Resilience of Multi-Robot Systems to Physical Masquerade Attacks

Kacper Wardega, Roberto Tron, and Wenchao Li
Boston University
{ktw,tron,wenchao}@bu.edu

***Abstract*—The advent of autonomous mobile multi-robot systems has driven innovation in both the industrial and defense sectors. The integration of such systems in safety- and security-critical applications has raised concern over their resilience to attack. In this work, we investigate the security problem of a stealthy adversary masquerading as a properly functioning agent. We show that conventional multi-agent pathfinding solutions are vulnerable to these *physical masquerade attacks*. Furthermore, we provide a constraint-based formulation of multi-agent pathfinding that yields multi-agent plans that are provably resilient to physical masquerade attacks. This formalization leverages inter-agent observations to facilitate introspective monitoring to guarantee resilience.**

***Index Terms*—Multi-robot systems; Multi-agent pathfinding; Observation planning; Physical masquerade attacks**

I. Introduction

Recent trends in industrial production automation indicate an ever-increasing adoption of autonomous mobile robots. Systems from Fetch Robotics (FetchCore [1]) and Amazon Robotics (Kiva [2]) are prime examples. These robots, distributed across a factory floor, aid production efficiency and lower human effort, but the security research community has begun to raise alarm over the security of these systems [3]. As a result, the factory floor is at risk from malicious actors aiming towards production shutdown [4] or causing human injury [5] through manipulation of the robots in the environment. These threats also extend to multi-agent systems (MAS) in a less structured environment such as unmanned aerial vehicles (UAVs) [6]. It is therefore important to devise new strategies that can preemptively address these threats.

In this paper, we consider a novel class of attacks called *physical masquerade attacks* – a compromised insider (robot) masquerading as a properly functioning robot and attempting to gain access into unauthorized locations without being noticed. We use the term *physical* to distinguish this type of attack from masquerade attacks typically considered in the network security literature [7]. In the multi-agent path finding (MAPF) context, this manifests as one of the agents deviating from its pre-planned path and moving into an unauthorized zone. We show that solutions to the traditional MAPF problem are susceptible to this type of attack.

We propose a novel defense mechanism through path planning by leveraging the physical-sensing capabilities of robots (e.g., cameras) to detect and mitigate these attacks. The key idea is that, even if the compromised robot can forge false location information, other uncompromised robots can detect the physical anomaly (i.e. a robot veering off from its designated path) if they are close enough. By specially crafting the multi-agent plan, the induced inter-agent observations can provide introspective monitoring guarantees – any adversarial agent that attempts to break the system-wide security specification must necessarily violate the induced observation plan. We show that our method can find a multi-agent plan with the guaranteed resilience (if one exists) under a strong attacker model where an agent is completely compromised and has full knowledge of the plan. Our work is inspired by the recent efforts on defending against Sybil or spoofing attacks in multi-robot systems [8], [9].

The contributions of this paper are summarized below.

- We introduce a new class of attack in the multi-agent planning domain called physical masquerade attacks.
- We show that conventional solutions to MAPF are vulnerable to physical masquerade attacks.
- We propose a novel automated detection mechanism by simultaneously constructing an observation plan during path planning.
- We show that an attack-proof plan can by synthesized via an encoding to an Exists/Forall Satisfiability Modulo Theory (EFSMT) problem.

The rest of the paper is organized as follows. In Section II we discuss related work. Section III introduces the scenario of attack-proof multi-agent path finding. Section IV summarizes our results and Section V concludes.

II. Related Work

Autonomous agents are increasingly being used to manage various physical systems. This has introduced a number of vulnerabilities. Quarta et. al. explore the vulnerabilities in robotic arms of the type used in factory assembly lines and also give a review of some notable exploits such as in automated blast furnaces and nuclear plants [3].

The interconnection and interaction of industrial robots with the physical world can also open up new attack surfaces. Bijani and Robertson provide a taxonomical treatment of attacks on multi-agent systems [10]. The common theme of these studies is that interconnected autonomous agents suffer from lack of effective monitoring. Our work provides introspective monitoring guarantees by crafting a multi-agent plan in such a way that an agent is required to be seen by other agents at specific locations and at specific times.

There is a large body of work on multi-robot path finding [11]–[16]. However, relatively scarce literature has taken security into consideration. Among those that consider security, existing works are primarily limited to patrol strategies for intrusion detection [17]–[21], secure communication [22], [23] and attack-resilient network protocols [9], [24]. More recently, approaches that leverage the physics of the environment to counter cyberattacks began to emerge. In [8], the authors propose an algorithm that uses the physics of wireless signals to defend against Sybil attacks in multi-robot networks. In [25], the authors propose a Sybil attack-resilient traffic estimation and routing algorithm that uses information from sensing infrastructure and the dynamics and proximities of vehicles. Our paper is similar to these in spirit in the use of physical channels. In addition, we consider novel attacker models that not only involve insider attacks but also involve maneuvers in the physical space.

In terms of multi-agent systems that consider observations made by the agents, a recent work by Lee and Winfield introduce mathematical tools to scrutinize the observations and claims made by agents in a multi-robot setting by formalizing strength of opinion and evidence [26]. Our work can be differentiated from the patrolling problem by the distinction that in our work the would-be attacker is taken to be one of the defenders. Another line of work that considers adversarial agents is Adversarial Cooperative Path Finding (ACPF) [27]. In ACPF two teams of robots are pitted against each other in a race to reach their goal positions first. In contrast to ACPF, in our work we do not know which, if any, robots are adversarial and must assume that any of the robots may attempt to foil the security property.

On the computational side, many MAPF problems such as MAPF for optimal makespan [28] and optimal MAPF with deadlines (maximizing the number of agents that can reach their goal locations within the deadline) [29] are known to be computationally intractable to solve. In our work we use formulations that stipulate an optimal multi-agent solution, as such we expect high runtimes and scalability is not the focus of this work. As our reader will uncover, we are interested only in specific solutions to the MAPF problem that are attack-free. Approaches exist for coping with the high complexity of optimal MAPF but with the sacrifice of completeness. These methods essentially achieve scalability by decoupling agents from each other, planning a single agent at a time and resolving conflicts as they arise – subsequently planned agents treat previously planned agents as moving obstacles. As is explained by Gabrielle and Helmert, suboptimal MAPF has been completely solved by Kornhauser in his 1984 master's thesis [30]. Although Kornhauser's work is mostly forgotten, results building on his thesis have been rediscovered independently, for example by Wang and Botea in their work on scalable MAPF [13]. Recent work on MAPF has attempted to bridge the gap between decoupled planning and dynamic planning by using reinforcement learning [31]. It is however not clear whether these incomplete methods can be applied to finding attack-free multi-agent plans since we might need to enumerate all possible MAPF solutions in the worst case.

There is also rich literature on fault detection and fault-tolerance in multi-agent systems [32]–[34]. Our method of simultaneously synthesizing an observation plan can also be seen as a way to detect "faults" (malfunctioning robots). In a similar light, an attack that defeats an observation plan can be viewed as "unobservable faults."

III. Observation Planning

In this section, we describe in detail how we reformulate the multi-agent path finding problem to directly incorporate security requirements. The main idea is that by scheduling the robots' paths concurrently with an observation plan, the overall system is able to detect when specific robots are not at assigned locations at predetermined times. We call this sort of multi-agent path finding *multi-agent observation planning*. A multi-agent observation plan entails sequences of planned observations between robots. By carefully constructing this multi-agent observation plan, the system can detect attacks (and faults) by detecting any difference between the planned observations and the actual observations reported by the robots. In fact, we would like to construct the multi-agent observation plan in a way that *if a faulty or attacking agent breaks the security specification then that agent would necessarily violate the observation plan*.

We begin by providing a formal definition of the multi-agent path finding with deadlines (MAPF-DL) problem, hereafter referred to as simply MAPF.

Definition 1: MAPF

The MAPF problem for R homogeneous robot agents is defined over a 6-tuple $M = (W, U, \delta, \{S_i\}_{i=1}^R, \{G_i\}_{i=1}^R, \Omega)$. The workspace, or world, W is the set of locations and U is the set of control inputs. $\delta : 2^W \times U \to 2^W$ encodes the dynamics of the homogeneous robots. $S_i \subseteq W$ is the initial set of locations occupied by agent i and $G_i \in W$ is the goal location of agent i. Obstacles in the environment are given as the $\Omega \subseteq W$. A solution $\mathbf{x} = \{\mathbf{x}_i\}_{i=1}^R$ to problem M is a set of T-length paths $\mathbf{x}_i = \langle \mathbf{x}_i^1, \dots, \mathbf{x}_i^T \rangle$ for each agent i satisfying the properties:

$$(\forall\, i \in \mathbb{N}_R)\left(\mathbf{x}_i^1 = S_i\right), \quad \mathbb{N}_R = \{1, \dots, R\} \tag{1}$$

$$(\forall\, i \in \mathbb{N}_R, t \in \mathbb{N}_T)\left(\mathbf{x}_i^t \subseteq W\right), \quad \mathbb{N}_T = \{1, \dots, T\} \tag{2}$$

$$(\forall\, i \in \mathbb{N}_R, t \in \mathbb{N}_{T-1}\, \exists\, u \in U)\left(\delta(\mathbf{x}_i^t, u) = \mathbf{x}_i^{t+1}\right) \tag{3}$$

$$(\forall\, i \in \mathbb{N}_R\, \exists\, t \in \mathbb{N}_T)\left(G_i \in \mathbf{x}_i^t\right) \tag{4}$$

$$(\forall\, i, j \in \mathbb{N}_R, t \in \mathbb{N}_T)\left(\mathbf{x}_i^t \cap \mathbf{x}_j^t \neq \emptyset \implies i = j\right) \tag{5}$$

$$(\forall\, i \in \mathbb{N}_R, t \in \mathbb{N}_T)\left(\mathbf{x}_i^t \cap \Omega = \emptyset\right) \tag{6}$$

Definition 1 allows us to work with both discrete and continuous workspaces W. In the discrete case of a 2D gridworld, we can take $W = \mathbb{N}_k^2$, $U = \{\cdot, \uparrow, \downarrow, \rightarrow, \leftarrow\}$, etc. to obtain a synchronous discrete MAPF problem of the kind explored in [13]. Solutions to the MAPF problem on a gridworld are shown in Figure 1 with solid lines.

Next, we describe the attacker model. In our scenario, an attacker aims to compromise the safety/security of a factory floor (e.g. entering an unauthorized zone) by replanning one of the robots. Replanning several robots and allowing coordination between attacking robots is also possible, but is not considered in this paper. We assume that regardless of how much additional power the attacker can gain over the system, at the very least the attacker can fully control the processes being run on the compromised robot including motion plans and robot intercommunication software. We further assume that the attacker can only move the compromised robot in the same fashion as an uncompromised robot can move; the control input set U for the attacker is inherited from an existing MAPF problem. There are two levels of information that an attacker can have about the motions of other agents.

1) The *full-information attacker* has knowledge of the full observation plan
2) The *partial-information attacker* knows only the motion plan for the compromised robot

In our motivating example we experiment with an attacker that has knowledge of the full observation plan, i.e. the attacker knows the motion plans for each of the robots and therefore knows when an attacking robot should be observed and by which observing robot. The alternative to the full information case is one where the attacker has imperfect information and knows with certainty only the initial position of the compromised robot and knows with uncertainty the motion plans of other uncompromised robots. As mentioned in Section II, this reduces to the attack-side of the patrolling problem.

Given a MAPF problem M with solution $\mathbf{x}$ we now consider the Attack-MAPF problem where an adversarial agent aims to reach a secure location undetected. The attacker knows that all of the robots are equipped with sensors for inter-robot communication and monitoring such as radios and cameras. Uncompromised agents will be reporting observations to a central controller for verification against the observation plan. The sensor properties are known to the attacker, i.e. the attacker knows which positions relative to uncompromised agents will result in observations being reported to the central controller.

Definition 2: Attack-MAPF

The Attack-MAPF problem is defined over a 4-tuple $A = (M, \mathbf{x}, \phi, \Xi)$. M is a MAPF problem, $\mathbf{x}$ is a solution to M, $\phi : 2^W \times 2^W \rightarrow \{\bot, \top\}$ is the observation function, and $\Xi \subseteq \Omega$ is the set of secure locations (they would appear as obstacles to normal agents in MAPF). A solution $\mathbf{y}$ to problem A is a T-length trace satisfying the properties:

$$(\exists\, i^* \in \mathbb{N}_R)\left(\mathbf{y}^1 = \mathbf{x}^1_{i^*}\right), \quad \text{call } i^* \text{ the attacking agent.} \tag{7}$$

$$(\forall\, t \in \mathbb{N}_T)\left(\mathbf{y}^t \subseteq W\right) \tag{8}$$

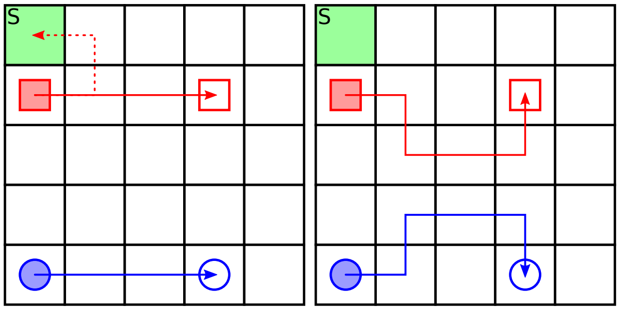
$$(\forall\, t \in \mathbb{N}_{T-1}\, \exists\, u \in U)\left(\delta(\mathbf{y}^t, u) = \mathbf{y}^{t+1}\right) \tag{9}$$

$$(\exists\, t \in \mathbb{N}_T)\left(\mathbf{y}^t \cap \Xi \neq \emptyset\right) \tag{10}$$

$$(\forall\, t \in \mathbb{N}_T)\left(\mathbf{y}^t \cap (\Omega \setminus \Xi) = \emptyset\right) \tag{11}$$

$$(\forall\, t \in \mathbb{N}_T, j \in \mathbb{N}_R \setminus i^*)\left(\mathbf{x}^t_i \cap \mathbf{y}^t = \emptyset\right) \tag{12}$$

$$(\forall\, t \in \mathbb{N}_T, j \in \mathbb{N}_R \setminus i^*)\left(\phi(\mathbf{x}^t_j, \mathbf{x}^t_{i^*}) \Leftrightarrow \phi(\mathbf{x}^t_j, \mathbf{y}^t)\right) \tag{13}$$



(a) Insecure (b) Secure

Fig. 1: (left) Solution to the MAPF problem (solid lines) for two agents in a 5×5 gridworld. This solution is not secure, since there is a solution to the corresponding Attack-MAPF problem for the square agent (dotted line). A compromised square agent can reach the secure location, marked S, undetected by the circle agent. (right) Solution to the APMAPF problem for two agents in a 5×5 gridworld. Neither agent can reach the secure location without breaking with the observations expected by the other agent.

Eqs. 8, 9, 11, 12 are related to the attacker motion and are analogous to the motion constraints in Definition 1. Eq. 7 ensures that the attacker is one of the robots present in the MAPF problem and inherits that robot's starting position. Eq. 10 stipulates that the successful attacker reaches one of the secure locations. Eq. 13 maintains that the attacker neither remove observations from nor introduce observations to any of the non-attacking robots. The observation function, $\phi(x^t_i, x^t_j)$, is read as "robot i observes robot j at time t"; ϕ need not be a symmetric function although in our experiments it always is symmetric. A solution to the Attack-MAPF problem on a gridworld is shown on the left of Figure 1 with dashed lines. ϕ for this example returns $\top$ only for adjacent robots. Now we can easily define the attack-proof MAPF (APMAPF) problem in terms of MAPF and Attack-MAPF.

Definition 3: APMAPF

The APMAPF problem is defined over 3-tuple $M_{ap} = (M, \phi, \Xi)$. M is a MAPF problem, ϕ an observation function, and Ξ a set of secure locations. A solution $\mathbf{x}_{ap}$ to M_{ap} is a solution to M such that the Attack-MAPF problem $(M, \mathbf{x}_{ap}, \phi, \Xi)$ has no solution.

Figure 1 shows a solution to the APMAPF problem alongside a corresponding MAPF problem. Definitions 1 to 3 can be easily modified to apply to non-homogeneous sets of agents by allowing agent-specific δ and ϕ. Non-homogeneous agents can be used to model a mixture of mobile robots with stationary security cameras.

We examine a toy discrete gridworld example and a more realistic continuous case with simple robot dynamics. For the discrete case we describe a procedure for the full APMAPF. For the continuous case we describe the primary obstacles to a full solution and just demonstrate Attack-MAPF, i.e. the single-agent planning problem to find a plan to enter a secure location undetected; we demonstrate possible attacks on MAPF solutions that result in weak observation plans.

The continuous case is significant especially when continuous dynamics are involved. Applications such as mapping or search-and-rescue using multiple UAVs fall into this category. We want to highlight that solving the APMAPF problem in the continuous case is much harder than in the discrete case.

A. $N \times N$ 4-connected grid

We encode the MAPF problem from Definition 1 as a Satisfiability Modulo Theories (SMT) proposition $\text{IsPlan}(x)$ following the 4-connected grid formulation of [13]–[16].

In the 4-connected grid environment, we decide that when robots are adjacent they can mutually observe one another. With this decision made, we encode the Attack-MAPF problem from Definition 2 as an SMT proposition $\text{Attacks}(y, x)$. We elect to use an SMT encoding for the Attack-MAPF problem because we would like to keep the approach general to other types of security specification or attack objective. Although in our Definition 2 version of Attack-MAPF the attacker is performing a reach-avoid task, we would like to have an approach that can also be used for attackers wishing to fulfill richer types of objectives such as those expressed in safe Linear Temporal Logic [35]. This means our approach can also be applied to settings where SAT/SMT-based motion planning is suitable, e.g. to satisfy additional dwell-time or sequence requirements [36].

Finally, We encode the APMAPF problem as an EFSMT problem (formulas of the form $\exists x.\forall y.\Phi(x, y)$ where $\Phi(x, y)$ is quantifier-free). The high-level EFSMT problem is shown in Eq. 14. In plain English Eq. 14 is saying that the satisfying solution x is a motion plan for all of the robots such that for all single-agent trajectories y, y is not a valid plan of attack for the full-information attacker.

$$(\exists x \, \forall y) \, (\text{IsPlan}(x) \wedge \neg \text{Attacks}(y, x)) \tag{14}$$

B. Continuous Case

In the continuous case the positions of the robots at each timestep become real-valued. Therefore we formulate the MAPF and Attack-MAPF problems as a Mixed Integer Quadratically Constrained Program (MIQCP) in the continuous case. Auxiliary integer-valued variables in the set $\{0, 1\}$ are used to handle disjunctions that are present in the EFSMT formulation of Section III-A using the standard mixed-integer programming tricks.

The workspace is now $W = [a, b]^2 \subset \mathbb{R}^2$. The secure location and obstacle sets $\Xi \subset \Omega \subset W$ consist of rectangles defined by extremal values, i.e. $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$.

Where in the 4-connected grid case the robots are constrained to movement between adjacent squares on the grid during one timestep, in the continuous case we adopt simple dynamics $x_i^t = x_i^{t-1} + u_i^{t-1}$ where u_i^t is the control input for robot i at timestep t with $\|u_i^t\|_2 \leq \bar{u}$. Convex optimization is used to obtain a solution $\mathbf{x}$ for this MAPF problem for the continuous case. The optimization objective we use is $\min \sum_{i \leq R, t \leq T} \|u_i^t\|_2$.

As opposed to the discrete case where adjacent robots in the grid are said to observe one another, in the continuous case we say that robots that are within a fixed radius r_{obs} are mutually observable. Given $\mathbf{x}$ we compute for each robot i a set of intervals $[t_{\text{start}}, t_{\text{end}}] \in [1, T]$ where i goes unobserved by all other robots. Since an unobserved attack can only take place during one of these intervals, we iterate over the intervals and secure locations and use convex optimization to solve for a feasible single-agent motion plan that reaches one of the secure locations.

The convex optimization objective in the Attack-MAPF step is $\min \sum_t \|y^t - \text{Center}(\xi)\|_2$ and is essentially a heuristic that encourages a single-agent plan that stays within the secure location ξ for as long as possible. The planning procedure returns a successful attack result if any of the y^t are within ξ.

IV. Results

The key metric for evaluating the danger posed by physical masquerading attacks is the *percentage of time that conventionally obtained MAPF solutions are vulnerable to the corresponding Attack-MAPF problem.*

First we evaluate the vulnerabilities in $N \times N$ 4-connected grids under varying grid size, number of agents, and number of obstacles. We use the `Z3` SMT solver [37] running on a Intel Core i7-7700 CPU @ 4.2GHz machine with 16GB RAM for solving the MAPF, Attack-MAPF and APMAPF problems. As in [38] we begin by generating 100 random 8×8 grids without obstacles. The first set of experiments allowed five minutes each for the MAPF step and Attack-MAPF step. The results, shown in the top half of Table I indicate that in over 90% of cases on average, conventional MAPF is vulnerable to the physical masquerade attack. The second set of experiments evaluate vulnerabilities in larger, more complicated grids with obstacles present in the environment. We generated 50 of these more complicated grids with five different settings and allow 20 minutes each for the MAPF step and Attack-MAPF step. The MAPF step is performed by an Enhanced Conflict-Based Search (ECBS) planner. ECBS is a conventional MAPF algorithm [39]. The results, shown in the bottom half of Table I, indicate that MAPF solutions returned by conventional planners are predominantly (over 95% on average) vulnerable to the physical masquerade attack.

The third set of experiments demonstrate the full APMAPF pipeline. We experiment with 50 instances for a variety of settings that we know from the prior two experiments that our SMT-based MAPF solver can handle in a reasonable amount of time. Now in addition to 20 minutes each for MAPF and Attack-MAPF we allow 2 hours for the APMAPF step. Because there is a tradeoff between total distance traveled and security with respect to Definition 2, we set the deadline for APMAPF to five timesteps more than the deadline from the MAPF step. The results of these trials in Table II demonstrate that for smaller grids we can obtain plans that are proof against physical masquerading attacks using the EFSMT-based approach described in Section III-A.

TABLE I: MAPF and Attack-MAPF results for the 4-connected grid case for varying grid size N, number of agents R, and number of obstacles O. The $N = 8$ trials correspond to the first set of experiments to set a baseline against [38]. The remaining trials have 20 minutes timeouts.

N	R	O	Attack UNSAT (%)	Vulnerable (%)
8	3	0	8	92
8	4	0	3	97
8	5	0	2	98
8	6	0	7	93
8	7	0	7	93
8	8	0	7	93
8	9	0	9	91
8	10	0	4	96
8	11	0	11	89
10	4	10	4	96
20	5	20	6	94
40	3	50	6	94
40	6	50	0	100
80	7	100	0	100

TABLE II: APMAPF results for the 4-connected grid case for varying grid size N, number of agents R, and number of obstacles O. 20 minutes each are allowed for MAPF and Attack-MAPF and two hours are allowed for APMAPF.

N	R	O	Attack UNSAT	Timeout	Secured
5	2	4	12	2	36
10	2	4	6	24	20
10	3	15	3	31	16
10	5	15	5	34	11
15	5	15	0	43	7

The fourth and final set of experiments demonstrate the physical masquerading attack on the continuous case described in Section III-B. We use GUROBI as our MIQCP solver [40] running on a Intel Core i7-6850K @ 12x 4GHz machine with 126GB RAM. We generate 20 random instances each for a varying number of agents and obstacles. The agents and obstacles are rectangular and the problem is scaled to be solvable in approximately 100 time steps, i.e. the workspace is a box of side length ten centered at $(0, 0)$ and the control input bound is $\bar{u} = 0.2$. The observation radius is $r_{\text{obs}} = 1.5$. The results of the continuous case experiments are shown in Table III and reinforce the findings from the 4-connected grid case; conventionally planned agents are vulnerable to attack by masquerading agents. An illustration of one of our continuous case experiments detailing the MAPF solution and corresponding Attack-MAPF solution is shown in Figure 2.

TABLE III: Attack-MAPF results for the continuous case with varying number of agents R and number of obstacles O.

R	O	Attack-MAPF UNSAT	Vulnerable
2	2	1	19
3	2	1	19
3	4	1	19
4	4	0	20
4	8	0	20
6	8	0	20

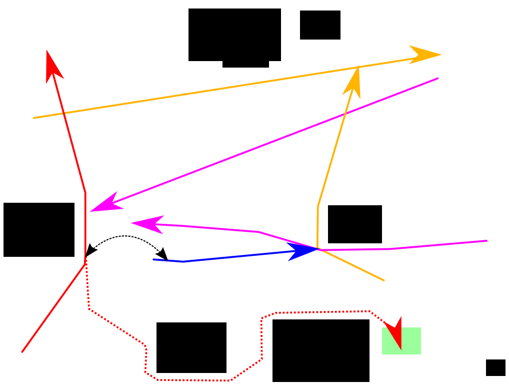


Fig. 2: Solution to the MAPF problem (solid lines) for six agents in a continuous workspace. This solution is not secure, since there is a solution to the corresponding Attack-MAPF problem for the red agent (dotted line). A compromised red agent can reach the secure location, shown in green, after being appropriately observed by the blue agent (double-headed black line) without any unplanned detections.

We observe that starting scenarios that result in congested plans with almost-collisions are more secure since congestion creates more observation points and leaves less opportunity for the compromised robot to deviate from its replanned path without violating the observation plan. We also experimented with special scenarios such as robots with goal positions in different rooms; for valid APMAPF solutions this necessitates that robots travel together for significant distances. We leave a systematic study of the relationship between attack-proof constraints and planning performances such as makespan and total distance traveled for future work.

V. Conclusion

This paper introduces a new class of attacks for multi-robot systems where a compromised robot can masquerade as a properly functioning agent and conduct clandestine maneuvers without being detected by other agents. We show that solutions to purely MAPF problems are susceptible to this type of attacks. Further, we propose a novel mechanism for detecting these physical masquerade attacks by simultaneous synthesizing observation constraints during path planning. In the future, we plan to study weaker attacker models such as attacker knowing only part of the plan and the security implication of these models. In the case where more than one agent are compromised, collusion between these agents are possible and new strategies will need to developed to detect and defend against masquerade attacks. Computationally, MAPF problems are in general NP-hard and APMAPF additionally requires the absence of potential attack paths in the solutions to the MAPF problems. A subject of current investigation is the exact complexity characterization of APMAPF. In addition, our EFSMT-based approach can be viewed as a centralized planning approach and this type of approaches often face scalability issues. We plan to investigate decoupled approaches to the APMAPF problem motivated by the high algorithmic complexity of the current approach.

References

[1] Fetch, "fetchcore: Cloud robotics platform," https://fetchrobotics.com/products-technology/fetchcore/, accessed: 2018-05-09.

[2] IEEE Spectrum, "Three engineers, hundreds of robots, one warehouse," https://www.spectrum.ieee.org/robotics/robotics-software/three-engineers-hundreds-of-robots-one-warehouse, accessed: 2018-04-02.

[3] D. Quarta, M. Pogliani, M. Polino, F. Maggi, A. M. Zanchettin, and S. Zanero, "An Experimental Security Analysis of an Industrial Robot Controller," *2017 IEEE Symposium on Security and Privacy (SP)*, pp. 268–286, 2017.

[4] M. Brunner, H. Hofinger, C. Krauss, C. Roblee, P. Schoo, and S. Todt, "Infiltrating critical infrastructures with next-generation attacks," http://publica.fraunhofer.de/documents/N-151330.html, 2010.

[5] C. Forrest, "Robot kills worker on assembly line, raising concerns about human-robot collaboration," https://www.techrepublic.com/article/robot-kills-worker-on-assembly-line-raising-concerns-about-human-robot-collaboration/, March 2017.

[6] A. Y. Javaid, W. Sun, V. K. Devabhaktuni, and M. Alam, "Cyber security threat analysis and modeling of an unmanned aerial vehicle system," in *2012 IEEE Conference on Technologies for Homeland Security (HST)*, Nov 2012, pp. 585–590.

[7] M. B. Salem, S. Hershkop, and S. J. Stolfo, *A Survey of Insider Attack Detection Research*. Boston, MA: Springer US, 2008, pp. 69–90.

[8] S. Gil, S. Kumar, M. Mazumder, D. Katabi, and D. Rus, "Guaranteeing spoof-resilient multi-robot networks," *Autonomous Robots*, vol. 41, no. 6, pp. 1383–1400, Aug 2017.

[9] V. Renganathan and T. Summers, "Spoof resilient coordination for distributed multi-robot systems," in *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, Dec 2017, pp. 135–141.

[10] S. Bijani and D. Robertson, "A review of attacks and security approaches in open multi-agent systems," *Artificial Intelligence Review*, vol. 42, no. 4, pp. 607–636, 2014.

[11] S. M. LaValle, *Planning Algorithms*. New York, NY, USA: Cambridge University Press, 2006.

[12] H. Choset, K. M. Lynch, S. Hutchinson, G. A. Kantor, W. Burgard, L. E. Kavraki, and S. Thrun, *Principles of Robot Motion: Theory, Algorithms, and Implementations*. MIT Press, 2005.

[13] K.-H. C. Wang and A. Botea, "A Scalable Multi-Agent Path Planning Algorithm with Tractability and Completenss Guarantees," *JAIR - Journal of Artificial Intelligence Research*, vol. 42, pp. 55–90, 2011.

[14] ——, "Tractable multi-agent path planning on grid maps," in *Proceedings of the 21st International Jont Conference on Artifical Intelligence*, ser. IJCAI'09. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2009, pp. 1870–1875.

[15] A. Ulusoy, S. L. Smith, X. C. Ding, C. Belta, and D. Rus, "Optimal multi-robot path planning with temporal logic constraints," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2011, pp. 3087–3092.

[16] A. Ulusoy, S. L. Smith, X. C. Ding, and C. Belta, "Robust multi-robot optimal path planning with temporal logic constraints," in *2012 IEEE International Conference on Robotics and Automation*, May 2012, pp. 4693–4698.

[17] J. M. Kim, J. S. Choi, and B. H. Lee, "Multi-agent coordinated motion planning for monitoring and controlling the observed space in a security zone," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 1679–1684, 2008, 17th IFAC World Congress.

[18] A. Fagiolini, G. Valenti, L. Pallottino, G. Dini, and A. Bicchi, "Decentralized intrusion detection for secure cooperative multi-agent systems," in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 1553–1558.

[19] A. Fagiolini, M. Pellinacci, G. Valenti, G. Dini, and A. Bicchi, "Consensus-based distributed intrusion detection for multi-robot systems," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 120–127.

[20] N. Agmon, S. Kraus, and G. A. Kaminka, "Multi-robot perimeter patrol in adversarial settings," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 2339–2345.

[21] N. Agmon, G. A. Kaminka, and S. Kraus, "Multi-robot adversarial patrolling: Facing a full-knowledge opponent," *J. Artif. Int. Res.*, vol. 42, no. 1, pp. 887–916, Sep. 2011.

[22] A. Bicchi, A. Danesi, G. Dini, S. L. Porta, L. Pallottino, I. M. Savino, and R. Schiavi, "Heterogeneous wireless multirobot system," *IEEE Robotics Automation Magazine*, vol. 15, no. 1, pp. 62–70, March 2008.

[23] S. Morante, J. G. Victores, and C. Balaguer, "Cryptobotics: why robots need cyber safety," https://www.frontiersin.org/articles/10.3389/frobt.2015.00023/full, September 2015.

[24] D. Gupta, J. Saia, and M. Young, "Proof of work without all the work," in *Proceedings of the 19th International Conference on Distributed Computing and Networking*, ser. ICDCN '18. New York, NY, USA: ACM, 2018, pp. 6:1–6:10.

[25] Y. Shoukry, S. Mishra, Z. Luo, and S. Diggavi, "Sybil attack resilient traffic networks: A physics-based trust propagation approach," in *Proceedings of the 9th ACM/IEEE International Conference on Cyber-Physical Systems*, ser. ICCPS '18. Piscataway, NJ, USA: IEEE Press, 2018, pp. 43–54.

[26] C. Lee, J. Lawry, and A. Winfield, "Combining Opinion Pooling and Evidential Updating for Multi-Agent Consensus," Tech. Rep., 2017.

[27] M. Ivanova and P. Surynek, "Adversarial Cooperative Path-Finding: Complexity and Algorithms," *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, vol. 2014-Decem, pp. 75–82, 2014.

[28] J. Yu, "Intractability of optimal multirobot path planning on planar graphs," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 33–40, Jan 2016.

[29] H. Ma, G. Wagner, A. Felner, J. Li, T. K. S. Kumar, and S. Koenig, "Multi-Agent Path Finding with Deadlines," no. July, 2018. [Online]. Available: http://arxiv.org/abs/1806.04216

[30] R. Gabriele and M. Helmert, "Non-Optimal Multi-Agent Pathfinding Is Solved (Since 1984)," *Symposium on Combinatorial Search*, no. Since 1984, pp. 173–174, 2012.

[31] G. Sartoretti, J. Kerr, Y. Shi, G. Wagner, T. K. S. Kumar, S. Koenig, and H. Choset, "PRIMAL: Pathfinding via Reinforcement and Imitation Multi-Agent Learning." [Online]. Available: https://arxiv.org/pdf/1809.03531.pdf

[32] H. Yang, M. Staroswiecki, B. Jiang, and J. Liu, "Fault tolerant cooperative control for a class of nonlinear multi-agent systems," *Systems & Control Letters*, vol. 60, no. 4, pp. 271 – 277, 2011.

[33] F. Arrichiello, A. Marino, and F. Pierri, "Observer-based decentralized fault detection and isolation strategy for networked multirobot systems," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 4, pp. 1465–1476, July 2015.

[34] P. Kouvaros, A. Lomuscio, and E. Pirovano, "Symbolic synthesis of fault-tolerance ratios in parameterised multi-agent systems," in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*. International Joint Conferences on Artificial Intelligence Organization, 7 2018, pp. 324–330.

[35] O. Kupferman and M. Y. Vardi, "Model checking of safety properties," *Formal Methods in System Design*, vol. 19, no. 3, pp. 291–314, October 2001.

[36] I. Saha, R. Ramaithitima, V. Kumar, G. J. Pappas, and S. A. Seshia, "Automated composition of motion primitives for multi-robot systems from safe ltl specifications," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2014, pp. 1525–1532.

[37] L. De Moura and N. Bjørner, "Z3: An efficient SMT Solver," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4963 LNCS, pp. 337–340, 2008.

[38] G. Sharon, R. Stern, A. Felner, and N. Sturtevant, "Conflict-Based Search For Optimal Multi-Agent Path Finding," pp. 563–569. [Online]. Available: www.aaai.org

[39] M. Barer, G. Sharon, R. Stern, and A. Felner, "Suboptimal variants of the conflict-based search algorithm for the multi-agent pathfinding problem," *Frontiers in Artificial Intelligence and Applications*, vol. 263, no. SoCS, pp. 961–962, 2014.

[40] Gurobi Optimization, LLC, "Gurobi optimizer reference manual," 2018. [Online]. Available: http://www.gurobi.com