Robotic Strategic Behavior in Adversarial Environments*

Noa Agmon
Department of Computer Science, Bar-Ilan University, Israel
agmon@cs.biu.ac.il

Abstract

The presence of robots in areas containing threats is becoming more prevalent, due to their ability to perform missions accurately, efficiently, and with little risk to humans. Having the robots handle adversarial forces in missions such as search and rescue, intelligence gathering, border protection and humanitarian assistance, raises many new, exciting research challenges. This paper describes recent research achievements in areas related to robotic mission planning in adversarial environments, including multi-robot patrolling, robotic coverage, multi-robot formation, and navigation, and suggests possible future research directions.

1 Introduction

In developing robots to fulfill a wide range of functional goals, it is necessary to address not only their ability to perform the tasks at hand, but also the ways in which they act within and respond to the various characteristics of their surroundings. As one of the foremost motives for replacing humans with robots in a task or mission involves proximity to dangerous or hostile entities, an emerging body of research is highlighting the need to account for the presence of adversaries in robotic environments. Dubbed *adversarial robotics*, this field is gaining traction, particularly in light of the growing reliance on unmanned vehicles and other robots in hazardous search and rescue missions and in security-related settings worldwide.

Research in adversarial robotics focuses on the considerations and challenges that arise when canonical robotic problems must be solved in the presence of opponents with conflicting goals or an intention to harm the robots involved. When an adversary is introduced into a system, these problems change from their original forms, and therefore require fundamentally different solutions. For example, as elaborated below, the aim of a robot charged with patrolling a certain threat-free area will be to optimize frequency criteria, while the aim of a robot patrolling the same area, now inhabited by an adversary, will be to maximize the detection of that adversary. Clearly, these two aims raise different considerations and call for different solutions.

*This research was funded in part by ISF grant 1337/15

Adversarial presence in robotic problems can be divided roughly into two: non-strategic, and strategic adversary. If the adversary is non-strategic, the risk (or threats) it poses exist in the environment and are fixed, thus the robots plan their task while accounting for those as given, static, factor. On the other hand, if the adversary adopts a strategic behavior, it reacts to the strategy chosen by the robots. Thus ultimately, alongside the strategic behavior of the robots themselves, problems in adversarial robotics require that the knowledge and behavior of the adversaries be modeled and addressed. To this point, research that considered adversarial presence mostly assumed (implicitly or explicitly) a non-strategic adversary. Many challenges and open questions exist assuming both adversary types, however the more general picture is the strategic adversary, as it includes within it the case of a non-strategic, fixed, adversary.

Adversarial robotics can be viewed from the perspective of game theory, with robots and adversaries each trying to maximize their self-perceived utility as opponents in a game. Comprehensive work on game theory in the past decades has indeed produced several models and tools (e.g., matrix game representation) that are theoretically and practically relevant to various aspects of adversarial robotics. However, the applicability of these pure game-theoretic methods is limited by the dynamic and continuous nature of robotic behavior and environments, as well as by the uncertainty of robot and adversary perceptions and actions. Taking these constraints into consideration, pioneering research has focused on the development of tailor-made approaches to account for adversarial presence in robotic environments.

A significant body of research has addressed the presence or involvement of opponents in robotic problems without explicitly modeling adversarial knowledge or behavior. Noteworthy examples include the issues of stealth navigation [Tews *et al.*, 2004] and covert path planning [Al Marzouqi and Jarvis, 2011], in which robots must navigate through their surrounding environment while minimizing detection by opponents. In the games of soccer [Stone *et al.*, 2000] and capture the flag [Huang *et al.*, 2011], though robots need to plan their behavior against the opposing team (with the purpose of scoring goals or capturing their flag), the work done to date on this paradigm has rarely involved strategic responses based on modeling of opponent behavior.

Overall, work in these areas has provided valuable insight

into robotic behavior in the face of an opponent, serving as the foundation for later attempts to model adversarial behavior and compute response strategies in accordance. This pattern, in which a known, non-reactive (static) adversary is initially assumed, and only then are the complications of strategic behavior in response to a goal-oriented and dynamic adversary addressed, is representative of a general approach to solving new problems in adversarial environments. A similar progression is demonstrated in the work on adversarial coverage and adversarial formation, as detailed in the corresponding sections below.

In the sections that follow, I review this work, which has, to date, focused on the problems of multi-robot patrolling, robotic coverage, multi-robot formation, and robot navigation in adversarial environments. While each of these problems entail a different set of considerations and solutions, the premise underlying all research in this novel field is that intelligent adversaries and dynamic threats are an inevitable component of virtually every real-world environment in which robots can be of benefit.

According to this conceptualization, which can be called *the paranoid conjecture*, the adversary is always there, though its behavior, and hence the response it requires, depends on what it knows. As demonstrated above, if the adversary knows nothing about the strategy of its opponent robots, it is best for them to behave as though it were not there at all.

2 Multi-robot Patrolling

Due largely to its immediate relevance to pressing security applications, the problem of multi-robot patrol has garnered much attention in recent years (e.g., [Chevaleyre, 2004; Ahmadi and Stone, 2006; Portugal and Rocha, 2011; Agmon *et al.*, 2011; 2012; Sless *et al.*, 2014]). In the absence of adversaries, the basic patrolling problem requires robots to repeatedly visit a target area in order to monitor changes in state [Chevaleyre, 2004]. The majority of studies in this field (e.g., [Elmaliach *et al.*, 2009]) involve frequency-based patrolling, in which the robots' goal is to optimize a particular frequency criterion, such as maximization of the minimal frequency of visits to a point in the environment or the average frequency of visits along the entire area. In such cases, efficient patrol is defined as guaranteeing a high frequency of visits to all parts of the designated area, which may be a continuous $1D$ or $2D$ environment or a discrete set of waypoints (graph).

As the visit frequency-optimization problem is NP-hard, several heuristic and/or approximation strategies have been proposed to solve it. These solutions may be based on the creation of one cyclic path to be followed by all robots or on the division of the patrol area between robots such that each patrols its own zone. Regardless of the specific method, strategies that optimize point-visit frequency criteria are deterministic by nature. Furthermore, it has been proven [Agmon *et al.*, 2011] that random robot behavior is suboptimal with respect to frequency criteria.

In patrol environments containing adversaries, the goal of robots changes from optimization of frequency criteria to optimization of adversary detection. Given that the robots face an adversary that may learn and adapt to their actions, they must adopt random behavior. However, this randomness should be smart, meaning that it should take the adversarial model into account. Thus, the adversary is assumed to be rational with respect to its knowledge of the patrolling robots, which it applies in choosing actions that maximize its chances of successful penetration.

If the adversary has full knowledge of the patrolling strategy, and knows where the robots currently reside, it will be able to take advantage of this knowledge and penetrate at the patrol's weakest spot (the point with the minimal probability that penetration will be detected). In this case, the robots' strategy should maximize the minimal probability of penetration detection. It has been shown [Agmon *et al.*, 2011] that such a strategy can be found in polynomial time in linear environments, i.e., along an open fence (line) or closed perimeter (circle). This is an example of a leader-follower game (also known as a Stackelberg game), in which the leader (our robots) fixes its strategy and the follower (the adversary) observes the strategy and responds to it.

When it cannot distinguish between two or more options, the adversary chooses either option at random, with uniform distribution. Therefore, on the other end of the adversarial knowledge continuum, if the adversary has no knowledge of the strategy of the patrolling robots, and thus does not know what they will do next, it will choose to penetrate through one of the locations in which they do not reside, at random, with uniform distribution. In this case, the robots will choose a patrol strategy that maximizes the expected probability of penetration detection. In [Agmon *et al.*, 2008] we have shown that a strategy maximizing the expected probability of penetration detection is one that maximizes the frequency of visits at each point in the environment, leading to the conjecture that the original problem of patrolling in neutral environments (without adversarial presence) is essentially a sub-case of the more general problem of patrolling in adversarial environments.

Considering that the adversary (being rational) chooses actions based on its knowledge on the patrolling robots, we have recently shown [Talmor and Agmon, 2017] that this knowledge can be *manipulated*. Examining two different types of manipulation, it is shown that in some cases it is possible to deceive the opponent into thinking that the patrolling robots have more power than they actually have. Specifically, we have shown that if the adversary can view only a part of the perimeter (a *window*), then it is possible to perfectly simulate a patrolling strategy along this window mimicking more robots than there actually are in the system. In addition, when using robots with no actual penetration detection capabilities (similar to scarecrows), then the vulnerability of the system is minimal. Common to both deception models, the ability to perfectly deceive the opponent (without changing any physical capabilities) is based on the stochasticity of the robots' actions.

3 Robotic Coverage

In the problem of robot coverage, a robot (one or more) must visit each point in a target area once. Deciding on a path that covers an area is one of the fundamental problems in robotics, which serves as the basis for problems spanning from search

and rescue, through cleaning, harvesting, and demining, to mapping [Galceran and Carreras, 2013].

In the original problem of coverage (without adversarial presence), the goal of the robot(s) is usually to minimize the time or energy required to complete the coverage. In a single-robot setting, this corresponds with minimizing the length of the path that hits each point at least once. In a multi-robot setting, various criteria can be selected for optimization. Common criteria include minimizing the maximal effort (path length, energy) a robot invests in the coverage, or minimizing the robot's average effort. The problem of robot coverage is NP-Hard, using a reduction from the Traveling Salesman Problem (TSP). Based on some assumptions about the size and the structure of the area with respect to the coverage tool, a popular method known as Spanning Tree Coverage (STC) [Gabriely and Rimon, 2001] produced an optimal solution for a single robot. This method was later generalized to multiple robots [Hazon and Kaminka, 2008] , though it is not optimal in this case. Other solutions perform a greedy heuristic assignment of sub-areas to robots such that the union of those areas covers the entire terrain, or use other TSP approximation algorithms to define a coverage path for the robot(s).

When robotic coverage is considered in adversarial environments, it is assumed that threats exist in the target area, such that they might stop the covering robot with some probability. In this case, a new goal is added: optimizing robot survivability, or the probability of finishing the coverage route safely. There is a tradeoff between the two goals, as optimizing survivability may cause the robot to revisit safe areas (areas with no threats), causing coverage time to grow, whereas minimizing coverage time may require the robots to adopt a pattern that involves visiting dangerous areas more often. Thus, "choosing your battle" becomes necessary. It should be noted that threats cannot be treated as obstacles; while obstacles should be avoided, threats must be visited, as they are part of the free-space. The question is, in what order. Optimizing survivability can be translated into different optimization functions. The problem has been initially investigated in a single-robot setting [Yehoshua *et al.*, 2016], where the optimization criteria was defined as either minimizing the total number of visits to dangerous areas, or maximizing the expected covered area (here, the order of visits matters: we will want to visit more dangerous areas later in the coverage process).

The problem of robot coverage in adversarial environments (or adversarial coverage, in short) is NP-hard in simple environments (grid world, no obstacles) as well, making it even more difficult than the original coverage problem (with no threats). This necessitates the use of heuristic or approximation algorithms for generating near-optimal coverage paths. Fortunately, a simple greedy algorithm has been shown empirically to perform close to optimally in practice, on a given map (offline), with theoretical bounds. Further MDP-based modeling for the offline case, as well as an online version, has also been shown to perform efficiently [Yehoshua and Agmon, 2015b; Yehoshua *et al.*, 2015].

Further research has suggested using a team of multiple robots for covering the adversarial environment [Yehoshua and Agmon, 2016]. When multiple robots are involved in the covering mission, then not only the solution becomes more complex by an order of magnitude (similar to the original coverage task in neutral environments), but determining the problem itself, that is the optimization criteria, becomes more complex and is no longer straightforward. Meaning, Survivability of the team is composed of the survivability of its individuals, but one can consider optimizing the minimal survivability among the teammates (by that strengthening the weakest link), maximizing the expected survivability (thus strengthening the average case), and more. The impact of losing a teammate can or cannot be considered when initially allocating the coverage mission among the robots.

The work described above has not taken into account strategic adversarial behavior. The threats are given, and assumed to be spread randomly around the work area. In contrast, in [Yehoshua and Agmon, 2015a] we describe a more sophisticated adversarial model, which represents a first step toward modeling strategic behavior. The problem is described from the adversary's perspective: given a map of the environment and k guards (threats), where should those guards be placed in order to maximize the probability of stopping a robot covering the environment. The strategy of the adversary changes with respect to the knowledge it has regarding the covering robot's path. If it knows exactly the (deterministic) covering path of the robot, it can easily optimize the probability of stopping the robot by placing the guards at points that are more frequently revisited by the robot. If it has no knowledge whatsoever regarding the covering path, the problem of finding an optimal placement of its guards is equivalent to finding a coverage path that minimizes point-revisit. In this case, since the problem is hard, the heuristic suggested solution is to identify points that the robot must revisit multiple times (using graph representation of the world, and its correlated graph-theoretic characteristics), and place the guards at those points in descending order from more frequently visited points to less frequently visited points. This has been shown empirically to significantly decrease the survivability of the covering robot.

An observant reader may have noted that if the adversary has no knowledge regarding the covering robot's path, the problem of finding optimal placement of guards is equivalent to finding the shortest covering path, which is the original (adversary-free) coverage problem. This supports the aforementioned paranoid conjecture, and extends it to the coverage problem as well.

4 Multi-robot Formation

One of the earliest problems in robotics is the problem of multi-robot formation: the need for a team of robots to travel in a connected form through an environment [Balch and Arkin, 1995]. This problem is motivated by natural phenomena: from a school of fish, to a flock of birds, to a pride of lions. Multi-robot formation is usually divided into two problems: (1) gathering in a certain shape (formation achievement) and (2) travelling in that shape through an environment (formation maintenance). The goals of the robots are usually to minimize gathering time and to minimize deviation from the desired formation while avoiding obstacles (respectively).

Research on multi-robot formation has concentrated on the means of maintaining a connected form. This involves determining which robot should monitor which (usually referred to as local leader), and how. The common way to model formations is by using a monitoring graph, representing the ability of each robot to monitor its teammates [Kaminka and Glick, 2006]. Based on this graph, an optimal assignment of local leader for each robot, as well as a formation leader (global leader) if one exists, can be deduced. The identity of leaders may change as the formation avoids obstacles or handles faults internal to the robots and/or between them.

While teams in natural phenomena are commonly generated to handle threats or other (possibly destructive) external forces, in robotic research these have been neglected until recently. In [Shapira and Agmon, 2015], the problem of path planning for optimizing the survivability of multi-robot formations in adversarial environments (or adversarial formation, in short) was introduced. In this problem, threats existing in the environment may stop the robots with some probability. The goal of the robots is therefore twofold: (1) to travel in a connected formation through the environment, and (2) to optimize the survivability of the teammates, i.e., their chances of leaving the area safely. It is thus necessary to determine both the path that the robots should travel and the shape that they must maintain to achieve their goals. In order to solve the problem, it is necessary to define the characteristics of the threat. This is achieved using two factors: time and space. The time factor refers to time-variant versus time invariant threats, essentially asking whether threats change over time and, if so, in what way. The space factor encompasses various characteristics, including: range of influence (fixed or monotonically decreasing with distance from the origin of the threat), shape (circular, e.g., radiation cloud; line, e.g., tsunami), and concealment property (whether one robot can conceal its teammates from the threat, e.g., a sniper, or not, e.g., an earthquake), among others.

The problem of finding an optimal path for a given formation traveling through an obstacle-free environment with time-invariant threats, without concealment, can be transformed to a graph problem, and solved both efficiently and optimally. If the shape of the formation is not fixed, and can change to adapt to the given threats, the optimal formation has been proven to be a convoy. This result, however, does not hold if concealment property exists, or in other complex environments.

5 Robot Navigation

The problem of path planning is one of the fundamental problems in the field of agents and robotics (e.g., [Stentz, 1994]). The goal in this problem is finding a sequence of world locations which allows the robot (or agent, in general) to arrive at its destination while avoiding collisions, and optimizing some given criteria (usually minimizing travel cost). Previous work has examined the problem of navigation in areas containing possible *static* threats (e.g., [Tews *et al.*, 2004]). Research on the pursuit-evasion problem considers mobile threats, though the evading robot's mission is to indefinitely avoid being captured by the pursuer(s).

In our recent work [Keidar and Agmon, 2017] we have described a new version of the navigation problem, *adversarial navigation*, in which an agent should navigate to a destination point (safehouse) while avoiding being captured by a mobile adversary. Given a model of the adversary, the agent's goal is to find a path to a safehouse such that its probability of being captured is minimized. The problem is modeled as a simultaneous game, where each player does not have full information of the opponent (its location and/or strategy). We have created a framework that models the environment, and allows the agent to choose its path at random (while avoiding various risk attitudes), such that it maximizes its expected probability of safe arrival. Moreover, the path can be adjusted on-the-fly, according to new information revealed to the player as the game progresses (for example, if it has viewed the opponent along the way).

6 The Challenges

Research accounting for adversarial presence in robotic problems is in its early stages. As such, many problems are still open and challenges are yet to be met. As stated above, much of the work to this point considers a non-strategic adversary. However, also under the assumption of such relatively impotent adversary, much is left unknown. In general, further study in adversarial robotics is crucial to maximizing the potential activity and contribution of robots with a wide range of functions and aims.

One of the main considerations that has received little attention in current research is the notion of uncertainty. Uncertainty may exist in various levels: uncertainty about the robots' perception (am I sensing what is actually there?) and action (am I doing or will I do what I was expected to be doing?), and about its opponent (what is my opponent really going to do? what is its type?). Similar uncertainties exist also from the perspective of the adversary.

There are other canonical problems from robotics that can benefit from adding an adversary to the environment. Also in the problems that have already been touched in the context of an adversary, many open questions exist. In the coverage problem, how should a robot plan its coverage path if facing a known dynamic threat? or a strategic opponent that plans its path according to the robots strategy? how will the robots benefit from adopting a stochastic behavior?

To this point, we have assumed that the adversary is rational with respect to its knowledge, i.e., it will optimize its personal utility based on the knowledge it has on the robots performing their mission, and the environment. A question remained to be explored is the optimal behavior of robots facing an irrational adversary. How this adversarial model change the problem, and the behavior of the robots, is yet to be explored.

One of the interesting challenges concerns the paranoid conjecture. Whether it can be mathematically proven in problems other than patrolling and coverage that in certain adversarial models (specifically, ones in which it has no knowledge on the active robots and/or their strategy) the robots should act as if there is no adversary, remains an open question.

References

[Agmon *et al.*, 2008] Noa Agmon, Vladimir Sadov, Gal A Kaminka, and Sarit Kraus. The impact of adversarial knowledge on adversarial planning in perimeter patrol. In *Proceedings of AAMAS*, pages 55–62, 2008.

[Agmon *et al.*, 2011] Noa Agmon, Gal A Kaminka, and Sarit Kraus. Multi-robot adversarial patrolling: facing a full-knowledge opponent. *Journal of Artificial Intelligence Research (JAIR)*, pages 887–916, 2011.

[Agmon *et al.*, 2012] Noa Agmon, Chien-Liang Fok, Yehuda Emaliah, Peter Stone, Christine Julien, and Sriram Vishwanath. On coordination in practical multi-robot patrol. In *Proceedings of ICRA*, pages 650–656, 2012.

[Ahmadi and Stone, 2006] Mazda Ahmadi and Peter Stone. A multi-robot system for continuous area sweeping tasks. In *Proceedings ICRA*, pages 1724–1729, 2006.

[Al Marzouqi and Jarvis, 2011] Mohamed Al Marzouqi and Ray A Jarvis. Robotic covert path planning: A survey. In *Proceedings of IEEE Conference on Robotics, Automation and Mechatronics*, pages 77–82, 2011.

[Balch and Arkin, 1995] Tucker R Balch and Ronald C Arkin. Motor schema-based formation control for multiagent robot teams. In *Proceedings of the International Conference on Multi Agent Systems*, pages 10–16, 1995.

[Chevaleyre, 2004] Yann Chevaleyre. Theoretical analysis of the multi-agent patrolling problem. In *Proceedings of IAT*, pages 302–308, 2004.

[Elmaliach *et al.*, 2009] Yehuda Elmaliach, Noa Agmon, and Gal A Kaminka. Multi-robot area patrol under frequency constraints. *Annals of Mathematics and Artificial Intelligence*, 57(3-4):293–320, 2009.

[Gabriely and Rimon, 2001] Yoav Gabriely and Elon Rimon. Spanning-tree based coverage of continuous areas by a mobile robot. *Annals of Mathematics and Artificial Intelligence*, 31(1-4):77–98, 2001.

[Galceran and Carreras, 2013] Enric Galceran and Marc Carreras. A survey on coverage path planning for robotics. *Robotics and Autonomous Systems*, 2013.

[Hazon and Kaminka, 2008] Noam Hazon and Gal A. Kaminka. On redundancy, efficiency, and robustness in coverage for multiple robots. *Robotics and Autonomous Systems*, 56(12):1102 – 1114, 2008.

[Huang *et al.*, 2011] Haomiao Huang, Jerry Ding, Wei Zhang, and Claire J Tomlin. A differential game approach to planning in adversarial scenarios: A case study on capture-the-flag. In *Proceedings of ICRA*, pages 1451–1456, 2011.

[Kaminka and Glick, 2006] Gal A Kaminka and Ruti Glick. Towards robust multi-robot formations. In *Proceedings of ICRA*, pages 582–588, 2006.

[Keidar and Agmon, 2017] Ofri Keidar and Noa Agmon. Safety first: Strategic navigation in adversarial environments (extended abstract). In *Proceedings of AAMAS*, pages 1581–1583, 2017.

[Portugal and Rocha, 2011] David Portugal and Rui Rocha. A survey on multi-robot patrolling algorithms. In *Doctoral Conference on Computing, Electrical and Industrial Systems*, pages 139–146. Springer, 2011.

[Shapira and Agmon, 2015] Yaniv Shapira and Noa Agmon. Path planning for optimizing survivability of multi-robot formation in adversarial environments. In *Proceedings of IROS*, pages 4544–4549, 2015.

[Sless *et al.*, 2014] Efrat Sless, Noa Agmon, and Sarit Kraus. Multi-robot adversarial patrolling: Facing coordinated attacks. In *Proceedings of AAMAS*, pages 1093–1100, 2014.

[Stentz, 1994] Anthony Stentz. Optimal and efficient path planning for partially-known environments. In *Proceedings of ICRA*, pages 3310–3317, 1994.

[Stone *et al.*, 2000] Peter Stone, Patrick Riley, and Manuela Veloso. Defining and using ideal teammate and opponent models. In *Proceedings of the Twelfth Annual Conference on Innovative Applications of Artificial Intelligence*, 2000.

[Talmor and Agmon, 2017] Noga Talmor and Noa Agmon. On the power and limitation of deception in multi-robot adversarial patrolling. In *Proceedings of IJCAI*, 2017.

[Tews *et al.*, 2004] Ashley D Tews, Gaurav S Sukhatme, and Maja J Mataric. A multi-robot approach to stealthy navigation in the presence of an observer. In *Proceedings of ICRA*, volume 3, pages 2379–2385. IEEE, 2004.

[Yehoshua and Agmon, 2015a] Roi Yehoshua and Noa Agmon. Adversarial modeling in the robotic coverage problem. In *Proceedings of AAMAS*, pages 891–899, 2015.

[Yehoshua and Agmon, 2015b] Roi Yehoshua and Noa Agmon. Online robotic adversarial coverage. In *Proceedings of IROS*, pages 3830–3835, 2015.

[Yehoshua and Agmon, 2016] Roi Yehoshua and Noa Agmon. Multi-robot adversarial coverage. In *Proceedings of ECAI*, pages 1493–1501, 2016.

[Yehoshua *et al.*, 2015] Roi Yehoshua, Noa Agmon, and Gal A Kaminka. Frontier-based rtdp: A new approach to solving the robotic adversarial coverage problem. In *Proceedings of AAMAS*, pages 861–869, 2015.

[Yehoshua *et al.*, 2016] Roi Yehoshua, Noa Agmon, and Gal A Kaminka. Robotic adversarial coverage of known environments. *The International Journal of Robotics Research (IJRR)*, 35(12):1419–1444, 2016.