



**NAVAL  
POSTGRADUATE  
SCHOOL**

**MONTEREY, CALIFORNIA**

**THESIS**

**A GAME-THEORETIC MODEL FOR REPEATED  
HELICOPTER ALLOCATION BETWEEN TWO SQUADS**

by

Clifton G. Lennon  
Jason M. McGowan

June 2006

Thesis Advisor:  
Second Reader:

Kyle Y. Lin  
Steven E. Pilnick

**Approved for public release; distribution is unlimited.**

THIS PAGE INTENTIONALLY LEFT BLANK

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved OMB No. 0704-0188</i>	
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.				
<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> June 2006	<b>3. REPORT TYPE AND DATES COVERED</b> Master's Thesis	
<b>4. TITLE AND SUBTITLE</b> A Game-Theoretic Model for Repeated Helicopter Allocation Between Two Squads			<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> Lennon, Clifton G. and McGowan, Jason M.				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Naval Postgraduate School Monterey, CA 93943-5000			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> N/A			<b>10. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> 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.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited.			<b>12b. DISTRIBUTION CODE</b>	
<b>13. ABSTRACT (maximum 200 words)</b> A platoon commander has a helicopter to support two squads, which encounter two types of missions—critical or routine—on a daily basis. During a mission, a squad always benefits from having the helicopter, but the benefit is greater during a critical mission than during a routine mission. Because the commander cannot verify the mission type beforehand, a selfish squad would always claim a critical mission to compete for the helicopter—which leaves the commander no choice but to assign the helicopter at random. In order to encourage truthful reports from the squads, we design a token system that works as follows. Each squad keeps a token bank, with tokens deposited at a certain frequency. A squad must spend either 1 or 2 tokens to request the helicopter, while the commander assigns the helicopter to the squad who spends more tokens, or breaks a tie at random. The two selfish squads become players in a two-person non-zero-sum game. We find the Nash Equilibrium of this game, and use numerical examples to illustrate the benefit of the token system.				
<b>14. SUBJECT TERMS</b> Game theory, Nash equilibrium, Markov Chain			<b>15. NUMBER OF PAGES</b> 53	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> UL	

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release; distribution is unlimited.**

**A GAME-THEORETIC MODEL FOR REPEATED HELICOPTER  
ALLOCATION BETWEEN TWO SQUADS**

Clifton G. Lennon  
Ensign, United States Navy  
B.S. United States Naval Academy, 2005

Jason M. McGowan  
Ensign, United States Navy  
B.S. United States Naval Academy, 2005

Submitted in partial fulfillment of the  
requirements for the degree of

**MASTER OF SCIENCE IN APPLIED SCIENCE  
(OPERATIONS RESEARCH)**

from the

**NAVAL POSTGRADUATE SCHOOL  
June 2006**

Authors: Clifton G. Lennon  
Jason M. McGowan

Approved by: Kyle Y. Lin  
Thesis Advisor

Steven E. Pilnick  
Second Reader

James N. Eagle  
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

## ABSTRACT

A platoon commander has a helicopter to support two squads, which encounter two types of missions—critical or routine—on a daily basis. During a mission, a squad always benefits from having the helicopter, but the benefit is greater during a critical mission than during a routine mission. Because the commander cannot verify the mission type beforehand, a selfish squad would always claim a critical mission to compete for the helicopter—which leaves the commander no choice but to assign the helicopter at random.

In order to encourage truthful reports from the squads, we design a token system that works as follows. Each squad keeps a token bank, with tokens deposited at a certain frequency. A squad must spend either 1 or 2 tokens to request the helicopter, while the commander assigns the helicopter to the squad who spends more tokens, or breaks a tie at random. The two selfish squads become players in a two-person non-zero-sum game. We find the Nash Equilibrium of this game, and use numerical examples to illustrate the benefit of the token system.

THIS PAGE INTENTIONALLY LEFT BLANK

## **THESIS DISCLAIMER**

The reader is cautioned that computer programs developed in this research may not have been exercised for all cases of interest. While every effort has been made, within the time available, to ensure that the programs are free of computational and logic errors, they cannot be considered validated. Any application of these programs without additional verification is at the risk of the user.

THIS PAGE INTENTIONALLY LEFT BLANK

# TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION.....</b>	<b>1</b>
1.1	<b>MATHEMATICAL MODEL .....</b>	<b>2</b>
1.2	<b>RELATED RESEARCH.....</b>	<b>3</b>
1.3	<b>CONTRIBUTION.....</b>	<b>4</b>
1.4	<b>THESIS ORGANIZATION.....</b>	<b>4</b>
<b>II.</b>	<b>SQUAD’S STANDPOINT.....</b>	<b>7</b>
2.1	<b>A MARKOV CHAIN MODEL.....</b>	<b>7</b>
2.2	<b>STEADY-STATE BEHAVIOR OF THE MARKOV CHAIN .....</b>	<b>12</b>
2.3	<b>THE NASH EQUILIBRIUM.....</b>	<b>13</b>
<b>III.</b>	<b>COMMANDER’S STANDPOINT .....</b>	<b>19</b>
3.1	<b>TOKEN REPLENISHMENT PROBABILITY .....</b>	<b>19</b>
3.2	<b>TOKEN BANK CAPACITY .....</b>	<b>22</b>
3.3	<b>SENSITIVITY ANALYSIS .....</b>	<b>23</b>
	3.3.1 <b>Adjusting Routine Mission Probability .....</b>	<b>26</b>
	3.3.2 <b>Adjusting Critical Mission Probability.....</b>	<b>26</b>
	3.3.3 <b>Adjusting Reward Values.....</b>	<b>27</b>
<b>IV.</b>	<b>CONCLUSION .....</b>	<b>29</b>
4.1	<b>FINDINGS .....</b>	<b>29</b>
4.2	<b>IMPROVEMENTS .....</b>	<b>29</b>
4.3	<b>EXTENSIONS.....</b>	<b>30</b>
	<b>LIST OF REFERENCES .....</b>	<b>31</b>
	<b>INITIAL DISTRIBUTION LIST .....</b>	<b>33</b>

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF FIGURES

Figure 1.	Transition diagram for a squad with $c = 2$ , $c_1 = 4$ , and $c_2 = 6$ .....	11
Figure 2.	Optimal policy for each squad when varying $m$ using the baseline example in Table 1. ....	14
Figure 3.	Optimal policy for each squad when varying $\mu$ using the baseline example in Table 1. ....	15
Figure 4.	Optimal policy for each squad when varying $p_1$ using the baseline example from Table 1.....	16
Figure 5.	Optimal policy for each squad when varying $p_2$ using the baseline example from Table 1.....	17
Figure 6.	Effect of varying $\mu$ on helicopter benefit when using parameters from baseline example in Table 1.....	21
Figure 7.	Optimal replenishment probabilities ( $\mu^*$ ) for $2 \leq m \leq 20$ when using parameters from baseline example in Table 1. ....	22
Figure 8.	Change in helicopter benefit as $m$ increases when using $\mu^*$ for each $m$ , individual optimum and social optimum also shown. ....	25
Figure 9.	Increase in helicopter benefit when using token system relative to the individual optimum and on the interval between the individual optimum and the social optimum. ....	25

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF TABLES

Table 1.	Baseline example parameters.....	13
Table 2.	Effect of critical reward on squad policy using the baseline example in Table 1. ....	16
Table 3.	Decrease in helicopter benefit as $m$ increases.....	23
Table 4.	Sensitivity analysis on $p_1$ . ....	26
Table 5.	Sensitivity analysis on $p_2$ . ....	27
Table 6.	Sensitivity analysis on $r_2$ .....	28

THIS PAGE INTENTIONALLY LEFT BLANK

## NOTATIONS

$p_0$	probability of no mission
$p_1$	probability of a routine mission
$p_2$	probability of a critical mission
$\mu$	probability of token replenishment
$m$	maximum token bank capacity
$r_1$	reward value for helicopter usage during a routine mission
$r_2$	reward value for helicopter usage during a critical mission
$c$	squad's cutoff for spending 2 tokens for a critical mission
$c_1$	squad's cutoff for spending 1 token for a routine mission
$c_2$	squad's cutoff for spending 2 tokens for a routine mission
$q_k(0)$	proportion of time squad $k$ spends 0 tokens
$q_k(1)$	proportion of time squad $k$ spends 1 token
$q_k(2)$	proportion of time squad $k$ spends 2 tokens
$\lambda_k(1)$	probability of squad $k$ spending 1 token and getting the helicopter
$\lambda_k(2)$	probability of squad $k$ spending 2 tokens and getting the helicopter

THIS PAGE INTENTIONALLY LEFT BLANK

## EXECUTIVE SUMMARY

This thesis addresses the problem of a platoon commander in charge of two squads which encounter two types of missions, critical or routine. The squads may request support in the form of the platoon's sole helicopter. The commander does not know each squad's current mission type and must assign the helicopter based on each squad's report. During a mission, a squad always benefits from having the helicopter, but the benefit provided by the helicopter is greater during a critical mission than during a routine mission. The platoon commander wishes to maximize the overall benefit provided by the helicopter to both squads.

The platoon commander must rely on the report of a squad that is more interested in its own benefit from helicopter usage than the overall benefit provided by the helicopter. Because a squad always benefits from helicopter usage during a mission, a selfish squad leader would always request the helicopter when facing any mission, which forces the platoon commander to frequently assign the helicopter at random. Random assignment significantly lowers the helicopter's overall benefit because quite often the helicopter is assigned to the squad with a routine mission while the other squad faces a critical mission.

To improve the overall benefit provided by the helicopter, we design a token system to encourage truth-telling from each squad. The mathematical model is formulated as follows: Each squad has a token bank with a finite capacity. In each time period, a squad first finds out its mission type, if it has one, and then decides whether to spend 1 or 2 tokens to request the helicopter. A request is granted if the other squad spends fewer tokens; in case of a tie, the platoon leader assigns the helicopter at random. At the end of each time period, each squad receives a token with some probability set by the platoon leader, provided that the number of tokens does not exceed the token bank capacity. Because tokens are limited, a squad needs to decide how to use them wisely. In addition, the commander needs to decide the frequency of new token deposits and the token bank capacity in order to maximize the overall benefit between the two squads. Ideally, the commander wants a policy to force the squads to spend 1 token on a routine

mission and 2 tokens on a critical mission, so that he can always assign the helicopter to the squad who needs it the most thus maximizing the helicopter's overall benefit. Because each squad acts as a selfish agent, we model the competition between the two squads as a two-person non-zero-sum game.

This thesis addresses a theoretical problem that could be adapted to model actual military problems. Although this study is not based on a previously observed problem, it has implications for any problem concerning repeated allocation of a resource to multiple parties when each party is only concerned with its own utility. When there are two squads, we show that the token bank system is extremely useful when a high probability of mission (sum of routine mission probability and critical mission probability) exists. In a typical combat situation, use of the token system allows the commander to achieve over 90% of the difference between the social optimum and the individual optimum. When there is a high probability of neither critical nor routine missions occurring, the increase in expected helicopter benefit provided by the token-bank system is very small.

Areas for future research include improving the runtime on our algorithm for finding the commander's optimal token replenishment probability, studying asymmetric squads that face different combat scenarios, and expanding the problem to incorporate more than two squads.

## I. INTRODUCTION

This thesis addresses the problem of a platoon commander in charge of two squads which encounter two types of missions, critical or routine. The squads may request support in the form of the platoon's sole helicopter. The commander does not know each squad's current mission type and must assign the helicopter based on each squad's report. During a mission, a squad always benefits from having the helicopter, but the benefit provided by the helicopter is greater during a critical mission than during a routine mission. The platoon commander wishes to maximize the long-run overall benefit provided by the helicopter to both squads.

The platoon commander must rely on the report of a squad which is more interested in its own long-run benefit than the overall benefit provided by the helicopter. Because a squad always benefits from helicopter usage during a mission, a selfish squad leader would request the helicopter every time the squad faces a mission, which forces the platoon commander to frequently assign the helicopter at random. Random assignment significantly lowers the helicopter's overall benefit because quite often the helicopter is assigned to the squad with a routine mission while the other squad faces a critical mission. We study a mechanism implemented by the platoon commander to improve the helicopter's overall benefit.

To improve the benefit provided by the helicopter, we design a token system to encourage truth-telling from each squad. The mathematical model is formulated as follows: Each squad has a token bank with a finite capacity. In each time period, a squad first finds out its mission type, if it has one, and then decides whether to spend one or two tokens to request the helicopter. A request will be granted if the other squad spends fewer tokens; in case of a tie, the platoon leader assigns the helicopter at random. At the end of each time period, each squad receives a token with some probability set by the platoon leader, provided that the number of tokens does not exceed the token bank capacity. Because tokens are limited, a squad needs to decide how to use them wisely. In addition, the commander needs to decide the frequency of new token deposits, and the token bank capacity in order to maximize the overall benefit between the two squads. Ideally, the commander wants a policy to force the squads to spend 1 token on a routine mission and

2 tokens on a critical mission, so that he can always assign the helicopter to the squad who needs it the most thus maximizing the helicopter's benefit.

From a squad's standpoint, the state can be defined as the number of tokens in its bank. The squad's policy is the rule that tells the squad whether to request the helicopter and how many tokens to spend based on its token bank balance and its mission type. We use a two-person non-zero-sum game to describe the competition between the two squads and find its Nash equilibrium. Finally, we look at the problem from the platoon commander's standpoint, and select the token bank capacity and token replenishment probability to maximize the overall benefit provided by the helicopter.

This study provides an answer to a theoretical problem that could be adapted to model actual military problems. Although this study is not based on a previously observed problem, it has implications for any problem concerning repeated allocation of a resource to multiple parties when each party is only concerned with its own utility. When there are two squads, we show that the token bank system is extremely useful when a high probability of mission (sum of routine mission probability and critical mission probability) exists. When there is a high probability of no mission, the increase in expected benefit provided by the token bank system is very small.

## 1.1 MATHEMATICAL MODEL

Consider a platoon leader equipped with a helicopter to support the missions of two squads, squad A and squad B, in a discrete-time model. In each time period, a squad faces a critical mission with probability  $p_2$ , a routine mission with probability  $p_1$ , or no mission with probability  $p_0$ , where  $p_0 + p_1 + p_2 = 1$ . The mission types between time periods are independent, as well as mission types between the two squads. A squad's reward value for completion of a routine mission with helicopter support is  $r_1$ , and the reward value for completion of a critical mission with helicopter support is  $r_2$ . Without loss of generality, the reward value for completion of either type of mission without helicopter support is 0. The difficulty of a critical mission and the increase in the helicopter's relative benefit causes  $r_2$  to be greater than  $r_1$ .

Each squad keeps a token bank with maximum capacity  $m$ . The commander awards each squad a token at the end of each time period with probability  $\mu$ , and whether

squad A receives a token is independent of whether squad B receives a token. At the beginning of each time period, a squad can spend 1 or 2 tokens to request the helicopter. For a given  $\mu$ , and  $m$ , a squad's policy is a function that maps from the decision space (mission type faced and number of tokens in the bank) to the action space (spend 0, 1, or 2 tokens). Because  $r_2 > r_1$ , we let a squad always spend at least 1 token on a critical mission unless it does not have a token, and we denote  $c$  the minimum number of tokens a squad must have to spend 2 tokens on a critical mission. When facing a routine mission, let  $c_1$  and  $c_2$  denote the minimum number of tokens a squad must have to request the helicopter with 1 and 2 tokens respectively.

The parameters  $p_0, p_1, p_2, r_1$ , and  $r_2$  are determined by the nature of the combat situation. The goal of each squad is to select  $c, c_1$ , and  $c_2$  to maximize its long-run average reward while competing for the same helicopter in a two-person non-zero-sum game. The goal of the platoon leader is to select  $\mu$  and  $m$  so that the overall long-run average benefit provided by the helicopter is maximized.

## 1.2 RELATED RESEARCH

Our research problem is similar to the classic prisoner's dilemma. If the two squads cooperate by always reporting truthfully, each squad's benefit is maximized. However, the individual optimal policy requires each squad to always request the helicopter when facing a mission. The novelty of our research is to design a mechanism to encourage truth-telling in a repeated assignment problem. To the best of our knowledge, our work is the first to study the repeated assignment problem in a game-theoretic framework.

Previous work concerning the repeated assignment problem studies a single decision maker, who assigns workers to jobs to maximize expected reward. For example, Righter (1989) considers the assignment of activities to resources which arrive according to a Poisson process. Derman (1972) considers the assignment of men to jobs with random values. Other examples include the work by Albright (1972, 1974). We consider a repeated assignment problem over an infinite-time horizon. The major distinction of our problem is that there are two squads competing for the same helicopter, so that each squad's optimal policy depends on the other's policy.

From the game-theoretic standpoint, our work fits in the category of one manager (platoon commander) versus multiple selfish agents (squads). This type of relationship has been studied primarily in the context of telecommunications. Chakravorti (1994) considers the problem of a manager of an M/M/1 queue who seeks optimal flow control of jobs arriving from selfish users with private information who are also myopic optimizers. Lin (2003) uses a game-theoretic approach to model admission control in a single server system with multiple gatekeepers. He uses an  $n$ -person non-zero-sum game in which each gatekeeper wishes to maximize its own long-run average reward. In these works, the manager can charge a fee for a service so that the individual optimality coincides with the social optimality. The mechanism we design does not rely on a service fee.

### **1.3 CONTRIBUTION**

The contribution of this thesis is twofold. First, we study a repeated assignment problem in a game-theoretic framework with multiple selfish agents. Second, we design a mechanism to encourage truth-telling that does not involve charging a fee to the agent. This problem proves relevant to any manager who must distribute a limited amount of some resource to a greater number of agents with the goal of optimizing that resource's benefit. Although our problem deals with a two-person game, it can be expanded to an  $n$ -person game. We believe that our token mechanism will become more effective as the number of squads increases relative to the number of helicopters.

### **1.4 THESIS ORGANIZATION**

In Chapter II, we discuss the interaction between the two squads and find the Nash equilibrium of the game. We do this by finding squad A's optimal policy assuming squad B does not exist. We then find squad B's optimal policy based on squad A's optimal policy. Squad B's new policy causes squad A to change its policy, and so on. This process continues until the game reaches the Nash equilibrium, and neither squad has any motivation to further change its policy.

In Chapter III, we find the platoon commander's optimal selection for token bank capacity and token replenishment probability. We develop an algorithm to compute this

optimal strategy. As the platoon commander adjusts these constraints, the policies of the squads again change. Therefore, the squads must reach a new Nash equilibrium each time the commander adjusts the token bank capacity or the replenishment probability. The goal of the platoon commander is to maximize the overall benefit provided by the helicopter.

We present our conclusions in Chapter IV, discuss some interesting findings, and present ideas for further research.

THIS PAGE INTENTIONALLY LEFT BLANK

## II. SQUAD'S STANDPOINT

This chapter analyzes the helicopter-sharing problem from the standpoint of a squad. The two squads are selfish agents participating in a two-person non-zero-sum game in which each squad wishes to maximize its own long-term benefit from helicopter usage. Each squad only controls its own cutoff values for spending tokens to request the helicopter; all other parameters are fixed by the commander or the nature of the combat situation. We assume both squads are rational players. Therefore, each squad chooses the policy that maximizes its own long-run average payoff. Since the policy of squad A affects the policy of squad B and vice versa, the choosing of a policy by one squad causes the other squad to choose a new policy. If at some point, each squad's policy is the best response to the other squad's policy, then no squad has motivation to further change its policy. A pair of such policies is called a Nash equilibrium.

The rest of this chapter is organized as follows: In Section 2.1, we use a Markov chain to describe the squad's behavior. In Section 2.2, we analyze this Markov chain and find its steady-state behavior. In Section 2.3, we find the Nash equilibrium between the two squads. The techniques used to analyze a Markov chain can be found in many textbooks such as Ross (2003).

### 2.1 A MARKOV CHAIN MODEL

Recall that a policy for a squad can be delineated by three parameters  $c$ ,  $c_1$ , and  $c_2$ . We define  $c$  as the minimum number of tokens a squad must have to spend 2 tokens on a critical mission. When facing a routine mission, let  $c_1$  and  $c_2$  denote the minimum number of tokens a squad must have to request the helicopter with 1 and 2 tokens respectively. We assume that a squad always spends at least 1 token on a critical mission.

Define a squad's state as the number of tokens in its token bank at the beginning of a period. For a given policy, the evolution of a squad's state satisfies the Markov property, because the future is conditionally independent of the past given the present. Hence, we model a squad's state evolution as a discrete-time Markov chain. We derive

the probabilities that a squad moves from one state to another during one time period called the one (time) step transition probabilities. These probabilities depend on the squad's policy, the mission probabilities, and the token replenishment probability. We use these transition probabilities to build an  $m+1 \times m+1$  transition matrix, where  $m$  is the token bank capacity. We use the transition probability matrix to find the limiting probability for each state, which is the long-run proportion of time the process is in that state.

Denote a squad's state in period  $n$  by  $X_n$ , and then  $\{X_n; n = 0, 1, \dots\}$  is a Markov chain. The state space of this Markov chain is  $\{0, 1, \dots, m\}$ . Since our process satisfies the Markov property, define  $P_{ij} = P\{X_{n+1} = j | X_n = i\}$ . The  $P_{ij}$  values are the one (time) step transition probabilities; therefore, they give the probability of the squad transitioning from state  $i$  to state  $j$  during one time period. Let  $P$  denote a square matrix consisting of entries  $P_{00}$  to  $P_{mm}$  where  $m$  is the maximum token bank capacity. Row  $n$  in the matrix contains entries  $P_{n0} \dots P_{nm}$ . Each row in  $P$  must sum to 1, and each entry must be between 0 and 1.

During one time period a squad can either remain in the same state (its token balance does not change), or it can transition to another state. We determine each transition probability from the squad's policy, the token replenishment probability, and the mission probabilities. The transition diagram in Figure 1 gives a generic example of each transition probability for a squad with  $c = 2$ ,  $c_1 = 4$ , and  $c_2 = 6$ . As stated earlier, we assume a squad always spends at least 1 token on a critical mission. We also assume that  $c_1 < c_2$  and  $c \leq c_2$ .

In state  $i$ , there are only four states the Markov chain can move to in the next time period, namely states  $i-2$ ,  $i-1$ ,  $i$ , and  $i+1$ . Four cases exist depending on a squad's policy.

Case 1:  $c_1 < c < c_2$

(i)  $i < c_1$ ,

$$P_{i,i-2} = 0$$

$$P_{i,i-1} = p_2(1-\mu)$$

$$P_{i,i} = (1-p_2)(1-\mu) + p_2\mu$$

$$P_{i,i+1} = (1-p_2)\mu$$

(ii)  $c_1 \leq i < c$ ,

$$P_{i,i-2} = 0$$

$$P_{i,i-1} = (p_1 + p_2)(1-\mu)$$

$$P_{i,i} = (1-p_1-p_2)(1-\mu) + (p_1 + p_2)\mu$$

$$P_{i,i+1} = (1-p_1-p_2)\mu$$

(iii)  $c \leq i < c_2$ ,

$$P_{i,i-2} = p_2(1-\mu)$$

$$P_{i,i-1} = p_1(1-\mu) + p_2\mu$$

$$P_{i,i} = (1-p_1-p_2)(1-\mu) + p_1\mu$$

$$P_{i,i+1} = (1-p_1-p_2)\mu$$

(iv)  $i \geq c_2$ ,

$$P_{i,i-2} = (p_1 + p_2)(1-\mu)$$

$$P_{i,i-1} = (p_1 + p_2)\mu$$

$$P_{i,i} = (1-p_1-p_2)(1-\mu)$$

$$P_{i,i+1} = (1-p_1-p_2)\mu$$

Case 2:  $c_1 = c < c_2$

(i)  $i < c_1 = c$ , same as (i) in case 1.

(ii)  $i = c = c_1$ , same as (iii) in case 1.

(iii)  $c < i < c_2$ , same as (iii) in case 1.

(iv)  $i \geq c_2$ , same as (iv) in case 1.

Case 3:  $c_1 < c = c_2$

(i)  $i < c_1$ , same as (i) in case 1.

(ii)  $c_1 \leq i < c = c_2$ , same as (ii) in case 1.

(iii)  $i = c = c_2$ , same as (iv) in case 1.

(iv)  $i > c_2$ , same as (iv) in case 1.

Case 4:  $c < c_1 < c_2$

(i)  $i < c$ , same as (i) in case 1.

(ii)  $c \leq i < c_1$ ,

$$P_{i,i-2} = p_2(1-\mu)$$

$$P_{i,i-1} = p_2\mu$$

$$P_{i,i} = (1-p_2)(1-\mu)$$

$$P_{i,i+1} = (1-p_2)\mu$$

(iii)  $c_1 \leq i < c_2$ , same as (iii) in case 1.

(iv)  $i \geq c_2$ , same as (iv) in case 1.



## 2.2 STEADY-STATE BEHAVIOR OF THE MARKOV CHAIN

The Markov chain developed in Section 2.1 is irreducible because all states communicate with each other. In addition, all states in the Markov chain are aperiodic. Hence, the Markov chain is regular, which implies that a unique positive limiting distribution exists. For each state  $j$ , let  $\pi_j$  denote its limiting probability. To find the limiting probabilities, we use *Matlab* to compute  $P^k$  for a large value of  $k$  until all rows converge to the same numbers.

Once we know the limiting probabilities, we can determine how often a squad spends 1 or 2 tokens to request the helicopter. For a given policy with  $c$ ,  $c_1$ , and  $c_2$  defined as before, the frequency squad  $k$  spends 1 token can be calculated as

$$q_k(1) = p_2 \sum_{i=1}^{c-1} \pi_i + p_1 \sum_{i=c_1}^{c_2-1} \pi_i. \quad (1)$$

In addition, the frequency the squad spends 2 tokens can be calculated as

$$q_k(2) = p_2 \sum_{i=c}^m \pi_i + p_1 \sum_{i=c_2}^m \pi_i. \quad (2)$$

It follows that

$$q_k(0) = 1 - q_k(1) - q_k(2). \quad (3)$$

Recall that each squad's goal is to maximize its own long-run average payoff. In order to calculate the long-run average payoff, we need to first calculate the probability a squad receives the helicopter when requesting it. Since the commander assigns the helicopter to the squad spending the most tokens or randomly breaks a tie, squad A receives the helicopter after spending 1 token only if squad B does not spend a token or squad B spends 1 token and the helicopter is randomly assigned to squad A. Therefore, the probability of squad A getting the helicopter when spending 1 token is

$$\lambda_A(1) = q_B(0) + \frac{q_B(1)}{2},$$

where  $q_B(0)$  and  $q_B(1)$  are squad B's probabilities of spending 0 and 1 tokens respectively as defined in Equations (3) and (1). Similarly, the probability of squad A getting the helicopter when spending 2 tokens is

$$\lambda_A(2) = q_B(0) + q_B(1) + \frac{q_B(2)}{2}.$$

Finally, we compute the long-run average payoff for squad A by conditioning on its state and whether squad A gets the helicopter according to its policy. Thus, squad A's long-term average payoff is

$$r_1 p_1 \left[ \left( \sum_{i=c_1}^{c_2-1} \pi_i \right) \left( q_B(0) + \frac{q_B(1)}{2} \right) + \left( \sum_{i=c_2}^m \pi_i \right) \left( q_B(0) + q_B(1) + \frac{q_B(2)}{2} \right) \right] + r_2 p_2 \left[ \left( \sum_{i=1}^{c-1} \pi_i \right) \left( q_B(0) + \frac{q_B(1)}{2} \right) + \left( \sum_{i=c}^m \pi_i \right) \left( q_B(0) + q_B(1) + \frac{q_B(2)}{2} \right) \right].$$

Squad B's payoff is calculated in the same manner. We can now determine a squad's optimal policy by searching through all feasible policies and finding the maximum payoff value.

### 2.3 THE NASH EQUILIBRIUM

The game's equilibrium is a pair of policies such that neither squad has motivation to change its policy. We start by finding squad A's optimal policy assuming squad B does not exist. Thus squad A's initial payoff would be

$$r_1 p_1 \left[ \left( \sum_{i=c_1}^{c_2-1} \pi_i \right) + \left( \sum_{i=c_2}^m \pi_i \right) \right] + r_2 p_2 \left[ \left( \sum_{i=1}^{c-1} \pi_i \right) + \left( \sum_{i=c}^m \pi_i \right) \right].$$

We then find squad B's optimal policy assuming that squad B has perfect knowledge of squad A's policy. Squad B's new policy causes squad A to change its policy, and so on. Usually both squads have the same optimal policy because the model is symmetric between two squads. We write a program in *Matlab* and usually can find the Nash equilibrium in seconds.

Table 1. Baseline example parameters.

$p_0$	$p_1$	$p_2$	$\mu$	$r_1$	$r_2$	$m$
0.30	0.50	0.20	0.90	1	8	20

We use the baseline example parameters from Table 1 to illustrate how our algorithm works to find the Nash equilibrium. Squad A's optimal policy assuming squad B does not exist is  $c_1 = 3$  (squad A never spends 2 tokens to request the helicopter since we assume squad B does not exist) which yields a payoff of 2.1000. Squad B's optimal response is  $c = 2$ ,  $c_1 = 7$ , and  $c_2 = 17$ , and squad B's payoff is 1.7347. Squad A responds to squad B by choosing a policy of  $c = 2$ ,  $c_1 = 7$ , and  $c_2 = 18$ , and squad A's payoff becomes 1.6852. Squad B responds with an identical policy of  $c = 2$ ,  $c_1 = 7$ , and  $c_2 = 18$  and has a payoff of 1.6879. Squad A does not change its policy, and it receives the same average payoff as squad B. Squad B then chooses to remain at the same policy, and the game has reached its Nash equilibrium with the helicopter providing an overall benefit of 3.3759.

Using the same baseline example from Table 1, we demonstrate the effects of varying some parameters on a squad's optimal policy. In most cases squad A and squad B have identical policies. However, in some cases the policies are slightly different. Figure 2 shows the change in the  $c$ ,  $c_1$ , and  $c_2$  cutoff values as  $m$  increases from 2 to 20. In Figure 3, we fix  $m = 20$  and increment  $\mu$  on  $[0.50, 1]$  by steps of 0.05. Table 2 shows the effect of varying  $r_2$  on the squad's policies. In Figure 4, we vary  $p_1$  while holding  $p_2$  constant, and we do the opposite in Figure 5.

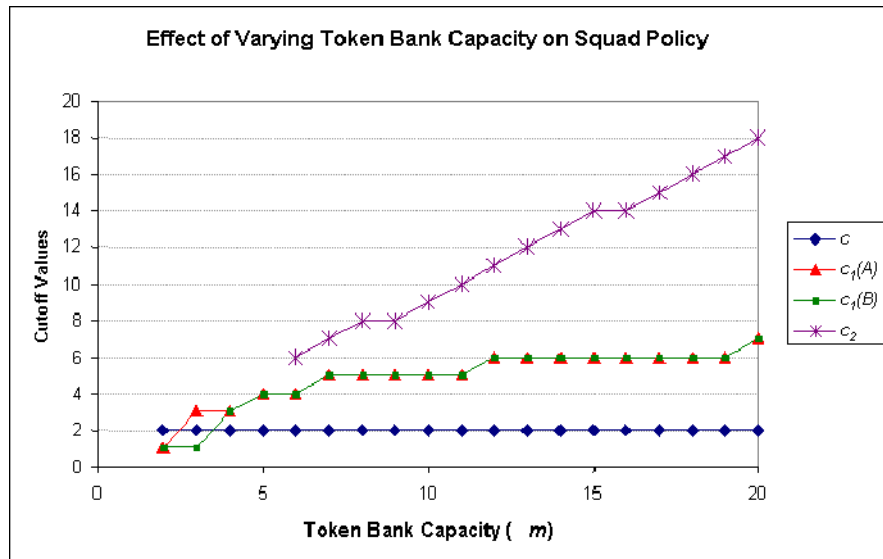


Figure 2. Optimal policy for each squad when varying  $m$  using the baseline example in Table 1.

Figure 2 shows that the squads are not willing to spend 2 tokens on a routine mission until  $m \geq 6$ , but they are always willing to spend 2 tokens on a critical mission. The routine cutoff values increase as  $m$  increases. The two squads have different policies when  $m = 3$ , otherwise the policies are identical. Usually the squads have identical policies since they are symmetric, but occasionally in the game's Nash equilibrium a squad's optimal response to the other squad's policy is a slightly different policy. The discrete nature of  $m$  and the cutoff values causes the squads' optimal policies to differ occasionally.

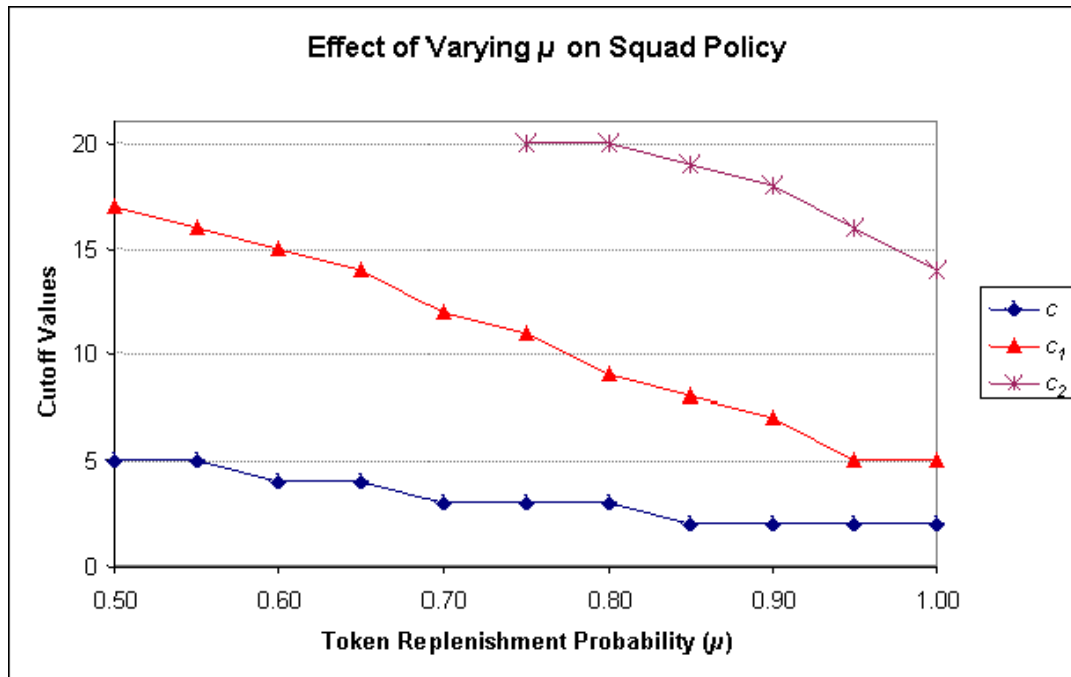


Figure 3. Optimal policy for each squad when varying  $\mu$  using the baseline example in Table 1.

As seen in Figure 3, the squads do not spend 2 tokens to request the helicopter during a routine mission until  $\mu \geq 0.75$ . The cutoff values decrease as  $\mu$  increases.

Table 2. Effect of critical reward on squad policy using the baseline example in Table 1.

$r_2$	$c$	$c_1$	$c_2$	Helicopter Benefit
2	3	5	18	1.2464
4	2	6	18	1.9566
8	2	7	18	3.3759
16	2	7	18	6.2190
32	2	8	18	11.8917

As seen in Table 2, an increase in the reward for helicopter usage during a critical mission makes the squads more willing to spend 2 tokens on a critical mission and less likely to request the helicopter for a routine mission.

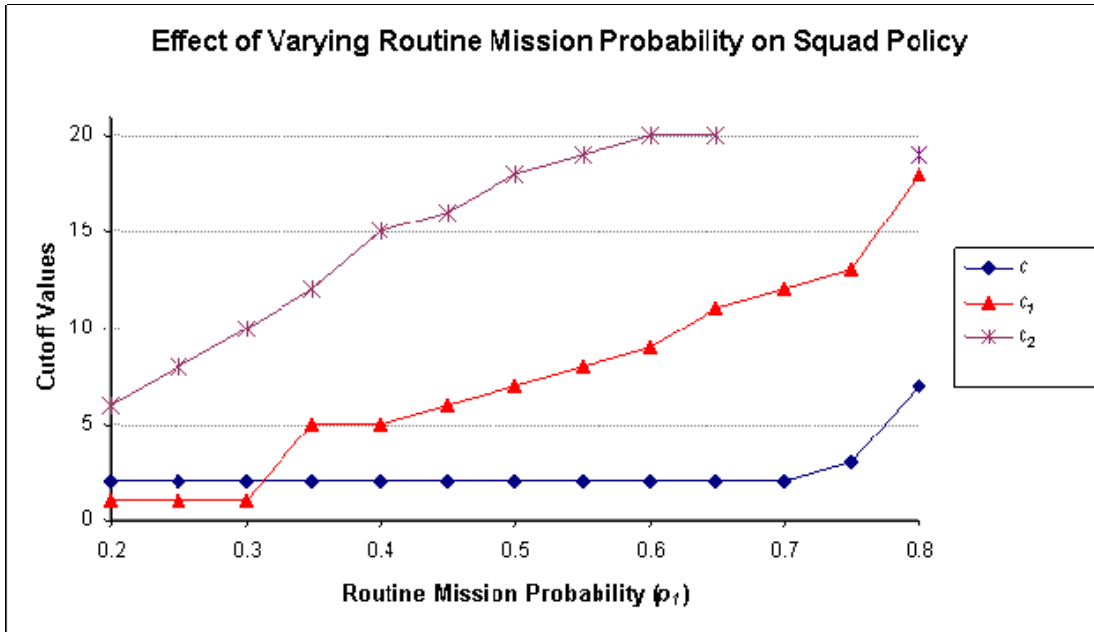


Figure 4. Optimal policy for each squad when varying  $p_1$  using the baseline example from Table 1.

As seen in Figure 4, the increase in  $p_1$  causes  $c_1$  and  $c_2$  to increase. For  $0.65 < p_1 < 0.80$ , the squads never spend 2 tokens on a routine mission. The squads always choose  $c = 2$  until  $p_1 \geq 0.75$ .

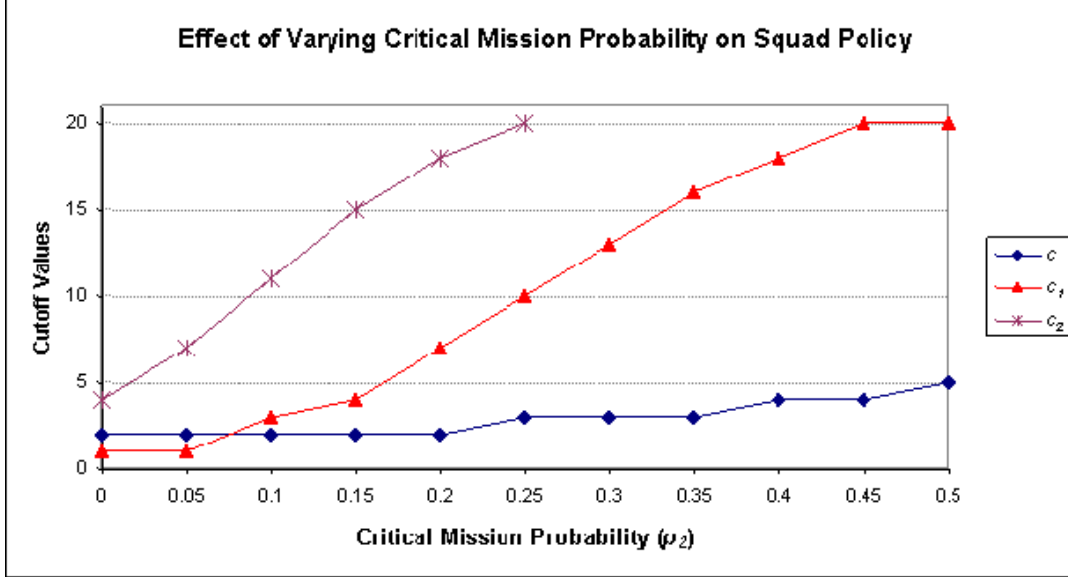


Figure 5. Optimal policy for each squad when varying  $p_2$  using the baseline example from Table 1.

As shown in Figure 5, an increase in  $p_2$  causes  $c$ ,  $c_1$ , and  $c_2$  to exhibit upward trends. The routine cutoff values increase such that the squads never spend 2 tokens on a routine mission when  $p_2 > 0.25$ , and they only spend 1 token on a routine mission with a full token bank when  $p_2 > 0.40$ . Once  $p_2 \geq 0.25$ ,  $c > 2$ .

As stated previously, the two policies in Nash equilibrium can be slightly different. For example, when  $p_0 = 0.30$ ,  $p_1 = 0.50$ ,  $p_2 = 0.20$ ,  $\mu = 0.90$ ,  $m = 3$ ,  $r_1 = 1$ , and  $r_2 = 8$  (as shown in Figure 2), these two policies form a Nash equilibrium: (A)  $c = 2$ , and  $c_1 = 3$  and (B)  $c = 2$ , and  $c_1 = 1$ . The squads do not spend 2 tokens on a routine mission in this example.

In a very rare occurrence, there does not exist a Nash equilibrium for the game. Such an occurrence typically involves three policies  $\alpha$ ,  $\beta$ , and  $\gamma$ , such that  $\beta$  is the best response to  $\alpha$ ,  $\gamma$  is the best response to  $\beta$ , while  $\alpha$  is the best response to  $\gamma$ . For example, when  $p_0 = 0.40$ ,  $p_1 = 0.40$ ,  $p_2 = 0.20$ ,  $\mu = 0.8874$ ,  $m = 9$ ,  $r_1 = 1$ , and  $r_2 = 4$ , the following cycle exists.

$$\alpha : c = 3, c_1 = 4, c_2 = 8$$

$$\beta : c = 2, c_1 = 4, c_2 = 7$$

$$\gamma : c = 3, c_1 = 4, c_2 = 7$$

THIS PAGE INTENTIONALLY LEFT BLANK

### III. COMMANDER'S STANDPOINT

This chapter analyzes the helicopter-sharing problem from the standpoint of the platoon commander. The commander wishes to maximize the overall average long-term benefit (sum of each squad's payoff) provided by the helicopter. Recall that once the commander decides on  $m$ , the token-bank capacity, and  $\mu$ , the replenishment probability, the two squads become players in the two-person non-zero-sum game described in Chapter II. The goal of the commander is to choose  $m$  and  $\mu$  such that the total benefit resulting from the Nash equilibrium in this two-person game is maximized.

The rest of the chapter is organized as follows: In Section 3.1, we fix  $m$  and find the value of  $\mu$  that maximizes the helicopter's benefit. In Section 3.2, we allow  $m$  to vary and discuss its effect on the helicopter's benefit. In Section 3.3, we present the game's individual optimum and social optimum, which are determined by the nature of the combat situation. We provide sensitivity analysis by changing the parameters of the combat situation and observing the effect on the commander's optimal policy.

#### 3.1 TOKEN REPLENISHMENT PROBABILITY

In this section we fix  $m$  and discuss the effect of varying  $\mu$ . The mission probabilities have the greatest effect on finding  $\mu^*$ , the optimal  $\mu$  that maximizes the total helicopter benefit. Ideally, the commander would like each squad to spend 2 tokens on a critical mission and 1 token on a routine mission so that the commander can always make the correct helicopter assignment. If a squad always requested truthfully, then the expected number of tokens that squad spends each time period is  $p_1 + 2p_2$  tokens. Since  $m$  is finite, the squad may have incentive to spend 2 tokens on a routine mission when its token bank is nearly full and to spend 1 token on a critical mission when its token bank has few tokens (in order to save tokens for possible future missions). As a consequence, the commander cannot force the squads to report truthfully no matter what values of  $m$  and  $\mu$  he chooses.

For a given  $m$ , we can evaluate the objective function—the total benefit provided by the helicopter between two squads—for  $\mu$  in  $[0,1]$  to find  $\mu^*$ . Because we assume the

objective function is unimodal in  $\mu$ , we use an algorithm employing the Golden Section search to find  $\mu^*$  more efficiently. Since  $\mu$  must be in  $[0,1]$ , we know that our algorithm provides an interval of width 0.0031 in which  $\mu^*$  can be found after 12 iterations. The algorithm goes as follows on the interval  $[a_l, b_l]$  for  $k = 1$ :

1. Set  $\alpha = \frac{\sqrt{5}-1}{2}$
2. Set  $\varphi_k = \mu_1 = a_k + (1-\alpha)(b_k - a_k)$
3. Set  $\rho_k = \mu_2 = a_k + \alpha(b_k - a_k)$
4. Each squad determines its optimal policy for  $\mu_1$  and  $\mu_2$ , and the commander compares the average helicopter benefit yielded by each  $\mu$ .  $(f(\varphi_k), f(\rho_k))$
5. Update
  - Case 1:  $f(\varphi_k) \geq f(\rho_k)$ 
    - i. Set  $a_{k+1} = a_k; \rho_{k+1} = \varphi_k; b_{k+1} = \rho_k$
    - ii. Set  $f(\rho_{k+1}) = f(\varphi_k)$
    - iii. Compute  $\varphi_{k+1} = a_{k+1} + (1-\alpha)(b_{k+1} - a_{k+1})$  and  $f(\varphi_{k+1})$
  - Case 2:  $f(\varphi_k) < f(\rho_k)$ 
    - i. Set  $a_{k+1} = \varphi_k; \varphi_{k+1} = \rho_k; b_{k+1} = b_k$
    - ii. Set  $f(\varphi_{k+1}) = f(\rho_k)$
    - iii. Compute  $\rho_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1})$  and  $f(\rho_{k+1})$
6. If  $b_{k+1} - a_{k+1} < \varepsilon$  end search,  $\mu^*$  is in  $[a_{k+1}, b_{k+1}]$ . Otherwise set  $k = k + 1$ , and go to Update.

Using the parameters given in Table 1, we investigate the effect of varying  $\mu$  on the helicopter's overall benefit. For this combat situation, we find  $\mu^* = 0.8773$ , and the average overall helicopter benefit is 3.3863. Figure 6 shows the helicopter's benefit improves as we increase  $\mu$  until  $\mu = \mu^*$ , then the overall benefit decreases.

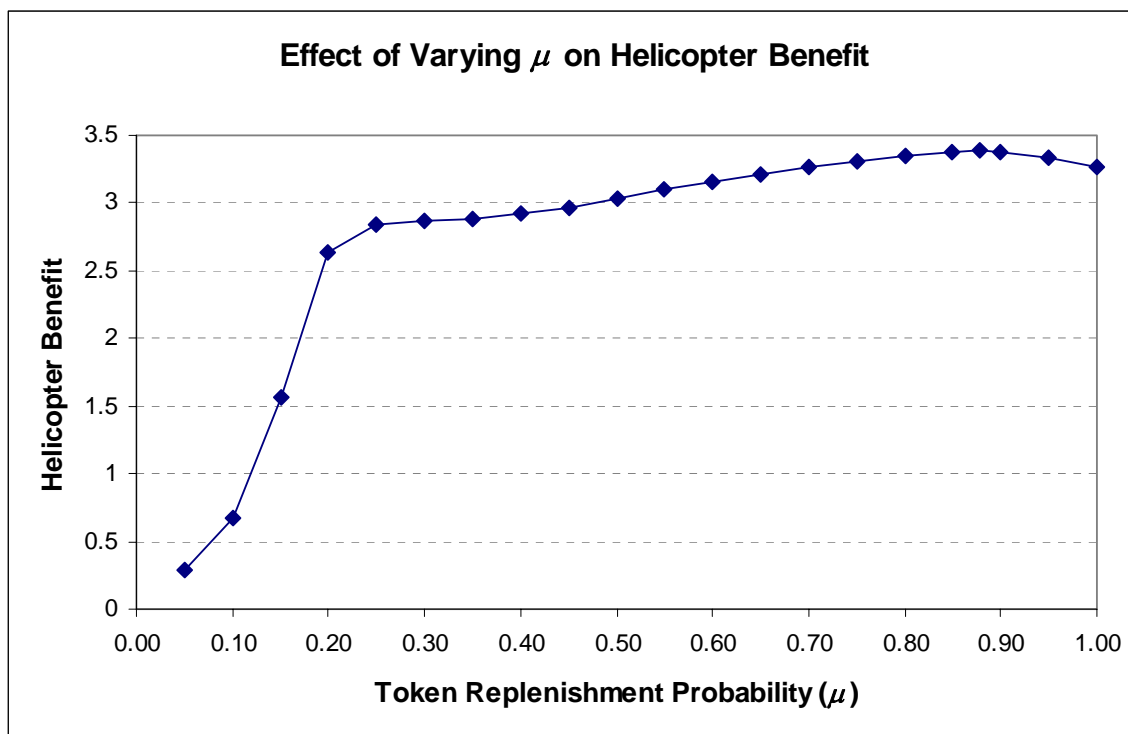


Figure 6. Effect of varying  $\mu$  on helicopter benefit when using parameters from baseline example in Table 1.

Using the parameters from Table 1, we increment  $m$  on  $[2, 20]$  and are able to find  $\mu^*$  using our Golden Section search algorithm for each  $m$ . Figure 7 shows  $\mu^*$  exhibiting a downward trend (it does not necessarily decrease monotonically) as it approaches a value slightly less than  $p_1 + 2p_2$ .

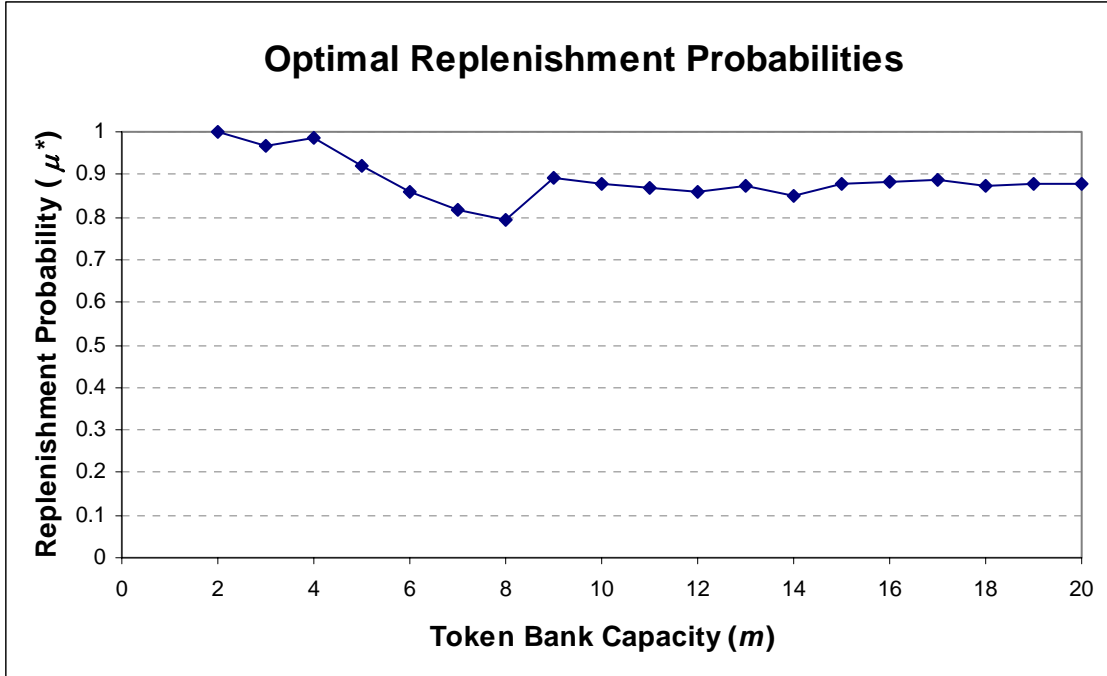


Figure 7. Optimal replenishment probabilities ( $\mu^*$ ) for  $2 \leq m \leq 20$  when using parameters from baseline example in Table 1.

### 3.2 TOKEN BANK CAPACITY

In this section we discuss how the total helicopter benefit changes as  $m$  changes. The overall long-term average benefit provided by the helicopter follows an upward trend as the commander raises  $m$ . However, it is not necessarily monotonically increasing. Eventually, as  $m$  continues to increase, the relative increase in helicopter benefit begins to decline. Since  $m$  must be finite, and it is unreasonable for it to be very large, the commander must develop a cutoff value for  $m$  based on the increase in the helicopter's benefit relative to  $m - 1$ .

Consider the baseline example from Table 1. Figure 8 shows overall helicopter benefit for each  $m$  on  $[0, 20]$  when the commander uses  $\mu^*$  for the given  $m$ . As stated earlier, helicopter benefit follows an upward trend as  $m$  increases.

Occasionally an increase in  $m$  causes a decrease in the overall helicopter benefit. This occasional decrease is attributed to the discrete nature of the cutoff values and that each squad has only a finite number of feasible policies. Table 3 shows the overall

helicopter benefit and each squad's policy when  $p_0 = 0.30$ ,  $p_1 = 0.50$ ,  $p_2 = 0.20$ ,  $r_1 = 1$ , and  $r_2 = 8$  for different  $m$  values. Both squads have the same policy in each example. Note that the commander can achieve a higher helicopter benefit by assigning  $m = 5$  than assigning  $m = 6$ .

Table 3. Decrease in helicopter benefit as  $m$  increases.

$m$	$\mu^*$	$c$	$c_1$	$c_2$	<i>Helicopter Benefit</i>
5	0.9187	2	4	6	3.3192
6	0.8572	2	4	7	3.3004
7	0.8154	3	5	8	3.3042
8	0.7945	3	6	9	3.3049
9	0.8936	2	5	9	3.3128
10	0.8792	2	5	10	3.3311

### 3.3 SENSITIVITY ANALYSIS

In this section we expand on the baseline example given in Table 1 by varying the combat parameters (mission probabilities and the critical mission reward value) and compare these results to the game's individual optimum and social optimum. If the commander does not employ some mechanism to encourage truth-telling, selfish squad leaders always request the helicopter when facing a mission. Therefore, the commander has no means of knowing the mission type of either squad. This lack of policy forces the commander to randomly assign the helicopter whenever both squads request it, which results in the game's individual optimum. This individual optimum can be calculated as the sum of each squad's long-run average payoff when the squads always request the helicopter for a mission:

$$2 \left[ \left( 1 - \frac{p_1 + p_2}{2} \right) (p_1 r_1 + p_2 r_2) \right].$$

To find the game's social optimum, we assume the squads are always truthful in their requests. A squad tells the commander the mission type it is facing, and the

commander assigns the helicopter to the squad that needs it most, or he randomly assigns the helicopter if both squads face the same mission type. The social optimum can be calculated as

$$2 \left[ p_1 \left( p_0 + \frac{p_1}{2} \right) r_1 + p_2 \left( p_0 + p_1 + \frac{p_2}{2} \right) r_2 \right].$$

We next compare the performance of our token bank policy with the individual and social optimum. We show that the token system greatly improves the helicopter's overall average benefit compared to the individual optimum during typical combat situations. As we increase the mission probabilities and the critical reward value, we show that the token system's benefit over the individual optimum increases. The usefulness of the token bank depends on the overall combat situation. If a very low probability of mission is coupled with a low critical reward value, the benefit provided by a token bank system may be trivial.

Using the baseline example given in Table 1, we calculate the individual optimum and social optimum as 2.73 and 3.43 respectively. Figure 8 shows the helicopter's overall benefit at  $\mu^*$  for each  $m$  and the individual optimum and social optimum as dictated by the combat situation. The token system always provides greater benefit than the individual optimum for these combat parameters. We can also compare the relative increase in the helicopter's overall benefit when the token system is employed. Figure 9 shows the increase in average helicopter benefit relative to the individual optimum and the increase in helicopter benefit on the interval between the individual optimum and the social optimum. When  $m = 20$ , the token system improves on the individual optimum by almost 25%, and it increases the helicopter's benefit over 90% of the feasible interval of improvement (region between individual optimum and social optimum). As we increase the mission probabilities and the critical reward value, we show in our sensitivity analysis that the token system provides even greater benefit relative to the individual optimum. In our sensitivity analysis we also study the effect of varying  $r_2$ ,  $p_1$ , and  $p_2$  on  $\mu^*$  and the optimal  $m$  ( $m^*$ ).

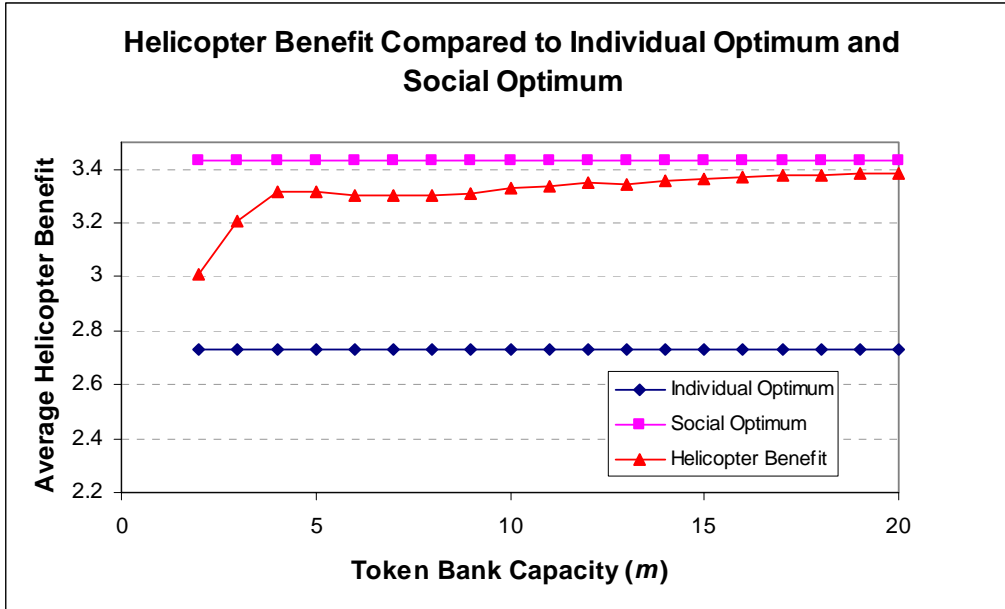


Figure 8. Change in helicopter benefit as  $m$  increases when using  $\mu^*$  for each  $m$ , individual optimum and social optimum also shown.

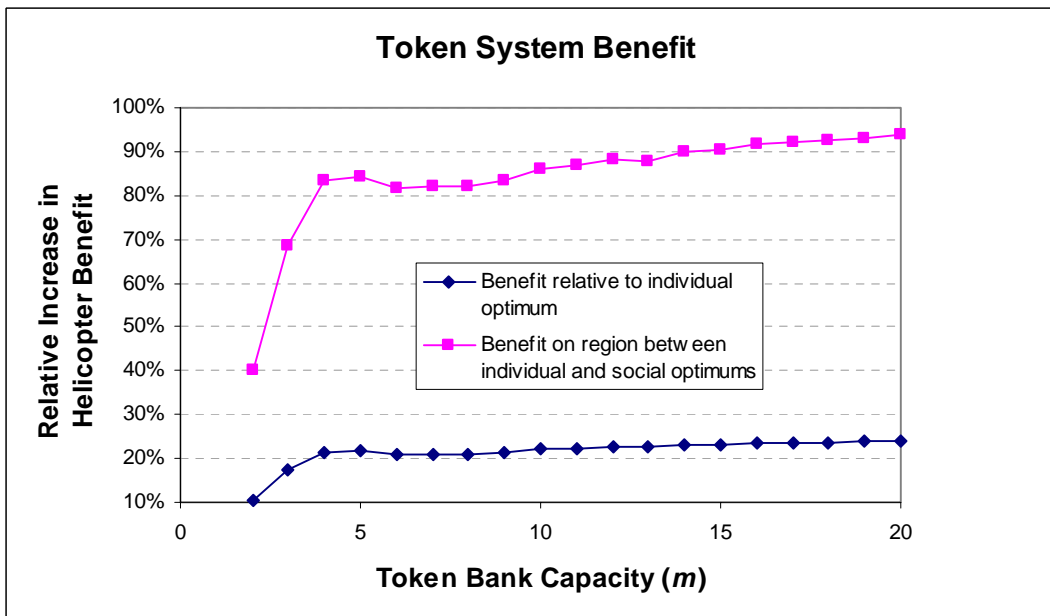


Figure 9. Increase in helicopter benefit when using token system relative to the individual optimum and on the interval between the individual optimum and the social optimum.

### 3.3.1 Adjusting Routine Mission Probability

Let  $p_2 = 0.20$ ,  $r_1 = 1$ ,  $r_2 = 8$ , and  $2 \leq m \leq 20$ . We adjust  $p_1$ , study the effect on  $\mu^*$  and  $m^*$ , and compare the results with the individual optimum and the social optimum. In Table 4, we show the results of this sensitivity analysis on  $p_1$ . The commander does not always choose  $m = 20$  as seen when  $p_1 = 0.20$ . For  $p_1 = 0.80$ ,  $m = 18, 19$ , or  $20$  all yield an equal average overall helicopter benefit. The commander would choose a larger  $m$  if allowed to do so because as shown earlier, helicopter benefit follows an upward trend as  $m$  increases. The optimal token replenishment probability,  $\mu^*$ , is near  $p_1 + 2p_2$  when  $p_1 + 2p_2 < 1$ , and it approaches 1 as  $p_1 + 2p_2$  becomes greater than 1. For  $p_1 = 0.80$ , the helicopter's benefit when using the token system is 45% greater than the individual optimum, and the token system increases the helicopter's benefit 96.38% on the feasible region of improvement (between the individual optimum and the social optimum).

Table 4. Sensitivity analysis on  $p_1$ .

$p_1$	$m^*$	$\mu^*$	Individual Optimum	Social Optimum	Helicopter Benefit with the Token System	Increased Benefit Relative to Individual Optimum	Increased Benefit Between Individual Optimum and Social Optimum
0.20	19	0.6200	2.88	3.16	3.0965	7.52%	77.32%
0.30	20	0.7020	2.85	3.27	3.2034	12.40%	84.14%
0.40	20	0.7891	2.80	3.36	3.3032	17.97%	89.86%
0.50	20	0.8773	2.73	3.43	3.3863	24.04%	93.76%
0.60	20	0.9718	2.64	3.48	3.4535	30.81%	96.85%
0.70	20	0.9988	2.53	3.51	3.4794	37.53%	96.88%
0.80	18-20	0.9988	2.40	3.52	3.4795	44.98%	96.38%

### 3.3.2 Adjusting Critical Mission Probability

Let  $p_1 = 0.50$ ,  $r_1 = 1$ ,  $r_2 = 8$ , and  $2 \leq m \leq 20$ . We now adjust  $p_2$ , study the effect on  $\mu^*$  and  $m^*$ , and compare the results with the individual optimum and the social optimum. We show our results in Table 5. The commander always chooses  $m = 20$  in these scenarios. For  $p_2 = 0.10$ ,  $\mu^*$  is near 0.70. As  $p_2$  increases,  $\mu^*$  is near  $p_1 + 2p_2$  until  $p_1 + 2p_2 > 1$  and  $\mu^*$  remains near 1. When comparing the token system's benefit to the

individual optimum, the increase in relative benefit is strictly increasing as  $p_2$  increases (approximately 33% when  $p_2 = 0.50$ ). The token system's increased benefit on the feasible region reaches approximately 95% when  $p_2 = 0.30$  then decreases slightly as  $p_2$  continues to increase.

Table 5. Sensitivity analysis on  $p_2$ .

$p_2$	$m^*$	$\mu^*$	Individual Optimum	Social Optimum	Helicopter Benefit with the Token System	Increased Benefit Relative to Individual Optimum	Increased Benefit Between Individual Optimum and Social Optimum
0.10	20	0.7001	1.82	2.17	2.1340	17.25%	89.71%
0.20	20	0.8773	2.73	3.43	3.3863	24.04%	93.76%
0.30	20	0.9988	3.48	4.53	4.4761	28.62%	94.87%
0.40	20	0.9988	4.07	5.47	5.3147	30.58%	88.91%
0.50	20	0.9888	4.50	6.25	6.0071	33.49%	86.12%

### 3.3.3 Adjusting Reward Values

Let  $p_1 = 0.50$ ,  $p_2 = 0.20$ ,  $r_1 = 1$ , and  $2 \leq m \leq 20$ . As stated earlier,  $r_2 > r_1$ . We increase  $r_2$  exponentially, study the effect on  $\mu^*$  and  $m^*$ , and compare the results with the individual optimum and the social optimum. We show our results in Table 6. The commander always chooses  $m = 20$  for these scenarios. His choice of  $\mu^*$  when  $r_2 = 2$  is approximately  $p_1 + 2p_2$  and decreases as  $r_2$  increases. In this example, the helicopter's benefit relative to the individual optimum, and the increased benefit on the region between the individual optimum and the social optimum are strictly increasing as  $r_2$  increases.

Table 6. Sensitivity analysis on  $r_2$ .

$r_2$	$m^*$	$\mu^*$	Individual Optimum	Social Optimum	Helicopter Benefit with the Token System	Increased Benefit Relative to Individual Optimum	Increased Benefit Between Individual Optimum and Social Optimum
2	20	0.9106	1.17	1.27	1.2475	6.62%	77.50%
4	20	0.8804	1.69	1.99	1.9581	15.86%	89.37%
8	20	0.8773	2.73	3.43	3.3863	24.04%	93.76%
16	20	0.8534	4.81	6.31	6.2472	29.88%	95.81%
32	20	0.8328	8.97	12.07	11.9895	33.66%	97.40%

We show in Section 3.2 that increasing  $m$  causes the average helicopter benefit to exhibit an upward trend. However, in Section 3.3 we only examine  $m$  such that  $2 \leq m \leq 20$ . This is because of the computing time required to run these scenarios with very large token bank capacities. When  $2 \leq m \leq 20$ , it takes several hours to find the corresponding  $\mu^*$  values. We further discuss this in Chapter IV when we suggest ideas for future research.

## IV. CONCLUSION

In this thesis we study the repeated assignment problem in a game-theoretic framework. The two squads are selfish agents in a two-person non-zero-sum game. As in the prisoner's dilemma, the socially optimal strategy yields a higher payoff for each player than the individually optimal strategy. We implement a token system to encourage the squads to truthfully report their mission type to the commander. We use discrete-time Markov chains to model a squad's state evolution. Other works which study a manager (platoon commander) versus multiple selfish agents (squads) from a game-theoretic framework require the manager to charge a service fee to encourage social optimality. We design a mechanism which does not rely on a service fee. The basis of our problem is theoretical, but its results can prove relevant for a manager repeatedly assigning a limited resource to multiple selfish agents.

### 4.1 FINDINGS

We develop an algorithm to find the commander's optimal token replenishment probability based on the combat situation and the size of the token bank. The commander cannot force the squads to always request truthfully. The desire of each squad to maximize its own payoff causes the Nash equilibrium of the game to always yield a lower average overall helicopter benefit than if the squads were truthful. For increasing  $m$ , the average helicopter benefit follows an upward trend. Numerical examples show that for typical combat scenarios, the benefit provided by the token bank system can be significant.

### 4.2 IMPROVEMENTS

We were unable to study the effects of a very large token bank capacity because of the required computing time to do so. Currently, the runtime on our algorithm for finding the optimal token replenishment probability increases exponentially as  $m$  increases. It takes several hours to find  $\mu^*$  for  $2 \leq m \leq 20$ . An improvement in the runtime of this algorithm would allow a more thorough examination of the effects of

raising  $m$ . We also assume that the helicopter's overall benefit is unimodal over all  $\mu$  for any given set of parameters. We came to this conclusion after working out numerous cases, but we did not prove this rigorously.

### **4.3 EXTENSIONS**

Several possible extensions to our work exist. The model could be modified for asymmetric squads such that each squad could have different mission probabilities and mission reward values. The problem could be expanded to an  $n$ -person non-zero-sum game. Other token systems are also possible. For instance, the commander could allow a squad to spend as many tokens as it wishes to request the helicopter. The commander could also deposit a new token with different probabilities depending on a squad's token balance. We expect these extensions to further shed light on repeated assignment problems with selfish agents.

## LIST OF REFERENCES

- Albright, S.C. 1974. Optimal sequential assignments with random arrival times. *Management Science* **21**, 60-67.
- Albright, S.C. and Derman, C. 1972. Asymptotic optimal policies for the stochastic sequential assignment problem. *Management Science* **19**, 46-51.
- Chakravorti, B. 1994. Optimal flow control of an M/M/1 queue with a balanced budget. *IEEE Transactions on Automatic Control* **39**, 1918-1921.
- Derman, C., Lieberman, G.J. and Ross, S.M. 1972. A sequential stochastic assignment problem. *Management Science* **18**, 349-355.
- Lin, K.Y. 2003. Decentralized admission control of a queueing system: A game-theoretic model. *Naval Research Logistics* **50**, 702-718.
- Righter, R. 1989. A resource allocation problem in a random environment. *Operations Research* **37**, 329-338.
- Ross, S. 2003. *Introduction to Probability Models*. Academic Press. 8<sup>th</sup> Edition.

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
3. Kyle Y. Lin  
Naval Postgraduate School  
Monterey, California
4. Steven E. Pilnick  
Naval Postgraduate School  
Monterey, California
5. Clifton G. Lennon  
Naval Postgraduate School  
Monterey, California
6. Jason M. McGowan  
Naval Postgraduate School  
Monterey, California