

Application of Game Theory to Sensor Resource Management

Sang “Peter” Chin

This expository article shows how some concepts in game theory might be useful for applications that must account for adversarial thinking and discusses in detail game theory’s application to sensor resource management.

HISTORICAL CONTEXT

Game theory, a mathematical approach to the analysis of situations in which the values of strategic choices depend on the choices made by others, has a long and rich history dating back to early 1700s. It acquired a firm mathematical footing in 1928 when John von Neumann showed that every two-person zero-sum game has a maximin solution in either pure or mixed strategies.¹ In other words, in games in which one player’s winnings equal the other player’s losses, von Neumann showed that it is rational for each player to choose the strategy that maximizes his minimum payoff. This insight gave rise to a notion of an equilibrium solution, i.e., a pair of strategies, one strategy for each player, in which neither player can improve his result by unilaterally changing strategy.

This seminal work brought a new level of excitement to game theory, and some mathematicians dared to hope that von Neumann’s work would do for the field of economics what Newtonian calculus did for physics. Most importantly, it ushered into the field a generation of young researchers who would significantly extend the outer limits of game theory. A few such young minds particularly stand out. One is John Nash, a Princeton mathematician who, in his 1951 Ph.D. thesis,² extended von Neumann’s theory to N -person noncooperative games and established the notion of what is now

famously known as Nash equilibrium, which eventually brought him Nobel recognition in 1994. Another is Robert Aumann, who took a departure from this fast-developing field of finite static games in the tradition of von Neumann and Nash and instead studied dynamic games, in which the strategies of the players and sometimes the games themselves change along a time parameter t according to some set of dynamic equations that govern such a change. Aumann contributed much to what we know about the theory of dynamic games today, and his work eventually also gained him a Nobel Prize in 2002. Rufus Isaacs deviated from the field of finite discrete-time games in the tradition of von Neumann, Nash, and Aumann (Fig. 1) and instead studied the continuous-time games. In continuous-time games, the way in which the players’ strategies determine the state (or trajectories) of the players depends continuously on time parameter t according to some partial differential equation. Isaacs worked as an engineer on airplane propellers during World War II, and after the war, he joined the mathematics department of the RAND Corporation. There, with warfare strategy as the foremost application in his mind, he developed the theory of such game-theoretic differential equations, a theory now called differential games. Around the same

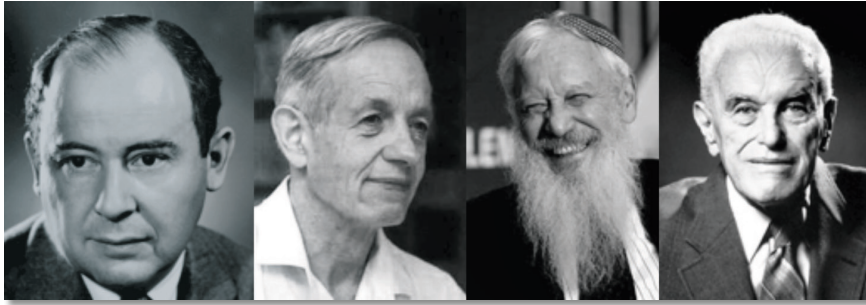


Figure 1. John von Neumann, John Nash, Robert Aumann, and John Harsanyi—pioneers of two-person, N -person, dynamic, and uncertain games, respectively.

time, John Harsanyi recognized the imperfectness and incompleteness of information that is inherent in many practical games and thus started the theory of uncertain games (often also known as games with incomplete information); he shared a Nobel Prize with Nash in 1994.

CURRENT RELEVANCE AND THE NEED FOR NEW APPROACHES

Despite their many military applications, two-person games, which were researched extensively during the two-superpower standoff of the Cold War, have limitations in this post-Cold War, post-September 11 era. Indeed, game theory research applied to defense-related problems has heretofore mostly focused on static games. This focus was consistent with the traditional beliefs that our adversary has a predefined set of military strategies that he has perfected over many years (especially during the Cold War) and that there is a relatively short time of engagement during which our adversary will execute one fixed strategy. However, in this post-September 11 era, there is an increasing awareness that the adversary is constantly changing his attack strategies, and such variability of adversarial strategies and even the variability of the game itself from which these strategies are derived, call for the application of dynamic games to address the current challenges. However, although solutions of any dynamic two-person game are known to exist by what is commonly termed “folk theorem,”³ the lack of an efficient technique to solve dynamic games in an iterative fashion has stymied their further application in currently relevant military situations. In our research in the past several years, we have developed such an online iterative method to solve dynamic games using insights from Kalman filtering techniques.

Furthermore, although the interaction between offense and defense can be effectively modeled as a two-person game, many other relevant situations—such as the problem of allocating resources among different assets of a sensor system, which can be modeled as an N -person game—involve more than two parties. Thanks to Nash’s seminal work in this area,² we know an equilibrium solution exists for any static N -person

game, although actually finding equilibrium is a highly nontrivial task (recall that Nash’s proof of his theorem is a nonconstructive proof). There are a number of heuristic methods of estimating and finding the equilibrium solutions of N -person games, but in our research, we have been developing a new method of approximately and efficiently solving N -person games using Brouwer’s fixed-point theorem (a crucial ingredient in

Nash’s doctoral thesis). In this article, we show how such methods can be used in various applications such as optimally managing assets of a sensor.

Finally, we sketch out a software architecture that brings the aforementioned ideas together in a hierarchical and dynamic way for the purpose of sensor resource management—not only may this architecture validate our ideas, but it may also serve as a tool for predicting what the adversary may do in the future and what the corresponding defensive strategy should be. The architecture we sketch out is based on the realization that there is a natural hierarchical breakdown between a two-person game that models the interaction between a sensor network and its common adversary and an N -person game that models the competition, cooperation, and coordination between assets of the sensor network. We have developed a software simulation capability based on these ideas and report some results later in this article.

NEW APPROACHES

Dynamic Two-Person Games

Many current battlefield situations are dynamic. Unfortunately, as important as von Neumann’s and Nash’s equilibrium results of two-person games are, they are fundamentally about static games. Therefore, their results are not readily applicable to dynamic situations in which the behavior of an adversary changes over time, the game is repeated many or even an infinite number of times (repeated games), the memory of players add dynamics to the game, or the rules of the game themselves change (fully dynamic games). It is true that there are theorems, such as folk theorems, that prove that an infinitely repeated game should permit the players to design equilibriums that are supported by threats and that have outcomes that are Pareto efficient. Also, at the limit case of the continuous-time case, instead of discrete time, we may bring to bear some differential game techniques, which are inspired mostly by well-known techniques of partial differential equations applied to the Hamilton–Jacobi–Bellman equation, as pioneered by Rufus Isaacs. However, unfortunately, these approaches do not read-

ily give insight into how to select and adapt strategies as the game changes from one time epoch to the next, as is necessary in order to gain battlefield awareness of all of the game's dynamics.

Therefore, on the basis of our experience in Kalman filtering techniques for estimating states of moving targets, we propose a different approach to solving dynamic games in order to help figure out the intent of an adversary to determine the best sensing strategies. The key observation in our approach is that the best estimate of the adversarial intent, which continues to change, can be achieved by combining the prediction of what the adversarial intent may next be, the most recent measurement of the adversarial intent (as in Kalman filtering), and a Nash equilibrium strategy for the adversary (the strategy that the adversary will most likely adopt without any further insight into the friendly force's intent). Furthermore, figuring out the adversary's intent makes it possible for a sensor network to select the best sensing strategy in response to this adversarial intent and helps achieve overall situational awareness.

Our approach is inspired by our experience with Kalman filtering, as shown in Fig. 2, where the y axis represents adversarial strategy [note that the pure strategies (S1, ..., S9) on the y axis are on the discrete points and that the in-between points correspond to the mixed strategies] and the x axis represents time. As in the Kalman filtering paradigm, our approach tries to find the right balance in combining the prediction of what the

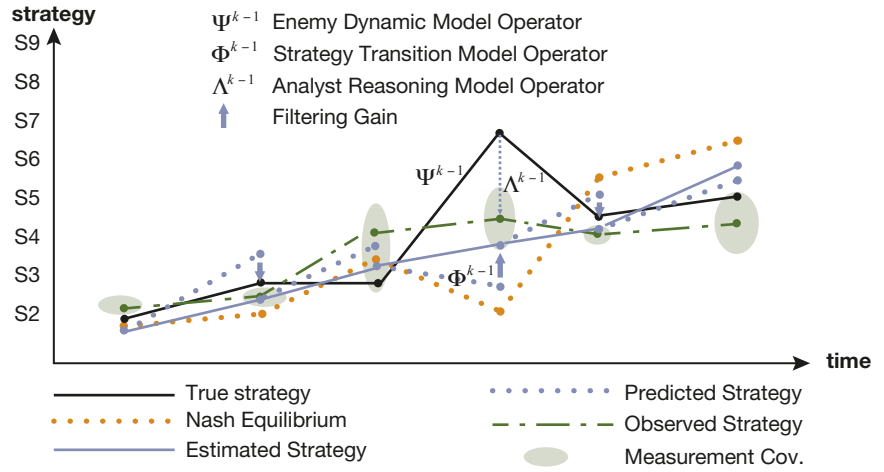


Figure 2. Filtering techniques for dynamic games. Cov., covariance.

adversarial intent may next be, governed by a model of adversarial strategy transition (operator Φ^{k-1} in the figure), and the most recent measurement of the adversarial intent, given by another model of how a decision node may process the sensor node data to map the measurement to a given set of adversarial strategies (operator Λ^{k-1} in the figure). However, unlike Kalman filtering, a third factor is also balanced with the prediction of adversarial strategy and the measurement of adversarial strategy: Nash equilibrium strategy for the adversary (the strategy that the adversary will most likely adopt without any further insight into the friendly entity's intent). Mathematically, this insight translates into adding a third term to the famous Kalman filtering equations. We start with the famous Kalman filtering equations:

$$\hat{y}_t^t = \hat{y}_t^{t-1} + K_t(x_t - B\hat{y}_t^{t-1}), \quad (1)$$

$$P_t^t = (I - K_t B)P_t^{t-1}, \text{ and} \quad (2)$$

$$K_t = P_t^{t-1} B^T (R + B P_t^{t-1} B^T)^{-1}. \quad (3)$$

From the Kalman filtering equations, we derive the following set of equations:

$$\hat{s}_t^t = \Phi^{t-1} \hat{s}_{t-1}^{t-1} + K_t(\Lambda^t s_t - \Lambda^{t-1} \Phi^{t-1} \hat{s}_{t-1}^{t-1}) + c_t(P_t^{t-1})N(G_t), \quad (4)$$

$$P_t^t = (I - K_t \Lambda^t)P_t^{t-1}, \text{ and} \quad (5)$$

$$K_t = P_t^{t-1} (\Lambda^t)^T (R + \Lambda^t P_t^{t-1} (\Lambda^t)^T)^{-1}, \quad (6)$$

where \hat{s}_t^t = estimated strategy at time t ; s_t = true strategy at time t ; $\Phi^t \in \text{HOM}(\Sigma^t, \Sigma^{t+1}; \mathfrak{R})$ models strategy transition; Σ^t = set of strategies at time t ; $\Lambda^t \in \text{HOM}(\Sigma^t, \Sigma^t; \mathfrak{R})$ models intelligence analyst's understanding of enemy strategy; G_t = game at time t ; $N(G_t)$ = a Nash equilibrium solution for the game G_t at time t ; c_t = a Nash equilibrium solution discount factor for the game G_t at time t ; P_t^t stands for the covariance of the moving object and K_t stands for its Kalman gain; and HOM stands for the group of homomorphisms; and R stands for the set of real numbers. The following main equation,

$$\hat{s}_t^t = \Phi^{t-1} \hat{s}_{t-1}^{t-1} + K_t(\Lambda^t s_t - \Lambda^{t-1} \Phi^{t-1} \hat{s}_{t-1}^{t-1}) + c_t(P_t^{t-1})N(G_t), \quad (7)$$

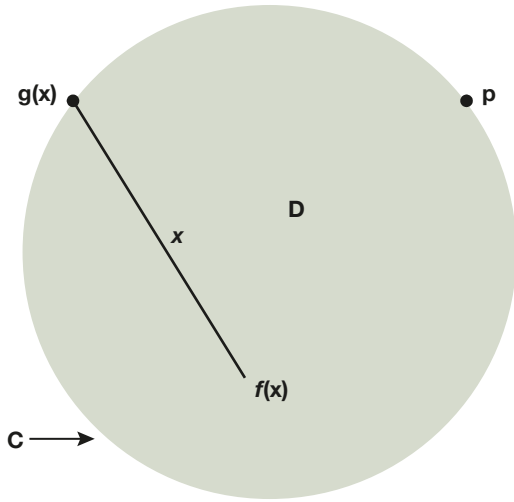


Figure 3. Brouwer’s fixed-point theorem: A continuous function from a ball (of any dimension) to itself must leave at least one point fixed. In other words, in the case of an N -dimensional ball, D , and a continuous function, f , there exists x such that $f(x) = x$.

shows how \hat{s}_t^t , the next estimate of the adversarial intent given by its next strategy, is a combination of the following three terms:

- $\Phi^{t-1} \hat{s}_t^{t-1}$: prediction of the next adversarial strategy given strategy transition model Φ^t ,
- $K_t (\Lambda^t s_t - \Lambda^{t-1} \Phi^{t-1} \hat{s}_t^{t-1})$: measurement of the current adversarial strategy given the model of how an intelligence analyst would process sensor measurement data, and
- $c_t (P_t^{t-1}) N(G_t)$: Nash equilibrium tempered by a discount factor, $c_t (P_t^t)$.

Furthermore, the next two equations describe how uncertainty of adversarial strategy given by the covariance term (P_t^t) and Kalman gain (K_t) grow under this dynamic system. We have empirical verification of the effectiveness of these techniques through extensive Monte Carlo runs, which we reported in Ref. 4, and we plan to solidify this theory of our filtering techniques for dynamic games to present a practical and computationally feasible way to understand adversarial interactions for future applications.

Solving N -Person Games

Unfortunately, there is no known general approach for solving N -person games because Nash’s famous result on equilibrium was an existence proof only and not a constructive

proof. Therefore, we need to further investigate innovative approaches to solve N -person games, which we will use to model cooperation among sensors or systems of sensors. Our key innovative idea for N -person games is not to resort to some heuristic approach in order to solve N -person games, as is usually done for static N -person games, but rather to use the techniques inspired by Brouwer’s fixed-point theorem used in Nash’s original Ph.D. thesis. We note that Nash’s arguments rest on the following equations, which describe a transformation from one strategy $\zeta = (s_1, s_2, s_3, \dots, s_n)$ to $\zeta' = (s'_1, s'_2, s'_3, \dots, s'_n)$ by the following mapping:

$$s'_i = \frac{s_i + \sum_{\alpha} \varphi_{i\alpha}(\zeta) \pi_{i\alpha}}{1 + \sum_{\alpha} \varphi_{i\alpha}(\zeta)}, \quad (8)$$

where $\varphi_{i\alpha} = \max(0, p_{i\alpha}(\zeta) - p_i(\zeta))$; ζ is an n -tuple of mixed strategies; $\pi_{i\alpha}$ is player i ’s α^{th} pure strategy; $p_i(\zeta)$ is the payoff of the strategy ζ to player i ; and $p_{i\alpha}(\zeta)$ is the payoff of the strategy ζ to player i if he changes to α^{th} pure strategy.

We note that the crux of Nash’s proof of the existence of equilibrium solution(s) lies in this fixed theorem (shown in Fig. 3) being applied to the following mapping $T : \zeta \rightarrow \zeta'$. By investigating the delicate arguments that Nash used to convert these fixed points into his famous Nash equilibrium solutions, we can use these arguments to construct an iterative method to find the approximate solutions of N -person games. One possible approach we have in mind is to first start from a set of sample points in the space of strategies and compute:

$$C(\zeta) = \frac{\|T(\zeta) - \zeta\|}{\|\zeta\|}, \quad (9)$$

where $C(\zeta)$ measures the degree to which a strategy ζ is changed by the map $T : \zeta \rightarrow \zeta'$. We can then look for a strategy ζ , where $C(\zeta)$ is close to 1, as the possible initial search points to look for equilibrium points to start such an iterative search process. Furthermore, we have incorporated such techniques into a recent work

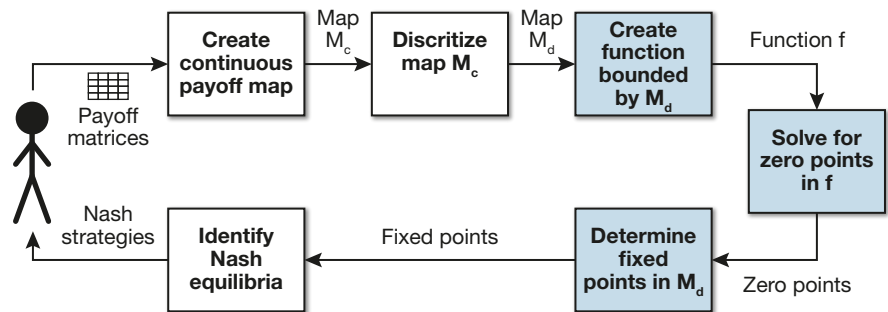


Figure 4. Flow chart for computing Nash equilibria for N -person games. M_c , continuous payoff map; M_d , discretized payoff map.

Table 1. Improved complexity on computing fixed points

Dimensions	Size	Complexity	Previous solve time (s)	New solve time (s)
3	$5 \times 6 \times 5$	36	0.015	0.000
5	$6 \times 6 \times 6 \times 6 \times 6$	1,296	2.532	0.406
3	$200 \times 200 \times 200$	40,400	26.922	1.109

by Chen and Deng,⁵ which has to date the best computational complexity result on the number of players and strategies for computing fixed points of direction preserving (direction-preserving maps are discrete analogues of continuous maps; see Fig. 4). We have developed a JAVA program that efficiently implements these new techniques (see Table 1 for the current computational time on one single dual-core PC running at 2 GHz given different numbers of players and strategies with various algorithmic complexities) and plan to further this technique for future applications in order to adjudicate resources among N sensors and systems of sensors. The advantage of using the Nash equilibrium solution is that it is an inherently safe solution for each actor (sensors or players) because it gives a solution that maximizes minimum utility of using each sensor given the current conditions and is robust against future changes in conditions. Moreover, whenever actors can trust and establish effective communi-

cation with each other, it allows us to use the theory of bargaining, which is well understood in the game theory community, to model the cooperation and coordination among sensors to achieve the most optimal solution (Pareto optimal). The optimal solution is not always achievable in noncooperative games, which we will use when the sensor nodes cannot trust each other as readily.

APPLICATIONS FOR SENSOR RESOURCE MANAGEMENT

Figure 5 shows how a possible approach could work. There are three levels at which game theory is being applied. First, there is a local two-person game that is defined by a set of sensing strategies for each sensor and the adversary who is aware of being sensed and thus tries to elude such sensing (which is being implemented

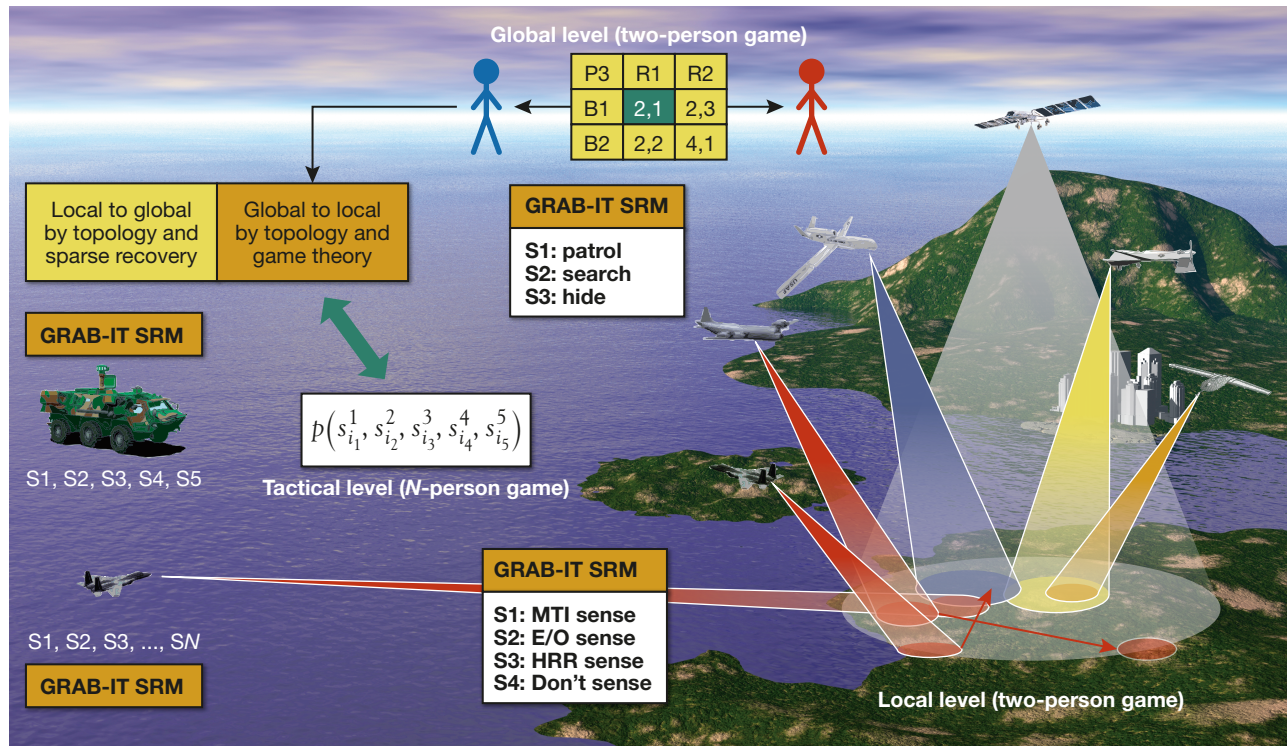


Figure 5. Game theory at three levels (global, tactical, and local). E/O, electro-optical; HRR, high-range resolution radar; MTI, moving target indicator.

by Game-theoretic Robust Anticipatory Behavior-leaning Individual Tracking Sensor Resource Manager, noted as GRAB-IT SRM in the figure). Then, among these sensors operating in a same area of interest, or among the systems of sensors operating in different areas of interest, there is an N -person game to allocate resources. Finally, as these local sensors or systems of sensors begin to infer the intents of adversaries, such knowledge may coalesce to define a global two-person game wherein the Blue force may be fighting against an overarching adversarial leader who is coordinating his subordinates' activities against the Blue force. None of these games are easy to define—they are all dynamic and uncertain (otherwise known as games of incomplete information), and such games will depend on many factors with inherent variabilities (environmental as well as human). And even if these games are defined, it is nontrivial to solve them. However, we believe our aforementioned new approaches in solving two-person, N -person, and dynamic games should overcome such difficulties and allow a game-theoretic approach to be effective in managing resources against adaptive adversaries.

We implemented a simple multilayer version of this in the software simulation environments below. As a finite game, it is defined by a set of strategies for each player and the payoff matrix, representing numerically how valuable each player views a particular set of strategy choices. The situation modeled here between the Blue and Red forces is described as a two-person, zero-sum game (in future work we will consider extensions to non-zero-sum games). Because the Blue force is uncertain of the enemy type, we define a game between the Blue force and each possible adversary type (offensive, defensive, and deceptive). To simplify the simulation and ensuing analysis without losing generality, we will assume that each enemy type has the same set of strategies. Namely, we have (for this simulation):

- Blue strategies = {act, wait}
- Red brigade types = {offensive, defensive, deceptive}
- Red strategies = {attack, defend, deceive}
- Payoff matrix:

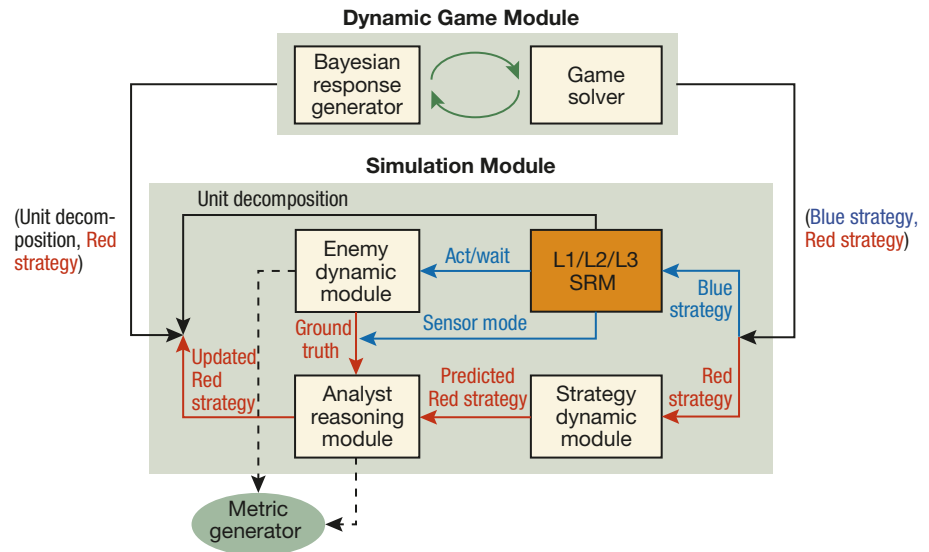


Figure 6. Putting game theory into the sensor resource allocation decision. L1/L2/L3, level 1/level 2/level 3 data fusion.

$$\begin{aligned}
 \circ A(\text{offensive}) &= \begin{bmatrix} 5 & -2 & 1 \\ -5 & 3 & 2 \end{bmatrix} \\
 \circ A(\text{defensive}) &= \begin{bmatrix} -1 & 2 & 2 \\ 1 & -1 & 1 \end{bmatrix}, \\
 \circ A(\text{deceptive}) &= \begin{bmatrix} 2 & -1 & 3 \\ -2 & 1 & -2 \end{bmatrix}
 \end{aligned} \tag{10}$$

where the first and second row of each payoff matrix corresponds to the Blue force act and wait strategies, respectively, and the first, second, and third columns correspond to the Red brigade attack, defend, and deceive strategies, respectively. For example, the first row of $A(\text{offensive})$ reflects that if an offense-minded adversary is attacking (column 1), a sensor should be in act mode (row 1) to sense the enemy's action, and therefore there should be a high payoff for using the sensor at the right time, resulting in a payoff of 5. However, if a sensor is in act mode (row 1) when the offense-minded enemy is defending (column 2), there is a sensor resource that gets wasted, resulting in a payoff of -2 . If the enemy is "deceiving" (neither fully attacking nor defending, as in column 3), it would still be of some use to put the sensor in act mode (row 1), resulting in a payoff of 1. The other cases are reasoned and modeled in a similar way. As described in Ref. 6, the equilibrium solution provides our next sensor decision (action or no action) as well as the probabilistic assessment of enemy strategy. If the decision is positive (action), we then use the level II valuation function to select a sensor mode. The assessed enemy strategy is used to predict the enemy strategy for the next time step. Such reasoning was put into the dynamic game module in the simulation architecture shown in Fig. 6.

We conduct a number of Monte Carlo simulations in which the initial enemy strategy is selected either randomly or based on the enemy unit composition. In

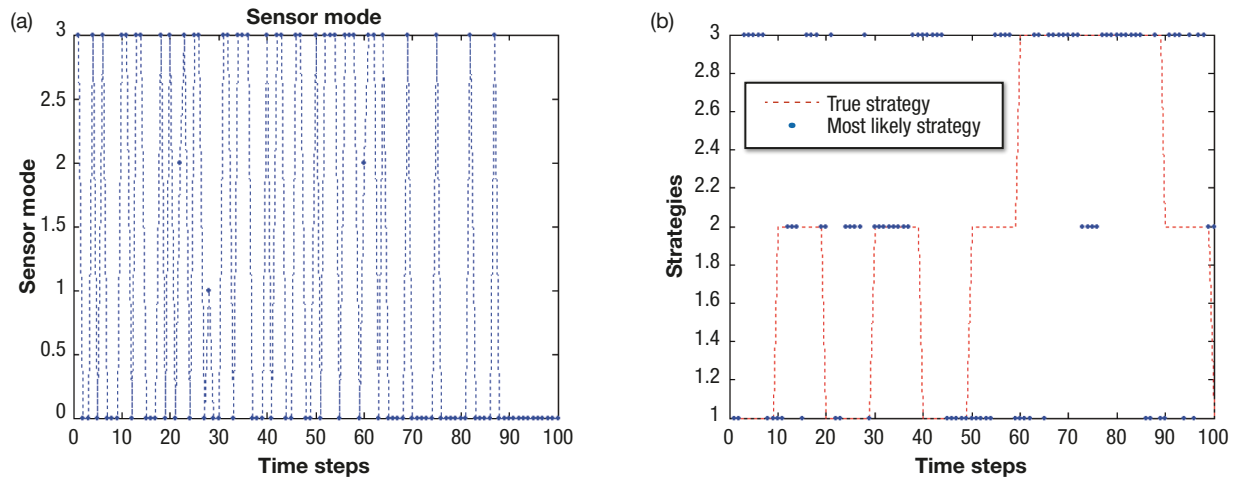


Figure 7. Sensor decision and probability of correct classification.

each trial, we conduct 100 time steps where a dynamic sensor decision is made at each step. In the simulation, we try three scenarios, with each consisting of a different composition of units (offensive, deceptive, and defensive). In the simulation, we run 50 Monte Carlo trials for each scenario. In each scenario, the window size for sensor action is set at 10 and the decision threshold is set differently in each test. For example, Fig. 7 shows the results of a typical trial with threshold equal to 40%. Figure 7a shows that approximately 36% of the time sensor is off (mode 0) and the rest sensor is on (mode 1 for ground moving target indicator, mode 2 for HRR on unit 1, and mode 3 for HRR on unit 2). Figure 7b shows that in this trial, the enemy's strategy has changed from 1 (offensive) to 2 (defense), then back to 1, then later to 3 (deceptive), and then back to 1. Figure 7b also shows that approximately 44% of the time, the most likely strategy of our assessment is the true one. [Note: in Figs. 7a and 7b, the y values

(sensor mode and strategies, respectively) only take on the integral values.]

In each test, we also compare the game solver with a heuristic algorithm where the sensor action/no-action decision was assigned on the basis of a prespecified probability. We have found that in general, the performance of the heuristic solver is significantly worse than the performance of the game solver (the heuristic solver determines when to or when not to use the sensor stochastically with the probability at $x\%$). This is understandable because the enemy strategy adapts to our sensor actions and changes accordingly, and thus it is much more difficult to assess enemy's strategy without an interactive game-theoretic strategy analyzer. Figure 8 shows the overall performance comparison between the heuristic approach and the game solver approach. In this test, we set the size of time window for enemy strategy policy at 10, and the decision threshold for change is 80% (i.e., our sensor will need to be on greater than 80% of

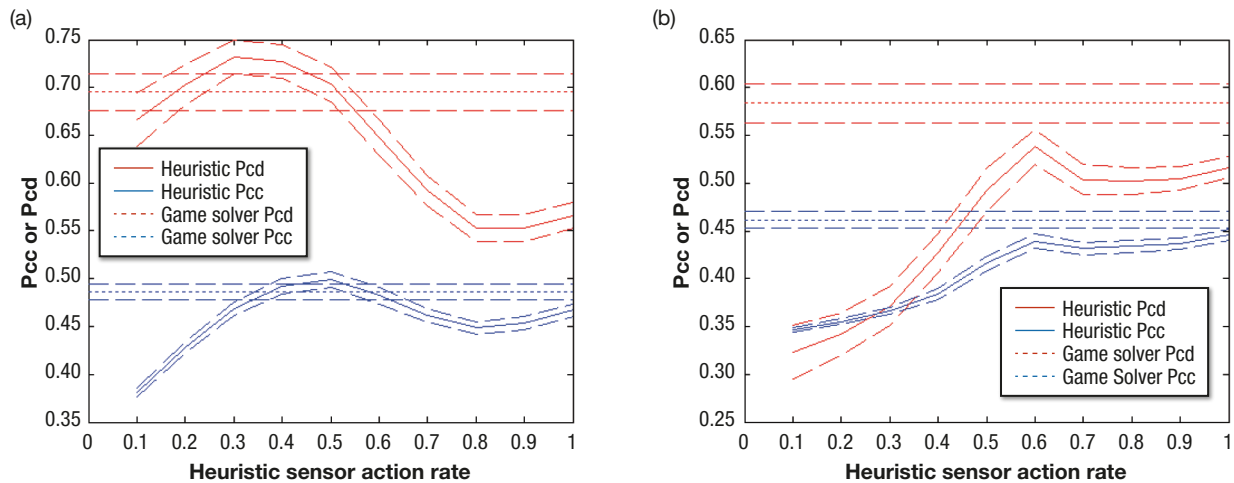


Figure 8. Effect of game theory on probability of correct classification (Pcc) and probability of correct decision (Pdd). (a) Offensive enemy; performance comparison, average of all cases. (b) Defensive enemy; performance comparison, average of all cases.

the time in the preceding time window of size 10 before the enemy will potentially change strategy). Note that in Fig. 8, we display both the mean and the 1-standard-deviation interval for both the Pcd and the Pcc. The x axis represents the probability of sensor action for the heuristic approach. For example, $x\%$ represents that for $x\%$ of the time without the game-theoretic reasoning, the sensor is being used. The corresponding y values represent that corresponding Pcc and Pcd of such action. The horizontal lines represent the Pcc and Pcd that are achieved when the game-theoretic approach is being incorporated. Figure 8a represents the case in which the adversary is of the offensive type. Figure 8b represents the case in which the adversary is of the defensive type. For the offensive adversary, the game-theoretic approach outperformed in most cases. For the defensive adversary, the game-theoretic approach seems to outperform any heuristic methods.

CONCLUSIONS

We have shown that game-theoretic reasoning can be used in the sensor resource-management problem to help identify enemy intent as the Blue force interacts and reacts against the strategies of the Red force. We have laid down a solid mathematical framework for our approach and built a Monte Carlo simulation environment to verify its effectiveness. Our approach uses game theory (two-person, N -person, cooperative/noncooperative, and dynamic) at different hierarchies of sensor planning, namely at the strategic/global level and the tactical/local level, treating each sensor node as a player in a game with a set of strategies corresponding to its set of sensing capabilities (location, geometry, modality, time availabilities, etc.) whose respective effectivenesses are numerically captured in a payoff function. As these sensors (players) in the sensor network come into the vicinity of each other (spatially, temporally, or both), they form either a noncooperative game (when the communication between sensors is minimal or nonexistent)

or a cooperative game (when sensors can communicate effectively enough to enter into a binding agreement). The Nash solution(s) of such games provides each sensor a strategy (either pure or mixed) that is most beneficial to itself and to its neighbor, and this strategy can then be translated into its most optimal sensing capability, providing timely information to the tactical users of the sensor network.

Beyond this proof of concept, we believe game theory has the potential to contribute to a myriad of current and future challenges. It is becoming increasingly important to estimate and understand the intents and the overall strategies of the adversary at every level in this post-September 11 era. Game theory, which proved to be quite useful during the Cold War era, now has even greater potential in a wide range of applications being explored by APL, from psychological operations in Baghdad to cyber security here at home. We believe our approach—which overcomes several stumbling blocks, including the lack of efficient game solver, the lack of online techniques for dynamic games, and the lack of a general approach to solve N -person games, just to name a few—is a step in the right direction and will allow game theory to make a game-changing difference in various arenas of national security.

REFERENCES

- ¹Von Neumann, J., and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ (1944).
- ²Nash, J. F., *Non-cooperative Games*, Ph.D. thesis, Princeton University (1951).
- ³Haurie, A., and Zaccours, G. (eds.), *Dynamic Games: Theory and Applications*, Springer, New York (2005).
- ⁴Chin, S., Laprise, S., and Chang, K. C., “Game-Theoretic Model-Based Approach to Higher-Level Fusion with Application to Sensor Resource Management,” in *Proc. National Symp. on Sensor and Data Fusion*, Washington, DC (2006).
- ⁵Chen, X., and Deng, X., “Matching Algorithmic Bounds for Finding a Brouwer Fixed Point,” *J. ACM* **55**(3), 13:1–13:26 (2008).
- ⁶Chin, S., Laprise, S., and Chang, K. C., “Filtering Techniques for Dynamic Games and their Application to Sensor Resource Management (Part I),” in *Proc. National Symp. on Sensor and Data Fusion*, Washington, DC (2008).

The Author

Sang “Peter” Chin is currently the Branch Chief Scientist of the Cyberspace Technologies Branch Office of the Asymmetric Operations Department, where he is conducting research in the areas of compressive sensing, data fusion, game theory, cyber security, cognitive radio, and graph theory. His e-mail address is sang.chin@jhuapl.edu.

The Johns Hopkins APL Technical Digest can be accessed electronically at www.jhuapl.edu/techdigest.