Optimality of the Greedy Shooting Strategy in the Presence of Incomplete Damage Information

Gad Manor,[1,2] **Moshe Kress**[3,4]
[1] *Center for Systems Analysis, IDF, Israel*

[2] *Faculty of Industrial Engineering and Management, Technion, Haifa, Israel*

[3] *Center for Military Analyses, P. O. Box 2250 (T1) Haifa 31021, Israel*

[4] *Department of Industrial and Manufacturing Engineering, University of Rhode Island, Kingston, Rhode Island 02881, USA*

Received August 1995; revised May 1997; accepted 14 May 1997

Abstract: Consider a situation where a single shooter engages, sequentially, a cluster of targets that may vary in terms of vulnerability and value or worth. Following the shooting of a round of fire at a certain target, the latter may either be killed or remain alive. We assume neither partial nor cumulative damage. If the target is killed, there is a possibility that the shooter is not aware of that fact and may keep on engaging that target. If the shooter recognizes a killed target as such, then this target is considered to be *evidently killed.* If the objective is to maximize the weighted expected number of killed targets, where the weight reflects the value of a target, then it is shown that a certain type of a shooting strategy, called a *Greedy Strategy,* is optimal under the general assumption that the more a target is engaged, but still not evidently killed, the less is the probability that the next round will be effective. If all weights are equal, then the greedy shooting strategy calls to engage, at each round, the least previously engaged target that is not evidently killed. © 1997 John Wiley & Sons, Inc. Naval Research Logistics **44:** 613–622, 1997

Keywords: optimal engagement; damage assessment; greedy strategy

1. INTRODUCTION

In the past few decades there has been a significant and clear trend towards tactical weapon systems with longer range that deliver more accurate and very expensive munitions. Precision Guided Missiles (PGMs) are typical examples. Consequently, the role of damage assessment has become an important factor in the engagement process. On the one hand, the shooter wishes to utilize his expensive munitions in a most efficient way and to minimize the number of ‘‘repeat kills’’ or *multikills.* On the other hand, in the absence of supporting

Correspondence to: M. Kress

 CCC 0894-069X/97/070613-10

surveillance systems, the capability to determine the damage that is inflicted to the target is reduced due to the long range of the weapon system and the lack of line-to-site. This handicap is prevalent in Precision Guided Missiles (PGMs) such as FOG-M.

FOG-M (Fiber Optics Guided Missile) is a weapon system that links an electrooptical sensor that is mounted on the missile, with an operator on the ground, via an optical fiber. This fiber-optics technology allows the shooter to engage the target from non-line-to-site positions, and to continuously watch it through the electrooptical sensor of the approaching missile. The special combination of homing device, control system, and type of warhead enables the shooter to concentrate his attention and to engage only one intended target at the time (hence, *Precision* Guided Missiles) with no effect on other targets. When the missile hits an object, the transmitted image disappears upon impact. Thus, while the shooter can see the area of the point of impact, prior to the time of impact, he cannot determine with certainty if the target was actually hit and killed. There are several sources (such as remote sensors) according which the hit result can be assessed in real time, but due to their limited reliability, it is only under rigid conditions and with the presence of very strong evidence that a target will be declared by the shooter as killed. This embedded ''risk aversion'' leads to situations where a killed target may not be detected as such, and therefore it will still be considered as a live target for future engagements (type-I error). On the other hand, it is safe to assume that no type-II error (declaring a target as killed while it is still alive) can occur.

Consider a single shooter whose objective is to engage a cluster of static targets using limited amount of ammunition. By the term *static* we simply mean that no new target appears and no target that was initially in the cluster disappears throughout the engagement. The shooter, who, we assume, can identify the targets (say, by their relative location in the cluster), selects for each round of fire (say, missile) a target to shoot. A round of fire may either hit and kill the target or cause no damage. In other words, neither partial damage nor cumulative damage is assumed, and a hit always results in a kill. The targets are not necessarily homogeneous, and each one of the targets may have different value or worth for the shooter (e.g., a command post may have a larger value to the shooter than, say, a truck). The tactical engagement that is discussed here is of a relatively short duration such that it can be assumed that the values of the targets remain unchanged during the entire engagement.

At each time step in the engagement process, the shooter selects a target, shoots a round upon it, and then observes, through the committed surveillance systems or other sensors, the result of this shooting on the certain target that was shot upon. Hence, a target is observed, at a certain time step, only if it was shot upon in that time step. If a target is killed by the round that was shot upon, then the shooter is either aware of that fact—in which case the target is said to be *evidently killed* (EK)—or the shooter is oblivious of that fact—in which case he may keep on engaging it (not necessarily continuously) until the target is EK or until he is out of ammunition. Thus, it is possible that ''repeat kills'' or *multikills* of a certain target will occur prior to declaring it evidently killed. As mentioned above, no type-II error is assumed.

The objective of the shooter is to maximize the expected weighted number of killed targets where the weights represent the values of the various targets. According to that objective, the shooter has to decide, for each round, which target to engage. His decision depends on the information available to him at that moment which includes the list of EK targets, the number of rounds delivered upon each target so far and the amount of ammunition that is left.

The problem of shoot-look-shoot strategies in the presence of incomplete damage information was first introduced by Aviv and Kress [2] where several shooting strategies were defined, analyzed, and compared with regard to typical engagement scenarios. This problem is a special case of general fire-allocation problems which were treated and analyzed in the past in the context of strategic defence [3], one-on-many fire allocation [4, 6], and air interdiction [1].

In the engagement sequence of a target, a certain round is defined to be *effective* if it killed a *live* target. In other words, a round is said to be *ineffective* with respect to a given target if it either caused no damage to the target or if the target had already been killed. Clearly, for each target, its engagement sequence may contain at most one effective round.

In this note we prove that a certain type of a shooting strategy—a *Greedy Shooting Strategy* (*GSS*)—is optimal with respect to the above-mentioned objective, under the assumption of fixed kill and damage assessment probabilities. It is shown that the result is true in general in the case where the conditional kill probabilities form a monotone nonincreasing sequence. That is, the more a target is engaged (and not EK), the less is the probability that the next round is effective. This assumption is applicable in real combat situations such as when the damage assessment capability of the shooter is low (in which case he may keep on engaging a killed target) or if the target can take protective measures and reduce its vulnerability by, say, digging in.

The proof of the optimality of GSS follows the ideas that are presented by Mandelbaum [7] in the analysis of the Multiarm Bandit problem. It is shown that the original shoot-look-shoot problem is equivalent, in terms of the expected number of kills, to a finite horizon deteriorating bandit problem.

If all weights are equal, then GSS amounts to engaging, at each round, the least previously engaged non-EK target.

In Section 2 we present some definitions and notation, and we state formally the assumptions that are made. In Section 3 we prove the optimality of GSS, and in Section 4 an extension of the result is presented, along with two examples.

2. DEFINITIONS, NOTATION, AND ASSUMPTIONS

Let M denote the number of targets in the cluster, and let N denote the number of rounds the shooter can use while engaging that cluster. Throughout the engagement process, each target may be in three possible states: NK (Not Killed), K (Killed but not evidently killed), and EK (Evidently Killed). We may assume, without loss of generality, that no shots towards EK targets are made, so that the effect of a shot on the state of the target may be characterized by one of the three following outcomes:

R_0—No hit. In this case the target remains in its previous state.

R_1—A hit (= kill) with no kill indication. If the target is already in the K state, it remains in that state. If it is in the NK state, then it changes it to K.

R_2—A hit with a kill indication. The target changes its state to EK.

Note that the shooter cannot discriminate between the outcomes R_0 and R_1.

For each target i define a series of random variables $X_i(1)$, $X_i(2)$, . . . , $X_i(N)$, where

$X_i(n) = j$ if the nth round that was shot upon target i resulted in outcome R_j, $j = 0, 1, 2$.

Let p_i denote the single shot hit probability for target i, $i = 1, \ldots, M$. Given a hit, let q_i denote the corresponding conditional probability of detecting a hit (= kill) by the shooter. Evidently, we assume here that these probabilities are constant along the shooting process on a certain target i. This assumption is relaxed later on. Formally,

$$p_i = \Pr[X_i(n) \neq 0], \tag{2.1}$$

$$q_i = \Pr[X_i(n) = 2/X_i(n) \neq 0]. \tag{2.2}$$

We assume that the shooting process comprises of statistically independent random variables, as stated in the following assumption.

ASSUMPTION 2.1: The set of variables $\{X_i(n): i = 1, \ldots, M, n = 1, 2, \ldots, N\}$ is a set of independent random variables.

For each target i two sets of indicator variables are defined: The $Z_i(n)$, $n = 1, 2, \ldots,$ indicate the first hit on target i, and the $U_i(n)$ indicate the (first) EK declaration of target i. Formally,

DEFINITION 2.1: Let $Z_i(0) = U_i(0) = 0$. For $n = 1, 2, \cdots$

$$Z_i(n) = \begin{cases} 1 & \text{if } X_i(k) = 0 \quad \text{for each } k = 0, \ldots, n-1 \quad \text{and} \quad X_i(n) \neq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

$$U_i(n) = \begin{cases} 1 & \text{if } X_i(k) \neq 2 \quad \text{for each } k = 0, \ldots, n-1 \quad \text{and} \quad X_i(n) = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

$Z_i(n)$ represents the effectiveness of the nth round shot at target i, and the vector $\mathbf{U}_i(n) = [U_i(0), \ldots, U_i(n)]$ represents the information that the shooter has about the state of target i after the nth shot upon it.

We define now a strategy as an M-dimensional stochastic process (a vector of M stochastic processes), which prescribes the decisions to be made in a cumulative manner. It states, for each time step t, the cumulative number of rounds to be shot at each of the M targets within the first t time steps.

DEFINITION 2.2: A *strategy* is a stochastic process $T(t) = \{(T_1(t), \ldots, T_M(t)), t = 0, 1, \ldots, N\}$ such that

(1) $T(0) = 0$ with probability 1.

(2) $T(t+1)$ is obtained from $T(t)$ such that $T_i(t+1) = T_i(t) + 1$ for some $i = 1, \ldots, M$, and $T_j(t+1) = T_j(t)$ for all other $j \neq i$.

(3) $T_i(t+1) - T_i(t)$ is determined by $\{(\mathbf{U}_i(T_i(t))), i = 1, \ldots, M\}$.

From Definition 2.2 (3), it is seen that the target to be shot upon is selected depending on the shooting pattern and the shooting results so far. Evidently, the strategy is a function of the past history.

Let W_i denote the reward for killing target i ($i = 1, \ldots, M$). For a given strategy T, let $R_t(T)$ be the reward for the decision made by the strategy at time t, which will be carried out at the $(t + 1)$th step. If the strategy chooses target i, then the next shot is the $(T_i(t) + 1)$th shot towards this target. This shot is effective (actually kills the target) if and only if the indicator $Z_i(T_i(t) + 1)$ equals 1 (see Definition 2.1). The reward for this shot is therefore $W_i Z_i(T_i(t) + 1)$.

Since the chosen target can be identified as the only target i for which $T_i(t + 1) - T_i(t) = 1$, one can easily verify that

$$R_t(T) = \sum_{i=1}^{M} W_i Z_i(T_i(t) + 1)[T_i(t + 1) - T_i(t)]. \tag{2.5}$$

The *total reward* $R(T)$ for strategy T is defined as the sum of the rewards $R_t(T)$ over the engagement process. That is,

$$R(T) = \sum_{t=0}^{N-1} \sum_{i=1}^{M} W_i Z_i(T_i(t) + 1)[T_i(t + 1) - T_i(t)]. \tag{2.6}$$

The objective of the shooter is to find a strategy that maximizes $E[R(T)]$.

3. OPTIMAL STRATEGIES

The above representation of the problem enables us to consider it as a finite-horizon multiarmed bandit problem. The multiarmed bandit problem (e.g., [5], [7], and [8]) is concerned with dynamic allocation of a single resource among a fixed number of projects (arms). The multiarmed bandit consists of a fixed set of M independent stochastic Markovian processes, which are called arms. At each instance on a discrete time scale, one armed is pulled and a random reward (depending on the Markovian state of the arm only) is obtained. The chosen arm changes its state according to a Markov transition rule, while the states of the other arms remain unchanged. The problem is how to choose the arms so as to maximize the expected total reward.

In the classic version of the problem, the rewards are discounted in time by a constant factor, and the time horizon is infinite. The classic version was solved by Gittins and Jones [5], who proved that the optimal strategy coincides with an *index strategy*. In an index strategy, a dynamic function called index is computed on the states of the arms, and the optimal (index) strategy is always to pull an arm with the highest index.

Our problem is a special case of the finite-horizon multiarmed bandit with no discount factor. Each target is represented by an arm and the nth shot at target i corresponds to the nth pull of arm i. The possible (Markovian) states of target i are $\{S_{ij}, j = 0, \ldots, N, \mathrm{EK}_i\}$, where S_{ij} means ''j rounds have been shot upon target i which is still not EK'' and EK_i means ''target i is EK.'' The random reward for pulling an arm i (shooting at target i) in state S_{ij} is W_i with probability a_{ij} and zero with probability $(1 - a_{ij})$, where a_{ij} is the probability that the $(j + 1)$th pull of arm i is successful, given state S_{ij}. The reward for pulling an arm in state EK is zero with probability 1. The independence of the arms is a direct conclusion from Assumption 2.1.

Unlike the infinite-horizon problem, the finite-horizon bandit problem cannot be solved using an index strategy, and therefore other methods, such as dynamic programming, may be used. Fortunately, however, our problem can be modified and be presented as a special case of the finite-horizon multiarm bandit known as the *deteriorating case* (see, e.g. [8]). The simplest version of this special case is prescribed by the two additional properties:

PROPERTY 3.1: The reward series for each arm, that is, the reward that is associated with each pull of the arm, is nonincreasing with probability 1.

PROPERTY 3.2: The reward for the nth pull of arm i is known to the player after the preceding pull [the $(n - 1)$th pull] of the same arm.

Property 3.2 is similar to assumption 2.1B in [7]. Notice that both properties do not hold in our representation of the problem as stated above.

DEFINITION 3.1: A multiarmed bandit problem is said to be *deteriorating bandit* if the above two properties hold.

In the case of the deteriorating bandit, it is trivial and well known that strategies choosing at each step of the game the arm with the highest reward are optimal. Note that such a strategy will always (with probability 1) choose the N highest-reward pulls in decreasing order. Therefore, its optimality continues to hold even when relinquishing the independence of the bandit's arms.

Our solution to this problem is based on a reduction of it to the deteriorating case. We define a new multiarmed bandit problem by replacing the rewards $W_i Z_i(n)$ with new rewards $W_i L_i(n)$, such that the new bandit problem is a deteriorating bandit but with the same expected total reward, as the original problem, for each strategy.

$L_i(n)$ is defined as the conditional expected effectiveness of the nth shot upon target i, given the information that was acquired during the $n - 1$ preceding shots upon that target. Formally,

DEFINITION 3.2:

$$L_i(0) = 0,$$

$$L_i(n) = E[Z_i(n)/\mathbf{U}_i(n-1)]. \tag{3.1}$$

Clearly $L_i(n)$ is a random variable.

The new total reward $L(T)$ is defined similarly to $R(T)$:

$$L(T) = \sum_{t=0}^{N-1} \sum_{i-1}^{M} W_i L_i(T_i(t) + 1)[T_i(t+1) - T_i(t)]. \tag{3.2}$$

Next we show that the newly defined bandit problem is a deteriorating bandit.

DEFINITION 3.3: Let $g_i(n)$ denote the conditional effectiveness of the nth round upon target i, given that the first $n - 1$ rounds resulted in no kill indication. That is,

$$g_i(n) = E[Z_i(n)/\mathbf{U}_i(n-1) = (0, 0, \ldots, 0)]. \tag{3.3}$$

LEMMA 3.1: $g_i(n)$ is a monotone nonincreasing function of n, for all $i = 1, \ldots, M$.

PROOF: From (2.3), (2.4), and (3.3) it follows that

$$g_i(n) = P(Z_i(n) = 1/X_i(1) \neq 2, \ldots, X_i(n-1) \neq 2)$$

or

$$g_i(n) = \frac{\Pr[X_i(1) = 0, \ldots, X_i(n-1) = 0, X_i(n) \neq 0]}{\Pr[X_1(1) \neq 2, \ldots, X_i(n-1) \neq 2]}.$$

From (2.1) and (2.2) we have that

$$g_i(n) = \frac{(1-p_i)^{n-1} p_i}{(1-p_i q_i)^{n-1}}$$

and the monotonicity follows since

$$\frac{(1-p_i)}{(1-p_i q_i)} \leq 1. \qquad \square$$

LEMMA 3.2: The newly defined multiarm bandit problem is a deteriorating bandit.

PROOF: We need to show first that

$$L_i(n+1) \leq L_i(n) \qquad \text{with probability 1 for all } i = 1, \ldots, M \text{ and } n = 1, 2, \ldots.$$

Clearly, if $\mathbf{U}_i(n) = (0, \ldots, 0)$, then $\mathbf{U}_i(n-1) = (0, \ldots, 0)$ too, and it follows that $L_i(n+1) = g_i(n+1)$ and $L_i(n) = g_i(n)$. According to Lemma 3.1, it follows that $L_i(n)$ is monotone nonincreasing.

If $\mathbf{U}_i(n) \neq (0, \ldots, 0)$, then $L_i(n+1) = 0 \leq L_i(n)$. Property 3.2 follows since $L_i(n)$ is defined as a conditional expectation, given the results of the first $n - 1$ rounds upon target i, as given in Definition 3.2. $\square$

We define now a *greedy strategy* as a strategy that always chooses the target with the highest weighted (i.e., multiplied by W_i) conditional expected effectiveness.

DEFINITION 3.4: A strategy $T(t)$ is called *greedy* if and only if for all $t = 0, \ldots, N - 1$ and $i = 1, \ldots, M$, $T_i(t+1) = T_i(t) + 1$ implies that

$$W_i L_i(T_i(t) + 1) = \operatorname*{Max}_{j=1,\ldots,M} W_j L_j(T_j(t) + 1). \tag{3.4}$$

LEMMA 3.3: Let T_g be a greedy strategy, and let T be any strategy; then

$$E[L(T_g)] \geq E[L(T)]. \tag{3.5}$$

PROOF: Since the new bandit is a deteriorating bandit, it is well known [5] that a strategy which chooses at each stage the most valuable arm is optimal. By Definition 3.4, T_g satisfies this condition and thus it is optimal, that is, its expected reward exceeds the expected reward of any other strategy. □

To complete the proof of our result, we need to show now that the expected rewards of the two reward functions—$W_i Z_i(n)$ and $W_i L_i(n)$—are equal.

LEMMA 3.4: For any shooting strategy T, $E\{R(T)\} = E\{L(T)\}$.

PROOF: From the linearity of the expectation operator, the expected value of (2.6) is

$$E[R(T)] = \sum_{t=0}^{N-1} \sum_{i=1}^{M} W_i E[Z_i(T_i(t) + 1)[T_i(t + 1) - T_i(t)]], \tag{3.6}$$

which may be written as

$$E[R(T)] = \sum_{t=0}^{N-1} \sum_{i=1}^{M} W_i E\{E[Z_i(T_i(t) + 1)(T_i(t + 1) - T_i(t)) | \mathbf{U}_i(T_i(t)), \mathbf{U}_j(N) \quad \text{for all } j \neq i]\}.$$

It follows from Definition 2.2(3) that $T_i(t + 1) - T_i(t)$ is determined by $\mathbf{U}_i(T_i(t))$ and $\mathbf{U}_j(T_j(t))$ for all $j \neq i$. In addition, $\mathbf{U}_j(T_j(t))$ is included in $U_j(N)$ for all $j \neq i$, and therefore $T_i(t + 1) - T_i(t)$ is determined by $\mathbf{U}_i(T_i(t))$ and $U_j(N)$ for all $j \neq i$. Consequently, $[T_i(t + 1) - T_i(t)]$ may be removed outside the conditional expectation sign. Hence

$$E[R(T)] = \sum_{t=0}^{N-1} \sum_{i=1}^{M} W_i E\{E[Z_i(T_i(t) + 1) | \mathbf{U}_i(T_i(t)), \quad \mathbf{U}_j(N)) \quad \text{for } j \neq i] \times [T_i(t + 1) - T_i(t)]\}. \tag{3.7}$$

Due to the independence of the targets (arms), Eq. (3.7) may be written as

$$E[R(T)] = \sum_{t=0}^{N-1} \sum_{i=1}^{M} W_i E\{E[Z_i(T_i(t) + 1) | \mathbf{U}_i(T_i(t))][T_i(t + 1) - T_i(t)]\}. \tag{3.8}$$

From Definition 3.2 and the linearity of the expectation operator, we obtain

$$E[R(T)] = E\left\{\sum_{t=0}^{N-1}\sum_{i=1}^{M} W_i L_i(T_i(t)+1)[T_i(t+1)-T_i(t)]\right\} = E[L(T)]. \quad (3.9)$$

□

THEOREM 1: Every greedy strategy is an optimal strategy with regards to the expected reward of a shooting sequence with incomplete damage information.

PROOF: It is sufficient to show that if T_g is a greedy strategy and T is some other strategy, then

$$E(R(T_g)) \geq E(R(T)),$$

but this follows directly from Lemma 3.3 and Lemma 3.4:

$$E(R(T_g)) = E(L(T_g)) \geq E(L(T)) = E(R(T)). \qquad \square$$

Note that Theorem 1 does not imply that greedy strategies are the only optimal strategies. For example, suppose that $N = 10$, $M = 2$, and the two targets are identical in terms of their probability attributes (p, q) and their worth (W). Consider a strategy T_2 in which the shooter always shoots twice on the first target, then twice on the second and then proceeds greedily. It is easy to show that $E[R(T_2)] = E[R(T_g)]$ for some greedy strategy T_g.

4. RELAXATION OF ASSUMPTIONS

We assumed so far that the hit probabilities p_i and the damage detection probabilities q_i remain constant throughout the engagement. We further assumed that the shooting outcomes are independent. It can be easily verified that these assumptions can be relaxed in the following manner:

ASSUMPTION 4.1: The probabilities $p_i(n)$ and $q_i(n)$ may change as a function of n so long as the function $g_i(n)$ is a monotone nonincreasing function.

ASSUMPTION 4.2: The M sets $\{X_1(n); n = 0, 1, \cdots\}, \ldots, \{X_M(n); n = 0, 1, \cdots\}$ are independent. The random variables within each one of the M sets may be dependent.

THEOREM 2: Any greedy strategy is an optimal shooting strategy for the case where Assumptions 4.1 and 4.2 hold.

PROOF: Clearly, Assumption 4.2 is sufficient for the independence requirement of the arms in the multiarm bandit setting. Assumption 4.1 is in fact Lemma 3.1, and therefore Lemma 3.2 and Lemma 3.3 hold. The proof of Lemma 3.4 holds since it does not depend directly on either of the relaxed assumptions. □

Example 1: As the engagement progresses, the shooter becomes more familiar with the battle area and therefore gains knowledge and experience that may improve its capability to detect a hit when such an event occurs. Thus, it is reasonable to assume that $q_i(n)$ are monotone nondecreasing functions of n for all $i = 1, \ldots, M$. This gain in the shooter's battlefield experience may be offset by the protective measures that the targets may apply. Thus it is less obvious what the behavior of p_i is, and therefore we leave the assumption that it is constant. As before, we assume that the rounds are independent. In this case $g_i(n)$ becomes

$$g_i(n) = \frac{(1 - p_i)^{n-1} p_i}{\prod_{j=1}^{n-1} (1 - p_i q_i(j))}$$

and

$$\frac{g_i(n+1)}{g_i(n)} = \frac{(1 - p_i) p_i}{(1 - p_i q_i(n)) p_i} < 1,$$

which means that $g_i(n)$ is monotone nonincreasing. Both Assumption 4.1 and Assumption 4.2 hold, and therefore the greedy shooting strategy applies to this situation too.

Example 2: Consider the situation where the (conditional) probability of a kill indication, given a hit, is set to be either one of two different values q_{i1} or q_{i2}—each with probability 0.5. These values are set to the various targets independently. The hit probability is p_i in both cases, and all shots are independent, given the values of q_i for all the targets. It can be shown that the random variables $X_i(0), X_i(1), \ldots, X_i(N)$ are now dependent but since the values of the kill detection probabilities are set independently, Assumption 4.2 holds. It can be verified that $1/g_i(n)$ is a sum of two nondecreasing functions, and therefore Assumption 4.1 holds too.

REFERENCES

[1] Armitage, J.V., ''Allocation of Sorties in Air Interdiction,'' *Operations Research,* **18,** 483–496 (1970).

[2] Aviv, Y., and Kress, M., ''Evaluating the Effectiveness of Shoot-Look-Shoot Tactics in the Presence of Incomplete Damage Information,'' CEMA Report 94/46, 1994 (to appear in *Military Operations Research* (*MOR*), 1997).

[3] Eckler, A.R., and Barr, S.A., ''Mathematical Models of Target Coverage and Missile Allocation,'' Military Operations Research Society (MORS), 1972.

[4] Friedman, Y., ''Optimal Strategy for the One-Against-Many Battle,'' *Operations Research,* **25,** 884–888 (1977).

[5] Gittins, J.C., and Jones, D.M., ''A Dynamic Allocation Index for the Sequential Design of Experiments,'' in J. Gani, et al. (Eds.), *Progress in Statistics.* North-Holland, Amsterdam, 1974, pp. 241–266.

[6] Kress, M., ''A Two-on-One Stochastic Duel with Maneuvering and Fire Allocation Tactics,'' *Naval Research Logistics,* **38,** 301–313 (1991).

[7] Mandelbaum, A., ''Discrete Multiarmed Bandits and Multiparameter Processes,'' *Probability Theory and Related Field,* **71,** 129–147 (1986).

[8] Weber, R., ''On the Gittins Index for Multi-Armed Bandits,'' *Annals of Applied Probability,* **2**(4), 1024–1035 (1992).