Poster: 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**

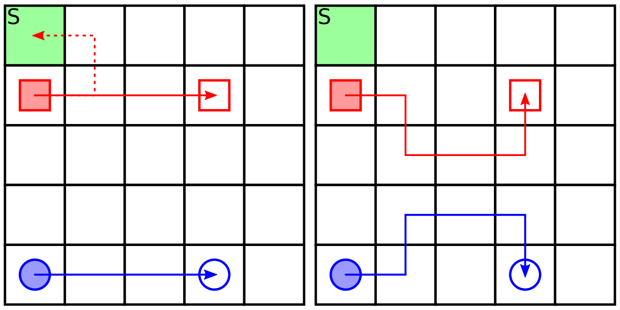
I. Introduction

Mobile robots, moving in synchrony across a factory floor, obviate the need for humans to muddy their hands with industrial work and multiply production efficiency. This is the *Industry 4.0* vision, where CPS-instrumented factories perform automated task scheduling based on sensory input within the factory and delegate teams of robots to carry out sub-objectives. Much work remains to be done to realize this vision, however. The implementation of an Industry 4.0 factory notwithstanding, there are a range of security concerns stemming from the interaction between a large-scale computing system and the physical world [1].

For example, what if a *physical* component behaves normally from the *virtual* perspective while simultaneously carrying out unplanned and potentially malicious actions in the physical space? This is exactly the concern we wish to address when we consider *physical masquerade attacks* – a compromised robot that masquerades as a properly functioning robot and attempts to enter unauthorized locations on the factory floor. Masquerade attacks will be familiar to our reader from the context of network security [2]. We modify with the term *physical* since we are specifically in the multi-agent path finding (MAPF) context, where masquerade attacks manifest as a robot changing its path from the preplanned course in order to reach an unauthorized zone without being detected.

II. Observation Planning

MAPF is the problem of collision-free routing for a set of agents through a transition system from their respective start locations to their respective goal locations. We extend this formulation to directly incorporate security requirements by bundling planned multi-agent routes with corresponding inter-agent observations, a so-called observation plan. These observations are made by way of cameras or other sensors fitted to the robots, and are recorded for runtime reporting. Our ultimate objective in this work is to carefully obtain a MAPF solution where the corresponding observation plan enables attack (and fault) detection via reasoning on deviations from the observation plan. The observation plan is an abstraction of the physical movement of the robots – we are not guaranteed that an appropriate observation plan exists, but if we can find one, we can guarantee attack detection in the physical space.



(a) Vulnerable (b) Attack-proof

Fig. 1: (left) Solution to the MAPF problem (solid lines) for two agents in a 5×5 gridworld. This solution is vulnerable, 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) An attack-proof solution to the MAPF 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.

To explain this more precisely, we begin with a MAPF problem instance and a candidate solution and then pose the *Attack-MAPF* problem. The Attack-MAPF problem asks if any of the agents can be re-routed to reach a secure location undetected. To reach the secure location undetected means that no uncompromised agent can come to know that the compromised agent has changed its path from the path detailed by the candidate solution. As a result, no external higher-level controller could come to know of any physical path deviation by studying the runtime observation reports. The existence of a solution to the Attack-MAPF problem indicates that the corresponding candidate MAPF solution is vulnerable to physical masquerade attack, and attack-proof otherwise. We illustrate these formulations in Fig. 1 in a 4-connected grid environment where each robot can observe adjacent squares.

III. Threat Model

In this work we consider an attacker that has full control over (any) one of the robots. We assume that the attacker inherits the dynamics of the non-compromised robot, i.e. that the attacker and system designer both have the same, complete model of the robot's physical capabilities. Furthermore, we allow the planned routes of all of the robots, the expected observations between robots, and the sensor properties to be common knowledge and therefore available to the attacker. The practical reason for common knowledge of the planned routes and observations is that this information aids in runtime collision avoidance, although this results in a relatively strong attacker. As a consequence, any technique that can defend against the strong attacker will also defend against a weaker attacker that perhaps knows only the planned route and observations for the compromised robot.

IV. Results

We evaluate the danger of physical masquerade attacks by measuring how often conventionally-obtained MAPF solutions turn out to be vulnerable to attack. In the 4-connected grid case, we obtain MAPF solutions using the Enhanced Conflict-Based Search algorithm (ECBS) [3] and solve the corresponding Attack-MAPF problem through a Satisfiability modulo theories (SMT) encoding. As is evident from Table I, in excess of 90% of all conventionally-obtained plans are vulnerable to physical masquerade attack. Additionally, we show that a complete algorithm for obtaining attack-proof MAPF solutions can be achieved with an Exists-Forall SMT encoding. In the continuous case with bounded-displacement dynamics, we obtain MAPF solutions through an encoding to a Mixed-Integer Quadratically-Constrained Program (MIQCP). The Attack-MAPF problem is similarly solved by MIQCP. A sample MAPF/Attack-MAPF solution pair is shown in Fig 2. Just as in the 4-connected grid case, we find that a vast majority of all MAPF solutions are vulnerable.

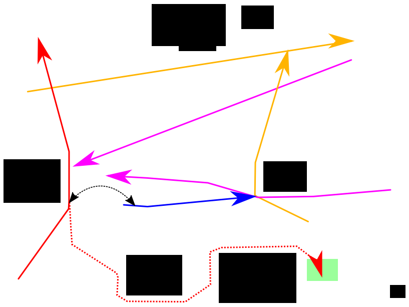


Fig. 2: Solution to the MAPF problem (solid lines) for six agents in a continuous workspace. This solution is not attack-proof, since there is a solution to the corresponding Attack-MAPF problem for the red agent (dotted line). The compromised red agent can reach the secure location, shown in green,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

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 [4]. 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

V. Conclusion

We introduce a new class of attacks on multi-robot systems whereby an attacker exploits discrepancies between the physical space and the virtual representation used to monitor the system. A compromised agent masquerades as a normal agent while conducting secret maneuvers. The primary conclusion of our work is that MAPF solvers are susceptible to this sort of attack. We propose to prevent physical masquerade attacks by simultaneously performing observation planning as part of the path planning procedure. A more rigorous treatment of our formulations is found in the full version of this paper (to appear in SafeThings 2019).

References

[1] 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. [Online]. Available: http://ieeexplore.ieee.org/document/7958582/

[2] M. B. Salem, S. Hershkop, and S. J. Stolfo, *A Survey of Insider Attack Detection Research.* Boston, MA: Springer US, 2008, pp. 69–90. [Online]. Available: https://doi.org/10.1007/978-0-387-77322-3_5

[3] 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.

[4] 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