OPTIMAL THRESHOLD POLICIES FOR HARD-KILL OF ENEMY RADARS WITH HIGH SPEED ANTI-RADIATION MISSILES (HARMS)

Arsalan Farrokh and Vikram Krishnamurthy

University of British Columbia
Department of Electrical and Computer Engineering
Vancouver, V6T 1Z4, Canada

ABSTRACT

In modern Network Centric Warfare (NCW) there is a dedicated platform (airplane) assigned to every group of aircraft that specializes in the hard-kill of the enemy guidance-radars by deploying High speed Anti-Radiation Missiles (HARM)s. In this paper we consider the problem of optimal launch control of the HARMs. We formulate the optimal trade-off between the cost of the HARMs and the latency in performing the hard-kill of the enemy radar as a Partially Observable Markov Decision Process (POMDP). Next, by reformulating this POMDP as a Markovian search problem, we prove that optimal missile launch control policies are threshold-based policies in nature. We then present optimal threshold policies that unlike their POMDP counterparts are computationally efficient and inexpensive to implement in real time combat systems. Numerical results demonstrate the effectiveness of these threshold based missile deployment algorithms.

1. INTRODUCTION

The primarily mission of an airplane such as EA-6B Prowler in the Missile Engagement Zone (MEZ) is to protect the accompanying aircraft by neutralizing the enemy Surface-to Air Missile (SAM) guidance-radars [1]. The Prowler is able to perform the "hard-kill" of enemy radars by destroying them with the High speed Anti-radiation Missile (HARM).

The HARM destroys the SAM guidance-radar by following the radar emission to its source. One common tactic by the enemy is launching the SAM without radar guidance to prevent the radar from being detected by the HARM. If the SAM guidance-radar is off, the launched HARM will not be able to track it and will explode in the air without hitting the target. However, with the radar off, it is less likely that the SAM hits the strike aircraft. In this view, the enemy may choose to activate and deactivate the radar randomly to enable the SAM guidance for some duration of time while hiding the radar from the HARM for another duration of time. The state of the enemy radar cannot be directly observed and the only information available is whether the hard-kill is successful or not. Associated with each HARM launch, there is a cost that represents the price of a HARM and limited resources of the system. On the other hand, the latency to perform a successful hard-kill incurs a cost representing the increased threat to the incoming strike fighters. Therefore, there is strong motivation to devise a novel launch control algorithm for attempting or suspending the hard-kill to minimize the cost up to the successful destruction of the enemy radar.

Main results: The main results of this paper are organized as follows:

(i) In Section 2, we introduce a stochastic dynamical model of the HARM engagement. We assume the Prowler decisions as whether to launch the HARM or suspend the attack are made at discrete times and the cost at each time is only dependent on the action taken for that time. The dynamics of the enemy SAM guidance-radar, i.e. the switching between On and Off states, is assumed to be a two-state Markov chain.

(ii) In Section 3, we use the model in Section 2 to formulate the launch control problem as an optimal search problem with POMDP framework. We show that in our case, the missile launch problem can be formulated as a search problem with a special structure described in [2], [3] for which the threshold policies are optimal.

(iii) In Section 4, we adopt the results of the Markovian search problem in [2] to present the optimal threshold policies for the launch control problem. We show that depending on the system parameters, the launch control systems can be categorized into three different Attack profiles. The optimal policy for each system has a different threshold level.

(iv) In Section 5, we use numerical examples to demonstrate the performance improvement that can be obtained by applying the optimal threshold policies as compared to heuristic algorithms.

Literature Review: Several papers consider the search problem formulation used in this paper. Ross [3] first conjectured the existence of threshold policies for this search problem. Recently, MacPhee and Jordan [2] prove the Ross' conjecture for an overwhelming proportion of transition rules and system parameters.

2. HARM ENGAGEMENT MODEL

In this section a stochastic model describing the HARM engagement scenario is presented. We outline our engagement model by the following five elements:

(i) Time: The time axis is divided into slots of equal duration denoted by ΔT. By convention, by time k, $k \in Z_+$, we mean the k^{th} time interval (corresponding to the physical time interval $[k\Delta T, (k+1)\Delta T)$). Here, Z_+ is the set of non-negative integers. We assume that the attack or suspend decisions by the Prowler plane are made in times $k \in Z_+$.

(ii) Markovian target: Target is the enemy SAM guidance-radar. As long as the radar is not destroyed by the Prowler plane, we refer to it as an *alive* radar. Assume the alive enemy radar is being activated (On state) and deactivated (Off) based on a two state Markov chain. Define the alive radar state space as $S = \{\text{On} := 1, \text{Off} := 2\}$. Also let $s_k \in S$ be the state of an alive enemy radar at time k. We assume s_k is a two-state irreducible Markov chain with the transition matrix A, where:

$$A = \begin{bmatrix} a & 1-a \\ 1-d & d \end{bmatrix} \tag{1}$$

Here, $a < 1$ is the probability that an active radar remains in the "On" state and $d < 1$ is the probability that a deactivated radar (hidden from the HARM) remains in the "Off" state. Note that all of the above probabilities are conditional on radar being alive.

(iii) Actions: At each time $k \in Z_+$, the decision is made whether to launch the HARM or suspend the attack. Define the action space U as: $U = \{\text{Launch the HARM} := \text{La}, \text{Suspend the Attack} := \text{Su}\}$. Also, define u_k as the action taken by the Prowler (controller) at time k.

(iv) We introduce a new "terminal" state $\mathcal{D}$ to represent the absorbing state of a "Destroyed" radar. The complete state space of the engagement dynamic can then be obtained by augmenting the radar state space S with this new terminal state. Define χ to be the system state space: $\chi = \{\text{On} := 1, \text{Off} := 2, \mathcal{D} := 3\}$. Let x_k be the system state at time k. We assume that if the enemy radar is off ($x_k = 2$) at time k or there is no hard-kill attempt at time k, the radar will remain alive at time $k+1$. On the other hand, if at time k the enemy radar is "On" ($x_k = 1$) and the HARM is launched ($u_k = \text{La}$), then the HARM will destroy the enemy radar with probability p_d which represents the precision of the attack. Consequently, $x_k \in \chi$ evolves according to a *controlled Markov Process* with the 3×3 transition probability matrix $R(u)$. Here, $u \in U$ denotes the action applied at the transition . Let $r_{ij}(u)$ be the ijth element of $R(u)$ so that when action u is applied, the system moves from state i to state j with probability $r_{ij}(u) = \mathbf{P}(x_{k+1} = j | x_k = i,\, u_k = u), \quad 1 \le i, j \le 3$. We can evaluate $r_{ij}(u)$ by partitioning the sample space into the events "radar destroyed" and "radar alive" for $u \in U$ and $1 \le i, j \le 3$. This gives (A is radar transition matrix):

$$R(\text{Su}) = \begin{bmatrix} A & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{1\times 2} & 1 \end{bmatrix}, \tag{2}$$

$$R(\text{La}) = \begin{bmatrix} a(1-p_d) & (1-a)(1-p_d) & p_d \\ 1-d & d & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

(iv) Observations: We assume that at any time interval the Prowler is able to observe whether the hard-kill is successful or not. Define $y_k \in \{\text{NAF, AFF}\}$ to be the observation by the Prowler at time k, where "AFF" and "NAF" denote that the hard-kill is "Affirmed" and "Not Affirmed", respectively. We have:

$$\begin{aligned} &\mathbf{P}(y_{k+1} = \text{NAF} | x_k = 1, u_k = \text{La}) = 1 - p_d \\ &\mathbf{P}(y_{k+1} = \text{NAF} | x_k = 2, u_k = \text{La}) = 1 \\ &\mathbf{P}(y_{k+1} = \text{NAF} | x_k = x, u_k = \text{Su}) = 1, \quad x \in \mathcal{X} \\ &\mathbf{P}(y_{k+1} = \text{NAF} | x_k = 3, u_k = u) = 0, \quad u \in U \end{aligned} \tag{4}$$

(v) Cost function: Associated with each HARM launch ($u_k = \text{La}$), there is a cost $c_1 > 0$ (independent of the current state or observation) that represents the price of the HARM, limitations in the system resources, and the risk of fratricide. On the other hand, as long as the enemy radar is functioning, each suspension of the attack also incurs a cost $c_2 > 0$. c_2 represents the latency cost in performing a successful hard-kill of the enemy radar. However, once the enemy radar is destroyed (terminal state $\mathcal{D}$), suspending the attack incurs no more cost. This ensures that the total cost over the infinite number of stages remains finite for all the stationary policies. Assume $g : \chi \times U \to \{c_1, c_2\}$ is a function that maps the states and the actions to the corresponding non-negative costs. We then have:

$$g(x_k, u_k) = \begin{cases} c_1 & , x_k \neq 3, u_k = \text{La} \\ c_2 & , x_k \neq 3, u_k = \text{Su} \\ 0 & , x_k = 3, u_k \in U \end{cases} \tag{5}$$

3. THE MISSILE LAUNCH CONTROL PROBLEM

Here, the goal is to minimize the expected cost up to the successful hard-kill of the enemy radar. In this section, we first formulate this problem as a POMDP. Next, we employ the special structure of this POMDP to reformulate it as a special Markovian search problem with two states and two observation outcomes.

Define *information vector* $I_k = (y_1, \ldots, y_k, u_1, \ldots, u_{k-1})$ with $I_1 = y_1$ as the information available to the controller at time $k \in Z_+$. This information consists of observations up to time k and actions up to time $k-1$. Define stationary policy ν as a measurable function that maps the information vector I_k to the action $u_k \in U$ as: $u_k = \nu(I_k)$ [5]. Also, let $\alpha_0 = [\alpha_0(1)\ \alpha_0(2)\ \alpha_0(3)]'$ be the initial distribution of the system state, i.e. $\alpha_0(i) := \mathbf{P}(x_0 = i),\ i = 1, 2, 3$. Let $J(\alpha_0, \nu)$ denote the total expected cost, having initial state distribution α_0 and applying stationary policy ν. We then have [5]:

$$J(\alpha_0, \nu) = \lim_{N \to \infty} \mathbf{E}_\nu \left\{ \sum_{k=1}^{N} g(x_k, u_k) | \alpha_0 \right\} \tag{6}$$

where $g(x_k, u_k)$ is defined in (5) and E_ν denotes the expected value conditional on applying the stationary policy ν. The missile launch control problem is then: $\min_{\nu \in \mathcal{U}} J(\alpha_0, \nu)$. It can be easily shown that the *conditional probability distributions* $P_{x_k|I_k}$ summarizes all the information available at time k [5]. Let $\alpha_k(x) := \mathbf{P}(x_k = x|I_k), x \in \mathcal{X}$, where $\mathbf{P}$ denotes the probability measure. The new *information state vector* α_k is then given by:

$$\alpha_k = [\alpha_k(1) \quad \alpha_k(2) \quad \alpha_k(3)]', \quad k = \{0, 1, 2, \ldots\} \tag{7}$$

The following Hidden Markov Model (HMM) predictor describes the evolution of the information state vector [5]:

$$\alpha_{k+1} = \frac{R(u_k)'T(u_k, y_k)\alpha_k}{\mathbf{1}'R(u_k)'T(u_k, y_k)\alpha_k} \tag{8}$$

where $T(u, y)$ is a 3×3 *diagonal* matrix describing the conditional probabilities of the observations. $T_{ij}(u, y)$, the *ij*th element of $T(u, y)$, is given by:

$$T_{ij}(u, y) = \begin{cases} \mathbf{P}(y_k = y|x_k = i,\, u_k = u) & , i = j \\ 0 & , i \neq j \end{cases} \tag{9}$$

Based on the above definition, it is easy to see $T(u_k, y_k)$ is independent of the action u_k. This is because knowing the state at time k, the observation y_k is completely known. In fact the consequence of the action u_k will appear on the observation y_{k+1} as given in (4). Simple calculation gives:

$$T(u_k, \text{NAF}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ T(u_k, \text{AFF}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{10}$$

Verifying from (8), information state update has the following structure:

$$\alpha_{k+1} = [p_{k+1}\ q_{k+1}\ 0]' I_{\{y_{k+1} = \text{NAF}\}} + [0\ 0\ 1]' I_{\{y_{k+1} = \text{AFF}\}} \tag{11}$$

where I_B is the indicator function of the event B and $p_{k+1} + q_{k+1} = 1$. Note that p_k is essentially the probability that the enemy radar is "ON" (state 1) at time $k+1$ conditional on the past history (I_k, u_k) and knowing $y_{k+1} = \text{NAF}$. Similarly, q_k denotes the conditional probability that the enemy radar is "OFF". At this stage the problem formulation as a POMDP with reduced dimension is completed.

In the following, we reformulate the optimal launch policy as a special Markovian search problem studied in [2], [3]. **Markovian search problem :** *Consider an object that moves between two sites based on a two-state Markov chain. One of the sites is searched at each times $k \in Z_+$ until the object is found. Associated with each search of site $i \in \{1, 2\}$ there is a cost C_i and an overlook probability β_i. The aim is to find the object with minimum average cost.*

In the rest of this section, we formally formulate the launch control problem as the above Markovian search problem. Upon the successful hard-kill of the enemy radar, the missile launch control is terminated. Therefore, as far as control is applicable, we can assume $y_k = \text{NAF}$ for $0 < k < N$, where N is the stopping time denoting the discrete-time when hard-kill is completed. In this case, the information state update in (11) will reduce to $\alpha_{k+1} = [p_{k+1}\ q_{k+1}\ 0]'$. Since $p_k + q_k = 1$, p_k contains all the information relevant to the launch control at time k. Let ϕ be a function that describes the evolution of the information state. We have from the HMM predictor in (8):

$$\phi(p_k, u_k) := p_{k+1} = \begin{cases} \frac{a(1-p_d)p_k + (1-d)q_k}{(1-p_d)p_k + q_k} & , u_k = \text{La} \\ \mu p_k + 1 - d & , u_k = \text{Su} \end{cases} \tag{12}$$

where $u_k \in U = \{\text{La}\ , \text{Su}\}$ is the action at time k, a and d are the radar transition probabilities defined in (1), and $\mu := a + d - 1$ is the transition memory. Now, we formulate the optimality equation and formally show that it has the same structure as the search problem in [2]. Let $\mathbf{u} = \{u_1, u_2, \ldots\}$ be a sequence of the actions $u_k \in U$ taken at times $k \in Z_+$. Define $\mathbf{u}_{(n)} = \{u_n, u_{n+1}, \ldots\}$. Let $V(p_1; \mathbf{u})$ be the average cost of completing the hard-kill of the enemy radar using the policy $\mathbf{u}$ with initial information state p_1. Also, define $V(p_1) := \inf_{\mathbf{u}} V(p_1; \mathbf{u})$ to be the *minimum* average cost starting with initial state p_1. $V(p_1)$ satisfies the Bellman dynamic programming functional equation $V(p_1) = \min\{V_1, V_2\}$, where:

$$V_1 = c_1 + V(\phi(p_1, \text{La}))((1 - p_d)p_1 + 1 - p_1)$$
$$V_2 = c_2 + V(\phi(p_1, \text{Su}))\},$$

and c_1 is given in (5) and $\phi(p_k, u_k)$ is given by (12). The above Bellman equation has the exact same structure as the the optimality equation in [2] with overlook probabilities $\beta_1 = 1 - p_d$ and $\beta_2 = 1$. We therefore conclude that the problem of optimal HARM launch control has been formulated as a Markovian search problem described in [2].

4. OPTIMAL THRESHOLD POLICIES

The following theorem states the existence of an optimal threshold policy for the missile launch problem:

Theorem 4.1 *Let p_k be the state information at time k in the missile launch control problem. Then there exists a threshold value, δ, such that for any $k \in Z_+$, if $p_k \geq \delta$, the optimal action at time k is to launch a HARM, and if $p_k < \delta$, the optimal action at time k is to suspend the attack.*

By observing the corresponding optimality equations, we established in Section 3 that our missile launch control problem is equivalent to a special form of a two-state Markovian search in [2]. The reader is then referred to [2] to see the details of the proof for the equivalent search problem.

In particular, which threshold level is applicable depends explicitly on the *fixed points* of the evolution equations in (12). Let P_{La} be the fixed point of $\phi(\cdot, \text{La})$ and P_{Su} be the fixed point of $\phi(\cdot, \text{Su})$ in (12). P_{La} and P_{Su} are then given by:

$$P_{\text{La}} = \frac{2 - (\beta a + d) - \sqrt{(\beta a + d)^2 - 4\beta\mu}}{2(1-\beta)}, \quad P_{\text{Su}} = \frac{1-d}{1-\mu},$$

where $\beta := 1 - p_d$. To express the main result of this section we need to define the following mappings of the information state by two different consecutive actions"

$$P_{\text{La,Su}}(\cdot) := \phi\left(\phi(\cdot, \text{La}), \text{Su}\right), \quad (13)$$

$$P_{\text{Su,La}}(\cdot) := \phi(\phi(\cdot, \text{Su}), \text{La}), \quad (14)$$

where $\phi(\cdot, u \in \{\text{La}, \text{Su}\})$ is defined in (12) and is given by equations in (12). The following proposition states the main result of [2] adapted to our missile launch control problem:

Proposition 4.1 *The missile launch control system is categorized into three different Attack profiles - Attack profile 1, Attack profile 2 and 3. Each Attack profile has a different threshold value δ_1, δ_2 and δ_3. Attack profile membership rule is as follows:*

1: $\delta_1 < P_{La}$ (15)

2: $P_{La} < \delta_2 < P_{Su}$ *and* $P_{Su,La}(\delta_2) < \delta_2 < P_{La,Su}(\delta_2)$

3: $P_{La} < \delta_3 < P_{Su}$ *and*$\{\delta_3 > P_{La,Su}(\delta_3)$ *or* $\delta_3 < P_{Su,La}(\delta_3)\}$,

where the fixed points P_{La} and P_{Su} are given in (13) and $P_{La,Su}$ and $P_{Su,La}$ are defined in (13). The Threshold levels for Attack profiles 1 and 2 are given by:

$$\delta_1 = \frac{(1-d)(c_1 - c_2)}{(1-\mu)c_1} = P_{Su}\frac{(c_1 - c_2)}{c_1}, \quad (16)$$

$$\delta_2 = \frac{(1-d)(c_1 - \mu c_2)}{(1-\mu)(c_1 + c_2)} = P_{Su}\frac{(c_1 - \mu c_2)}{c_1 + c_2}, \quad (17)$$

where μ is the transition memory. The threshold level for Attack profile 3, δ_3, cannot be obtained in closed form but as explained in [2], δ_3 can be numerically computed by applying multiple compositions of $\phi(\cdot, La)$ and $\phi(\cdot, Su)$.

5. NUMERICAL EXAMPLES

The purpose of this section is to evaluate by numerical experiments the performance of the optimal threshold policy in terms of the incurred average cost up to the successful completion of the hard-kill. We examine three different schemes for the launch control policies: optimal threshold policy as outlined in proposition 4.1, persistent attack and periodic attack. The persistent attack is the most aggressive method where the controller chooses to attack at each time slot until the enemy radar is destroyed. In periodic attack the controller attacks and suspend alternately in consecutive time slots. In Fig 1 the average cost is depicted in terms of the transition memory. It is clear that the threshold policy gives the best performance in all cases. One observation is that the performance of the persistent attack significantly degrades when μ increases. This is because when the radar has little or no memory, the controller cannot significantly exploit the dynamic of the engagement. In this case the persistent attack may show a relative good performance. However, as the memory increases, the persistent attack is less cost effective.

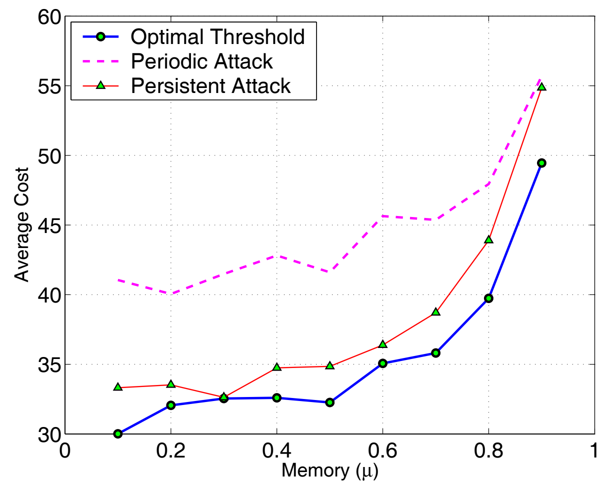

Fig. 1. Attack profile 1 : c_1=3, c_2=1, β=0.9, p_d=0.1.

6. CONCLUSION

We have derived stochastic control algorithms to achieve the optimal trade-off between the HARM's launch cost (attack cost) and the latency in performing the hard-kill of the enemy radar. Structural results in Markovian target search have been used to show that optimal launch policies are threshold in nature. We have shown by numerical examples that these polices outperform non-optimal heuristic algorithms in terms of the average cost of a successful hard-kill.

7. REFERENCES

[1] B. Sweetman, *Radar Jammers: The EA-6B Prowlers (War Planes)*. Capstone High-Interest Books, 2002.

[2] I. MacPhee and B. Jordan, "Optimal search for a moving target," *Probability in the Engineering and Information Sciences*, vol. 9, pp. 159–182, 1995.

[3] S. Ross, *Introduction to Stochastic Dynamic Programming*. Academic Press, 2000.

[4] L. A. Johnston and V. Krishnamurthy, "Optimality of threshold transmission policies in Gilbert Elliott fading channels," in *IEEE International Conference on Communications, ICC '03,*, vol. 2, pp. 1233–1237, May 2003.

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2st ed., 2000.