ARTICLE IN PRESS
JID: EOR [m5G;November 29, 2016;16:10]

European Journal of Operational Research 000 (2016) 1–14

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Discrete Optimization

Approximate dynamic programming for missile defense interceptor fire control

Michael T. Davis, Matthew J. Robbins*, Brian J. Lunday

Air Force Institute of Technology, Department of Operational Sciences, 2950 Hobson Way, Wright-Patterson AFB, OH 45433, United States

ARTICLE INFO

Article history:
Received 27 April 2016
Accepted 10 November 2016
Available online xxx

Keywords:
Approximate dynamic programming
Least squares temporal differences
Markov decision processes
Military applications
Weapon target assignment problem

ABSTRACT

Given the ubiquitous nature of both offensive and defensive missile systems, the catastrophe-causing potential they represent, and the limited resources available to countries for missile defense, optimizing the defensive response to a missile attack is a necessary national security endeavor. For a single salvo of offensive missiles launched at a set of targets, a missile defense system protecting those targets must determine how many interceptors to fire at each incoming missile. Since such missile engagements often involve the firing of more than one attack salvo, we develop a Markov decision process (MDP) model to examine the optimal fire control policy for the defender. Due to the computational intractability of using exact methods for all but the smallest problem instances, we utilize an approximate dynamic programming (ADP) approach to explore the efficacy of applying approximate methods to the problem. We obtain policy insights by analyzing subsets of the state space that reflect a range of possible defender interceptor inventories. Testing of four instances derived from a representative planning scenario demonstrates that the ADP policy provides high-quality decisions for a majority of the state space, achieving a 7.74% mean optimality gap over all states for the most realistic instance, modeling a longer-term engagement by an attacker who assesses the success of each salvo before launching a subsequent one. Moreover, the ADP algorithm requires only a few minutes of computational effort versus hours for the exact dynamic programming algorithm, providing a method to address more complex and realistically-sized instances.

1. Introduction

Currently, over 30 countries have inventories of theater ballistic missiles (George C. Marshall & Claremont Institutes, 2015) while an additional 50 employ multiple launch rocket systems (Global Firepower, 2015). Both of these weapon systems are capable of causing large amounts of damage and of inflicting a high number casualties on their targets. The proliferation of these weapon systems has increased their destructive potential to a worldwide scale while continued research and development on them has led to the creation of even more capable systems that can be used by their developers to threaten neighboring countries or demand concessions in exchange for halting their production. Even U.S. officials concede that, because of the country's recent focus on counter-terrorism, other world powers have closed the gap on guided munitions technology, and the U.S. is now facing the uncertainty of being able to win a "guided munitions salvo competition" (Goure, 2015).

The threat from these weapons has led to the development and spread of missile defense systems. The U.S.-developed Patriot system has been in service for over 30 years, seeing use in both Gulf Wars as well as other conflicts (Davenport, 2015), and it has been fielded by 12 other countries (Raytheon, 2015). One of the best known systems from recent conflicts is Israel's Iron Dome. Developed by Israel and funded mostly by the U.S., the Iron Dome boasts a 90% success rate of destroying incoming rockets headed towards civilian populations, intercepting over 500 rockets during Operation Protective Edge (The Jerusalem Post, 2015). Israel has exported its Iron Dome technology to Canada (Harress, 2015) and continues to work closely with India to develop cutting-edge surface-to-air missiles (SAM) (IBC News Bureau, 2015). Moreover, Israel is currently negotiating the export of the interceptor technology underlying the recent success of Iron Dome for joint production and use by the U.S. (Opall-Rome, 2016). Still more countries, like Turkey, are seeking to acquire long-range missile defense systems (Reuters, 2015), and the U.S. continues to push ahead with missile defense planning and coordination for Europe and Africa (Mindock, 2015).

The security these defense systems may provide comes at a significant financial cost. Initial acquisition costs can be billions of dollars, depending on the size and scope of the order. For example,

---

* Corresponding author.
E-mail addresses: matthew.robbins@afit.edu (M.J. Robbins), brian.lunday@afit.edu (B.J. Lunday).

http://dx.doi.org/10.1016/j.ejor.2016.11.023
0377-2217/© 2016 Published by Elsevier B.V.

Please cite this article as: M.T. Davis et al., Approximate dynamic programming for missile defense interceptor fire control, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.11.023

the cost to equip Qatar with the Patriot missile defense system in late 2014 was $2.4 billion (Tomkins, 2014). Once the system is in place, it must be modernized periodically to counter the evolution of missile threat systems. South Korea paid $770 million for a recent upgrade to its missile defense system (PR Newswire, 2015). Finally, the cost of the interceptor missiles themselves is a large part of the ongoing price of missile defense. The U.S. recently awarded a $1.5 billion contract to Lockheed Martin for an order of its latest interceptors (Brown, 2015), and Saudi Arabia has purchased 600 of the same missiles for $5.4 billion (Dillow, 2015).

Given the ubiquitous nature of both offensive and defensive missile systems, the catastrophe-causing potential they represent, and the limited resources available to countries for missile defense, optimizing the defensive response to a missile attack is a valuable endeavor. For a single salvo of offensive missiles launched at a set of targets, a missile defense system protecting those targets must decide how many interceptors to fire at each incoming missile. This decision is the well-studied static weapon-target assignment problem. However, missile engagements between an attacker and defender typically extend over many salvos of missile launches by the attacker. That is, the attacker does not launch all of its missiles at once. Instead, it launches subsets of its inventory at selected targets in discrete time periods. Hence, the defense cannot fire all of its interceptors at once; it must reserve some number of its inventory in consideration of subsequent attack salvos. This component of time is a distinguishing characteristic of the dynamic weapon-target assignment problem (DWTAP).

In this paper, we formulate an asset-based, defensive variant of the DWTAP wherein we explicitly model the status of the protected assets and interceptor inventory levels over time. We consider a sequence of “look-shoot” engagements in which the defender may fire only one interceptor salvo at the incoming salvo of missiles fired by the attacker. The size and intended targets of the incoming attacker salvo are not known by the defender prior to its launch by the attacker. Moreover, the total number of engagements (i.e., total number of salvos fired by the attacker) is also unknown to the defender. We develop a Markov decision process (MDP) model of this DWTAP. The large size of the motivating problem instance of interest yields a high-dimensional state space, suggesting that classical dynamic programming methods are inappropriate. As such, we apply approximate dynamic programming (ADP) methods to attain high-quality interceptor fire control policies. We develop and test an approximate policy iteration (API) algorithm that utilizes least-squares temporal differences (LSTD) for policy evaluation. We define a set of basis functions within a linear architecture to approximate the value function around the post-decision state. To demonstrate the applicability of our MDP model and to examine the efficacy of our proposed solution methodology, we construct a notional, representative missile defense planning scenario consisting of four related problem instances. We design and conduct a computational experiment to determine how selected problem features and algorithmic features affect the quality of solutions attained by our ADP policies as compared to the optimal policy.

This paper makes the following contributions. First, we examine an asset-based variant of the DWTAP in which the overlapping nature of the interceptor engagement envelopes and the presence of heterogeneous-valued assets creates an interesting, fundamental tension: should a particular SAM battery expend interceptors to protect a higher-valued asset, which is also protected by other SAM batteries, or reserve interceptors for protecting a lower-valued asset that only it can protect? Second, we propose an ADP method that represents a new solution approach for the DWTAP; the closest related work is that of Bertsekas, Homer, Logan, Patek, and Sandell (2000), who also utilize an ADP solution methodology to solve a DWTAP, but employ a neural network-based value function approximation scheme as opposed to our linear architecture approximation scheme with LSTD learning. Third, our research informs the continuing development of a larger, more comprehensive game theoretic perspective of the missile defense situation.

Previous work by Han, Lunday, and Robbins (2016) provides an initial game theoretic framework for the examination of the missile defense situation. Han et al. (2016) formulate and solve a defender-attacker-defender extensive-form game wherein the defender first decides where to locate its SAM batteries, the attacker subsequently decides which targets to engage with a *single salvo* (within which multiple missiles can target a single defender asset), and then the defender makes fire control decisions to intercept the incoming attacker missiles. In contrast to Han et al. (2016), our paper considers SAM batteries at affixed locations and provides a formulation to optimize a defender’s fire control decisions in response to a *multiple-salvo* missile engagement by an attacker; thus, we decrease the emphasis on the rarely visited, long-term location decisions and improve the realism of missile engagements considered. Moreover, although our proposed MDP model is inherently a construct for a single decision maker (in this case, the defender), we incorporate a “smart” attacker into the formulation to better inform the resulting fire control policy. In a sequel to this work, we are working to address the optimization of the attacker’s actions within a multiple-salvo missile engagement setting, utilizing a “smart” defender that acts according to the ADP policies developed herein. Once ADP techniques are developed for modeling both the defender’s and attacker’s respective policies, we intend to incorporate them within a broader, game theoretic framework to examine the larger, more realistic missile defense situation.

The remainder of this paper is organized as follows. Section 2 presents a review of pertinent literature concerning the weapon-target assignment problem. Several ADP papers that inform the development of our model and solution methodology are also reviewed. Section 3 presents a description of the DWTAP variant considered herein. Section 4 describes the MDP model formulation of the DWTAP and presents our ADP approach. In Section 5, we demonstrate the applicability of our model and examine the efficacy of our proposed solution methodology. Section 6 provides conclusions and directions for future research.

2. Literature review

Two streams of literature inform our work. The first stream of literature relates to the weapon-target assignment problem (WTAP). The second stream of literature involves selected works concerning approximate dynamic programming (ADP).

2.1. WTAP

The WTAP is a classical operations research problem of great importance to defense-related applications. Simply stated, the WTAP seeks an optimal assignment of a fixed number of weapons to a fixed number of targets to maximize the total damage inflicted on the set of targets. Research on the WTAP has increased through the years as threat systems and platforms proliferate in type and number, to the extent that a weapon-target assignment system that can efficiently solve a WTAP is now a key component of battlefield planning (Athans, 1987; Roux & Van Vuuren, 2007).

Research on the WTAP began in the 1950s. Manne (1958) developed a linear programming approximation, after which Bradford (1961) and Day (1966) studied WTAP modeling issues including its decomposition into subproblems with subsequent reconstitution.

Eckler and Burr (1972), Matlin (1970), Murphey (2000b, Chapter 3), and Cai, Liu, Chen, and Wang (2006) provide extensive surveys of the WTAP literature. Matlin (1970) reviews the literature

Please cite this article as: M.T. Davis et al., Approximate dynamic programming for missile defense interceptor fire control, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.11.023