Deception in Non-Cooperative Games with Partial Information[†]

Technical Report

João P. Hespanha[‡] hespanha@usc.edu
Yusuf S. Ateşkan ateskan@usc.edu
Hüseyin H. Kızılocak kiziloca@usc.edu

Dept. of Electrical Engineering—Systems, Univ. of Southern California
3740 McClintock Ave., Room 318, MC 2563, Los Angeles, CA 90089-2563
phone: (213) 740-9137, fax: (213) 821-1109

February 14, 2001[§]

Abstract

In this report we illustrate how deception can be used by rational players in the context of non-cooperative stochastic games with partial information. We show that, when one of the players can manipulate the information available to its opponents, deception can be used to increase the player's payoff by effectively rendering the information available to its opponent useless. However, this is not always the case. Paradoxically, when the degree of possible manipulation is high, deception becomes useless against an intelligent opponent since it will simply ignore the information that has potentially been manipulated. This study is carried out for a prototype problem that arises in the control of military operations, but the ideas presented are useful in other areas of applications, such as price negotiation, multi-object auctioning, pursuit-evasion, etc.

[†]This work was supported by the Defense Advanced Research Projects Agency, JFACC Program.
[‡]Corresponding author.
[§]This report supersedes two previous reports with the same title, dated April 2000 and June 2000.

1 Introduction

Competitive games are usually classified as either having full or partial information. In *full-information* games both players know the whole state of the game when they have to make decisions. By state, we mean all information that is needed to completely describe the future evolution of the game, when the decision rules used by both players are known. Examples of full information games include Chess, Checkers, and Go. *Partial-information* games differ from these in that at least one of the players does not know the whole state of the game. Poker, Bridge, and Hearts are examples of such games. In full information games, as a player is planning its next move, it only needs to hypothesize over its and the opponent's future moves to predict the possible outcomes of the game [1]. This is key to using dynamic programming to solve full-information games. Partial information games are especially challenging because this reasoning may fail. In many partial information games, to predict the possible outcomes of the game, a player must hypothesize not only on the future moves of both players, but also on the past moves of the opponent. This often leads to a tremendous increase in the complexity of the games. In general, partial information stochastic games are poorly understood and the literature is relatively sparse. Notable exceptions are games with lack of information for one of the players [2, 3] and games with particular structures such as the Duel game [4], the Rabbit and Hunter game [5], the Searchlight game [6, 7], etc.

Another issue that makes partial information games particularly interesting is the fact that a player can obtain future rewards by either one of two possible mechanisms:

1. Choosing an action that will take the game to a more favorable state;
2. Choosing an action that will make the other player act in our own advantage by making it believe that the game is in a state other than the actual one.

The latter corresponds to a deception move and is only possible in partial information games.

The potential use of deception has been recognized in several areas, such as price negotiation [8, 9], multi-object auctioning [10], pursuit-evasion [11, 12], human relations [13], and card games [14]. In [8, 9], the authors analyze a negotiation where the players don not know each other's payoffs, but receive estimates from their opponents. In order to increase its gain each player may bias the estimate given. In [9], an advising scheme is proposed to make deception mostly useless. In [10], it is analyzed how a bidder can use deception to lower the price of an item sold in a multi-object auction. A pursuit-evasion game is analyzed in [11], where the evader corrupts the information available to its opponent to gain an advantage. It is assumed that the evader can jam the pursuer's sensor and therefore induce measurement errors, produce false targets, or interrupt the observations. Deception has also been studied in the context of military operations [11, 15, 16, 17, 18, 19]. A notable historical event was operation Overlord during the Second World War that culminated with the D-day invasion of France in June 1944 by the allied forces. The success of this operation relied heavily on deceiving the German command regarding the time and place of the sea-borne assault [19]. It is also widely recognized that, by the end of the cold war, the trend in Soviet naval electronic warfare was changing toward an independent type of combat action instead of a purely support role. This new role emphasized radio deception and misinformation at all levels of command [15]. The detection of false targets or decoys is now an important area of research in radar systems [16, 18].

In this report , we analyze the use of deception in the framework of non-cooperative stochastic games with partial information. We take as a working example a prototype problem that arises in the control of military operations. In its simplest form, this game is played by an attacker that has to select one of several alternative targets and a defender that must distribute its defensive assets among them. This is a partial information game because the attacker has to make a decision without knowing precisely how the defense units have been distributed among the potential targets. We explore several variations of this game that differ in the amount of information available to the attacker. This can range from no information at all to perfect information provided, e.g., by intelligence, surveillance, or reconnaissance. The interesting cases happen between these two extremes because, in practice, the information available is not perfectly accurate and is often susceptible to manipulation by the opponent. It turns out that when the defender can manipulate the information available to its opponents—e.g., by camouflaging some of its defensive units and not camouflaging others—deception can be used to increase its payoff by effectively rendering the information available to the attacker useless.

The remaining of this report is organized as follows. In Section 2, we formally introduce the simplest version of the prototype game where both attacker and defender have no information available to use in their decisions. This will serve as the baseline to compare the deception games that follow. In Section 3, we consider the extreme situation where the defender completely controls the information available to the attacker. We show that none of the players profits from this new information structure and deception is useless. This situation changes in Section 4 where the defender may profit from using deception. Paradoxically, when the degree of possible manipulation is high, deception becomes useless against an intelligent opponent since it will simply ignore the information that has potentially been manipulated. Section 5 contains some concluding remarks and directions for future research. .

2 A Prototype Non-Cooperative Game

Consider a game between two players that pursue opposite goals. The *attacker* must choose one of two possible targets (A or B) and the *defender* must decide how to better defend them. We assume here that the defender has a finite number of assets available that can be used to protect the targets. To make these assets effective, they must be assigned to a particular target and the defender must choose how to distribute them among the targets. To raise the stakes, we assume that the defender only has three defense units and is faced with the decision of how to distribute them among the two targets. We start by assuming that both players make their decisions independently and execute them without knowing the choice of the other player. Although it is convenient to regard the players as "attacker" and "defender," this type of games also arise in non-military applications. For example, the "attacker" could be trying to penetrate a market that the "defender" currently dominates.

The game described above can be played as a zero-sum game defined by the cost below, which the attacker tries to minimize and the defender tries to maximize:

$$J := \begin{cases} c_0 & \text{no units defending the target attacked} \\ c_1 & \text{one unit defending the target attacked} \\ c_2 & \text{two units defending the target attacked} \\ c_3 & \text{three units defending the target attacked.} \end{cases}$$

Without loss of generality we can normalize these constants to have $c_0 = 0$ and $c_3 = 1$. The values for the constants c_1 and c_2 are domain specific. Here, we consider arbitrary values for c_1 and c_2, subject to the reasonable constraint that $0 < c_1 \leq c_2 < 1$. Implicit in the above cost is the assumption that both targets have the same strategic value. We only make this assumption for simplicity of presentation.

As formulated above, the attacker has two possible choices (attack A or attack B) and the defender has a total of four possible ways of distributing its units among the two targets. Each choice available to a player is called a *pure policy* for that player. We will denote the pure policies for the attacker by α_i, $i \in \{1, 2\}$, and the pure policies for the defender by δ_j, $j \in \{1, 2, 3, 4\}$. These policies are enumerated in Tables 1(a) and 1(b), respectively. In Table 1(b), each "$\circ$" represents one defensive unit. The defender polices δ_1 and δ_2 will be called *3-0 configurations*, whereas the policies δ_3 and δ_4 will be called *2-1 configurations*.

The game under consideration can be represented in its *extensive form* by associating each policy of the attacker and the defender with a row and column, respectively, of a matrix $G \in \mathbb{R}^{2\times 4}$. The entry g_{ij}, $i \in \{1, 2\}$, $j \in \{1, 2, 3, 4\}$ of G corresponds to the cost J when the attacker chooses policy α_i and the defender chooses policy δ_j. The matrix G for this game is given by

$$G := \begin{array}{c} \begin{array}{cccc} \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{array} \\ \begin{bmatrix} 1 & 0 & c_2 & c_1 \\ 0 & 1 & c_1 & c_2 \end{bmatrix} \end{array} \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \end{array} \tag{1}$$

policy	target assigned
α_1	A
α_2	B

(a) Attacker policies

policy	target A	target B
δ_1	∘ ∘ ∘	
δ_2		∘ ∘ ∘
δ_3	∘∘	∘
δ_4	∘	∘∘

(b) Defender policies

Table 1: Pure policies

In the context of non-cooperative zero-sum games, such as the one above, optimality is usually defined in terms of a saddle-point or Nash equilibrium [20]. A *Nash equilibrium* in pure policies would be a pair of policies $\{\alpha_{i^*}, \delta_{j^*}\}$, one for each player, for which

$$g_{i^*j} \leq g_{i^*j^*} \leq g_{ij^*}, \qquad \forall i, j.$$

Nash policies are chosen by rational players since they guarantee a cost no worst than $g_{i^*j^*}$ for each player, *no matter what the other player decides to do.* As a consequence, playing at a Nash equilibrium is "safe" even if the opponent discovers our policy of choice. They are also reasonable choices since a player will never do better by unilaterally deviating from the equilibrium. Not surprisingly, there are no Nash equilibria in pure policies for the game described by (1). In fact, all the pure policies violate the "safety" condition mentioned above for Nash equilibria. Suppose, for example, that the attacker plays policy α_1. This choice is certainly not safe in the sense that, if the defender guesses it, he can then choose the policy δ_1 and subject the attacker to the highest possible cost. Similarly, α_2 is not safe and therefore cannot also be in a Nash equilibrium pair.

To obtain a Nash equilibrium, one needs to enlarge the policy space by allowing each player to randomize among its available pure policies. In particular, suppose that the attacker chooses policy α_i, $i \in \{1, 2\}$, with probability a_i and the defender chooses policy δ_j, $j \in \{1, 2, 3, 4\}$, with probability d_j. When the game is played repeatedly, the expected value of the cost is then given by

$$\mathrm{E}[J] = \sum_{i,j} a_i g_{ij} d_j = a' G d.$$

Each vector $a := \{a_i\}$ in the 2-dimensional simplex[1] is called a *mixed policy for the attacker*, whereas each vector $d := \{d_j\}$ in the 4-dimensional simplex is called a *mixed policy for the defender*. It is well know that at least one Nash equilibrium in mixed policies always exists for finite matrix games (cf. Minimax Theorem [21, p. 27]). In particular, there always exists a pair of mixed policies $\{a^*, d^*\}$ for which

$$a^{*\prime} G d \leq a^{*\prime} G d^* \leq a' G d^*, \qquad \forall a, d.$$

Assuming that both players play at the Nash equilibrium the cost will then be equal to $a^{*\prime} G d^*$, which is called the *value of the game*. It is straightforward to show that the unique Nash equilibrium for the matrix G in (1) is given by

$$a^* := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}', \qquad d^* := \begin{cases} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}' & c_1 + c_2 \leq 1 \\ \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}' & c_1 + c_2 > 1 \end{cases} \tag{2}$$

with value equal to

$$a^{*\prime} G d^* = \max\left\{ \frac{c_1 + c_2}{2}, \frac{1}{2} \right\}.$$

[1] We call the set of all vectors $x := \{x_i\} \in \mathbb{R}^n$ for which $x_i \geq 0$ and $\sum_i x_i = 1$, the *n-dimensional simplex*.

This equilibrium corresponds to the intuitive solution that the attacker should randomize between attacking targets A or B with equal probability, and the defender should randomize between placing most of its units next to A or next to B also with equal probability. The optimal choice between 3-0 or 2-1 configurations (policies δ_1/δ_2 versus δ_3/δ_4) depends on the parameters c_1 and c_2. From (2) we conclude that 3-0 configurations are optimal when $c_1 + c_2 \leq 1$, otherwise the 2-1 configurations are preferable.

In the game described so far, there is no role for deception since the players are forced to make a decision without any information. We will change that next.

3 Full Manipulation of Information

Suppose now that the game described above is played in two steps. First the defender decides how to distribute its units. It may also disclose the position of some of its units to the attacker. On the second step the attacker decides which target to strike. To do this, it may use the information provided by the defender. For now, we assume that this is the only information available to the attacker and therefore the defender completely controls the information that the attacker uses to make its decision.

The rationale for the defender to voluntarily disclose the position of its units is to deceive the attacker. Suppose, for example that the attacker uses two units to defend target A and only one to defend B. By disclosing that it has units next to B, the defender may expect the opponent to attack A and, consequently, suffer a heavier cost.

In this new game, the number of admissible pure policies for each player is larger than before. The attacker now has 8 distinct pure policies available since, for each possible observation (no unit detected, unit detected defending target A, or unit detected defending target B), it has two possible choices (strike A or B). These policies are enumerated in Table 2(a). In policies α_1, α_2 the attacker ignores any available information and always attacks target A or target B. These policies are therefore called *blind*. In policies α_3 and α_4, the attacker never selects the target where it detects a defense unit. These policies are called *naive*. In policies α_5 and α_6 the attacker chooses the target where it detects a defending unit. These policies are called *counter-deception* since they presume that a unit is being shown close to the least defended target.

The defender has ten distinct pure policies available, each one corresponding to a particular configuration of its defenses and a particular choice of which units to disclose (if any). These are enumerated in Table 2(b), where "$\circ$" represents a defense unit whose position has not been disclosed and "$\bullet$" a defense unit whose position has been disclosed. Here, we are assuming that the defender will, at most, disclose the placement of one unit because more than that would never be advantageous. In policies δ_1 through δ_4 nothing is disclosed about the distribution of the units. These are called *no-information* policies. In policies δ_9 and δ_{10} the defender shows units placed next to the target that has fewer defenses. These are *deception* policies. Policies δ_5 through δ_8 are *disclosure* policies, in which the defender is showing a unit next to the target that is better defended.

This game can be represented in extensive form by the following 8×10 matrix

$$G := \begin{array}{c} \begin{array}{cccccccccc} \delta_1 & \delta_2 & \delta_3 & \delta_4 & \delta_5 & \delta_6 & \delta_7 & \delta_8 & \delta_9 & \delta_{10} \end{array} \\ \left[\begin{array}{cccccccccc} 1 & 0 & c_2 & c_1 & 1 & 0 & c_2 & c_1 & c_2 & c_1 \\ 0 & 1 & c_1 & c_2 & 0 & 1 & c_1 & c_2 & c_1 & c_2 \\ 0 & 1 & c_1 & c_2 & 0 & 0 & c_1 & c_1 & c_2 & c_2 \\ 1 & 0 & c_2 & c_1 & 0 & 0 & c_1 & c_1 & c_2 & c_2 \\ 1 & 0 & c_2 & c_1 & 1 & 1 & c_2 & c_2 & c_1 & c_1 \\ 0 & 1 & c_1 & c_2 & 1 & 1 & c_2 & c_2 & c_1 & c_1 \\ 1 & 0 & c_2 & c_1 & 0 & 1 & c_1 & c_2 & c_1 & c_2 \\ 0 & 1 & c_1 & c_2 & 1 & 0 & c_2 & c_1 & c_2 & c_1 \end{array}\right] \end{array} \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \end{array} \qquad (3)$$

Just as before, for this game to have Nash equilibria one needs to consider mixed policies. However, this

policy	target assigned when ...		
	no obs.	unit detected at A	unit detected at B
α_1	A	A	A
α_2	B	B	B
α_3	B	B	A
α_4	A	B	A
α_5	A	A	B
α_6	B	A	B
α_7	A	B	B
α_8	B	A	A

(a) Attacker policies

policy	target A	target B
δ_1	○ ○ ○	
δ_2		○ ○ ○
δ_3	○○	○
δ_4	○	○○
δ_5	○ ○ ●	
δ_6		○ ○ ●
δ_7	○●	○
δ_8	○	○●
δ_9	○○	●
δ_{10}	●	○○

(b) Defender policies

Table 2: Pure policies

particular game has multiple Nash equilibria, one of them being

$$a^* := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}', \qquad d^* := \begin{cases} \begin{bmatrix} \frac{1}{2} \; \frac{1}{2} \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \end{bmatrix}' & c_1 + c_2 \leq 1 \\ \begin{bmatrix} 0 \; 0 \; \frac{1}{2} \; \frac{1}{2} \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \end{bmatrix}' & c_1 + c_2 > 1 \end{cases}$$

with value equal to

$$a^{*\prime} G d^* = \max\left\{ \frac{c_1 + c_2}{2}, \frac{1}{2} \right\}.$$

This shows that (i) the attacker can ignore the information available and simply randomize among the two blind policies; and (ii) the defender gains nothing from disclosing information and can therefore randomize among its no-information policies. It should be noted that there are Nash equilibria that utilize different policies. For example, when $c_1 + c_2 > 1$, an alternative Nash equilibrium is

$$\bar{a}^* := \begin{bmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}', \qquad \bar{d}^* := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}'. \tag{4}$$

In this case, the defender randomizes between deception and disclosure policies with equal probability and the attacker between the naive and counter-deception policies. However, in zero-sum games all equilibria yield the same value, so the players have no incentive to choose this equilibrium that is more complex in terms of the decision rules. Finally, it should also be noted that, because of the equilibrium interchangeability property for zero-sum games, the pairs $\{a^*, \bar{d}^*\}$ and $\{\bar{a}^*, d^*\}$ are also Nash equilibria [21, p. 28].

We have just seen that the attacker gains nothing from using the measurements available, even though these measurements give precise information about the position of some of the defense units. At an intuitive level, this is because the information available to the attacker is completely controlled by its opponent. And, if the defender chooses to disclose the position of some of its units, this is done solely to get an advantage. This can be seen, for example, in the equilibrium given by (4). We shall consider next a version of the game where the defender no longer has complete control over the information available to the attacker. For the new game, the attacker may sometimes improve its cost by using the available information.

4 Partial Manipulation of Information

In practice, when the defender decides to "show" one of its units it simply does not camouflage it, making it easy to find by the surveillance sensors used by the attacker. In the previous game we assumed that shown units are always detected by the attacker and hidden ones are not. We will deviate now from this ideal situation and assume that (i) shown units may not be detected and, more importantly, (ii) hidden units may sometimes be detected by the attacker. We consider here a generic probabilistic model for the attacker's surveillance, which is characterized by the conditional probability of detecting units next to a particular target, given a specific total number of units next to that target and how many of them are being shown. In particular, denoting by $\mathfrak{D}_A$ the event that defenses are detected next to target A, we have that

$$\mathrm{P}(\mathfrak{D}_A \mid \mathbf{n}_A = n_A,\ \mathbf{s}_A = s_A) = \chi(n_A, s_A),$$

where $\chi(\cdot)$ is the *characteristic function of the sensor*, $\mathbf{n}_A$ denotes the total number of units defending target A, and $\mathbf{s}_A$ the number of these that are shown. We assume here that $\mathfrak{D}_A$ is conditionally independent of any other event, given specific values for $\mathbf{s}_A$ and $\mathbf{n}_A$. Since there is no incentive for the defender to show more than one unit $\mathbf{s}_A \in \{0, 1\}$, whereas $\mathbf{n}_A \in \{0, 1, 2, 3\}$. For simplicity of notation, we assume that the surveillance of target B is identical and independent, with

$$\mathrm{P}(\mathfrak{D}_B \mid \mathbf{n}_B = n_B,\ \mathbf{s}_B = s_B) = \chi(n_B, s_B),$$

where the symbols with the B subscript have the obvious meaning.

For most of the discussion that follows, the characteristic function $\chi(\cdot, \cdot)$ can be arbitrary, provided that it is monotone non-decreasing with respect to each of its arguments (when the other is held fixed). The monotonicity is quite reasonable since more units (shown or not) should always result in a higher probability of detection. However, it will be sometimes convenient to work with a specific characteristic function $\chi(\cdot, \cdot)$. One possible choice is:

$$\chi(n, s) = 1 - (1-p)^s (1-q)^{n-s}, \qquad n \in \{0, 1, 2, 3\},\ s \in \{0, 1\}, \tag{5}$$

where $0 \leq q \leq p \leq 1$. This model assumes that (i) the surveillance sensors will provide a positive reading (i.e., announce that units were detected next to a particular target) when they are able to detect, at least, one unit, and (ii) the probability of detecting a particular defense unit that is hidden is equal to q, whereas the probability of detecting a unit that is shown is equal to p. A few particular cases should be considered:

- When $\chi(n, s) = 0$, $\forall\ n, s$ (or when $p = q = 0$ in (5)) we have the first game considered in this report , since the defense units are never detected.
- When

  $$\chi(n, s) = \begin{cases} 1, & s > 0 \\ 0, & s = 0 \end{cases} \qquad \forall n, s$$

  (or when $p = 1$ and $q = 0$ in (5)) the defense units are detected only when they are shown. This corresponds to the second game considered here, where the defender has full control of the information available to the attacker.

Another interesting situation occurs when placing more units next to a particular target makes that target more likely to be detected by the surveillance sensors *regardless of how many units are shown.* In such a case the attacker's sensors are said to be *reliable.* This can be formally expressed by the condition

$$\chi(n_1, s_1) \geq \chi(n_2, s_2), \tag{6}$$

for all $n_1 > n_2$ such that $n_1 + n_2 = 3$, and all s_1,s_2. Because the characteristic functions of the sensors are monotone non-decreasing with respect to each of its arguments (when the other is held fixed), it is straightforward to show that this reliability condition is equivalent to

$$\chi(2, 0) \geq \chi(1, 1). \tag{7}$$

For the characteristic function in (5), this corresponds to

$$2q - q^2 \geq p.$$

We will see below that the attacker can choose naive policies, when the sensors are reliable. A special case of reliable sensors arises when $\chi(n, s)$ is independent of s for all values of n, (or when $p = q$ in (5)). In this case we have sensors that cannot be manipulated by the defender since the detection is independent of the number of units "shown" be the defender.

In terms of the policies available, the game considered in this section is very similar to the one in Section 3. The only difference is that, in principle, the attacker may now detect defense units next to both targets. In practice, this means that Table 2(a) should have a forth column entitled "units detected at A and B," which would result in 16 distinct pure policies. It turns out that not detecting any unit or detecting units next to both targets is essentially the same. Because of this we shall consider for this game only the 8 policies in Table 2(a), with the understanding that when units are detected next to both targets, the attacker acts as if no units were detected. It is not hard to prove that this introduces no loss of generality. The defender's policies are the same as in Section 3, and are given in Table 2(b).

This game can also be represented in extensive form by an 8×10 matrix. To compute the entries of the matrix G, let us define the events $\{\emptyset, \mathfrak{A}, \mathfrak{B}\}$ as follows: [2]

$$\mathfrak{A} := \mathfrak{D}_A \cap \bar{\mathfrak{D}}_B \qquad \mathfrak{B} := \bar{\mathfrak{D}}_A \cap \mathfrak{D}_B \qquad \emptyset := \mathfrak{A} \cap \mathfrak{B}.$$

Let $p_j(o)$ denote the probability of the event $o \in \{\emptyset, \mathfrak{A}, \mathfrak{B}\}$ given the defender chooses the policy δ_j, $j = 1, 2, \cdots, 10$. These probabilities can be computed as

$$p_j(\mathfrak{A}) = \chi(n_A, s_A)(1 - \chi(n_B, s_B)) \qquad p_j(\mathfrak{B}) = \chi(n_B, s_B)(1 - \chi(n_A, s_A))$$
$$p_j(\emptyset) = 1 - p_j(\mathfrak{A}) - p_j(\mathfrak{B})$$

It is then straightforward to show that

$$g_{ij} = \sum_o c_{ij}(o) p_j(o) \tag{8}$$

$i = 1, 2, \cdots, 8$ and $j = 1, 2, \cdots, 10$ where $c_{ij}(o)$ is the cost incurred when the players choose policies α_i and δ_j given the event $o \in \{\emptyset, \mathfrak{A}, \mathfrak{B}\}$. The value of $c_{ij}(o)$ is equal to the cost c_k where $k \in \{0, 1, 2, 3\}$ is the number of units at the target selected by the policy α_i, given that the defender chooses policy δ_j. As before the game has Nash equilibria in mixed policies, but now the policies of the two players depend on the values of $\chi(n, s)$.

4.1 Finding a Nash Equilibrium with the Reduced of Number of Policies

To make the computation of the Nash Equilibrium simpler, we will reduce the size of the matrix of the game, using the symmetry of the policies. Then we will find a Nash Equilibrium for the reduced game, and construct a solution for the actual game. We will show that this solution is in fact a Nash Equilibrium for the actual game.

Using the intuitive notion that the optimal policies will be symmetric, we will consider three symmetric policies for the attacker: $\bar{\alpha}_1$ that randomizes equally between the blind policies α_1 and α_2; $\bar{\alpha}_2$ that randomizes equally between the naive policies α_3 and α_4; $\bar{\alpha}_3$ that randomizes equally between the counter-deception policies α_5 and α_6.

As for the defender we will consider five symmetric policies: $\bar{\delta}_1$ that randomizes equally between 3-0 no-information policies δ_1 and δ_2; $\bar{\delta}_2$ that randomizes equally between 2-1 no-information policies δ_3 and δ_4; $\bar{\delta}_3$ that randomizes equally between 3-0 disclosure policies, δ_5 and δ_6; $\bar{\delta}_4$ that randomizes equally between 2-1 disclosure policies δ_7 and δ_8; and $\bar{\delta}_5$ that randomizes equally between 2-1 deception policies δ_9 and δ_{10}.

[2] Given an event $\mathfrak{D}$, we denote by $\bar{\mathfrak{D}}$ the complement of $\mathfrak{D}$

Now the game can be represented by a matrix $G_R \in \mathbb{R}^{3\times 5}$:

$$G_R := \begin{array}{c} \\ \\ \\ \\ \end{array} \begin{array}{ccccc} \bar\delta_1 & \bar\delta_2 & \bar\delta_3 & \bar\delta_4 & \bar\delta_5 \end{array}$$

$$G_R := \begin{array}{c|c} \begin{array}{ccccc} \bar\delta_1 & \bar\delta_2 & \bar\delta_3 & \bar\delta_4 & \bar\delta_5 \end{array} & \\ \begin{bmatrix} \frac{1}{2} & c & \frac{1}{2} & c & c \\ \frac{1}{2}-\frac{k_1}{2} & c-bk_2 & \frac{1}{2}-\frac{k_3}{2} & c-bk_4 & c-bk_5 \\ \frac{1}{2}+\frac{k_1}{2} & c+bk_2 & \frac{1}{2}+\frac{k_3}{2} & c+bk_4 & c+bk_5 \end{bmatrix} & \begin{array}{c} \bar\alpha_1 \\ \bar\alpha_2 \\ \bar\alpha_3 \end{array} \end{array} \tag{9}$$

where

$$\begin{aligned} k_1 &= \chi(3,0) \geq 0, & k_2 &= \chi(2,0) - \chi(1,0) \geq 0, \\ k_3 &= \chi(3,1) \geq 0, & k_4 &= \chi(2,1) - \chi(1,0) \geq 0, \\ k_5 &= \chi(2,0) - \chi(1,1), & & \\ c &= \frac{c_1 + c_2}{2}, & b &= \frac{c_2 - c_1}{2}. \end{aligned}$$

For the choice of probability functions in (5), we have

$$\begin{aligned} k_1 &= 3q - 3q^2 + q^3, & k_2 &= q - q^2, \\ k_3 &= p + 2q - 2pq - q^2 + pq^2, & k_4 &= p - pq, \\ k_5 &= -p + 2q - q^2. & & \end{aligned}$$

The constants k_1, k_2, k_3 and k_4 are nonnegative $\forall\ 0 \leq q \leq p \leq 1$. Depending on the values of the parameters, we have three different cases for the Nash Equilibrium solution:

1. When the sensors are reliable, (i.e. $\chi(2,0) \geq \chi(1,1)$), row 2 dominates [3] over rows 1 and 3. When there is a dominating row, the policies corresponding to the dominated rows always have zero probability in the Nash Equilibrium solution[21]. This means that the policy corresponding to row 2 is always more advantageous than the policies corresponding to rows 1 and 3 for the attacker. Therefore the attacker can always trust its observations and play *naive*. The defender will choose the policy which gives the largest cost against a naive policy. Since

$$k_3 \geq k_1 \geq k_2 \geq k_5, \qquad\qquad k_4 \geq k_2 \geq k_5 \tag{10}$$

and $0 \leq c_2 - c_1 \leq 1$, the largest of the elements in row 2 is $\max\{\frac{1}{2} - \frac{k_1}{2}, c - bk_5\}$. Therefore, the attacker chooses either the 3-0 no information, or the 2-1 deception policy depending on the values of the parameters. So the Nash equilibrium for the reduced game, when $\chi(2,0) \geq \chi(1,1)$, is given by

$$\bar{a}^* := \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}', \qquad \bar{d}^* := \begin{cases} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}' & \frac{1}{2} - \frac{k_1}{2} \geq c - bk_5 \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}' & \frac{1}{2} - \frac{k_1}{2} < c - bk_5 \end{cases} \tag{11}$$

with the value of the reduced game equal to

$$\begin{aligned} \bar{a}^{*\prime} \bar{G}_R \bar{d}^* &= \max\left\{\frac{1}{2} - \frac{k_1}{2}, c - bk_5\right\} \\ &= \max\left\{\frac{(1-\chi(3,0)}{2}, \frac{c_1 + c_2}{2} - \frac{e_1}{2}\right\}, \end{aligned}$$

where $e_1 = (c_2 - c_1)(\chi(2,0) - \chi(1,1))$.

[3] In a matrix $A \in \mathbb{R}^{m\times n}$, row i dominates row k, if $a_{ij} \leq a_{kj}$, $\forall j = 1, 2, \ldots, n$ with a strict inequality holding for at least one j. Rows i and k are called *dominating* and *dominated rows*, respectively.

2. When $\chi(2,0) < \chi(1,1)$, and $c_1 + c_2 \geq 1$, straightforward computation along the lines of [21, Sect. 2.3] gives one of the Nash Equilibrium solutions as

$$\bar{a}^* := \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}', \qquad \bar{d}^* := \begin{bmatrix} 0 & e_2 & 0 & 0 & 1-e_2 \end{bmatrix}', \tag{12}$$

where

$$e_2 := \frac{-k_5}{k_2 - k_5} = \frac{\chi(1,1) - \chi(2,0)}{\chi(1,1) - \chi(1,0)},$$

with the value of the game equal to

$$\bar{a^*}'\bar{G}_R\bar{d}^* = c = \frac{c_1 + c_2}{2}.$$

This means that the attacker ignores the sensor information, and plays *blind*; whereas the defender randomizes among 2-1 no information, and deception policies with probabilities e_2 and $1 - e_2$, respectively.

3. When $\chi(2,0) < \chi(1,1)$ and $c_1 + c_2 < 1$, one Nash Equilibrium solution is given by

$$\bar{a}^* := \begin{bmatrix} 1-e_3 & e_3 & 0 \end{bmatrix}', \qquad \bar{d}^* := \begin{bmatrix} e_4 & 0 & 0 & 0 & 1-e_4 \end{bmatrix}', \tag{13}$$

where

$$e_3 := \frac{c - \frac{1}{2}}{bk_5 - \frac{k_1}{2}} = \frac{1 - c_1 - c_2}{\chi(3,0) - e_1}, \qquad e_4 := \frac{bk_5}{bk_5 - \frac{k_1}{2}} = \frac{e_1}{e_1 - \chi(3,0)},$$

with the value of the game equal to

$$\begin{aligned} \bar{a^*}'\bar{G}_R\bar{d}^* &= \frac{ck_1 - bk_5}{(k_1 - 2bk_5)} \\ &= \frac{(c_1+c_2)\chi(3,0) - e_1}{2(\chi(3,0) - e_1)}. \end{aligned}$$

Therefore, the attacker randomizes among blind and naive policies with probabilities $1 - e_3$, and e_3, whereas the defender randomizes among 3-0 no information and 2-1 deception policies with probabilities e_4, and $1 - e_4$.

4.2 Computing the Nash Equilibrium for the Actual Set of Policies

Using a Nash Equilibrium solution for the reduced matrix, we can construct a Nash Equilibrium solution for the actual set of policies.

Theorem 1. *Let*

$$\bar{a}^* = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}' \qquad \bar{d}^* = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 & d_5 \end{bmatrix}'$$

be a Nash Equilibrium solution for $\bar{G}_R \in \mathbb{R}^{3\times 5}$ *as computed in* (11), (12), *and* (13). *Then,*

$$a^* = \begin{bmatrix} \frac{a_1}{2} & \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_2}{2} & \frac{a_3}{2} & \frac{a_3}{2} & 0 & 0 \end{bmatrix}' \qquad d^* = \begin{bmatrix} \frac{d_1}{2} & \frac{d_1}{2} & \frac{d_2}{2} & \frac{d_2}{2} & \frac{d_3}{2} & \frac{d_3}{2} & \frac{d_4}{2} & \frac{d_4}{2} & \frac{d_5}{2} & \frac{d_5}{2} \end{bmatrix}'$$

is a Nash Equilibrium solution for $G \in \mathbb{R}^{8\times 10}$.

Proof of Theorem 1. We will show below that the proposed pair (a^*, d^*) satisfies

$$a^{*\prime}Gd \leq a^{*\prime}Gd^* \leq a'Gd^*, \qquad \forall a, d. \tag{14}$$

and is therefore a Nash Equilibrium for G. We consider the same three cases given in Section 2.

1. For $\chi(2,0) \geq \chi(1,1)$, the proposed policies for the attacker and the defender are

$$a^* := \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}', \qquad d^* := \begin{cases} \left[\, \frac{1}{2}\ \frac{1}{2}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \,\right]' & \frac{1}{2} - \frac{k_1}{2} \geq c - bk_5 \\ \left[\, 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \frac{1}{2}\ \frac{1}{2} \,\right]' & \frac{1}{2} - \frac{k_1}{2} < c - bk_5 \end{cases} \tag{15}$$

and straightforward computation gives

$$\begin{aligned} a^{*\prime} G d^* &= \max\{\frac{1}{2} - \frac{k_1}{2}, c - bk_5\} \\ a^{*\prime} G &= \begin{bmatrix} \frac{1}{2} - \frac{k_1}{2} & \frac{1}{2} - \frac{k_1}{2} & c - bk_2 & c - bk_2 & \frac{1}{2} - \frac{k_3}{2} & \frac{1}{2} - \frac{k_3}{2} & c - bk_4 & c - bk_4 & c - bk_5 & c - bk_5 \end{bmatrix} \\ G d^* &= \begin{cases} \left[\, \frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}-\frac{k_1}{2}\ \frac{1}{2}-\frac{k_1}{2}\ \frac{1}{2}+\frac{k_1}{2}\ \frac{1}{2}+\frac{k_1}{2}\ \frac{1}{2}\ \frac{1}{2} \,\right] & \frac{1}{2} - \frac{k_1}{2} \geq c - bk_5 \\ \left[\, c\ c\ c-bk_5\ c-bk_5\ c+bk_5\ c+bk_5\ c\ c \,\right] & \frac{1}{2} - \frac{k_1}{2} < c - bk_5 \end{cases} \end{aligned}$$

Since we have (10),

$$\begin{aligned} a^{*\prime} G d &\leq \max\{\frac{1}{2} - \frac{k_1}{2}, c - bk_5\} = a^{*\prime} G d^* \\ a' G d^* &\geq \max\{\frac{1}{2} - \frac{k_1}{2}, a - bk_5\} = a^{*\prime} G d^* \end{aligned}$$

Since the policies in (15) satisfy (14), they form a Nash Equilibrium pair.

2. For $\chi(2,0) < \chi(1,1)$ and $c_1 + c_2 \geq 1$, the proposed policies for the attacker, and the defender are

$$a^* := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}' \tag{16}$$

$$d^* := \begin{bmatrix} 0 & 0 & \frac{e_1}{2} & \frac{e_1}{2} & 0 & 0 & 0 & 0 & \frac{1-e_1}{2} & \frac{1-e_1}{2} \end{bmatrix}' \tag{17}$$

and straightforward computation gives

$$\begin{aligned} a^{*\prime} G d^* &= c. \\ a^{*\prime} G &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & c & c & \frac{1}{2} & \frac{1}{2} & c & c & c & c \end{bmatrix} \\ G d^* &= \begin{bmatrix} c & c & c & c & c & c & c & c \end{bmatrix}' \end{aligned}$$

therefore we have

$$\begin{aligned} a^{*\prime} G d &\leq c = a^{*\prime} G d^* \\ a' G d^* &\geq c = a^{*\prime} G d^* \end{aligned}$$

and (16) and (17) satisfy (14), and form a Nash Equilibrium pair.

3. For $\chi(2,0) < \chi(1,1)$ and $c_1 + c_2 < 1$, the proposed policies for the attacker and the defender are

$$a^* := \begin{bmatrix} \frac{1-e_2}{2} & \frac{1-e_2}{2} & \frac{e_2}{2} & \frac{e_2}{2} & 0 & 0 & 0 & 0 \end{bmatrix}', \tag{18}$$

$$d^* := \begin{bmatrix} \frac{e_3}{2} & \frac{e_3}{2} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-e_3}{2} & \frac{1-e_3}{2} \end{bmatrix}' \tag{19}$$

By straightforward computation we get

$$\begin{aligned} a^{*\prime} G d^* &= \frac{ck_1 - bk_5}{k_1 - 2bk_5} \\ a^{*\prime} G &= \left[\, \tfrac{1}{2} - e_2 \tfrac{k_1}{2}\ \ \tfrac{1}{2} - e_2 \tfrac{k_1}{2}\ \ c - e_2 bk_2\ \ c - e_2 bk_2\ \ \tfrac{1}{2} - e_2 \tfrac{k_3}{2}\ \ \tfrac{1}{2} - e_2 \tfrac{k_3}{2}\ \ c - e_2 bk_4\ \ c - e_2 bk_4\ \ c - e_2 bk_5\ \ c - e_2 bk_5 \,\right] \\ G d^* &= \begin{bmatrix} \frac{e_3}{2} - c(1-e_3) & \frac{e_3}{2} - c(1-e_3) & \frac{e_3}{2} - c(1-e_3) & \frac{e_3}{2} - c(1-e_3) \end{bmatrix}' \end{aligned}$$

therefore we have

$$a^{*\prime}Gd \le \max\{\frac{1}{2} - e_2\frac{k_1}{2}, c - e_2bk_5\}$$

In fact, one can show that,

$$\frac{1}{2} - e_2\frac{k_1}{2} = c - e_2bk_5 = \frac{ck_1 - bk_5}{k_1 - 2bk_5}.$$

which leads to

$$a^{*\prime}Gd \le \frac{ck_1 - bk_5}{k_1 - 2bk_5}$$

On the defender's side, we have

$$a'Gd^* \ge \frac{e_3}{2} - c(1 - e_3) = \frac{ck_1 - bk_5}{k_1 - 2bk_5} = a^{*\prime}Gd^*$$

So (18) and (19) satisfy (14), and form a Nash equilibrium pair for the actual game. ■

Having computed the Nash Equilibrium solutions for the matrix, we can conclude the following:

1. When the sensors are reliable (i.e., $\chi(2,0) \ge \chi(1,1)$), the attacker randomizes among its naive policies and the defender either randomizes among the deception policies or the 3-0 no-information configurations. The latter occurs when the attacker only incurs in significant cost when 3 units are in its path and therefore 2-1 configurations are not acceptable for the defender. The value of the game is
$$a^{*\prime}Gd^* = \frac{\max\{1 - \chi(3,0), c_1 + c_2 - e_1\}}{2} \le \frac{c_1 + c_2}{2}.$$
This value for the cost is smaller than the one obtained in the previous two games, making it more favorable to the attacker, which is now able to take advantage of the surveillance information.

2. When the sensors are not reliable (i.e., $\chi(2,0) < \chi(1,1)$) and $c_1 + c_2 \ge 1$, the attacker randomizes among its blind policies and the defender randomizes between deception and no-information in 2-1 configurations. The probability distribution used by the defender is a function of the several parameters. However, the value of the game is always
$$a^{*\prime}Gd^* = \frac{c_1 + c_2}{2}.$$
This means that the surveillance sensors of the attacker are effectively rendered useless by the defender's policy. This happens because the sensors are not reliable and therefore the defender can significantly manipulate the information available to the attacker. For the characteristic function of the sensors in (5), this occurs when the probability p of detecting a unit that is shown is significantly large when compared to the probability q of detecting a unit that is hidden. It is interesting to note that the region of the (p, q) parameter space where this happens is actually quite large (cf. Figure 1). This means that such situations are likely to occur in practice.

3. When the sensors are not reliable (i.e., $\chi(2,0) < \chi(1,1)$) and $c_1 + c_2 < 1$, the attacker randomizes between its blind and naive policies, whereas the defender randomizes between deception and no-information in 3-0 configurations. The value of the game is
$$a^{*\prime}Gd^* = \frac{1}{2} - \frac{(1 - c_1 - c_2)\chi(3,0)}{2(\chi(3,0) - e_1)} \le \frac{1}{2}.$$
Therefore, the attacker can attain a cost smaller than $\frac{1}{2}$ which would be obtained by only using blind policies.

Note that both in cases 2 and 3 the sensors are not reliable and the defender has sufficient power to manipulate the information available to the attacker so as to make it effectively useless. However, in case 3, the 2-1 configurations required for deception are very costly to the defender and deception is no longer a very attractive alternative.

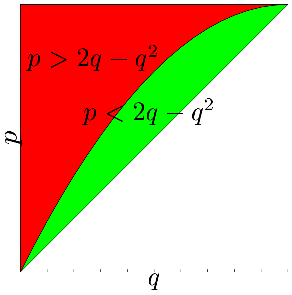

Figure 1: (p, q) Parameter Space

5 Conclusions

We demonstrated that, when one of the players in a competitive game can manipulate the information available to its opponents, deception can be used to increase the player's payoff. We showed that an intelligent player can effectively render the information available to its opponent useless by carefully using deception. This study was carried out for a prototype problem in asset distribution in military operations. The ideas presented here can be applied to devise optimal strategies that use and counteract deception in many other problems. This is the subject of our current research.

References

[1] R. Bellman, *Dynamic Programming.* Princeton, NJ: Princeton University Press, 1957.

[2] S. Sorin and S. Zamir, ""Big Match" with lack of information on one side (III)," in Raghavan *et al.* [22], pp. 101–112.

[3] C. Melolidakis, "Stochastic games with lack of information on one side and positive stop probabilities," in Raghavan *et al.* [22], pp. 113–126.

[4] G. Kimeldorf, "Duels: An overview," in *Mathematics of Conflict* (M. Shubik, ed.), pp. 55–72, Amsterdam: North-Holland, 1983.

[5] P. Bernhard, A.-L. Colomb, and G. P. Papavassilopoulos, "Rabbit and hunter game: Two discrete stochastic formulations," *Comput. Math. Applic.*, vol. 13, no. 1–3, pp. 205–225, 1987.

[6] G. J. Olsder and G. P. Papavassilopoulos, "About when to use a searchlight," *Journal of Mathematical Analysis and Applications*, vol. 136, pp. 466–478, 1988.

[7] G. J. Olsder and G. P. Papavassilopoulos, "A Markov chain game with dynamic information," *Journal of Optimization Theory and Applications*, vol. 59, pp. 467–486, Dec. 1988.

[8] I. Greenberg, "The effect of deception on optimal decisions," *Operations Research Lett.*, vol. 1, pp. 144–147, Sept. 1982.

[9] S. Matsubara and M. Yokoo, "Negotiations with inaccurate payoff values," in *Proc. of the Int. Conf. on Multi-agent Syst.*, pp. 449–50, 1998.

[10] D. B. Hausch, "Multi-object auctions: Sequential vs. simultaneous sales," *Management Science*, vol. 32, pp. 1599–1610, Dec. 1986.

[11] Y. Yavin, "Pursuit-evasion differential games with deception or interrupted observation," *Computers and Mathematics with Applications*, vol. 13, no. 1–3, pp. 191–203, 1987.

[12] J. P. Hespanha, M. Prandini, and S. Sastry, "Probabilistic pursuit-evasion games: A one-step Nash approach," in *Proc. of the 39th Conf. on Decision and Contr.*, Dec. 2000.

[13] C. Reimold, "Games of deception," in *Proc. of the Int. Professional Communication Conf.*, pp. 96–101, 1989.

[14] D. Billings, D. Papp, J. Schaeffer, and D. Szafrone, "Poker as a testbed for machine intelligence research," in *Advances in Artificial Intell.: Proc. of the 12th Biennial Conf. of the Canadian Soc. for Computational Studies*, pp. 228–238, 1998.

[15] M. Vego, "Soviet naval electronic warfare," *Signal*, vol. 44, no. 4, pp. 96–99, 1989.

[16] C. L. C. Oliveira, M. A. Grivet, and E. R. Pantoja, "Radar-ECM simulation system," in *Proc. of the Int. Microwave Conf.*, vol. 1, pp. 255–260, 1993.

[17] M. Tambe, "Recursive agent and agent-group tracking in real-time dynamic environment," in *Proc. of the Int. Conf. on Multi-agent Syst.*, pp. 368–375, 1995.

[18] S. Z. Yanglin, "Identification of false targets in bistatic radar system," in *Proc. of the IEEE National Aerospace and Electronics Conf.*, vol. 2, pp. 878–883, July 1997.

[19] R. W. Burns, "Deception, technology and the D-day invasion," *Engineering Science and Education J.*, vol. 4, pp. 81–88, Apr. 1995.

[20] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.

[21] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* No. 23 in Classics in Applied Mathematics, Philadelphia: SIAM, 2nd ed., 1999.

[22] T. E. S. Raghavan, T. S. Ferguson, and T. Parthasarathy, eds., *Stochastic Games and Related Topics: In Honor of Professor L. S. Shapley*, vol. 7 of *Theory and Decision Library, Series C, Game Theory, Mathematical Programming and Operations Research.* Dordrecht: Kluwer Academic Publishers, 1991.