Poster: A moving target defense scheme with overhead control

Ashley McAbee, Murali Tummla, and John McEachen
Electrical and Computer Engineering Department
Naval Postgraduate School
Monterey, CA, USA
{asmcabee1, mtummala, mceachen}@nps.edu

***Abstract*—We present a scheme for harnessing advantages of moving target defense while controlling overhead. We leverage a partially observable Markov decision process with an absorbing state to incorporate attacker dynamics, overhead and effectiveness of defenses, and limitations of intrusion detection in a single model that can be solved to produce a policy graph. We demonstrate the utility of the policy graph as the backbone of a cyber defense system that thwarts the threat just enough, just in time, defending against attacks while also preserving the defender's own system availability.**

***Index Terms*—Moving target defense, optimality, POMDP**

I. Problem Summary

As it gained attention in 2010, moving target defense (MTD) was touted as a game changer capable of seizing the advantage from attackers [1]. Today, MTD encompasses more than 90 individual techniques that defend systems by creating uncertainty for attackers, yet these options can carry considerable negative side effects [2]. Techniques are needed to harness the advantages of MTD while controlling overhead, especially when costs accumulate per-event such as in availability losses when reconfigurations are deployed. This poster describes an in-progress effort to develop an optimization scheme for MTD leveraging partially observable Markov decision processes (POMDP) with an absorbing state. Because the model encompasses attacker dynamics, defensive measures, the per-event overhead of each, and the uncertainty inherent in detecting attacker progress, it can be used to develop defensive policies that trigger defenses in balance with competing requirements.

II. Proposed Solution

The MTD system we propose, illustrated in Figure 1, monitors indications of adversary activity to trigger optimal defensive actions. Within this system, the overhead of MTD is controlled by leveraging POMDP to develop a complete model of attack-defense interactions under realistic conditions that can be used to find an optimal defense implementation policy. POMDP underpin the dominant techniques for developing optimal decisions in sequential decision processes with state uncertainty, which in our case stems from missed and false attack detection by the upstream intrusion detection sysetm (IDS). POMDP are described via (S, A, T, R, Ω, O) [3]. The formulation for an n-stage cyber attack process is outlined in Table I.

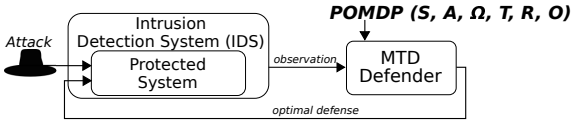

Fig. 1. MTD scheme with overhead control via POMDP

TABLE I
POMDP formulation to facilitate MTD optimization

Component	Description	Size
States (S)	attack phases (ex: reconnaissance)	$1 \times n$
Actions (A)	MTD options (ex: IP address hopping)	$1 \times m$
Observations (Ω)	IDS indication of attack (also by phase, i.e., $S = \Omega$)	$1 \times n$
Transitions (T)	likelihood attacker moves between any two phases under given MTD	$n \times n \times m$
Costs/Rewards (R)	overhead incurred by phase and MTD ($-$ to denote cost)	$n \times m$
Observation Probabilities (O)	likelihood IDS indication aligns with true attack phase (p_D), earlier phase (p_M), or later phase (p_{FA})	$n \times n$

Within this system, the defender selects between m actions based on an observation-based belief regarding the current state with the goal of maximizing rewards. We formulate the reward matrix R with negative rewards for both attack state and defensive overhead to ensure solutions offer a balanced approach to accomplishing attack suppression. POMDP has previously been employed for optimal dynamic cyber defense selection [4]. Additionally, frameworks have been developed to manually optimize MTD implementation schemes based on cost factors of each defense [5]. To our knowledge, we are the first to use POMDP with an absorbing state to autonomously optimize MTD with respect to accumulated per-event cost.

The state space S includes all attack stages. State s_0 aligns with perfect defense, and s_n is an absorbing state representing the attacker's ultimate goal. An absorbing final state was adopted because each attack recovery process is unique and thus not well represented stochastically. Additionally, we found the absorbing state created desirable defensive performance in that defensive aggression increases in the late stages of the attack. Transition probabilities $T(s, s', a_0)$ represent the likelihood an attacker moves between any two states of their own volition as sourced from study of past attacks. MTD

techniques are incorporated as additional actions, a_1 through a_m, defined in terms of the state transition probabilities and cost incurred if enacted. The model captures likelihood of attack detection as a state-aligned ($S = \Omega$) observation function describing the performance of an upstream intrusion detection system. Probability of detection, p_D, is the likelihood observation and state match, with missed detection, p_M and false alarm p_{FA} representing detection of the left and right neighboring states, respectively.

In practice, POMDP formulation requires analysis of attack patterns, MTD effectivness and cost, and the defender's overhead and attack tolerances. These requirements are a drawback of implementing the proposed model-based decision system, but we believe the potentially significant benefit of optimization justifies steep initial investment.

For sufficiently small state and action spaces, POMDPs can be solved via dynamic programming techniques for a policy graph that maximizes the expected discounted value and thus optimizes the overhead over the infinite horizon. Incoming observations are coupled with the current node in the polciy graph to look up the next node which prescribes the next action. Operating under the policy accrues rewards as close as possible to those expected under ideal condition $p_D = 1.0$, which simplifies to a Markov decision process with optimal policy Π prescribing actions by state [6].

Validation of our system requires metrics of attack suppression, Φ, and availability. The first is quantified as $\Phi = 1 - \frac{\tau_0}{\tau}$, which compares τ, the expected number of state transitions before s_n is reached, to τ_0, the same quantity without MTD. By extending absorbing Markov chain theory to accommodate the formulated POMDP, the metrics can be predicted to within 3% of the values measured in simulation, which is useful in exploring the impact of policy changes without the need for additional policy graph computation or simulation.

III. Validation

To validate our proposed system, we measured Φ and availability during simulated defense against a five-stage attack process (*Start, Target Scan, Vulnerability Scan, Exploit Launch, Attacked*) exhibiting stochastic behavior $T(s, s', a_0)$ based on published honeypot data [7]. The defender chooses between three MTD options with specifications listed in Table II in accordance with policy graphs developed via incremental pruning within *pomdp-solve*, a publicly available software package [8]. For comparison, availability was also measured in a system where reconfigurations deploy randomly with exponential interarrival times calibrated for equivalent Φ. State transitions are assumed to occur every 10 seconds in accordance with data from [7], which facilitates extension of availability loss to an overall availability percentage.

The POMDP-based scheme maintains attack suppression and availability at greater than 96% for $p_D \geq 0.8$, as shown in Figure 2(a)-(b). Compared to randomly deployed reconfigurations, the POMDP-based scheme offers significant overhead control. As indicated in Figure 2(c), the case for the proposed system becomes increasingly compelling as p_D improves. A random system calibrated to match Φ suffers availability values up to 73 percentage points lower.

TABLE II
Available defenses, specifications developed from data in [5].

Reconfig. Basis	Impact (P[$s' = s_0$])	Impact (%)	Avail. Loss (sec)	Avail (%)
No Action	n/a	n/a	0.000	100%
IP address	$\frac{range-1}{range} = \frac{255}{256}$	99.6%	9.590	4.1%
Service	$\frac{ser-ser_{vuln.}}{ser} = \frac{2}{3}$	66.7%	0.635	93.7%

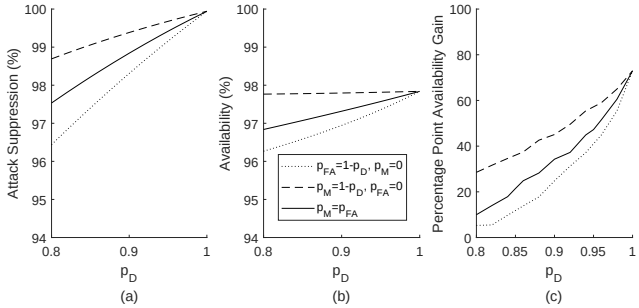

Fig. 2. Attack suppression (a) and availability (b) under simulated operation of POMDP-based MTD system over approximately three million action selections. Availability improvement (c) is measured against an MTD system where reconfigurations occur with exponential inter-arrival times calibrated to achieve equivalent attack suppression to (a).

Initial results strongly support the ability of the proposed system to harnesses MTD advantages *and* control overhead. Future work will incorporate on-line POMDP solution techniques to expand the state and action space capacity before intractability becomes a concern. We also recognize that defense hinged on attack detection can introduce risk that requires assessment before this system could be practically adopted, as well as an assessment of performance if there are inaccuracies in the POMDP formulation.

References

[1] U.S. Cyber Security and Information Assurance Interagency Working Group Networking and Information Technology Research and Development Subcommittee, "Cybersecurity game-change research and development recommendations," 2010. [Online]. Available: https://www.nitrd.gov/Publications/PublicationDetail.aspx?pubid=24

[2] B. Ward, S. Gomez, R. W. Skowyra, D. Bigelow, J. Martin, J. Landry, and H. Okhravi, "Survey of cyber moving targets," MIT Lincoln Laboratory, Lexington, Massachusetts, Tech. Rep. 1228, Januaray 2018.

[3] M. J. Kochenderfer, *Decision making under uncertainty: theory and application*. Cambridge, Massachusetts: MIT Press, 2015.

[4] E. Miehling, M. Rasouli, and D. Teneketzis, "A POMDP approach to the dynamic defense of large-scale cyber networks," *IEEE Trans. on Inform. Forensics and Security*, vol. 13, no. 10, pp. 2490–2505, Oct 2018.

[5] W. Connell, L. H. Pham, and S. Philip, "Analysis of concurrent moving target defenses," in *Proc. of the 5th ACM Workshop on Moving Target Defense*. New York, NY, USA: ACM, 2018, pp. 21–30.

[6] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[7] S. Panjwani, S. Tan, K. M. Jarrin, and M. Cukier, "An experimental evaluation to determine if port scans are precursors to an attack," in *Proc. 2005 Int. Conf. on Dependable Systems and Networks (DSN '05)*, Jun. 2005, pp. 602–611.

[8] A. R. Cassandra, "The POMDP page." [Online]. Available: http://www.pomdp.org/