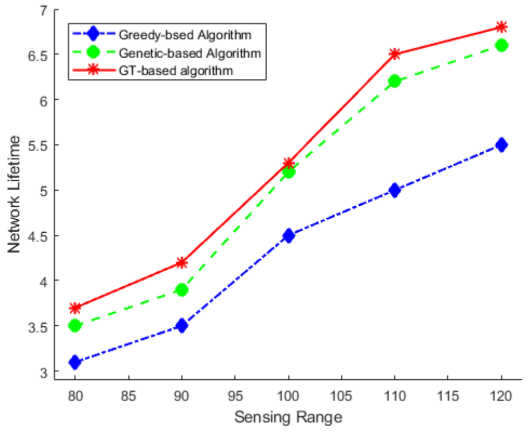

Fig. 8. Effect of sensing range on the network lifetime.

6- Conclusion

In this paper, we presented a new game theory-based algorithm for target coverage in networks containing sensors with multiple directions and sensing ranges to extend network lifetime. Due to the energy limitations in sensor networks, we formulate the target coverage problem as a finite strategic game in which a utility function is formulated to consider the tradeoff between energy consumption and coverage quality. To achieve an efficient action profile, we present a distributed payoff-based learning algorithm. The performance of our proposed algorithm was evaluated via simulations and compared to the greedy-based and genetic-based algorithms. The simulation results demonstrated the performance of our proposed algorithm and its superiority over previous approaches in terms of increasing the network lifetime.

References

[1] B. D. Deebak and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Networks,* vol. 97, p. 102022, 2020.

[2] K. Evbarunegbe, "Coverage in Wireless Sensor Networks: A Survey," Texas A&M University-Kingsville, 2020.

[3] S. Harizan and P. Kuila, "Evolutionary algorithms for coverage and connectivity problems in wireless sensor networks: a study," in *Design Frameworks for Wireless Networks*: Springer, 2020, pp. 257-280.

[4] S. Rahili, J. Lu, W. Ren, and U. M. Al-Saggaf, "Distributed Coverage Control of Mobile Sensor Networks in Unknown Environment Using Game Theory: Algorithms and Experiments," *IEEE Transactions on Mobile Computing,* vol. 17, no. 6, pp. 1303-1313, 2018.

[5] J. R. Marden, H. P. Young, and L. Y. Pao, "Achieving pareto optimality through distributed learning," *SIAM Journal on Control and Optimization,* vol. 52, no. 5, pp. 2753-2770, 2014.

[6] T. AlSkaif, M. G. Zapata, and B. Bellalta, "Game theory for energy efficiency in wireless sensor networks: Latest trends," *Journal of Network and Computer Applications,* vol. 54, pp. 33-61, 2015.

[7] E. Golrasan, H. SHIRAZI, and K. Dadashtabar Ahmadi, "Sensing Radius Adjustment in Wireless Sensor Networks: A Game Theoretical Approach," *INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY RESEARCH,* vol. 11, no. 4, pp. -, 2019.

[8] F. Hajjej, M. Hamdi, R. Ejbali, and M. Zaied, "A distributed coverage hole recovery approach based on reinforcement learning for Wireless Sensor Networks," *Ad Hoc Networks,* vol. 101, p. 102082, 2020.

[9] J. R. Marden and J. S. Shamma, "Game theory and control," *Annual Review of Control, Robotics, and Autonomous Systems,* vol. 1, pp. 105-134, 2018.

[10] N. Mehendale, "Game Theory-Based Planning of Nodes in Wireless Sensor Networks for Optimum Coverage With Maximum Battery Life," *Available at SSRN 3639170,* 2020.

[11] H. Mostafaei, M. U. Chowdhury, and M. S. Obaidat, "Border surveillance with WSN systems in a distributed manner," *IEEE Systems Journal,* vol. 12, no. 4, pp. 3703-3712, 2018.

[12] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE Journal of Selected Topics in Signal Processing,* vol. 7, no. 2, pp. 230-242, 2013.

[13] B. S. Pradelski and H. P. Young, "Learning efficient Nash equilibria in distributed systems," *Games and Economic behavior,* vol. 75, no. 2, pp. 882-897, 2012.

[14] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge University Press, 2007.

[15] J. Habibi, H. Mahboubi, and A. G. Aghdam, "A gradient-based coverage optimization strategy for mobile sensor networks," *IEEE Transactions on Control of Network Systems,* vol. 4, no. 3, pp. 477-488, 2016.

[16] J. Ai and A. A. Abouzeid, "Coverage by directional sensors in randomly deployed wireless sensor networks," *Journal of Combinatorial Optimization,* vol. 11, no. 1, pp. 21-41, 2006.

[17] H. Mohamadi, S. Salleh, and M. N. Razali, "Heuristic methods to maximize network lifetime in directional sensor networks with adjustable sensing ranges," *Journal of Network and Computer Applications,* vol. 46, pp. 26-35, 2014.

[18] A. Alibeiki, H. Motameni, and H. Mohamadi, "A new genetic-based approach for maximizing network lifetime in directional sensor networks with adjustable sensing ranges," *Pervasive and Mobile Computing,* vol. 52, pp. 1-12, 2019.

[19] J. Yu, S. Wan, X. Cheng, and D. Yu, "Coverage contribution area based $ k $-coverage for wireless sensor networks," *IEEE Transactions on Vehicular Technology,* vol. 66, no. 9, pp. 8510-8523, 2017.

[20] J.-P. Sheu and H.-F. Lin, "Probabilistic coverage preserving protocol with energy efficiency in wireless sensor networks," in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, 2007, pp. 2631-2636: IEEE.

[21] C.-I. Weng, C.-Y. Chang, C.-Y. Hsiao, C.-T. Chang, and H. Chen, "On-supporting energy balanced $ k $-barrier coverage in wireless sensor networks," *IEEE Access,* vol. 6, pp. 13261-13274, 2018.

[22] L. Jing, W. Ruchuan, H. Huang, and S. Lijuan, "Voronoi-based coverage optimization for directional sensor networks," *Wireless Sensor Network,* vol. 1, no. 05, p. 417, 2009.

[23] T.-W. Sung and C.-S. Yang, "Distributed Voronoi-based self-redeployment for coverage enhancement in a mobile directional sensor network," *International Journal of Distributed Sensor Networks,* vol. 9, no. 11, p. 165498, 2013.

[24] J. R. Marden and J. S. Shamma, "Game Theory and Control," *Annual Review of Control, Robotics, and Autonomous Systems,* no. 0, 2017.

[25] S. Rahili, "Distributed Optimization in Multi-Agent Systems: Game Theory Based Sensor Coverage and Continuous-Time Convex Optimization," UC Riverside, 2016.

[26] M. Movassagh and H. S. Aghdasi, "Game theory based node scheduling as a distributed solution for coverage control in wireless sensor networks," *Engineering Applications of Artificial Intelligence,* vol. 65, pp. 137-146, 2017.

[27] M. Varposhti, P. Saleh, S. Afzal, and M. Dehghan, "Distributed area coverage in mobile directional sensor networks," in *2016 8th International Symposium on Telecommunications (IST)*, 2016, pp. 18-23: IEEE.

[28] M. Varposhti, V. Hakami, and M. Dehghan, "Distributed coverage in mobile sensor networks without location information," *Autonomous Robots,* vol. 44, no. 3, pp. 627-645, 2020.

[29] M. Hasanbeig and L. Pavel, "Distributed Coverage Control by Robot Networks in Unknown Environments Using a Modified EM Algorithm," *World Academy of*

Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, vol. 11, no. 7, pp. 721-729, 2017.

[30] R. Cerulli, R. De Donato, and A. Raiconi, "Exact and heuristic methods to maximize network lifetime in wireless sensor networks with adjustable sensing ranges," *European Journal of Operational Research,* vol. 220, no. 1, pp. 58-66, 2012.

[31] J. R. Marden, "Learning in large-scale games and cooperative control," CALIFORNIA UNIV BERKELEY DEPT OF MECHANICAL ENGINEERING2007.

[32] H. P. Young, *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton University Press, 2001.

[33] D. P. Foster and H. P. Young, "Regret testing: learning to play Nash equilibrium without knowing you have an opponent," *Theoretical Economics,* vol. 1, no. 3, pp. 341-367, 2006.

[34] I. Ariel and Y. Babichenko, *Average testing and the efficient boundary*. Center for the study of Rationality, 2011.

[35] D. Fudenberg and E. Maskin, "The folk theorem in repeated games with discounting or with incomplete information," in *A Long-Run Collaboration On Long-Run Games*: World Scientific, 2009, pp. 209-230.

[36] H. P. Young, "Learning by trial and error," *Games and economic behavior,* vol. 65, no. 2, pp. 626-643, 2009.

[37] H. P. Young and B. S. Pradelski, "Learning Efficient Nash Equilibria in Distributed Systems," 2010.

[38] A. Menon and J. S. Baras, "Convergence guarantees for a decentralized algorithm achieving Pareto optimality," in *2013 American Control Conference*, 2013, pp. 1932-1937: IEEE.

Elham Golrasan received her M.Sc. degree in Computer Engineering from Sharif University of Technology, Tehran, Iran, in 2013 and her PhD degree from Malek-Ashtar University of Technology, Tehran, Iran, in 2021. Her research interests include Wireless Sensor Networks, Game Theory and Learning Algorithms.

Marzieh Varposhti received her B.Sc. and M.Sc. from Isfahan University, Isfahan, Iran in 2005 and 2008 respectively, and Ph.D. from Amirkabir University of Technology (AUT), Tehran, Iran in 2015. She is currently an assistant professor in the department of computer engineering at Shahrekord University. Her research interests are in wireless networks, sensor networks, and IOT.