NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

OPTIMAL POSITIONING OF REMOTELY PILOTED FUEL BLADDERS TO SUPPORT DISTRIBUTED MARITIME OPERATIONS

by

Jeremy T. Tan

December 2020

Thesis Advisor: Jefferson Huang
Second Reader: Michael P. Atkinson

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE				*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.				
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** December 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis		
4. TITLE AND SUBTITLE OPTIMAL POSITIONING OF REMOTELY PILOTED FUEL BLADDERS TO SUPPORT DISTRIBUTED MARITIME OPERATIONS				**5. FUNDING NUMBERS**
6. AUTHOR(S) Jeremy T. Tan				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000				**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A				**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.				**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This thesis research examines the problem of optimally routing a remotely piloted fuel bladder (RPB) to effectively serve distributed maritime forces. In response to changes in the global threat environment, the U.S. Navy is developing new concepts that involve distributed surface forces operating in large threat areas over prolonged periods at sea. An idea that has been identified to support increasingly distributed forces is the use of minimally manned or unmanned prepositioned bulk fuel storage systems as part of a larger fuel distribution network. While current U.S. defense maritime logistics forces can continue to be called upon to resupply surface forces, they were not designed to support distributed maritime operations. Doing so may, in turn, affect mission effectiveness and operational outcomes. The problem is modeled as a dynamic facility location problem—how to relocate the RPB over discrete-time periods relative to the locations of the distributed surface forces or supported units (SUs). A Markov decision process model is formulated and analyzed with the objective of minimizing the total cost to serve the SUs, whose movements can be stochastic in nature.

14. SUBJECT TERMS dynamic facility location, stochastic, logistics, distributed forces, remotely piloted fuel bladder, RPB, supported units, SUs			**15. NUMBER OF PAGES** 59
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

OPTIMAL POSITIONING OF REMOTELY PILOTED FUEL BLADDERS TO SUPPORT DISTRIBUTED MARITIME OPERATIONS

Jeremy T. Tan
Commander, Republic of Singapore Navy
BS, National University of Singapore, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
December 2020

Approved by: Jefferson Huang
Advisor

Michael P. Atkinson
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis research examines the problem of optimally routing a remotely piloted fuel bladder (RPB) to effectively serve distributed maritime forces. In response to changes in the global threat environment, the U.S. Navy is developing new concepts that involve distributed surface forces operating in large threat areas over prolonged periods at sea. An idea that has been identified to support increasingly distributed forces is the use of minimally manned or unmanned prepositioned bulk fuel storage systems as part of a larger fuel distribution network. While current U.S. defense maritime logistics forces can continue to be called upon to resupply surface forces, they were not designed to support distributed maritime operations. Doing so may, in turn, affect mission effectiveness and operational outcomes. The problem is modeled as a dynamic facility location problem—how to relocate the RPB over discrete-time periods relative to the locations of the distributed surface forces or supported units (SUs). A Markov decision process model is formulated and analyzed with the objective of minimizing the total cost to serve the SUs, whose movements can be stochastic in nature.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AFPs	adaptive force packages
AO	area of operations
CLF	Combat Logistics Force
MDP	Markov decision process
RPBs	remotely piloted fuel bladders
SUs	supported units
USN	United States Navy

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

In response to changes in the global threat environment, the U.S. Navy (USN) is developing operating concepts such as "Distributed Lethality" (Rowden et al. 2015). As the USN shifts toward distributed maritime operations, changes in its logistics capabilities are required to support new operating concepts (Walton et al. 2019). An idea that has been identified to support distributed maritime forces is the use of minimally manned or unmanned prepositioned bulk fuel storage systems, or remotely piloted fuel bladders (RPBs), as part of a larger fuel distribution network.

Prior studies on maritime logistics to support distributed forces included analysis of having a resupply vessel positioned at a fixed location, outside of the combat zone where adaptive force packages (AFPs) operate within (Atkinson et al. 2016). The choice of the refueling point (i.e., the fixed location of the resupply vessel) affects the distance AFPs have to travel to the resupply vessel for refueling as well as the distance the resupply vessel has to travel to a port for its own replenishment. Fixing the refueling point closer to the combat zone increases the amount of time AFPs spent in their assigned stations, or on-station time. However, doing so may increase the risk of the resupply vessel being detected/attacked and can result in instances where AFPs wait in line for refueling as the resupply vessel requires more travel time during its own replenishment cycle. The employment of RPBs, in particular those that have low signature and are attritable (low-cost) by design, can mitigate the effects of these tradeoffs. Low-signature RPBs can potentially be placed within the combat zone to reduce the distance between the AFPs and a refueling point while keeping the risk of detection low. Moreover, RPBs can be routed to keep distance from AFPs as they move around within the combat zone, thus making the refueling support rendered by RPBs more responsive vis-à-vis a fixed refueling point.

In this report, we examine the problem of optimally routing a RPB to effectively serve distributed maritime forces, or supported units (SUs), whose movements can be stochastic in nature. This was modeled as a dynamic facility location problem (Rosenthal et al. 1978) where the decision-maker determines the relocations of the RPB as part of a discrete-time Markov decision process (MDP). Each of these decisions has an associated decision cost. It comprises the RPB-relocation cost and the service cost, which represents a measure of the distance between the RPB and the SUs. Using linear programming (LP), an optimization model was formulated with the objective to minimize the expected decision cost. The studies were focused on the single-RPB, single-SU case based on notional values. An optimal RPB-relocation policy is one where the RPB loosely follows the anticipated movements of the SU; it balances between the extent of how tightly the RPB should follow the SU in order to minimize relocation costs and the positioning of the RPB to serve the SU in order to minimize service costs.

The "value" of an optimal RPB-relocation policy was quantified by comparing its expected discounted decision cost incurred over an infinite time horizon against that of the fixed-RPB policy, where the RPB is stationary and fixed at a given location. To ensure a fair comparison, we first fix an initial position of the SU and then determine the optimal initial position of the RPB that minimizes the expected total discounted decision cost under each of the two policies. Based on the notional values used in the numerical analysis, an optimal RPB-relocation policy is more cost-effective than the fixed-RPB policy, regardless of the initial position of the SU. This difference in the expected total discounted decision cost incurred can be viewed as the "value" of an optimal RPB-relocation policy or more simply put, mobility. However, should relocation costs become much larger relative to service costs, the linear program can yield an optimal RPB-relocation policy that prescribes the RPB to operate as though it is under the fixed-RPB policy. This highlights the tradeoffs involved when mobility comes at a much greater cost and such situations can apply when we wish to impose speed restrictions on the RPB. The "value" of mobility can also be seen through the wider range of initial deployment options that are available to the decision-maker under an optimal RPB-relocation policy vis-à-vis the fixed-RPB policy.

The optimization model developed for the single-RPB cases is flexible enough to support follow-on studies or future work, including the extension of the problem to the multiple-RPB, multiple-SU case. While the single-RPB cases assume that the RPB is assigned to serve the SU(s), we need to address how each RPB is assigned to serve each SU in the multiple-RPB, multiple-SU case in order to achieve an optimal RPB-relocation policy for the system as a whole. Not only must the decision taken at each discrete-time period of the MDP address the relocations of the RPBs, it must also capture the accompanying assignment relationships between each RPB and each SU such that the overall expected decision cost is minimized. Modeling such assignment relationships can be achieved through the use of additional indicator variables, without the need for mixed-integer LP. An outline on how the LP formulation developed for the single-RPB cases can be used to obtain an optimal RPB-relocation policy for the multiple-RPB, multiple-SU case is given.

References

Atkinson M, Kress M, Szechtman RS (2016) Logistic model for distributed lethality. Naval Research Program (NRP) report. http://hdl.handle.net/10945/56307

Rosenthal RE, White JA, Young D (1978) Stochastic dynamic location analysis. *Management Science* 24(6):645-653, https://doi.org/10.1287/mnsc.24.6.645

Rowden T, Gumataotao P, Fanta P (2015) “Distributed Lethality”. *Proceedings* Vol. 141/1/1, 343, https://www.usni.org/magazines/proceedings/2015/january/distributed-lethality

Walton TA, Schramm H, Boone R (2019) Sustaining the fight: Resilient maritime logistics for a new era. Center for Strategic and Budgetary Assessments, https://csbaonline.org/research/publications/sustaining-the-fight-resilient-maritime-logistics-for-a-new-era/publication/1

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

It was a challenging but fulfilling experience having to complete my graduate studies at the Naval Postgraduate School via distance learning amid the global COVID-19 situation. I would like to acknowledge the contributions of the Department of Operations Research, the International Graduate Programs Office and the Thesis Processing Office for facilitating the successful completion of my master's degree as we all navigate through and adapt to the "new normal" under challenging times.

It was my privilege to have worked with Professor Jefferson Huang and Professor Michael Atkinson on this thesis research. In particular, Professor Huang has been patient and helpful in guiding me through every step of my thesis work. I enjoyed the discussions we had as a team, and it has been a great learning journey for me overall. Thank you as well goes to Professor Huang and Professor Atkinson for your invaluable advice and guidance on this thesis research.

Last but not least, I would like to thank the Republic of Singapore Navy for giving me the unique opportunity to pursue a master's degree at the Naval Postgraduate School. More importantly, to my wife, Eleanor, thank you for being my pillar of strength and supporting me steadfastly in everything that I do.

May God bless us all.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

In response to changes in the global threat environment, the U.S. Navy (USN) is developing operating concepts such as "Distributed Lethality" (Rowden et al. 2015). These concepts involve the employment of increasingly distributed forces, such as surface action groups operating over large contested areas. While current U.S. defense maritime logistics forces can continue to be called upon to resupply surface forces, they were not designed to support distributed maritime operations (Walton et al. 2019). Doing so may, in turn, affect mission effectiveness and operational outcomes. Innovative changes in the USN maritime logistics architecture are required to support distributed forces effectively. An idea that has been identified to support distributed maritime forces is the use of minimally manned or unmanned systems for refueling at sea, or remotely piloted fuel bladders (RPBs), as part of a larger fuel distribution network. Such low-signature systems will provide attritable (low-cost) refueling options in a contested environment, as well as responsive refueling support over a wide area by having refueling points located closer to forward-operating combatants.

B. MODEL AND APPROACH

This thesis research examines the problem of optimally routing a RPB to effectively serve distributed maritime forces, or supported units (SUs). The routing of a RPB concerns the relocation of the RPB over time. This will be modeled as a dynamic facility location problem, where a decision-maker seeks to make dynamic relocation decisions for a server that must interact with customers whose relocations are stochastic processes (Rosenthal et al. 1978). Here, the RPB will take on the server role and each SU will take on the customer role. The relocations of the RPB are choice-determined by the decision-maker while the successive locations of the SUs are described by a discrete-time Markov chain. Together, their successive locations are modeled as states of a discrete-time Markov decision process (MDP), which can be viewed as a "controlled" Markov chain where a certain degree of influence is exerted over the states of a system. The decision-maker seeks to find a policy that prescribes the relocation of the RPB based on the current state of the MDP. An expected decision cost, which is location-dependent, is incurred at each time period. It consists of two components: the RPB-relocation cost and the service cost, which represents a measure of the distance between the RPB and the SUs. The service cost can be perceived as the opportunity cost of having the SU depart from its current station and travel to the RPB for refueling. Using notional values, an optimization model will be formulated with the objective of obtaining an optimal RPB-relocation policy that minimizes the expected total discounted decision costs incurred over an infinite time horizon.

C. RESEARCH QUESTIONS

The following research questions will be addressed in this thesis:

(1) How do we to determine an optimal relocation policy for the RPB?

(2) How can we quantify the "value" of an optimal RPB-relocation policy?

(3) What are the implications or tradeoffs involved for an optimal RPB-relocation policy as the number of RPBs and SUs increases or as the value of the model parameters changes?

D. METHODOLOGY

First, we examine the single-RPB, single-SU case in greater detail as we build understanding toward the general case involving multiple RPBs and multiple SUs. Markov decision models are built to solve the RPB-relocation problem. These are analytical models that can be solved via linear programming (LP). Notional values are used for the model parameters, including decision costs. An RPB-relocation policy prescribes RPB relocations based on the current state of the MDP. Solving the linear program yields an optimal RPB-relocation policy that minimizes the expected total discounted decision costs incurred over an infinite time horizon.

Next, we use the fixed-RPB policy, where the RPB is stationary and fixed at a given location, as the common reference that an optimal RPB-relocation policy is evaluated against. In doing so, we are able to quantify the "value" of a dynamic RPB-relocation policy or more simply put, mobility, by taking the difference between the expected total discounted decision costs incurred for the fixed-RPB policy and that of the RPB-relocation policy.

Finally, we examine the implications for an optimal RPB-relocation policy as we generalize the single-RPB, single-SU case to involve multiple RPBs and/or multiple SUs in the model. We also study the tradeoffs involved for an optimal RPB-relocation policy through solving the linear program with different notional values for the model parameters.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A. GAS STATION RESUPPLY MODEL

A previously proposed logistic model for distributed maritime forces concerns the use of a resupply vessel as a gas station (Atkinson et al. 2016). The gas station is located in the communication zone, outside of the combat zone where adaptive force packages (AFPs) operate within. AFPs will travel to the gas station for refueling when required, while the resupply vessel will travel to a port for its own replenishment when required. The choice of the refueling point (i.e., the fixed location of the gas station) affects the distance AFPs have to travel to the resupply vessel for refueling as well as the distance the resupply vessel has to travel to a port for its own replenishment. Assuming the resupply vessel has sufficient fuel capacity, fixing the refueling point closer to the AFPs reduces the travel time to the gas station based on some given speed of the AFPs. Consequently, this increases the amount of time AFPs spend in their assigned stations within the combat zone, or on-station time and thus, can have a positive impact on their mission effectiveness.

However, doing so involves a couple of tradeoffs. First, fixing the refueling point closer to the AFPs requires the resupply vessel to operate closer to the combat zone. This may increase the risk of the resupply vessel being detected or attacked by enemy forces. Second, fixing the refueling point closer to the AFPs increases the distance that the resupply vessel has to travel to a port for its own replenishment. Consequently, the resupply vessel spends a longer time away from the refueling point, assuming its speed is kept constant. This can result in instances where AFPs wait in line at the gas station as the resupply ship has not completed its replenishment cycle and returned to the refueling point. To reduce the time required for the resupply vessel to be replenished, the model addressed the scenario where a shuttle ship, which traverses between a port and the refueling point, is used to refuel the resupply ship without having the resupply ship leave the refueling point.

The employment of RPBs, in particular those that have low signature and are attritable (low-cost) by design, can mitigate the effects of these tradeoffs. Low-signature

RPBs can potentially be placed within the combat zone to reduce the distance between the AFPs and a refueling point while keeping the risk of detection low. Moreover, RPBs can be routed to keep distance from AFPs as they move around within the combat zone, thus making the refueling support rendered by RPBs more responsive vis-à-vis a fixed refueling point. When an RPB is low on fuel, there is a lesser regard for it to be refueled as it can be replaced by another RPB that is deployed ahead of time such that there is minimal disruption of the refueling support to AFPs. In this thesis, we assume the replacement of RPBs to be seamless, such that the refueling support to AFPs is uninterrupted and a RPB can be modeled to have infinite fuel capacity.

An extension of this gas station resupply model was proposed to include additional commodities besides marine fuel (e.g., naval aviation fuel and ordnance). The multi-commodity logistic model for distributed lethality (Mannila 2018) studied the use of smaller resupply ships that have a lower signature than the traditional, larger resupply ships (e.g., Combat Logistics Force [CLF] units) to provide logistics support to AFPs. While positioned outside of the combat zone, these smaller resupply ships, or mini-CLF ships, can operate closer to the AFPs but are subject to missile threats within an anti-access environment. The study recommended a capacity and number of mini-CLF ships required to support a given number of AFPs. Given the similar nature of operations, these recommendations can potentially be used to inform the design and force level of RPBs.

B. DYNAMIC FACILITY LOCATION ANALYSIS

This thesis research examines the problem of optimally routing a RPB to effectively serve distributed maritime forces, or supported units (SUs). The routing of a RPB concerns the relocation of the RPB over time. This will be modeled as a dynamic facility location problem, where a decision-maker seeks to make dynamic relocation decisions for a server that must interact with customers whose relocations are stochastic processes (Rosenthal et al. 1978). Both the server and customers are allowed to change positions. The server's location is under the control of a decision-maker while successive locations visited by the customers are described by a discrete-time Markov chain. Costs, which are location-dependent, are incurred in two ways: when the server is relocated and when the server interacts with customers. Such a system is described by a discrete-time MDP, whose states are the locations of the server and customers. A Markov decision model seeks to solve for a server relocation policy that minimizes the expected costs incurred over an infinite time horizon.

In the context of routing an RPB, the RPB will take on the server role and each SU will take on the customer role. An expected decision cost, which is location-dependent, is incurred at each time period of the MDP. It consists of two components: the RPB-relocation cost and the service cost, which represents a measure of the distance between the RPB and the SUs. The service cost can be perceived as the opportunity cost of having the SU depart from its current station and travel to the RPB for refueling. The decision-maker seeks to find a policy that prescribes the RPB relocations based on the current state of the MDP. In the following chapter, an optimization model will be formulated with the objective of obtaining an optimal RPB-relocation policy that minimizes the expected decision costs incurred over an infinite time horizon.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DYNAMIC RELOCATION MODELS

A. SINGLE-RPB, SINGLE-SU CASE

1. Markov Decision Process Model

According to the discounted MDP model, including the notations, specified by Rosenthal et al. (1978), we define the following variables that will be used to model the special case where a single RPB is assigned to serve a single SU:

$N = \{1, 2, ..., n\} =$ set of possible locations for the RPB and SU
$X_t =$ RPB's location at time t, $X_t \in N$
$A_t =$ SU's location at time t, $A_t \in N$
$P = n \times n$ Markov transition probability matrix for location of the SU
$F = n \times n$ RPB-relocation cost matrix
$G = n \times n$ service cost matrix
$\beta =$ discount factor, $0 < \beta < 1$

Suppose the AO is finite and consists of n locations. The successive locations of the RPB and SU are described by a discrete-time MDP (see Figure 1). The states of the MDP are given by (X_t, A_t), where X_t is choice-determined and A_t is chance-determined, based on P. $P[j,l]$ gives the conditional probability that the SU will move to location l in the next time period given that it is at location j in the current time period. The MDP begins with some initial state (X_0, A_0) at time $t = 0$. At each discrete-time period, the decision-maker observes the current state $(X_t = i, A_t = j),\ i, j \in N$ and then determines $X_{t+1} = k,\ k \in N$. A decision cost is incurred with the determination of X_{t+1}. It is the sum of $F[X_t = i, X_{t+1} = k]$, which gives the cost incurred for relocating the RPB from location i in the current time period to location j in the next time period and $G[X_{t+1} = k, A_{t+1} = l]$, which gives the service cost incurred based on the locations of the RPB and SU in the next time period. The service cost can be perceived as the opportunity cost of having the SU depart from its current station and travel to the RPB for refueling. When the RPB is "close" to the SU, a lower service cost is incurred as the amount of time that the SU is off-station, if it requires refueling, is expected to be small; a higher service cost is incurred when the RPB is "farther" from the SU. The discount factor β, which is applied to the decision cost in the form of β^t, can be thought of as a "tuning parameter" that reflects the decision-maker's preference for near-term costs (smaller discount for smaller values of t) over far-term costs (larger discount for larger values of t). We assume that operational circumstances allow for the situation where the RPB is able to meet the refueling demand of the SU at all times, and so the MDP can have an infinite time horizon.

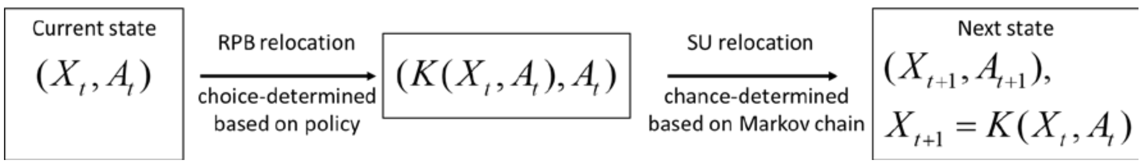

Figure 1. Markov Decision Process for the Single-RPB, Single-SU Case

Recall that the decision cost consists of the known relocation cost, which is based on the decision taken to relocate the RPB in the current time period, and the future service cost, which is dependent on the chance-determined location of the SU in the next time period. While there are alternative ways to define the decision cost (e.g., by replacing the future service cost with the known service cost in the current time period), such a cost formulation is consistent with the discounted MDP model specified by Rosenthal et al. (1978). Although the actual service cost incurred is contingent on the next state, we can determine the "present worth" of the decision taken in the current time period by taking the expected discounted decision cost. Define K to be an RPB-relocation policy that takes the current state of the MDP as input, and outputs the next location of the RPB (i.e., $X_{t+1} = K(X_t, A_t)$). Let κ denote the set of all such policies. Then, the problem is to find an optimal policy K^* such that the expected discounted decision cost incurred over an infinite time horizon is minimized, that is,

$$\min_{K \in \kappa} \mathrm{E}[\sum_{t=0}^{\infty} \beta^t (F[X_t, X_{t+1}] + G[X_{t+1}, A_{t+1}]) \mid X_{t+1} = K(X_t, A_t)] \tag{1}$$

2. Linear Programming

According to the theory of MDP (Puterman 1994, p. 223), we can solve for an optimal RPB-relocation policy K^*, that is, the minimization problem in (1), via LP. The linear program is set up as follows.

Sets and Indices. S represents the state space of the MDP; $(i, j) \in S$ indicates that the RPB is at location i while the SU is at location j. $D_{(i,j)}$ represents the decision space for the relocation of the RPB based on state (i, j); $d \in D_{(i,j)}$ represents the decision to relocate the RPB to location d at the next discrete-time period.

$$S = \{(i, j) \in N \times N\} \subseteq N^2$$
$$D_{(i,j)} = \{d \in N \mid (i, j) \in S\} \subseteq N$$

Parameters. Let $s = (i, j) \in S$ and $s(d) \in S$ denote the state after the RPB has been relocated to location d given state s (i.e., $s(d) = (d,l)$, $d \in D_s$). Define Q to be the state-decision transition probability matrix of the MDP. Then, $Q[s, s(d)] = Q[(i, j),(d,l)] = P[j,l]$, that is, the transition probability from state s to $s(d)$ is equivalent to the transition probability of the SU moving from location j to location l. Define $c(s,d)$ to be the expected decision cost to relocate the RPB to location d given state s (i.e., $c(s,d) = \mathrm{E}[F[i,d] + G[d,l]]$); it comprises the deterministic relocation cost $F[X_t = i, X_{t+1} = d]$ and the expected service cost $\mathrm{E}[G[X_{t+1} = d, A_{t+1} = l]]$, $l \in N$.

$$\begin{aligned}
Q[s, s(d)] &= Q[(i, j),(d,l)] \\
&= P[j,l] \\
c(s,d) &= \mathrm{E}[F[i,d] + G[d,l]] \\
&= F[i,d] + \mathrm{E}[G[d,l]] \\
&= F[i,d] + \sum_{l \in N} P[j,l]G[d,l]
\end{aligned}$$

Decision Variables. The decision variable $\rho(s,d)$ represents a measure of how frequent a given state s is visited, along with the decision taken being $d \in D_s$. There are n^3 decision variables in total.

$$\rho(s,d)\ \forall s \in S, \forall d \in D_s$$

Formulation. In an optimal solution to the linear program, for a given state s, there is only one non-zero $\rho(s,d)\ \forall d \in D_s$; the value of the non-zero $\rho(s,d)$ is the total discounted number of times that an optimal policy K^* will visit state s and prescribe the decision to relocate the RPB to location d.

$$\min \sum_{s \in S} \sum_{d \in D_s} c(s,d)\rho(s,d)$$

subject to:

$$\sum_{d \in D_s} \rho(s,d) - \beta \sum_{\substack{s' \in S \\ s' \neq s}} \sum_{d \in D_{s'}} Q[s', s'(d) = s]\rho(s',d) = 1 \qquad \forall s \in S$$

$$\rho(s,d) \geq 0 \qquad \forall s \in S, \forall d \in D_s$$

3. Numerical Example

Notional values are used for the model parameters in the numerical example. Suppose the RPB and SU operate within a finite AO which can be represented by a 5×5 grid (i.e., the number of locations $n = 25$, see Figure 2).

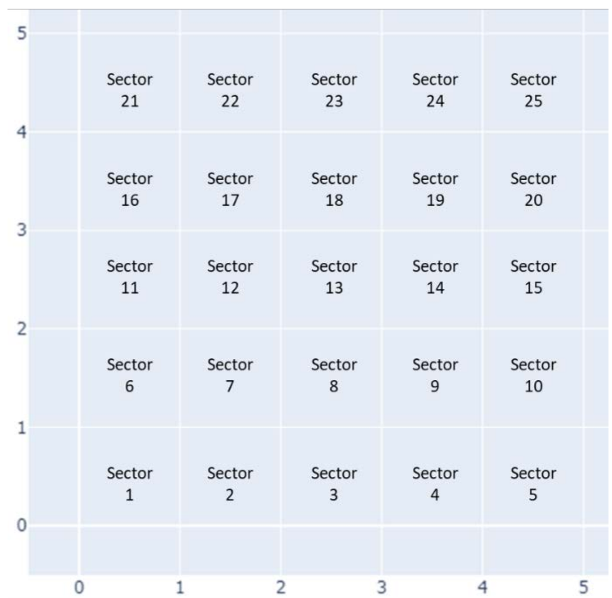

Figure 2. AO for the RPB and SU

The movements of the SU are assumed to be known in advance. The SU covers the AO in a sequential and cyclical manner: if it is at Sector 1, then it will be at Sector 2 next; if it is at Sector 2, then Sector 3 next; if it is at Sector 25, then Sector 1 next.

$$
P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ 23 \\ 24 \\ 25 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} 1 & 2 & 3 & 4 & \ldots & 23 & 24 & 25 \end{array} \\
\begin{pmatrix}
0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

A cost of 300 is incurred whenever the RPB moves to a different location; no cost is incurred if the RPB remains at its current location. Here, the same relocation cost is incurred regardless of how far the RPB is relocated between successive discrete-time periods; the RPB can move to an adjacent location (e.g., from Sector 1 to Sector 2) or across multiple locations (e.g., from Sector 1 to Sector 25) and still incur the same relocation cost (e.g., $F[1,2] = F[1,25]$). This implicitly suggests that no speed restrictions are imposed on the RPB, an assumption that we might not be prepared to make. To model speed restrictions for the RPB (e.g., the RPB can only move to an adjacent location between successive discrete-time periods), we can specify certain relocation costs to be very much larger (e.g., $F[1,25] >> F[1,2]$). Doing so will ensure that the associated decision (e.g., to relocate the RPB from Sector 1 to Sector 25) is too costly to be considered for an optimal RPB-relocation policy.

$$F = \begin{array}{c|cccccccc|} & 1 & 2 & 3 & 4 & \ldots & 23 & 24 & 25 \\ \hline 1 & 0 & 300 & 300 & 300 & \cdots & 300 & 300 & 300 \\ 2 & 300 & 0 & 300 & 300 & \ldots & 300 & 300 & 300 \\ 3 & 300 & 300 & 0 & 300 & \ldots & 300 & 300 & 300 \\ 4 & 300 & 300 & 300 & 0 & \ldots & 300 & 300 & 300 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \\ 23 & 300 & 300 & 300 & 300 & \ldots & 0 & 300 & 300 \\ 24 & 300 & 300 & 300 & 300 & \ldots & 300 & 0 & 300 \\ 25 & 300 & 300 & 300 & 300 & \ldots & 300 & 300 & 0 \end{array}$$

A service cost of 200 is incurred if the RPB is not at a location adjacent to that of the SU; 100 if the RPB is at a location adjacent to that of the SU; no cost is incurred if the RPB and SU are at the same location. For example, given that Sector 2 is adjacent to Sectors 1, 3, 6, 7 and 8 (see Figure 2), we have $G[2,1] = G[2,3] = G[2,6] = G[2,7] = G[2,8] = 100$.

$$G = \begin{array}{c|ccccccccccccccc|} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \ldots & 21 & 22 & 23 & 24 & 25 \\ \hline 1 & 0 & 100 & 200 & 200 & 200 & 100 & 100 & 200 & 200 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 2 & 100 & 0 & 100 & 200 & 200 & 100 & 100 & 100 & 200 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 3 & 200 & 100 & 0 & 100 & 200 & 200 & 100 & 100 & 100 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 4 & 200 & 200 & 100 & 0 & 100 & 200 & 200 & 100 & 100 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 5 & 200 & 200 & 200 & 100 & 0 & 200 & 200 & 200 & 100 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 6 & 100 & 100 & 200 & 200 & 200 & 0 & 100 & 200 & 200 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 7 & 100 & 100 & 100 & 200 & 200 & 100 & 0 & 100 & 200 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 8 & 200 & 100 & 100 & 100 & 200 & 200 & 100 & 0 & 100 & \ldots & 200 & 200 & 200 & 200 & 200 \\ 9 & 200 & 200 & 100 & 100 & 100 & 200 & 200 & 100 & 0 & \ldots & 200 & 200 & 200 & 200 & 200 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\ 21 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & \cdots & 0 & 100 & 200 & 200 & 200 \\ 22 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & \ldots & 100 & 0 & 100 & 200 & 200 \\ 23 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & \ldots & 200 & 100 & 0 & 100 & 200 \\ 24 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & \ldots & 200 & 200 & 100 & 0 & 100 \\ 25 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & 200 & \ldots & 200 & 200 & 200 & 100 & 0 \end{array}$$

To solve for an optimal RPB-relocation policy via the LP formulation, we need to define (1) the state space S and (2) the decision space D_s, $s=(i,j)\in S$, as well as compute (3) the state-decision transition probabilities $Q[s,s(d)]=Q[(i,j),(d,l)]$ and (4) the expected decision costs $c(s,d)$. Assuming no locality restrictions are imposed, both the state space and the decision space will cover the entire AO (i.e., we define $S=N\times N$ and $D_s=N\ \forall s\in S$). The computation of the state-decision transition probabilities and expected decision costs are extensive and will not be shown in its entirety. Instead, the following examples are shown to illustrate how these quantities are computed.

State-Decision Transition Probabilities. Suppose the decision-maker observes that the RPB is at Sector 3 and the SU is at Sector 1 in the current discrete-time period and relocates the RPB to Sector 2 at the next period. Then, the probability that the current state $(3,1)$ will transit to the next state $(2,l)$, $l\in N$ is given by:

$$Q[s=(3,1),s(d)=(2,l)]=P[1,l]=\begin{cases}1 & \text{if } l=2 \ \because P[1,2]=1\\ 0 & \text{if } l\neq 2 \ \because P[1,l]=0\end{cases}$$

Expected Decision Costs. Suppose the decision-maker observes that the RPB is at Sector 5 and the SU is at Sector 1 in the current discrete-time period and relocates the RPB to Sector 8 at the next period. Then, the expected decision cost is given by:

$$\begin{aligned}c(s=(5,1),d=8) &= F[5,8]+\sum_{l\in N}P[1,l]G[8,l]\\ &= 300+\sum_{l=1}^{n}P[1,l]G[8,l] && \because F[5,8]=300\\ &= 300+P[1,2]G[8,2] && \because P[1,l]=0 \text{ if } l\neq 2\\ &= 300+(1)100 && \because P[1,2]=1 \text{ and } G[8,2]=100\\ &= 400\end{aligned}$$

Now, we can solve for an optimal RPB-relocation policy via the LP formulation. Let K^* be an optimal RPB-relocation policy, where $K^*[i, j] = d^*$ represents the prescribed decision to relocate the RPB when the decision-maker observes that the RPB is at location i and the SU at location j. For this numerical example, with $\beta = 0.99$, an optimal policy is given by:

$K^* =$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1	9	1	1	1	1	9	9	1	17	18	19	19	19	17	18	19	19	24	1	1	1	1	1	1
2	2	2	2	2	2	2	2	9	2	17	18	19	19	19	17	18	19	19	24	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	17	18	19	19	19	17	18	19	19	24	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	17	18	19	19	19	17	18	19	19	24	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	17	18	19	19	19	17	18	19	19	24	22	23	24	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	19	19	19	17	18	19	19	24	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	19	19	17	18	19	19	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	19	17	18	19	19	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	17	18	19	19	24	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	17	18	19	10	10	17	18	19	19	24	22	23	24	10	10	7
11	8	11	11	11	11	11	11	11	11	11	11	19	11	11	11	11	19	19	24	22	23	24	11	11	7
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	19	24	22	23	24	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	19	24	22	23	24	13	13	7
14	8	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	22	23	24	14	14	7
15	8	9	9	15	12	15	15	15	15	15	15	15	15	15	17	18	19	15	15	22	23	24	15	15	7
16	8	9	9	16	12	13	16	16	16	16	16	16	16	16	16	16	19	16	16	16	16	24	16	16	7
17	8	9	9	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	7
18	8	9	9	18	12	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	7
19	8	9	9	19	12	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	7
20	8	9	9	20	12	13	9	9	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	7
21	8	9	9	21	12	13	9	9	21	17	18	19	21	21	21	21	21	21	21	21	21	24	21	21	7
22	8	9	9	22	12	13	9	9	22	17	22	22	22	22	22	22	22	22	22	22	22	22	22	22	7
23	8	9	9	23	12	13	9	9	23	17	18	23	23	23	23	23	23	23	23	23	23	23	23	23	7
24	8	9	9	24	12	13	9	9	24	17	18	19	24	24	24	24	24	24	24	24	24	24	24	24	7
25	8	9	9	25	12	13	9	9	25	17	18	19	19	25	25	25	25	25	25	25	25	25	25	25	7

For example, suppose the RPB and SU are at Sector 1 and 7 respectively in the current discrete-time period, then $K^*[1,7] = 9$ prescribes the decision to relocate the RPB to Sector 9 in the next period. The next state of the MDP would be (9,8), that is, the RPB at Sector 9 (as per the prescribed decision) and the SU at Sector 8 (since $P[7,8]=1$). We then refer to $K^{*}[9,8]=9$ and obtain the prescribed decision to keep the RPB at Sector 9 in the following period (see Figure 3 for an illustration). If we were to continue with the MDP, K^{*} prescribes that the RPB should remain at Sector 9 until the SU has visited Sector 15, instead of tightly following the movements of the SU. In general, an optimal RPB-relocation policy is one where the RPB loosely follows the anticipated movements of the SU. An optimal RPB-relocation policy thus balances between the extent of how tightly the RPB should follow the SU in order to minimize relocation costs and the positioning of the RPB to serve the SU in order to minimize service costs.

Figure 3. Illustration of Movements of the RPB and SU under an Optimal RPB-Relocation Policy K^{*}

Using the game of soccer as an analogy, if the movements of the SU are likened to that of the soccer ball, then the movements of the RPB can be likened to that of the referee, who seeks to take up the “best” position in the field in order to make decision calls through observing the game play. To do so, the referee should not simply follow the direction of where the soccer ball is heading toward without considering how the anticipated game play may otherwise require him/her to move to a different position or remain at his/her current location.

B. SINGLE-RPB, MULTIPLE-SU CASE

We generalize the RPB-relocation problem where a single RPB is assigned to serve multiple SUs. Suppose there are M SUs that are indexed by $m = 1, 2, \ldots, M$. The location of each SU in the MDP is given by A_t^m with transition probabilities based on P_m. The LP formulation for the multiple-SU case is similar to the single-SU case in that we replace the single element j with a vector $J = (j_1, j_2, \ldots, j_M)$, where j_m represents the location of SU m. A state of the MDP is given by $(i, j_1, j_2, \ldots, j_M)$ or $(i, J) \in S$. The decision space is indexed by (i, J) (i.e., we have $D_{(i,J)} = \{d \in N \mid (i, J) \in S\} \subseteq N$). We assume that each SU moves independently, and so the state-decision transition probability $Q[(i, J), (d, L)]$ is given by $\prod_m P_m[j_m, l_m]$ (i.e., the product of the transition probabilities of each SU). The expected decision cost $c((i, J), d)$ is given by $F[i, d] + \sum_m \sum_{l \in N} P_m[j_m, l] G[d, l]$. There are n^{M+2} decision variables $\rho((i, J), d)$ in the linear program. The solution to this LP formulation yields an optimal RPB-relocation policy K^*, where $K^*[i, J]$ represents the prescribed decision to relocate the RPB when the decision-maker observes that the RPB is at location i and the SUs are at locations given by J.

IV. ANALYSIS AND RESULTS

A. VALUE OF RPB-RELOCATION POLICY

The motivation to obtain an optimal RPB-relocation policy stems from the intuition that a mobile RPB which can be relocated would be more cost-effective than a stationary RPB that is fixed at a given location (i.e., the fixed-RPB policy K_{fixed}, where $K_{fixed}[i, j] = i \; \forall j \in N$). Again, using the game of soccer as an analogy, if the referee is only allowed to take up a fixed position in the field, there will be instances where the referee will be “too far” from the soccer ball and is unable to observe the game play clearly. Under such circumstances, we can expect that the quality or “value” of the referee’s decision calls to be worse-off as compared to if the referee were to move and take up different positions, such that the referee is “close enough” to the soccer ball and able to better observe the game play.

We seek to determine the “value” of an optimal RPB-relocation policy vis-à-vis the fixed-RPB policy by considering the difference in the minimum expected discounted decision cost incurred over an infinite time horizon. To ensure a fair comparison, the minimum cost incurred for each policy will be based on the same initial position of the SU $A_0 = j$. For example, with a general RPB-relocation policy K, the problem is to determine the initial position for the RPB $X_0 = i^* \in N$ such that given $A_0 = j$, the expected total discounted decision cost is minimized, that is, $\min_{i \in N} \mathrm{E}[\sum_{t=0}^{\infty} \beta^t (F[X_t, X_{t+1}] + G[X_{t+1}, A_{t+1}]) \mid X_0 = i, A_0 = j]$. On the other hand, with the fixed-RPB policy K_{fixed}, $X_t = i' \in N,\; t \geq 0$ and $F[X_t, X_{t+1}] = 0$ (i.e. no relocation cost is incurred). Here, the problem is simplified to finding the initial (and fixed) position for the RPB $X_0 = (i')^* \in N$ such that given $A_0 = j$, the expected total discounted service cost is minimized, that is, $\min_{i' \in N} \mathrm{E}[\sum_{t=1}^{\infty} \beta^t G[X_t = i', A_t] \mid X_0 = i', A_0 = j]$. While the SU can have the same initial location $A_0 = j$ under both the relocation and fixed policies, it is not necessary that we have $i^* = (i')^*$.

B. FORMULATION OF COST INCURRED

The formulation of the cost incurred will be developed for the general case, where the RPB is mobile; the cost incurred for the case where the RPB is fixed can be evaluated by setting the relocation cost to zero.

For any given policy K and initial positions of the RPB and SU, we can evaluate the expected discounted decision cost incurred over an infinite time horizon via the expression in (1). Define $v_K(i, j)$ to be the expected discounted decision cost incurred over an infinite time horizon with $X_0 = i, A_0 = j$ under policy K, where $X_{t+1} = K[X_t, A_t]$.

$$\begin{aligned} v_K(i,j) &= \mathrm{E}[\sum_{t=0}^{\infty} \beta^t (F[X_t, X_{t+1}] + G[X_{t+1}, A_{t+1}]) \mid X_0 = i, A_0 = j] \\ &= \sum_{t=0}^{\infty} \beta^t \mathrm{E}[(F[X_t, K[X_t, A_t]] + G[K[X_t, A_t], A_{t+1}]) \mid X_0 = i, A_0 = j] \end{aligned} \tag{2}$$

Let $h_K(i,j) = \mathrm{E}[F[i, K[i, j]] + G[K[i, j], A_{t+1}] \mid X_t = i, A_t = j]$ denote the expected decision cost incurred based on the observed state (i, j) and policy K. Then, we can rewrite (2) as follows:

$$\begin{aligned} v_K(i,j) &= \sum_{t=0}^{\infty} \beta^t \mathrm{E}[(F[X_t, K[X_t, A_t]] + G[K[X_t, A_t], A_{t+1}]) \mid X_0 = i, A_0 = j] \\ &= \beta^0 h_K(i,j) + \mathrm{E}[\sum_{t=1}^{\infty} \beta^t (F[X_t, X_{t+1}] + G[X_{t+1}, A_{t+1}]) \mid X_0 = i, A_0 = j] \\ &= h_K(i,j) + \beta \mathrm{E}[\sum_{t=1}^{\infty} \beta^{t-1} (F[X_t, X_{t+1}] + G[X_{t+1}, A_{t+1}]) \mid X_0 = i, A_0 = j] \\ &= h_K(i,j) + \beta \mathrm{E}[v_K(X_1, A_1) \mid X_0 = i, A_0 = j] \end{aligned}$$

By considering all possible states $(i, j) \in N \times N$, we can generalize $h_K(i, j)$ to be an $n^2 \times 1$ column vector h_K indexed by (i, j). Recall that Q is the state-decision transition probability matrix of the MDP where $Q[s, s(d)] = Q[(i, j), (d, l)] = P[j, l]$ is the transition probability from state (i, j) to state (d, l). Then, $\mathrm{E}[v_K(X_1, A_1 \mid X_0 = i, A_0 = j] = \sum_{(d,l)\in N\times N} Q[(i, j),(d,l)]v(d,l)$. We generalize $v_K(i, j)$ to be an $n^2 \times 1$ column vector v_K indexed by (i, j), as follows:

$$\begin{aligned} v_K &= h_K + \beta Q v_K \\ v_K - \beta Q v_K &= h_K \\ (I_{n^2\times n^2} - \beta Q)v_K &= h_K \\ v_K &= (I_{n^2\times n^2} - \beta Q)^{-1} h_K \end{aligned} \quad (3)$$

We state without proof that the matrix $(I_{n\times n} - \beta Q)$ is always invertible for any matrix Q and $0 < \beta < 1$ (Kemeny and Snell 1974, p. 22). Hence, the solution for v_K always exists. Then, for each initial position of the SU $A_0 = j$, we solve $\min_{i\in N} v_K(i, j)$ and determine $i^* \in N$.

C. NUMERICAL EXAMPLE (CONTINUED)

Suppose P, Q, F, G, β and K^* were as given in the numerical example in Chapter III. We use (3) to evaluate the expected total discounted cost of an optimal RPB-relocation policy K^* vis-à-vis the fixed-RPB policy K_{fixed}.

First, we determine the index for each state of the MDP $(i, j) \in N \times N$ by mapping each state to a non-negative integer in the set $\{0,1,\ldots,n^2 - 1\}$. Given that $n = 25$, we have $(i = 1, j = 1) \to 0,\ (i = 1, j = 2) \to 1, \ldots,\ (i = 25, j = 25) \to 624$ and so, both column vectors h_{K^*} and $h_{K_{fixed}}$ are indexed from 0 to 624.

Next, compute h_{K^*} and $h_{K_{fixed}}$. For example, $h_{K^*}[1]$ gives the expected decision cost based on state (1,2) and policy K^*, that is,

$$
\begin{aligned}
h_{K^*}[1] &= \mathrm{E}[F[1,K^*[1,2]]+G[K^*[1,2],l]] \\
&= F[1,9]+\sum_{l\in N} P[2,l]G[9,l] \qquad \because K^*[1,2]=9 \\
&= 300+P[2,3]G[9,3] \qquad \because P[2,l]=0 \text{ if } l\neq 3 \\
&= 300+(1)100 \\
&= 400
\end{aligned}
$$

On the other hand, since $K_{fixed}[i,j]=i\ \forall j\in N$, we have $F[X_t=i, X_{t+1}=i]=0\ \forall i\in N$ and $h_{K_{fixed}}$ gives the expected service cost instead of the expected decision cost. For example, $h_{K_{fixed}}[1]$ gives the expected service cost based on state (1,2) and policy K_{fixed}, that is,

$$
\begin{aligned}
h_{K_{fixed}}[1] &= \mathrm{E}[G[K_{fixed}[1,2],l]] \\
&= \sum_{l\in N} P[2,l]G[1,l] \qquad \because K_{fixed}[1,2]=1 \\
&= P[2,3]G[1,3] \qquad \because P[2,l]=0 \text{ if } l\neq 3 \\
&= (1)200 \\
&= 200
\end{aligned}
$$

Last but not least, solve for v_{K^*} and $v_{K_{fixed}}$ via (3). The table below summarizes the initial positions of the RPB i^* under an optimal RPB-relocation policy K^* and initial positions of the RPB $(i')^*$ under the fixed-RPB policy K_{fixed} that minimize the expected total discounted decision cost for each initial position of the SU j.

Table 1. Comparison of Initial Positions of RPB and Minimum Expected Total Discounted Decision Cost under Optimal RPB-Relocation Policy K^* and Fixed-RPB Policy K_{fixed}

	Optimal RPB-relocation policy K^*		Fixed-RPB policy K_{fixed}	
SU's initial position $A_0 = j$	RPB's initial position $X_0 = i^*$	Expected total discounted decision cost $v_{K^*}(i^*, j)$	RPB's initial position $X_0 = (i')^*$	Expected total discounted decision cost $v_{K_{fixed}}((i')^*, j)$
1	8	14675.559	7	15759.168
2	9	14705.426	8	15759.168
3	9	14752.955	9	15759.168
4	12	14718.341	9	15817.342
5	12	14664.990	12	15801.577
6	13	14709.257	12	15759.168
7	9	14744.883	13	15759.168
8	9	14792.811	14	15759.168
9	17	14753.938	14	15817.342
10	17	14700.947	17	15801.577
11	18	14656.939	17	15759.168
12	19	14612.486	18	15759.168
13	19	14659.077	19	15759.168
14	19	14706.138	19	15817.342
15	17	14740.174	19	15876.103
16	18	14693.897	17	15852.931
17	19	14647.153	18	15852.931
18	19	14694.094	19	15852.931
19	24	14746.460	19	15912.051
20	22	14786.523	7	16007.341
21	23	14737.862	7	15967.011
22	24	14688.709	7	15926.274
23	24	14736.070	7	15885.125
24	7	14687.919	7	15843.561
25	7	14634.262	7	15801.577

Based on the notional values used in the numerical example, we have shown that an optimal RPB-relocation policy is more cost-effective than the fixed-RPB policy (6% to 8% lower in cost), regardless of the SU's initial position. Of note, the difference in the minimum expected total discounted decision cost between an optimal RPB-relocation policy and the fixed-RPB policy can be seen a way for the decision-maker to quantify the "value" of the RPB's mobility.

However, should relocation costs become too costly relative to service costs, it is plausible that the "value" of an optimal RPB-relocation policy be diminished to the extent where the fixed-RPB policy is more cost-effective instead. For example, suppose a relocation cost of 1000 is incurred (instead of 300) while keeping all other variables constant, solving the linear program yields an optimal RPB-relocation policy K' that is similar to the fixed-RPB policy. In particular, we have $K'[i, j] = i \ \forall j \in N$, for $i \in N_{in} = \{7,8,9,12,13,14,17,18,19\}$, where N_{in} denotes the set of interior sectors of the AO (see Figure 4).

$K' =$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1	1	1	1	1	1	14	14	1	17	18	19	19	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	14	2	17	18	19	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	17	18	19	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	17	18	19	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	17	18	19	19	19	17	18	19	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	19	19	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	17	18	10	10	10	17	18	19	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11	19	11	11	11	11	19	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
15	8	9	15	15	15	15	15	15	15	15	15	15	15	15	17	15	15	15	15	15	15	15	15	15	7
16	8	9	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	7
17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
20	8	9	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	7
21	8	9	9	21	12	13	14	14	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	7
22	8	9	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	7
23	8	9	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	7
24	8	9	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	7
25	8	9	9	25	12	13	14	14	25	17	25	25	25	25	25	25	25	25	25	25	25	25	25	25	7

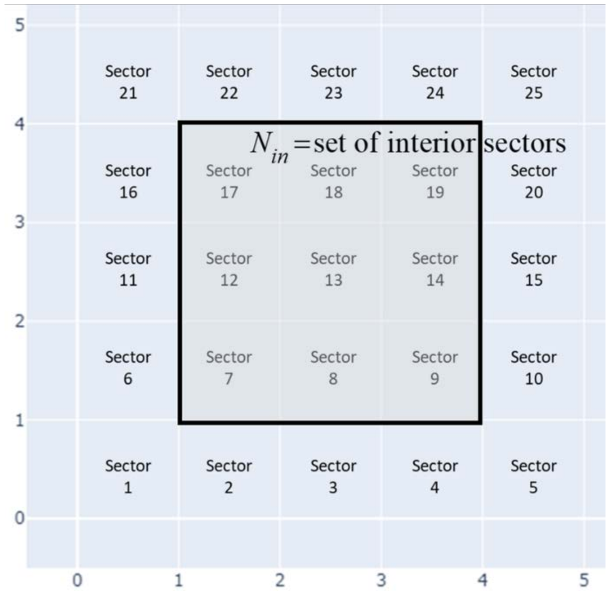

Figure 4. Interior Sectors of AO

While it is not necessary that $i^* = (i')^*$ given the same initial position of the SU, it is interesting to note that $(i')^* \in N_{in}$ and there exists $i^* \notin N_{in}$. This shows that although it is not cost-effective for the RPB to be positioned beyond the interior sectors of the AO under the fixed-RPB policy, such an action can be favorable under an optimal RPB-relocation policy. This gives the decision-maker a wider range of initial deployment options for the RPB and reinforces the "value" of the RPB's mobility.

The calculations for the numerical example were performed on my personal laptop (4-core CPU 1.6 GHz, 8GB RAM). The size of the RPB-relocation problem is characterized by the number of decision variables which depends on the value of n as well as the number of RPBs and SUs. For the single-RPB, single-SU case with $n = 25$ or a 5×5 AO, there is a total of $25^3 = 15,625$ decision variables and it took my personal laptop 15 minutes to solve for an optimal RPB-relocation policy. The computational performance in solving for an optimal RPB-relocation policy for various single-RPB cases are summarized in the table below.

Table 2. Computational Performance in Solving for an Optimal RPB-Relocation Policy Using Personal Laptop

n	No. of RPBs	No. of SUs	No. of decision variables	Approximate runtime to solve for an optimal RPB-relocation policy via LP
25	1	1	$25^3 = 15{,}625$	15 minutes
30			$30^3 = 27{,}000$	More than 2 hours
10		2	$10^4 = 10{,}000$	12 minutes
12			$12^4 = 20{,}736$	More than 2 hours
6		3	$6^5 = 7{,}776$	15 minutes
7			$7^5 = 16{,}807$	More than 2 hours

THIS PAGE INTENTIONALLY LEFT BLANK

V. SUMMARY AND CONCLUSIONS

As the USN shifts toward distributed maritime operations, changes in its logistics capabilities are required to support new operating concepts (Walton et al. 2019). An idea that has been identified to support distributed maritime forces is the use of minimally manned or unmanned systems for refueling at sea, or RPBs, as part of a larger fuel distribution network.

Prior studies on maritime logistics to support distributed forces included analysis of having a resupply vessel positioned in an area that is outside of the combat zone where AFPs operate within (Atkinson et al. 2016). The choice of the refueling point entails a tradeoff between the distance AFPs have to travel to the resupply vessel for refueling and the distance the resupply vessel has to travel to a port for its own replenishment. The employment of RPBs, in particular those that have low signature and are attritable (low-cost) by design, can enable the refueling point to be located within contested areas without considerable regard for the need for the RPBs themselves to be replenished. Doing so thus reduces the distance AFPs have to travel to the resupply vessel and, in turn, enhances their mission effectiveness.

In Chapter III, we examined the problem of optimally routing a RPB to effectively serve distributed maritime forces, or SUs, whose movements can be stochastic in nature. This was modeled as a dynamic facility location problem (Rosenthal et al. 1978) where the decision-maker determines the relocations of the RPB as part of a discrete-time MDP. Each of these decisions has an associated decision cost. It comprises the RPB-relocation cost and the service cost, which represents a measure of the distance between the RPB and the SUs. Using LP, an optimization model was formulated with the objective to minimize the expected decision cost. The studies were focused on the single-RPB, single-SU case based on notional values. An optimal RPB-relocation policy is one where the RPB loosely follows the anticipated movements of the SU; it balances between the extent of how tightly the RPB should follow the SU in order to minimize relocation costs and the positioning of the RPB to serve the SU in order to minimize service costs.

In Chapter IV, we quantified the "value" of an optimal RPB-relocation policy by comparing its expected discounted decision cost incurred over an infinite time horizon against that of the fixed-RPB policy. To ensure a fair comparison, we first fix an initial position of the SU and then determine the optimal initial position of the RPB that minimizes the expected total discounted decision cost under each of the two policies. Based on the notional values used in the numerical analysis, an optimal RPB-relocation policy is more cost-effective than the fixed-RPB policy, regardless of the initial position of the SU. This difference in the expected total discounted cost incurred can be viewed as the "value" of an optimal RPB-relocation policy or more simply put, mobility. However, should relocation costs become much larger relative to service costs, the linear program can yield an optimal RPB-relocation policy that prescribes the RPB to operate as though it is under the fixed-RPB policy. This highlights the tradeoffs involved when mobility comes at a much greater cost and such situations can apply when we wish to impose speed restrictions on the RPB. The "value" of mobility can also be seen through the wider range of initial deployment options that are available to the decision-maker under an optimal RPB-relocation policy vis-à-vis the fixed-RPB policy.

The optimization model developed for the single-RPB cases is flexible enough to support follow-on studies or future work, including the extension of the problem to the multiple-RPB, multiple-SU case, which is broadly covered in Chapter VI.

VI. FUTURE WORK

While the extension of the RPB-relocation problem is relatively straightforward from the single-RPB, single-SU case to the single-RPB, multiple-SU case, generalizing it to the multiple-RPB, multiple-SU case is more complicated as doing so adds a new "dimension" to the problem—how each RPB is "assigned" to each SU affects how the linear program solves for an optimal RPB-relocation policy.

To better understand how assignment relationships between the RPBs and SUs affect the solution to the linear program, consider the following examples. Suppose each RPB serves all the SUs at the same time. Then, an optimal RPB-relocation policy is one where the same decision is prescribed for all the RPBs. In other words, the RPBs will move synchronously as a single group under such an optimal policy and the same expected decision cost is incurred for each RPB at each discrete-time period. This effectively reduces the multiple-RPB problem to the single-RPB, multiple-SU case and sub-optimizes the employment of multiple RPBs. On other hand, suppose no more than one RPB is "assigned" to a SU. Then, the multiple-RPB problem can be "decomposed" into multiple instances of the single-RPB, multiple-SU case or the single-RPB, single-SU case. In other words, each RPB is serving a subset of the SUs where these subsets are pair-wise disjoint and the expected decision cost incurred is different for each RPB at each discrete-time period. Under such a situation, an optimal RPB-relocation policy will, in general, prescribe different decisions for each RPB that are tailored based on the assigned SU's movements. From an operational standpoint, these two examples illustrate the extreme options in terms of refueling support provided to the SUs; the former provides the maximum redundancy in terms of the number of RPBs serving each SU while the latter provides zero redundancy.

The decision-maker can determine the assignment relationships via two approaches – decentralized and centralized. The decentralized (and more trivial) approach requires the assignment to be hardwired for the entire duration of operations (e.g., RPB 1 is assigned to SU 1, RPB 2 to SU 2 etc.). Under this approach, we can obtain an optimal relocation policy for each RPB individually by solving the linear program for the single-RPB, single-SU case or the single-RPB, multiple-SU case as appropriate. The decentralized approach is itself a heuristic approach in the sense that although optimality is achieved by considering the RPBs individually, combining the individual optimal RPB-relocation policies into a single policy may not necessarily be an optimal solution for the system as a whole. On the other hand, the centralized approach takes into account the movements of all the SUs collectively, and then determines the assignment relationships dynamically at each discrete-time period such that the overall expected decision cost is minimized. In other words, the centralized approach gives an optimal RPB-relocation policy for the system as a whole.

Here, we outline how the LP formulation can be modified to solve for an optimal RPB-relocation policy under the centralized approach. To simplify the problem, consider the case where there are $R = 2$ RPBs serving $M = 2$ SUs. Define $a_1 = (a_1^1, a_1^2),\ a_1^m \in \{0,1\},\ m = 1,2$ to be the assignment vector for RPB 1; $a_1^m = 1$ indicates that RPB 1 is assigned to SU m while $a_1^m = 0$ indicates that RPB 1 is not assigned to SU m. A state of the MDP is represented by (i_1, i_2, j_1, j_2) or $(I, J) \in S$, where $I = (i_1, i_2)$ represents the locations of RPB 1 and RPB 2 respectively. The decision taken at each discrete-time period will not only include the relocation of each RPB, but also the assignment of each RPB. To ensure that SU m is being served by at least y_m RPBs, $y_m \le R$, we need to ensure $\sum_{r=1}^{2} a_r^m \ge y_m\ \forall m$. Thus, the decision sets are given by $D_{(I,J)} = \{(d,a) = (d_1, d_2, a_1, a_2) \mid d = (d_1, d_2) \in N \times N,\ \sum_{r=1}^{2} a_r^m \ge y_m\ \forall m\}$. Assuming the SUs move independently, the state-transition probability $Q[(I,J),(d,L)]$ is given by $\prod_{m=1}^{2} P_m[j_m, l_m]$. The expected decision cost $c((I,J),(d,a))$ is given by $\sum_{r=1}^{2} (F[i_r, d_r] + \sum_{m=1}^{2} a_r^m (\sum_{l \in N} P_m[j_m, l] G[d_r, l]))$. In general, there are $n^{M+2R} 2^{MR}$ decision variables $\rho((I,J),(d,a))$ in the linear program. The solution to this LP formulation yields an optimal RPB-relocation policy K^*, where $K^*[I,J] = (d^*, a^*)$ represents the prescribed decisions d^* to relocate the RPBs, as well as the accompanying assignment relationships given by a^*, when the RPBs are at locations given by I and the SUs are at locations given by J.

An optimal RPB-relocation policy was solved on my personal laptop for the case involving two RPBs and two SUs, with the condition that the assignment relationship is one RPB to one SU. In other words, in each discrete-time period, we have either RPB 1 assigned to SU 1, RPB 2 assigned to SU 2 (i.e., $a_1^1 = 1,\ a_1^2 = 0,\ a_2^1 = 0,\ a_2^2 = 1$) or RPB 1 assigned to SU 2, RPB 2 assigned to SU 1 (i.e., $a_1^1 = 0,\ a_1^2 = 1,\ a_2^1 = 1,\ a_2^2 = 0$). The largest value of n for which the linear program can be readily solved using my personal laptop is 4; the corresponding number of decision variables is $4^6 \times 2 = 8{,}192$. For $n = 5$, the number of decision variables is $5^6 \times 2 = 31{,}250$; a more powerful computer is more suited to readily solve larger RPB-relocation problems.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Atkinson M, Kress M, Szechtman RS (2016) Logistic model for distributed lethality. Naval Research Program (NRP) report. http://hdl.handle.net/10945/56307

Kemeny JG, Snell JL (1960) *Finite Markov Chains* (Van Nostrand, Princeton, New Jersey).

Mannila SJ (2018) Multi-commodity logistic model for distributed lethality. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, California. http://hdl.handle.net/10945/58334

Puterman ML (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, Inc., Hoboken, New Jersey).

Rosenthal RE, White JA, Young D (1978) Stochastic dynamic location analysis. *Management Science* 24(6):645-653, https://doi.org/10.1287/mnsc.24.6.645

Rowden T, Gumataotao P, Fanta P (2015) "Distributed Lethality". *Proceedings* Vol. 141/1/1, 343, https://www.usni.org/magazines/proceedings/2015/january/distributed-lethality

Walton TA, Schramm H, Boone R (2019) Sustaining the fight: Resilient maritime logistics for a new era. Center for Strategic and Budgetary Assessments, https://csbaonline.org/research/publications/sustaining-the-fight-resilient-maritime-logistics-for-a-new-era/publication/1

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California