Tambe, Developing the Science and Applications of Security Games: Machine Learning, Uncertainty and Preference Elicitation in Game Theory for Security

Having successfully founded the research area of security games, which has led to real-world applications in scheduling the deployment of limited resources (patrols, checkpoints, inspections, etc.), we now provide fundamental advances by incorporating machine learning to enhance real-world security applications, new models of opportunistic security games, robust methods for handling uncertainty, and novel techniques for preference elicitation techniques.

Project Technical Description

1. Theme Area: Risk Management/Operations Research
2. Principal Investigator: Milind Tambe
3. Institution: USC
4. Co-investigator: None
5. Lead POC: Milind Tambe
6.Keywords: Game theory, security games, machine learning, mathematical programming

An illegal fishing vessel (lancha) captured by the US Coast Guard: US Coast Guard flights that have flown missions based on our software for lancha interdiction; LA Metro trains where data will help us build models of opportunistic crime.

7.Brief Description:

In the past, we successfully founded the research area of *security games*. This research focuses on using game theory – often defender-attacker Stackelberg games -- to optimize the use of limited security resources, accounting for adversaries who conduct surveillance and then react; it also led to one of the most successful set of applications of computational game theory. Among our past application successes include ARMOR, to randomize vehicle checkpoints and canine unit patrols at the Los Angeles International Airport (LAX); IRIS, deployed by the Federal Air Marshals Service (FAMS); PROTECT, deployed by the US Coast Guard in Boston, Houston, Los Angeles, and New York for port protection; and several others. New applications have recently emerged however, including US Coast Guard's protection of fisheries, protection against crimes in the LA Metro System, US Coast Guard's interdiction of drugs and protection of wildlife against poaching (which isn't strictly a DHS mission but is in the same family of security games, and note that illegal poaching funds terrorist groups like Al-Shabab). These new set of applications bring to fore some fundamental new research challenges arising from key novel characteristics such as: (i) these are games involving repeated interaction with adversaries; (ii) there is not a single but an entire heterogeneous population of adversaries; (iii) the adversaries may act more opportunistically rather than fully strategically; (iv) there is significant adversary data available. The original security game model doesn't apply as a result, since it assumes a single shot game with a highly strategic adversary, with limited or no data about adversaries.

We propose several directions to address these new challenges. The repeated nature of the interactions and the presence of significant data, allows for machine learning techniques; the opportunistic nature of crime requires formulating opportunistic security games; the presence of data (if not in significant amounts) about adversaries leads to solving Robust Bayesian games; finally we address preference elicitation techniques, a universal challenge applicable in many domains.

8. Research Objectives: (Clear expression of purpose and goals.)

These are fundamental research objectives in addressing challenges at the intersection of game theory and machine learning:

a) *Machine learning in repeated games*: A key assumption of previous applications of Stackelberg games is that the adversary is perfectly rational. While there has been research on bounded rationality models, model parameters were just hand crafted; given the presence of data in domains such as

fishery protection and drug interdiction, we propose to derive machine learning approaches to solve repeated Stackelberg games where attackers follow a boundedly rational model, e.g., QR/SUQR (Quantal Response/Subjective Utility Quantal Response).

b) *Bayesian-robust approach for addressing boundedly rational attackers in security games:* In domains such as fishery protection, drug interdiction, crime prevention or wildlife protection, there is not just a single adversary but a whole group of adversaries. Thus, we need to learn bounded rationality models for a whole population of adversaries. How do we handle such populations given we may not have enough data to model the entire population accurately? We will typically not have enough data to model the problem as a Bayesian game, but a robust approach only produces a conservative strategy by generating the best resource allocation against the worst type of attacker. Here, we propose a fundamental research advance integrating both approaches so as to unify their strengths and overcome their limitations.

c) *Learning models of opportunistic criminals*: In domains such as urban security and urban crime prevention, the majority of adversaries are opportunistic criminals, who conduct little planning or surveillance before attacking. These adversaries capitalize on local opportunities and react to real-time information. However, in traditional applications of security games such as counter-terrorism, we focus on adversaries with significant planning and little execution. We propose Opportunistic Security Games (OSG) to address such criminals who are less strategic in planning attacks and more flexible in executing them. Rather than simply hypothesizing a model as done in previous work, we propose to develop new machine learning approaches to learn the model of opportunistic criminals from real-world criminal activity data in OSG.

d) *Preference elicitation for security games*: In domains such as fishery and wildlife protection and drug interdiction, information about vast geographical areas is not available to security agencies. Even in traditional applications of security games such as counter-terrorism, we may not know the value of the different targets with precision. However, in domains with a large number of targets, it is impractical for the defender to gather information at all targets due to limited information gathering resources. Therefore, we aim at determining targets at which reducing uncertainty would improve the defender's patrols the most. We propose to augment our game-theoretic resource allocation approaches using an information/preference elicitation algorithm, marrying the security allocation algorithm with a preference/information elicitation mechanism that allows reducing uncertainty only over those targets that will improve the resource allocation the most.

9. Research Transition Objectives: Our past research has led to five major software products for security resource optimization: the ARMOR software for LAX security scheduling, the IRIS software for scheduling deployments of Federal Air Marshals, the PROTECT software for the US Coast Guard, the Compass software for the US Coast Guard, and the TRUSTS software for the LASD. While this first generation of security games software has provided benefits to these agencies, our proposed research will be able to enhance these software products by addressing key issues faced such as in uncertainty in target values and adversary behavior modeling. For example, the US Coast Guard fishery protection efforts will be improved by the machine learning research mentioned in this proposal to better estimate adversary action; whereas PROTECT could be improved by the enhancements mentioned in preference elicitation.

10.Interfaces to CREATE Projects:

We will interface with CREATE's projects on preference elicitation and behavioral modeling.

11. Previous or current work relevant to the proposed project, why is DHS interested, identify/who are your expected DHS users: Above, we have discussed several previous applications we have developed (ARMOR, IRIS, PROTECT, Compass, TRUSTS) which are in actual use by a variety of DHS agencies such as the US Coast Guard, the Federal Air Marshals and others. Security resource optimization is a fundamental challenge faced by these agencies, and therefore we expect

significant DHS interest. Specifically, in this new work, we are working with the US Coast Guard for protecting fisheries, and with USC Dept of Public Safety to improve Campus Security (which would also be applicable to many other police departments).

12. Major Deliverables, Research Transition Products and Customers: (What are the major products, and who are the primary clients that are interested in the results.) One of our major deliverables will include research papers published in top international conferences based on four research objectives we have mentioned above. In addition, our research will enhance several of the software applications that we have delivered to DHS agencies in the past as mentioned above. In particular, as mentioned above, past customers of our work include LAX airport, US Coast Guard, Federal Air Marshals Service, LA Sheriff's Department and several others. These customers have expressed an interest in many of the outcomes of the research discussed here.

13.Technical Approach:

We outline technical approach for each of the challenges below:

(a) Machine learning in repeated games: In learning bounded rationality models, e.g., QR/SUQR (Quantal Response/Subjective Utility Quantal Response) models for the adversary, we exploit data about the behavior of the adversary, such as when and where illegal fishing occurs. There are two main challenges that must be addressed. One main challenge is that adversary data is only available from locations where the defender visits while conducting patrols. Given this fact, it is critical to avoid confirmation bias from over relying on previously collected data, which can lead to self-reinforcing behavior and undesired consequences. For example, in the early rounds of the game, the defender could observe multiple instances of illegal fishing in particular regions of the patrol area. As the defender updates its model of the adversary it will naturally focus more patrols on these regions. By patrolling these regions more frequently, the defender is more likely to observe any instances of illegal fishing that occur there. Thus, the defender can end up in situations where certain regions are patrolled with a high frequency due to the perceived high volume of illegal fishing and a high volume of illegal fishing is perceived due to the high frequency of patrols. To prevent such situations, we propose to incorporate the concept of exploitation (patrolling where significant adversary data has been previously collected) and exploration (patrolling where little to no adversary has been previously collected) into the defender's patrolling strategy. The main idea is that the defender should visit all regions of the patrol area with at least some probability to ensure the data collected is representative of the adversary's behavior. By considering this tradeoff between exploitation and exploration, we can take advantage of available data and avoiding over-fitting, in order to learn a more accurate model of the adversary over time.

A second main challenge is to protect against losses in initial rounds. The overall idea in machine learning in repeated games is to estimate and predict the behavior of the adversary by learning its behavior through carefully collected data *but without allowing the adversary to inflict damage while the learning is in progress*. This is an important issue because existing approaches for learning bounded rationality models in repeated Stackelberg games assume that enough data will be collected in the initial rounds to learn a reliable adversary model. Our analysis reveals that the issue is not the amount of data, *but insufficient exposure of attack surface* in initial rounds to gather sufficient information about adversary responses to various strategies and learn a reliable model. Learning is biased towards the limited available information and hence the defender incurs significant losses until enough of the right kind of data becomes available. This degraded performance in initial rounds may have severe consequences for three reasons: (i) In domains like fisheries protection or drug interdiction or wildlife crime, each round lasts for weeks or potentially months and so initial round losses (if massive) could imply irrecoverable losses of resources (e.g., fish stocks). (ii) Given the nature of these domains, re-initialization of the game may periodically be necessary and thus initial

rounds may be repeated. One approach to address the problem of limited exposure to significant portions of the attack surface in initial rounds is to reason about similarity between exposed and unexposed areas of the attack surface to build a better prediction model of the adversary's preferences. In addition, we would incorporate a discounting parameter, computed based on the amount of attack surface exposed at the end of each round, to mitigate the adversary's lack of exposure to enough of the attack surface in the initial rounds. This combined approach would help us to not only exploit regions where we have data but also identify potential areas of attack where we have no data and thus improve defender performance.

(b) Bayesian-robust approach for addressing boundedly rational attackers in security games: The challenge faced in addressing this research objective is that each attacker type results in a particular bounedly-rational behavioral model of attacker. However, learning a Bayesian distribution over the attacker types requires a significant amount of data in order to achieve a precise estimation of that distribution. The lack of data will lead to an inaccurate prediction, thus would cause a severe deterioration in computing the optimal strategy for the defender; which maximizes the defender's expected utility over the (inaccurately) predicted distribution. On the other hand, solely relying on maximizing the defender's expected utility against the worst type of attacker is an overly conservative approach which wastes available data, which partly reveals information w.r.t the attacker population. Our idea is to apply a robust approach that deals with the worst outcome from the prediction inaccuracy of a Bayesian distribution w.r.t attacker types, which allows us to leverage the available data while providing a robust solution to a prediction inaccuracy.

To generate a hybrid of Bayesian and Robust approaches, one approach is to relax the requirement of a known exact distribution in the Bayesian approach by considering a hyper-rectangle in the space of probability distributions centered at the estimated distribution of adversary types. All distributions lying within this hyper-rectangle are considered as a possible Bayesian distribution of attacker types. We then attempt to deal with the worst-case scenario for the defender under this type of uncertainty by maximizing the defender's utility against the worst-case distribution of attacker type. When the size of intervals goes to infinity, we obtain the robust approach. On the other hand, when interval size shrinks to a single point, our method becomes the Bayesian approach. Another idea is to solely focus on predicting the probability of attack by the adversary at every target. The intuition behind this idea is that in any approach for addressing boundedly rational attackers, what finally matters is the probability of attack by the adversary, which determines the averaged expected utility of the defender. Thus, we directly predict the probability of attack by first learning a single attacker behavioral model that captures the decision making of the whole attacker population. Then we apply a relaxation method which assumes all attack probability distributions to lie within a hyper-rectangle. We then maximize the defender's utility against the worst-case attack probability distribution of the attacker within that uncertainty hyper-rectangle.

(c) Learning models of opportunistic criminals: Instead of building an exact criminal behavior model with constraints, we propose a novel approach in Opportunistic Security Games (OSG): we aim to learn the behavior model from real data (we have received significant amounts of crime and patrol data from USC DPS). We do so by modeling the interaction between the criminal and patrol officers as a Dynamic Bayesian Network (DBN). In every time-step of the DBN we capture the following actions: the patrol officers are assigned to protect patrol areas and criminals react to such allocation by committing crimes opportunistically. Across time-steps the criminal can move from one area to another. The criminals' payoff is influenced by the attractiveness of targets and the assignment of the officers. These payoffs drive the behavior of the criminals, but rather than learning these payoff our DBN learns how the criminal moves from area to area and how likely is he to commit crime.

Given a DBN, we apply the well-known Expectation Maximization (EM) algorithm to learn the unknown parameters in DBN. However, the number of unknown variables is much larger than the available data in our case, which results in over-fitting. Therefore, we modify our basic DBN by marginalizing states in the DBN using approximation technique from the Boyen-Koller algorithm and exploiting structure of this problem. These modifications lead to a compact representation of the DBN model, in which the number of parameters is smaller than the available data. The compact model results in better learning accuracy and increased speed of learning. Finally, we propose to plan and deploy the optimal officer's strategy based on the criminal model. However, the criminal behavior would change as he observes and reacts to the deployment of a new strategy. Hence, we propose to frequently update our adversary model as we obtain new training data, which leads to an online mechanism. In this mechanism, we update criminal's model based on real-time crime/patrol data and dynamically plan our allocation strategy.

(d) *Preference elicitation for security games*: Whether in domains such as fishery protection, drug interdiction and wildlife protection, or in traditional applications of security games such as for counter-terrorism, we may not precisely know target values. To address this shortcoming, we propose new preference/information elicitation algorithms married to the security allocation algorithm, so that we focus our attention only on gathering information and reducing uncertainty over those targets that maximally reduce our *regret* (see next paragraph) with respect to the security game resource allocation. Given quantitative benefit of reducing uncertainty over particular targets, we could potentially design a path that will be covered by a data gathering mobile sensor such as an unmanned aerial vehicle (a UAV), which will then be used to reduce the uncertainty over the game.

Specifically, we focus on addressing payoff uncertainty (e.g., animal or fishery density) at every target. Our idea is to use the robust solution concept of minimax regret which attempts to compute the optimal strategy for the defender that minimizes the regret or the loss in terms of the defender's utility w.r.t the actual optimal strategy over all possible realizations of the payoffs under uncertainty. The minimax regret method provides us not only a robust strategy for the defender against uncertainty but also a quantitative measurement of the patrolling quality in terms of distance to the optimal strategy. We then try to integrate this minimax regret solution into a payoff elicitation process for reducing uncertainty, which finds the best path for UAVs to cover on a flight based on minimax regret solution.

14. Major Milestones and Dates:

A paper in Computer Science conferences is a major milestone. Therefore an estimated timeline is given in terms of CS publications and tasks mentioned above:

- December 2015: Completed one paper on task (c), begin paper on task (d)
- April 2016: Completed a paper on task (d), begin paper on task (b)
- August 2016: Completed a paper on task (a) and (b)

15. Selected References

1. *F.M. Delle Fave*, A. Jiang*, Z. Yin, C. Zhang, M. Tambe, S. Kraus, J. Sullivan (* Both Delle Fave and Jiang are first authors of this article.) Game-theoretic Security Patrolling with Dynamic Execution Uncertainty and a Case Study on a Real Transit System In Journal of Artificial Intelligence Research (JAIR), 50:321-367, 2014.*
2. *T. Nguyen, A. Jiang, M. Tambe Stop the Compartmentalization: Unified Robust Algorithms for Handling Uncertainties in Security Games In International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2014.*
3. *T. Ngueyn, A. Yadav, B. An, M. Tambe, C. Boutilier Regret-based Optimization and Preference Elicitation for Stackelberg Security Games with Uncertainty In Conference on Artificial Intelligence (AAAI), July 2014*

4. *C. Zhang, A. Jiang, M. Short, J. Brantingham, M. Tambe Defending Against Opportunistic Criminals: New Game-Theoretic Frameworks and Algorithms In Conference on Decision and Game Theory for Security (GameSec), 2014*
5. *M. Brown, W. Haskell, M. Tambe Addressing Scalability and Robustness in Security Games with Multiple Boundedly Rational Adversaries In Conference on Decision and Game Theory for Security (GameSec), 2014*
6. *M. Brown, S. Saisubramanian, P. Varakantham , M. Tambe STREETS: Game-Theoretic Traffic Patrolling with Exploration and Exploitation In Conference on Innovative Applications of Artificial Intelligence (IAAI), 2014.*
7. *Y. Qian, W. Haskell, A. Jiang, M. Tambe Online Planning for Optimal Protector Strategies in Resource Conservation Games In International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2014.*

17. Brief Bio: Milind Tambe

Professional Preparation:

Carnegie Mellon University	Computer Science	PhD, 1991
Carnegie Mellon University	Postdoctoral Fellow, CS	1991-1993

Appointments:

Helen N and Emmett H Jones Professor in Engineering	2012-- present
Professor, Computer Science & Industrial Systems Engineering Dept University of Southern California(USC)	2006 – present

Five Key papers:

1. T. Nguyen, A. Jiang, M. Tambe Stop the Compartmentalization: Unified Robust Algorithms for Handling Uncertainties in Security Games In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), May 2014.
2. T. Ngueyn, A. Yadav, B. An, M. Tambe, C. Boutilier Regret-based Optimization and Preference Elicitation for Stackelberg Security Games with Uncertainty In Proceedings of the Conference on Artificial Intelligence (AAAI), July 2014.
3. R. Yang, B. Ford, M. Tambe, A. Lemieux Adaptive Resource Allocation for Wildlife Protection against Illegal Poachers In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), May 2014.
4. Y. Qian, W. Haskell, A. Jiang, M. Tambe Online Planning for Optimal Protector Strategies in Resource Conservation Games In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), May 2014.
5. Jiang, T. Nguyen, M. Tambe, A. Procaccia Monotonic Maximin: A Robust Stackelberg Solution Against Boundedly Rational Followers In Proceedings of the Conference on Decision Theory and Game Theory for Security (GameSec), November 2013.

Major Awards

1. Major awards in Computer Science
   a. *Fellow of the Association for Computing Machinery (2013)*
   b. *Fellow of the Association for Advancement of Artificial Intelligence (2007)*
   c. *Winner of the ACM/SIGART Autonomous Agents Research Award (2005)*

   Major awards in Operations Research
   a. *INFORMS Wagner Prize (2012)*
   b. *David Rist Prize, Military Operations Research Society (2011)*

   Awards for real-world impact of research
   a. *Meritorious Team Commendation from Commandant, US Coast Guard (2013)*
   b. *IBM Faculty award (2012)*
   c. *Certificate of Appreciation, the US Federal Air Marshals Service* (2011)
   d. *Commander, First Coast Guard District's Operational Excellence Award* (2011)
   e. *Christopher Columbus Fellowship Foundation Homeland Security Award* for Border (2010)
   f. *Commendation, City of Los Angeles, Los Angeles World Airports Police Department* (2009)

   Awards from USC for real-world impact of research
   a. *USC Associates award for creativity in research and scholarship (2014)*
   b. *Inaugural USC Viterbi School of Engineering Use-inspired research award* (2009).
2. 11 separate best paper awards, influential paper awards from international conferences & workshops including Autonomous Agents and Multiagent Systems (AAMAS), Intelligent Virtual Agents (IVA).