Adversarial Uncertainty in Multi-Robot Patrol

Noa Agmon, Sarit Kraus, Gal A. Kaminka and Vladimir Sadov *
Department of Computer Science
Bar Ilan University, Israel
{segaln, sarit, galk, sadovv}@cs.biu.ac.il

Abstract

We study the problem of multi-robot perimeter patrol in adversarial environments, under uncertainty of adversarial behavior. The robots patrol around a closed area using a nondeterministic patrol algorithm. The adversary's choice of penetration point depends on the knowledge it obtained on the patrolling algorithm and its weakness points. Previous work investigated full knowledge and zero knowledge adversaries, and the impact of their knowledge on the optimal algorithm for the robots. However, realistically the knowledge obtained by the adversary is neither zero nor full, and therefore it will have uncertainty in its choice of penetration points. This paper considers these cases, and offers several approaches to bounding the level of uncertainty of the adversary, and its influence on the optimal patrol algorithm. We provide theoretical results that justify these approaches, and empirical results that show the performance of the derived algorithms used by simulated robots working against humans playing the role of the adversary is several different settings.

1 Introduction

The problem of multi-robot patrol has gained interest in recent years [Ahmadi and Stone, 2006; Chevaleyre, 2004; Agmon *et al.*, 2008a], mainly due to its applicability in various security applications. In this problem, robots are required to repeatedly visit a target area, to monitor it. Many researches have focused on a frequency-based approach, guaranteeing some point-visit frequency criteria are met by the patrol algorithm. Others [Paruchuri *et al.*, 2007b; Agmon *et al.*, 2008a; 2008b; Amigoni *et al.*, 2008] have advocated an adversarial approach, in which the robots' goal is to patrol in a way that maximizes their chances of detecting an adversary trying to penetrate through the patrol path. This problem is inherently different from the frequency driven patrol problem, mainly in the need to add randomization to the robots' behavior.

*This research was supported in part by ISF grant #1357/07 and #1685/07.

Recent investigations have examined the optimality of patrol algorithms in two extreme adversarial settings, which vary in the knowledge of the adversary on the patrol algorithm and its parameters. Agmon et al. [Agmon *et al.*, 2008b] has explored optimal algorithms for a zero-knowledge adversary which chooses its penetration point arbitrarily. More commonly, a worst-case full-knowledge adversary is investigated, which is assumed to know the randomization parameters (e.g., heading change probability), and therefore select the optimal penetration point [Paruchuri *et al.*, 2007b; Agmon *et al.*, 2008a; Amigoni *et al.*, 2008]. However, realistically, most adversaries would have neither perfect knowledge nor zero knowledge, but *partial knowledge*. Unfortunately, optimal algorithms for either extreme case fail in partial-knowledge cases.

This paper provides a theoretical discussion of the case in which the adversary lies somewhere along the knowledge continuum, between full and zero knowledge. Specifically, we focus on the influence of the adversary's partial knowledge on its uncertainty of its choice of action, and the impact of this uncertainty on the robots' optimal patrol algorithm.

We describe two approaches for bounding the uncertainty of the adversary in its choices. In the first approach, we assume the adversary will choose to penetrate through one of the v weakest spots, i.e., the v spots with minimal probability of penetration detection. In the second approach, we assume the adversary will choose to penetrate through the set of v points surrounding the weakest spot. Both cases generalize the knowledge continuum; for both, maximal v corresponds to zero knowledge (thus maximal uncertainty), and minimal v is equivalent to full knowledge adversary (no uncertainty).

For both approaches, we present *optimal patrol algorithms*, which have polynomial run-time complexity. We prove that in some cases the v physical neighborhood and v weakest spots algorithms are equivalent. However, in many cases, their predictions of the adversary's actions are different.

We therefore provide an empirical evaluation of the two patrol algorithms, using 71 human subjects that played as adversaries against simulated robots. We compare the two approaches to two other algorithms: MaxiMin, proven optimal for a full-knowledge adversary [Agmon *et al.*, 2008a], and a novel heuristic algorithm MidAvg that averages the MaxiMin and the zero-knowledge algorithm.

Results show that given partial knowledge (correspond-

ing to limited observation time), the two v-algorithms outperformed the others. Moreover, the MaxiMin algorithm—optimal for full-knowledge adversary—performed poorly. We discuss the results of the game, and the possible reasons for reaching such results.

2 Background

Multi-robot patrolling algorithms have been studied in various contexts. Many of these focused on optimizing frequency criteria [Ahmadi and Stone, 2006; Chevaleyre, 2004; Elmaliach *et al.*, 2008], without any reference to the existence of an adversary. In this paper we consider the problem of multi-robot patrol in adversarial environments [Agmon *et al.*, 2008a; 2008b; Sak *et al.*, 2008], which is inherently different from all frequency-driven patrol approaches.

Agmon et. al. [Agmon *et al.*, 2008a; 2008b] investigated multi-robot *adversarial* perimeter patrolling algorithms for full- and zero-knowledge cases. They introduced the robot motion model that we also utilize. They describe the MaxiMin polynomial-time algorithm that maximizes the minimal probability of penetration detection along the perimeter (i.e., improves the weakest point of penetration). This algorithm is proven optimal for a full-knowledge adversary. For the partial knowledge case, Agmon et al. propose a heuristic algorithm, yet provided no theoretical discussion of this case. In contrast, in this paper we focus on the partial knowledge case, and provide both theoretical and empirical analysis.

Sak et. al. [Sak *et al.*, 2008] considered the case of multi-agent patrol in general graphs (rather than perimeters, as is our focus here). They concentrated on an empirical evaluation (using a simulation, with no human subjects involved) of several non-deterministic patrol algorithms that can be roughly divided into two: Those that divide the graph between the patrolling agents, and those that allow all agents to visit all parts of the graph. They considered three types of adversaries: random adversary, an adversary that always chooses to penetrate through a recently-visited node and an adversary that uses statistical methods to predict the chances that a node will be visited soon. They concluded that there is no patrol metric that outperformed the other in all the domains they have checked, but the results depend on the environment. In contrast to this investigation, we provide empirical results from tests with human subjects, and theoretic proofs of optimality for different settings.

Other closely related work is the work by Paruchuri et. al. [Paruchuri *et al.*, 2007b; 2007a], which considered the problem of placing security checkpoints in adversarial environments. They use policy randomization for the agents' behavior in order to maximize their rewards. In their work, the adversary has full knowledge of the agents' behavior, therefore it can use it in order to minimize its probability of being caught in some checkpoint. They again do not consider sensorial scenarios which depend on different sensorial models of the robots. Pita et. al. [Pita *et al.*, 2009] continued this research to consider the case in which the adversary makes their choice based on their bounded rationality or uncertainty, rather than make the optimal game-theoretic choice. They considered three different types of uncertainty over the adversary's choices, and provide new algorithms that deal with these types of uncertainties. In our work we discuss other aspects of uncertainty in adversary's choice, and provide *optimal polynomial-time* solutions.

Amigoni et. al. [Amigoni *et al.*, 2008] also used a game-theoretic approach for determining the optimal strategy for patrolling agents, using leader-follower games. They consider an environment in which a robot can move between any two nodes in a graph, as opposed to the perimeter model we focus on. Their solution is suitable for one robot, and since the computation of the optimal strategy is exponential, they described a heuristic approach for finding a solution.

3 Robot and environment model

We are given a team of k homogenous robots, required to patrol around a closed area (perimeter). The perimeter is divided into N segments, where the travel time of each robot through a segment is uniform, i.e., all robots travel through one segment per time cycle . Hence the segments' length is uniform in time, but not necessarily in distance.

The robots have directionality associated with their movement, i.e, if they go backwards they physically turn around. We model the cost of turning around in time, and denote time it takes the robots to turn around and stabilize in their new direction by τ. In this paper we focus on the case of $\tau = 1$.

The system of perimeter patrol is linear, meaning that at each time step the robots have one of two options: go straight or turn around. Therefore the robot's patrol algorithm is characterized by a probability p, i.e., at each time step go straight forward with probability p, or turn around with probability $q = 1 - p$.

We consider coordinated robotic systems in the sense that if the robots decide to turn around, they do it simultaneously. Moreover, we require that the robots are placed uniformly along the perimeter, with distance of $d = N/k$ segments between every two consecutive robots along the path. The coordination and uniform-distance requirements are derived from the optimality proofs given in [Agmon *et al.*, 2008b; 2008a], which have shown that the probability of penetration detection is optimized under these conditions. These optimality proofsare based on the fact that the probability of penetration detection decreases as the distance from the robot increases [Agmon *et al.*, 2008a]. Therefore it is best to minimize the maximal distance between every two consecutive robots. This is done by guaranteeing that the distance in time between every two consecutive robots is equal, maintained so by the requirement that the robots are coordinated.

In the adversarial models considered here, the adversary decides at time 0 through which segment to penetrate, and the time it takes it to penetrate is not instantaneous, and lasts t time units.

The chances of the robots to detect an adversary passing through a segment s_i is defined as the *probability of penetration detection* at the segment, and denoted by ppd_i. This is the probability that some robot will pass through segment s_i during t time units, and is a function of the probability of the robots to continue straightforward, p.

4 Uncertainty in the adversary's perspective

In most cases, it is realistic to assume that the adversary's knowledge on the patrol algorithm lies somewhere along the knowledge continuum, between full and zero knowledge. Usually, the adversary does not gain enough information on the patrol algorithm in order to derive the exact algorithm (i.e., probability p) or the exact weakest spots of the algorithm. Therefore we explore theoretically two directions in handling partial knowledge of the adversary. In the first, the adversary might have some estimation of the probability p characterizing the patrol algorithm. In the second, the adversary might have some estimation of the weakest spot of the algorithm. In both cases, we wish to use the region of possible beliefs of the adversary in order to find an optimal patrol algorithm for the patrolling robots.

A common way of handling uncertainties of systems is to assume that when having no knowledge, a random choice, with uniform probability, is made. In this domain, this approach was proven to be useful in an empirical evaluation in [Agmon *et al.*, 2008b], where a patrol algorithm proven to be optimal for a random adversary performed substantially better than other algorithms for humans playing the role of an adversary that had no knowledge of the patrolling robots. We will use a similar approach here, i.e., within the region of estimation of the adversary — either of the patrol algorithm p or of the weakest spots — the adversary will be assumed to choose its actions at random.

We first examine the approach according to which the adversary estimates the probability p characterizing the patrol algorithm with some error. Unfortunately, we show that it is impossible to find an optimal patrol algorithm in this case.

We then discuss two alternative approaches, in which the uncertainty is reflected by the choice of *penetration spot*. In this case, we do not necessarily assume that the adversary calculates the probability p, but tries to estimate the weakest spot using two estimation methods - physical proximity, or probability proximity to the minimal ppd.

4.1 Estimating p - negative result

In this section we discuss the case in which the adversary estimates the probability p characterizing the patrol algorithm.

The problem of estimating the probability p can be considered as observing a Bernoulli trial, where a success is an event of going straight with probability p, and a loss is turning around with probability $1-p$. We can use the Central Limit Theorem [Devore, 1991] that bounds the expected error from the real value of p after viewing it for t_v trials. The average of successes after viewing t_v trials is inside the boundaries $[p-\delta, p+\delta]$ with probability p_{conf}, where δ is a function of t_v and depends on p_{conf}. Therefore the adversary can estimate the real value of p inside some interval around p, and we will try to use this interval in order to optimize the patrol algorithm of the robots. Consider the following problem.

P-Interval **problem definition:**
Let p be the probability characterizing the perimeter patrol algorithm of a team of robots. Assume the adversary estimates that the real value of p is inside the interval $[p-\delta, p+\delta]$. Therefore it chooses its believed p_b at random with uniform probability inside this interval. The algorithm needs to find the probability p characterizing the patrol of the robots such that it maximizes the expected ppd throughout the perimeter.

Unfortunately, we prove that this problem is unsolvable unless $\delta = 0$. We prove it by showing that the expected ppd function inside the interval $[0,1]$ is monotonically increasing, i.e., as p grows the expected ppd grows, hence the optimal p does not converge unless $\delta = 0$ (since maximization of the expected ppd is obtained in the right bound of the interval, which is also the mid point of the interval only when $\delta = 0$).

We omit the proof of the following Lemma due to space constraints.
Let $E_{\text{ppd}}(p)$ be the expected ppd for probability $p \in [0,1]$

Lemma 1. *The expected* ppd*, as a function of p, is a monotonically increasing function in the range $[0,1]$, i.e., for all $0 \leq p' < p \leq 1$, $E_{\text{ppd}}(p') < E_{\text{ppd}}(p)$*

Theorem 2. P-Interval *is unsolvable unless $\delta = 0$.*

Proof. Assume, towards contradiction that $\delta > 0$, yet there exists $p*$ that maximizes the expected ppd throughout the perimeter. By the definition of P-Interval, the adversary deduces an interval around $p*$ in which it chooses its believed p at random inside the interval $[p*-\delta, p*+\delta]$. By Lemma 1, the expected ppd function is monotonically increasing, therefore the maximal expected ppd inside this interval is obtained in $p*+\delta$. This contradicts the assumption that $p*$ maximizes the expected ppd, unless $\delta = 0$. □

4.2 Uncertainty in the choice of penetration spot

In this section we explore the case in which the patrial knowledge of the adversary on the patrol algorithm is translated into different possible options of penetration spots. For several reasons, the adversary might not choose to penetrate through the *exact* weakest spot. We present herein two deviations from the weakest spots, hence two possible corresponding optimal ways of choosing the patrol algorithm in such cases.

The adversary, after studying the robots' patrol for a period of time, could result in several reasonable segments which the ppd values, as it believes, are small enough. In this case it could choose to penetrate through one of the v weakest spots at random, with some probability distribution (for example uniform). Hence the robots should choose p such that the expected ppd along the v segments with minimal ppd is maximal. We refer to this approach by v-Min.

The second case is that the adversary might not choose to penetrate through the segment with the minimal ppd, but either through that segment, or through one of its neighboring segments at random. Hence the robots should choose p such that the minimal expected ppd along v neighboring segments is maximized. This approach is referred to as v-Neighbor.

Note the difference between the two cases - in v-Min we are looking for the value $0 \leq p \leq 1$ such that the weighted average of the v minimal ppd's is maximized, and in v-Neighbor case we are looking for p such that the minimal weighted average of v neighboring segments is maximized.

In both cases, the two extremities of uncertainties—full knowledge adversary (no uncertainty) and zero knowledge adversary (complete uncertainty)—match the results obtained

by [Agmon *et al.*, 2008a; 2008b], respectively. If $v = 1$, i.e., there is no uncertainty in the choice of the weakest spot, then the algorithms are required to return exactly the value p such that the minimal ppd is maximized, similar to the MaxiMin algorithm presented in [Agmon *et al.*, 2008a]. On the other hand, if $v = d$ and the probability distribution is uniform, then the algorithms will return the value p that maximized the expected ppd throughout the perimeter (=average ppd). As proven in [Agmon *et al.*, 2008b], the optimal algorithm in this case is $p = 1$, i.e., the deterministic algorithm.

The algorithms for finding an optimal patrol uses the ppd_i function for each segment s_i. These ppd_i are functions of p, and are calculated in polynomial time using a dynamic-programming algorithm described in [Agmon *et al.*, 2008a].

Optimality of the patrol algorithm using the v-Min approach

We present herein Algorithm ComputeMinV that finds the optimal patrol algorithm, corresponding to the probability p of going straight at each time step under the v-Min scenario. Specifically, Algorithm ComputeMinV computes the value p such that the minimal v ppd's are maximized, given a probability distribution $V = \{v_i, v_2, \ldots, v_v\}$, where v_i is the probability that the adversary will choose to penetrate through the i'th weakest spot, $\sum_{i=1}^{v} v_i = 1$. This distribution can be used to further manipulate the impact of the extent of knowledge of the adversary on its choice of penetration spot, for example after obtaining more knowledge v_1 may increase to more than the uniform distribution $(1/v)$.

Algorithm 1 ComputeMinV$(v, V, \{\text{ppd}_1, \ldots, \text{ppd}_d\})$

1: Set $BufP \leftarrow \{0, 1\}$ {initialize list of all intersection points}
2: **for** every pair $\text{ppd}_i, \text{ppd}_j$, $1 \leq i, j \leq d$, $i \neq j$ **do**
3: $Intersect_{i,j} \leftarrow$ intersection points between ppd_i and ppd_j.
4: $BufP \leftarrow BufP \bigcup Intersect_{i,j}$
5: Sort $BufP$ in ascending order
6: $Res_f, Res_p \leftarrow 0$ {initialize maximin value and its p }
7: **for** $j \leftarrow 1$ to $|BufP|$ **do**
8: Find v functions $f_{j_1}, \ldots, f_{j_v}$ such that $f_{j_i}(p') < f_n(p')\ \forall p' \in [BufP(j), BufP(j+1)], 1 \leq i \leq v$, $f_n \neq f_{j_i}$
9: $f_{avg} \leftarrow \sum_{i=1}^{v} v_i \times f_{j_i}$
10: $m \leftarrow f_{avg}(p*)$ such that $\forall p \in [BufP(j), BufP(j+1)]$, $f_{avg}(p*) \geq f_{avg}(p)$
11: **if** $m > Res_f$ **then**
12: $Res_f \leftarrow m$; $Res_p \leftarrow p*$
13: Return Res_p

The algorithm works as follows. First, it identifies all intersection points between every pair of ppd_i, ppd_j functions $(1 \leq i, j \leq d, i \neq j)$. Then it divides the range $[0, 1]$ to sections according to all the intersection points. For each section $[p_a, p_b]$, the algorithm identifies the minimal v curves between $[p_a, p_b]$, and finds their average curve, f_{avg}. Since the adversary chooses to penetrate through one of the v segments with lowest ppd at random with the given distribution V, the *weighted average* (given weight v_i to the i'th minimal curve) of the v curves represent the *expected* ppd in that section. Last, ComputeMinV calculates the maximal value of $f_{avg}(a, b)$ in the section $[p_a, p_b]$, and reports the point p_{opt} that is maximal among all minimal points of the average functions. An illustration of this algorithm is shown in Figure 1.

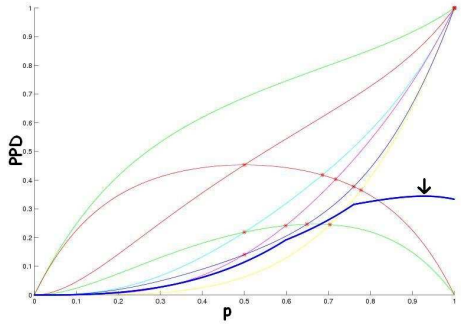

Figure 1: *An illustration of Algorithm* ComputeMinV *for* $d = 8, t = 6, v = 3$. *The small stars mark the intersection points, and the bold curve is the average of the* 3 *minimal curves at each section. The arrow marks the maximal point computed by* ComputeMinV.

The time complexity of ComputeMinV is $\mathcal{O}(d^4 + d^3 \log d^3)$ (compared to time complexity of $\mathcal{O}(d^3)$ of the MaxiMin algorithm for full knowledge adversary [Agmon *et al.*, 2008b]).

Optimality of the patrol algorithm using the v-Neighbor approach

As stated previously, the adversary might attempt to penetrate not only through the weakest segment, but through one of its neighboring segments. Therefore this can be used in order to find a patrol algorithm (p value) more suitable for the situation. Algorithm ComputeNeighborV computes the weighted average of v neighboring segments according to a distribution $V = \{v_1, \ldots v_v\}$, then finds the maximin point of the new curves. Note that if the robot currently resides inside the v-neighborhood of a segment s_i (i.e., $v - i < 0$ or $v + i > d$), its current location is excluded, i.e., we average fewer segments for that case. The probability distribution can be used to express the fact that the adversary tends, for example, to try and penetrate through the segments further away from the robot in its current position. Figure 2 illustrates the algorithm for $d = 8$, $t = 6$ and $v = 3$.

Algorithm 2 ComputeNeighborV$(v, V, \{\text{ppd}_1, \ldots, \text{ppd}_d\})$

1: Set $FuncSet \leftarrow \emptyset$
2: **for** $i \leftarrow 1$ to d **do**
3: $i_e = \min(d, i + v)$
4: $FuncSet \leftarrow \sum_{j=i}^{i_e} v_{j-i+1} \times \text{ppd}_j \bigcup FuncSet$
5: $p_{opt} \leftarrow$ MaxiMin$(FuncSet, d)$
6: Return p_{opt}

The time complexity of Algorithm ComputeNeighborV is $\mathcal{O}(d^3)$ (similar to the complexity of MaxiMin).

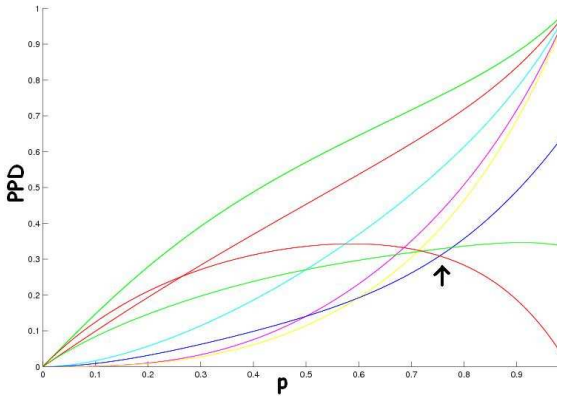

Figure 2: *An illustration of Algorithm* ComputeNeighborV *for* $d = 8, t = 6, v = 3$*. The curves are* not *the original* ppd_i *functions, but the average of* $v-$*neighborhood of each segment. The arrow points to the maximin point of the new curves.*

Comparing v-Min **and** v-Neighbor

The two approaches of v-Min and v-Neighbor towards bounding the uncertainty of the adversary in its choice of penetration spot might seem to be inherently different. Consider for example the case in which $d = 8$, $t = 6$ and $v = 3$ (Figures 1 and 2). The optimal p in case of v-Neighbor is $p = 0.7359$, and the optimal p for v-Min is $p = 0.9273$. The result returned by the MaxiMin algorithm (used in case of a full knowledge adversary, i.e., $v = 1$) is $p = 0.7037$.

However, in some cases they coincide, as proven in the following Theorem.

Theorem 3. *The optimal choice of* p *according to* v-Neighbor *coincides with the optimal* p *according to* v-Min *if* $t = \lfloor d/2 \rfloor + 1$.

Proof. The optimal p in the $v-$neighborhood and in the $v-$minimal, is the one maximizing the minimal ppd of the average of the $v-$ neighborhood and $v-$ minimal ppd_i functions, correspondingly. Therefore it is enough to show that along this optimal point p_o, the $v-$ minimal ppd_i functions are also all neighbors. Formally, we need to show that $\mathsf{ppd}_{i_1}, \ldots, \mathsf{ppd}_{i_v}$ are minimal, where $i_l = j + l$ for some index $1 \leq j \leq d - v$.

Consider the section of d segments between two consecutive robots R_a and R_b. First, assume d is odd. In this case, ppd_i for $1 \leq i \leq t$ is influenced only by R_a, and ppd_i for $t + 1 \leq i \leq d$ is influenced only by R_b. Moreover, every ppd_i function for $1 \leq i \leq t$ equals 0 if $p = 0$, and equals 1 if $p = 1$. On the other hand, every ppd_i function for $t + 1 \leq i \leq d$ equals 0 in both $p = 0$ and $p = 1$.

Note that if a robot is headed clockwise, then any ppd function of a segment s_i of distance i to its right is larger than a ppd function of segment which is in the same distance, but to the left. For example, $\mathsf{ppd}_1 > \mathsf{ppd}_d$, $\mathsf{ppd}_2 > \mathsf{ppd}_{d-1}$ and so on. The reason lies in the fact that the probability of reaching a segment of distance i in the opposite direction is equivalent to the probability of reaching a segment in distance i in the same direction, but multiplied by $(1 - p)$. Since we assume that $p \leq 1$, this it always true.

Combining all known facts together, from Lemma 1 in [Agmon *et al.*, 2008a] we see that $\mathsf{ppd}_{t+1} \leq \mathsf{ppd}_{t+2} \leq \ldots \leq \mathsf{ppd}_{d-1} \leq \mathsf{ppd}_d$ and $\mathsf{ppd}_t \leq \mathsf{ppd}_{t-1} \ldots \leq \mathsf{ppd}_1$. Also, as shown herein, $\mathsf{ppd}_1 \geq \mathsf{ppd}_d$, $\mathsf{ppd}_2 \geq \mathsf{ppd}_{d-1}, \ldots, \mathsf{ppd}_{t-1} \geq \mathsf{ppd}_{t+1}$. It follows that, necessarily, the minimal function is ppd_{t+1}, the function above it is ppd_t and ppd_{t-1}, followed by ppd_{t+2} and ppd_{t-2} and so on (see an example in Figure 3). Therefore, necessarily, when considering the $v-$ minimal segments, for all v, we remain in the $v-$ neighborhood of ppd_{t+1}.

If d is even, then the only difference is that function ppd_t receives components from both R_a and R_b, hence it is not straightforward that it is smaller than ppd_{t-1}. Calculating the exact value of ppd_t shows us that $\mathsf{ppd}_t = p^t + (1 - p)p^{t-1} = p^{t-1}$. On the other hand, $\mathsf{ppd}_{t-1} = p^{t-1}$, i.e., $\mathsf{ppd}_{t-1} = \mathsf{ppd}_t$, and the rest of the proof follows directly as in the case of an odd d. □

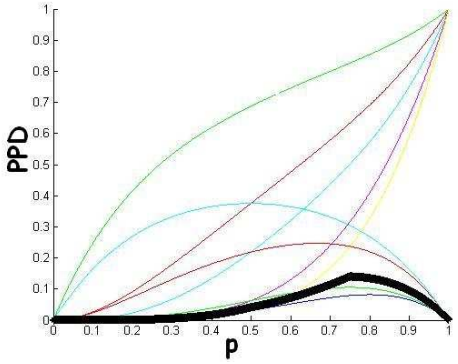

Figure 3: *An illustration of proof of Theorem 3, in which the* $v-$*neighborhood and* $v-$ *minimal coincide (here* $d = 9$, $t = 5$ *and* $v = 3$*). The bold line represents the average of* $v-$ *minimal /* $v-$ *neighboring segments.*

5 Evaluation

In order to evaluate the performance of the suggested algorithms when working against different adversaries, we created a variation of the PenDet-game that was described in [Agmon *et al.*, 2008b]. In this game, the adversary is played by a human subject, working against simulated robots in a Web-based environment. Using humans as adversaries models the realistic requirement of such a system, i.e., the robots will perform in real world against human adversaries in various environments (similar evaluation method is also used in [Pita *et al.*, 2009; Agmon *et al.*, 2008b]).

The patrol algorithms executed by the robots were calculated according to the following:

v-Min (with several v values)

v-Neighbor (with several v values)

MaxiMin, which maximizes the minimal ppd along the perimeter, proven by [Agmon *et al.*, 2008a] to be optimal

against a full-knowledge adversary
MidAvg, a novel heuristic algorithm that averages between the p value of the optimal algorithms against full and zero knowledge adversaries (MaxiMin and deterministic algorithms, respectively).

5.1 Experimental setting

The game was played by 71 human subjects, all undergraduate Computer Science students, playing the role of the adversary that tries to penetrate through the simulated robots. The subjects received both an oral presentation explaining the rules of the game, and were handed additional explanation sheets.

Each trial was composed of 6 subgames. Each such subgame starts with an observation phase of 60 seconds, in which the player studies the patrol algorithm by observing the actions of the robots in order to choose a penetration spot.

Preliminary experiments that included observation periods of 5 and 30 seconds have shown that these observation periods were not long enough to enable learning of the robots' patrol algorithm, hence the choices made by the subjects were arbitrary. Thus we focused on an observation period of 60 seconds.

After the observation phase, the players chose a penetration spot through which they assumed to maximize their chance of penetrating without being detected by the patrolling robots. Each player played 6 subgames, however the player did not get feedback on whether the penetration attempt was successful.

We checked two pairs of (d, t) values: $(8, 6)$, in which the resulted patrol algorithm was different for the v-Min and v-Neighbor for $v = 2, 3$. In the second pair, $(16, 9)$, as proven also by Theorem 3, the results of v-Min and v-Neighbor coincide, hence we checked the following v values: $v = 3, 5, 7, 9$. We considered only the uniform distribution of V, i.e., $v_i = 1/v$.

Each set of (d, t) values and algorithm (characterized by probability p) was played by 34 to 37 subjects. The order of the subgames in the trial was randomly selected, and there were no repetitions of sets in one trial.

5.2 Experimental results and discussion

Figures 4, 5 describes the *expected* probability of penetration detection given the players' choice of penetration locations for $d = 8, t = 6$ and $d = 16, t = 9$ (respectively) for all algorithms described above, given an observation time of 60 seconds, where the patrol algorithm was unknown to the player. The bars represent the expected penetration detection ratio given the actual choices of the players' penetration spots. In order to compare the performance results obtained by the different algorithms, we used the Mann-Whitney U-test [Mann and Whitney, 1947], which is a non-parametric test, suitable for data with no normal distribution (like the data in our case).

For the first case in which $d = 8, t = 6$ we can clearly see that the best-performing algorithm, i.e., the algorithm that achieved the highest expected probability of penetration detection based on the choices of the players, was v-Min for $v = 3$ (denoted by $3-$Min). Specifically, the results of $3-$min were statistically significantly better than v-Neighbor and

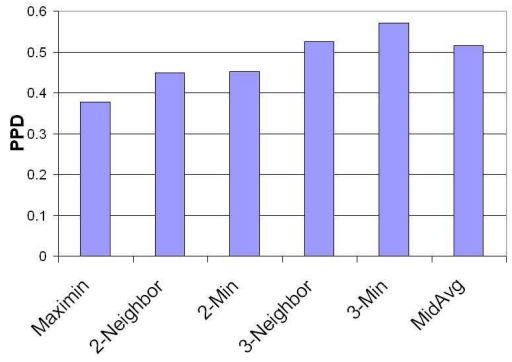

Figure 4: Results of the experiment for $d = 8, t = 6$. The bars represent the expected penetration detection ratio of the robots given the actual choices of the players.

v-Min for $v = 2$ ($p - value = 0.001$ and $p - value = 0.01$, respectively), MaxiMin ($p - value = 0.003$) and MidAvg ($p - value < 0.002$). However, the results of $3-$min were not significantly better than v-Neighbor for $v = 3$ (denoted by $3-$Neighbor).

In order to explain the advantage of using $3-$Min, we inspected the actual choices the players made concerning their penetration spots. Approximately 50% of the players decided to penetrate through one of the 3 segments with minimal ppd, and the expected ppd in these segments is 34%. In contrast, only approximately 29% of the players detected the weakest spot when executing the MaxiMin algorithm (in this case having an expected ppd of 24%). This means that the $3-$Min algorithm indeed had better predictions concerning the penetration spots.

Another reason for the $3-$Min's good performance lies in the fact that the other 50% of the players who didn't choose to penetrate through the weakest spots, had better chances of getting caught by the $3-$Min algorithm also in the other segments. The MaxiMin algorithm attempts to strengthen the weakest spot, and thus it substantially decreases the probability of penetration detection in the other segments. For example, for $d = 8, t = 6$, the expected ppd in the non-weakest segments using the MaxiMin algorithm is 49%, whereas with the $3-$Min algorithm it is 83%. The minimal ppd, though, decreases from 24% to 11% with the $3-$Min algorithm.

Since the players did not obtain enough information to identify the exact weakest spots and enter through those spots, the use of the MaxiMin algorithm was not worthwhile. The $2-$Min and $2-$Neighbor algorithms suffer from the same problem, though not as profoundly as the MaxiMin does. Therefore they did not perform as well as the 3-min or 3-neighbor algorithms. The $3-$Min algorithm performed better than the $3-$Neighbor algorithm, yet not significantly better, since they both assume similar uncertainty level (3 segments).

Note that it may be worthwhile to enlarge the level of uncertainty, i.e., the v value in order to capture more choices of penetration spots. However, if $d = 8, t = 6$, for $v > 3$ the optimal algorithm is deterministic, which is highly predictable, and as shown in [Agmon *et al.*, 2008b], it is easily

manipulated by an adversary with even a small amount of information.

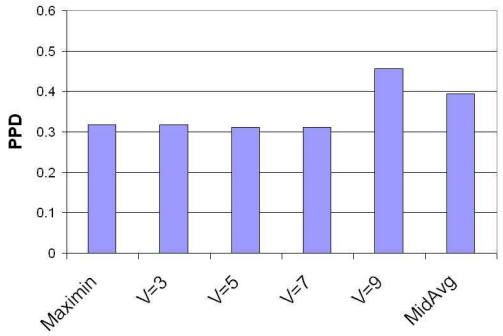

Figure 5: Results of the experiment for $d = 16, t = 9$. The bars represent the expected penetration detection ratio of the robots given the actual choices of the players.

For $d = 16, t = 9$ we see from the expected ppd values that v-Min / v-Neighbor for $v = 9$ (denoted by $V - 9$) performs considerably better than algorithms with smaller v's (tested for $v = 3, 5, 7$), MaxiMin and MidAvg. However, we were not able to obtain statistical significance using the Mann-Whitney U-test. The reason lies in the fact that $V - 9$ expects the players to penetrate through one of the 9 weakest segments, and indeed 68% of the players chose to penetrate through these segments. Therefore the expected ppd consists of a large range of possible values that are not normally distributed, therefore even though the $V - 9$ produced highest levels of ppd, we could not show significance.

6 Conclusions and future work

In this paper, we considered the problem of multi-robot perimeter patrol in adversarial environments, under uncertainty of adversarial behavior. In this case, the adversary has some knowledge on the patrolling robots — on the scale between zero to full knowledge, yet it is uncertain in determining its best option for penetration spot. We bounded the uncertainty in such a way that enabled us to find an optimal patrol algorithm for the robots that will maximize their chances of detecting the adversary. We described two approaches for dealing with this problem. In the first, we assumed that the adversary will choose to penetrate through some spot which is in some physically proximity to the weakest spot of the patrol, and in the second we assumed the adversary would penetrate through one of the weakest spots. We described a polynomial time algorithm for determining the optimal patrol algorithm for each case, and have shown when these two approaches coincide. We performed a massive experiment, using 71 human subjects playing the role of the adversary against simulated robots, and have shown significant benefits in using our suggested algorithms.

For future work we consider the following points. We intend to further investigate uncertainties in the adversary's choice. As a first step, we would like to perform additional empirical evaluation with a longer learning phase of the patrolling robots, and try to extract the transition points between possible v values. We would also like to examine possible uncertainties in the internal robotic system, originated for example in faulty movement and faulty sensing.

References

[Agmon *et al.*, 2008a] N. Agmon, S. Kraus, and G. A. Kaminka. Multi-robot perimeter patrol in adversarial settings. In *ICRA*, 2008.

[Agmon *et al.*, 2008b] N. Agmon, V. Sadov, S. Kraus, and G. A. Kaminka. The impact of adversarial knowledge on adversarial planning in perimeter patrol. In *AAMAS*, 2008.

[Ahmadi and Stone, 2006] M. Ahmadi and P. Stone. A multi-robot system for continuous area sweeping tasks. In *ICRA*, 2006.

[Amigoni *et al.*, 2008] F. Amigoni, N. Gatti, and A. Ippedico. Multiagent technology solutions for planning in ambient intelligence. In *Proc. of Agent Intelligent Technologies (IAT)*, 2008.

[Chevaleyre, 2004] Y. Chevaleyre. Theoretical analysis of the multi-agent patrolling problem. In *Proc. of IAT*, 2004.

[Devore, 1991] J. L. Devore. *Probability and Statistics for Engineering and the Sciences*. Brooks/Cole Publishing Company, 1991.

[Elmaliach *et al.*, 2008] Y. Elmaliach, A. Shiloni, and G. A. Kaminka. A realistic model of frequency-based multi-robot fence patrolling. In *AAMAS*, 2008.

[Mann and Whitney, 1947] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 1947.

[Paruchuri *et al.*, 2007a] P. Paruchuri, J. P. Pearce, M. Tambe, F. Ordonez, and S. Kraus. An efficient heuristic approach for security against multiple adversaries. In *AAMAS*, 2007.

[Paruchuri *et al.*, 2007b] P. Paruchuri, M. Tambe, F. Ordonez, and S. Kraus. Security in multiagent systems by policy randomization. In *AAMAS*, 2007.

[Pita *et al.*, 2009] J. Pita, M. Jain, F. Ordonez, , M. Tambe, S. Kraus, and R. Magorii-Cohen. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties, to appear. In *AAMAS*, 2009.

[Sak *et al.*, 2008] T. Sak, J. Wainer, and S. K. Goldenstein. Probabilistic multiagent patrolling. In *SBIA*, pages 124–133, 2008.