The Generalized Location Routing Problem with Profits for Planetary Surface Exploration and Terrestrial Applications

by

Jaemyung Ahn

B.S. in Aerospace Engineering, Seoul National University, 1997
M.S. in Aerospace Engineering, Seoul National University, 1999

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2008

Author ....
Department of Aeronautics and Astronautics
December 7, 2007

Certified by ....
Olivier L. de Weck
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Certified by ....
David Simchi-Levi
Professor of Civil and Environmental Engineering and Engineering Systems

Certified by ....
Jeffrey A. Hoffman
Professor of the Practice of Aeronautics and Astronautics

Certified by ....
Richard. H. Battin
Senior Lecturer of Aeronautics and Astronautics

Accepted by ....
David L. Darmofal
Associate Professor of Aeronautics and Astronautics
Chairman of the Graduate Committee

The Generalized Location Routing Problem with Profits for Planetary Surface Exploration and Terrestrial Applications

by

Jaemyung Ahn

Submitted to the Department of Aeronautics and Astronautics
on December 7, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

As the scale of space exploration gets larger, planning of planetary surface exploration becomes more complex and campaign-level optimization becomes necessary. This is a challenging profit maximization problem whose decisions encompass selection of bases, technological options, routes, and excursion methods under constraints on a route, a mission, and a whole campaign.

The Generalized Location Routing Problem with Profits (GLRPP) is developed in this thesis as a framework to solve this campaign optimization problem. A mathematical formulation for the GLRPP is developed and two solution methods to solve the GLRPP - a single phase method and a three-phase method - are presented. Numerical experiments for these two solution methods are carried out and their performance in terms of efficiency and effectiveness are analyzed.

Two case studies are carried out. The first case study is a global Mars surface exploration campaign optimization. Problem instances for 100 potential bases and 1000 potential exploration sites are successfully solved using a three-phase solution method. A methodology to express the incremental value of a technology using exploration profits is demonstrated to evaluate an orbiting depot and in-situ resource utilization (ISRU). The second case study is a college football recruiting problem. A GLRPP instance is created out of the NCAA football division I-A schools and airports from which the schools can be reached. The problem is successfully solved using the three-phase solution method within a very small optimality gap.

Thesis Supervisor: Olivier L. de Weck
Title: Associate Professor of Aeronautics and Astronautics and Engineering Systems

Thesis Supervisor: David Simchi-Levi
Title: Professor of Civil and Environmental Engineering and Engineering Systems

Thesis Supervisor: Jeffrey A. Hoffman

Title: Professor of the Practice of Aeronautics and Astronautics

Thesis Supervisor: Richard. H. Battin
Title: Senior Lecturer of Aeronautics and Astronautics

Acknowledgments

First and foremost, I would like to express my special gratitude to my advisor, Prof. Olivier de Weck. This thesis would never have been completed without his guidance, patience, and confidence in me.

I would like to thank my committee members, Prof. David Simchi-Levi, Prof. Jeffery Hoffman, and Prof. Richard Battin, and thesis readers, Prof. Diego Klabjan and Dr. Robert Shishko. Their invaluable advice matured this thesis from both an Astronautics and an Operations Research perspectives.

It was a wonderful opportunity for me to participate in the Space Logistics project and work with excellent colleagues. Brian, Christine, Christoph, Erica, Gene, James, Jason, Joe, Matt, Nii, and Sarah - I thank you all and feel proud of what we have achieved. I also thank Dr. Martin Steele (NASA Kennedy Space Center), who was the COTR of the project.

Luckily a lot of great friends have been with me during my time at MIT. I would like to thank Bill, Chiwon, Dan, Eun-Suk, Gergana, Han-Lim, Jin-Ho, Jin-Oh, Paul, Ryan, Sang-Il, Sang-Won, Seungbum, Seung-Wook, Wilfried, and Young-Han for sharing thoughts on research, talking about the future, and having fun with me. I would also like to thank Ariane, Byung-Hoon, Cornel, Dong-Chul, Dong-Geun, Dong-Gun, Dong-Jin, Do-Nyun, Gunwoo, Gyu-Boong, Ho-Hyun, Jennifer, Jhong-Woo, Ji-Hyun, Ji Jun, Joo-Eun, Julien, Jun-Geun, Jungwoo, Kang-Hyun, Kyungryeol, Rania, Sang-Jin, Sang-Mok, Sung-Joong, Soon-Jo, Taeho, Tairin, Yoshi, Youjin, and Young-June for making my life in Cambridge enjoyable.

Prof. Youdan Kim, Prof. Changdon Kee, and Dr. Jeongjoo Park encouraged and helped me to start my Ph.D. program. I am profoundly grateful to them all.

My Ph.D. program was supported by the Cambridge Fellowship program operated by Cambridge Culture Foundation in Seoul, Korea. I sincerely appreciate the sponsorship that enabled me to complete this thesis without any financial difficulty.

My sisters and brother-in-law, Sung-Yeon, Hye-Young, and Sung-Min, deserve many thanks for their concern and support. I thank them all.

I would like to thank Eunice for always being with me, loving me, and believing in me. I love you, Eunice.

Finally, I would like to thank my parents, Mr. Seung-Jik Ahn and Ms. Kwang-Yeon Park, for their unconditional love and support. This thesis is dedicated to them.

Contents

1 Introduction: Global Planetary Surface Exploration **21**

2 Problem Description and Mathematical Formulation **27**

2.1 Concepts used for Problem Description 28

2.1.1 Agent, Site and Profit 28

2.1.2 Base, Route, Routing Tactic, and Single-Route Constraint . . 29

2.1.3 Mission and Mission Strategy 31

2.1.4 Campaign and Budget Constraint 32

2.2 Optimization of a Global Planetary Surface Exploration Campaign . 33

2.3 Literature Review for Routing Problems 35

2.4 Unique Features of the Global Planetary Surface Exploration Campaign Optimization Problem 40

2.5 Abstract Description of the GLRPP 41

2.6 Mathematical Formulation of the GLRPP 44

2.7 Sample Problem 46

2.8 Chapter Summary 50

3 Single-Phase Method **51**

3.1 Single-Phase Solution Method: Introduction 51

3.2 Formulating LP Relaxation of the GLRPP and Its Dual Problem . . . 52

3.3 Solving the GLRPPLR Using Column Generation 53

3.4 Obtaining a Near-Optimal Solution for the GLRPP and an Optimality Gap 59

3.5 Sample Problem Using the Single-Phase Method 62
3.6 Single-Phase Method Summary 66

4 Three-Phase Method 67

4.1 Three-Phase Solution Method: Introduction 67
4.2 Phase I - Divide 72
4.3 Phase II - Conquer 75
4.4 Phase III - Synthesize 77
4.5 Sample Problem Using the Three-Phase Method 80
4.5.1 *Divide* 81
4.5.2 *Conquer* 81
4.5.3 *Synthesize* 85
4.6 Numerical Experiments 86
4.6.1 GLRPP Instance Generation and Basic Statistics for Numerical Experiments 86
4.6.2 Estimating Performance Metrics 88
4.6.3 Summary of Numerical Experiments 96
4.7 Three-Phase Method Summary 98

5 Space Application: Global Mars Surface Exploration 99

5.1 Introduction 99
5.2 Global Mars Surface Exploration Campaign Optimization 99
5.3 Selection of Potential Bases and Exploration Sites 100
5.4 Mission Strategies and Routing Tactics 101
5.5 Design of the Mars Orbiting Depot 107
5.5.1 Concept of the Orbiting Depot 107
5.5.2 Orbit Selection 109
5.5.3 Individual Supply Unit Design 111
5.5.4 Characteristics of the Orbiting Depot Strategy 115
5.6 Numerical Examples 117
5.6.1 Problem Instances 117

5.6.2 Results 117
5.6.3 Value of a Technology 122
5.7 Mars Surface Exploration Summary 124

6 Terrestrial Application: College Football Recruiting 127
6.1 Introduction 127
6.2 Problem Description 128
6.3 School and Airport Selection and Campaign Characteristics 129
6.4 Numerical Result 133
6.5 College Football Recruiting Summary 134

7 Conclusions 137
7.1 Thesis Summary 137
7.2 Contributions 138
7.3 Suggested Future Work 139
7.3.1 Improved Solution Procedure 139
7.3.2 Balanced Optimization Scheme - Consideration of Profit Vectors 140
7.3.3 Potential Applications of the GLRPP 142

List of Figures

1-1 Superimposed Plot for Historical Planetary Surface Explorations [1] . 24

1-2 Thesis Roadmap 26

2-1 Mars Candidate Exploration Sites [40] 28

2-2 Examples of Routing Tactics 30

2-3 Resource Consumption and Feasibility Check for a Route 31

2-4 Mission Strategy Example 32

2-5 Resource Consumption for a Mission and Collective Constraint 33

2-6 Planetary Surface Exploration Campaign Example 34

2-7 Lineage Tree of Routing Problem Family 36

2-8 Selected Problems in Routing Problem Family 39

2-9 Decision and Parameter Hierarchy for the GLRPP 43

2-10 Sample GLRPP Instance 47

3-1 Structures of the columns associated with $x_j^{b,s,k}$ 57

3-2 Flow Chart for the Single-Phase Solution Method 61

3-3 Sample Problem Solution using a Single-Phase Method 63

3-4 Columns associated with $y^{1,1}$, $x_1^{1,1,2}$, and $x_{18}^{1,1,1}$ 65

4-1 GLRPP Instance and Two Clusters for the Instance 69

4-2 Solving the Instance Using the *Full Factorial Strategy Method* 70

4-3 Solving the Instance Using the *Modified Method* 71

4-4 Three-Phase Method for Solving the GLRPP 72

4-5 Flow Chart for the MDVRPP Solution Procedure 77

4-6 Proximity Sets for the Sample Problem 81
4-7 Clustering Result for Sample Problem 83
4-8 Clusters and Cluster Strategies for the Sample Problem (Divide Phase) 84
4-9 Result Summary of the Conquer Phase for the Sample Problem . . . 84
4-10 Sample Problem Solution Using a Three-Phase Method 85
4-11 Single-Phase Method Computation Time - Actual vs. Fitted 90
4-12 Three-Phase Method Computation Time - Actual vs. Fitted 92
4-13 Performance Boundary 94
4-14 Performance Boundary Shift with Problem Size Change 95
4-15 Optimality Gap, Single-Phase Method and Three Phase Method . . . 96

5-1 Surface Exploration Vehicle - UPV 103
5-2 Surface Exploration Vehicle - Camper 104
5-3 Surface Exploration Vehicle - UPV/Camper Assembly 104
5-4 Reference Scenario for a 7-day Excursion 105
5-5 Conceptual Diagram for the Orbiting Depot and Its Functionality . . 108
5-6 Shape of the Supply Unit 114
5-7 Instance 4 - Clustering Result 118
5-8 Instance 4 - Solution 119
5-9 Routes for the Mission Using Base 78 (Standard Strategy) 120
5-10 Routes for the Mission Using Base 99 (Orbiting Depot Strategy) . . . 121
5-11 Value of Technology - Orbiting Depot and ISRU 125

6-1 College Football Recruiting - Clustering Result 133
6-2 College Football Recruiting - the GLRPP Solution 134
6-3 Used Airports and Visited Schools 136

List of Tables

1.1 Historical Planetary Surface Explorations 23

2.1 Parameters, Decisions, Constraints, and an Objective for the Problem 35

2.2 Sample GLRPP Instance 48

2.3 Strategies, Tactics and Budget for the Sample Problem 49

3.1 Column Generation Procedure for Solving the GLRPPLR 56

3.2 Sample Problem Solution Summary - The Single-Phase Method . . . 64

4.1 Pseudocode - Clustering 75

4.2 Clustering Iteration for Figure 4-6 82

4.3 Clustering-derived Sets in Figure 4-7 83

4.4 Characteristics of Instances for Numerical Experiments 87

4.5 Basic Statistics for the Numerical Experiment Results 88

4.6 Estimation of the Computation Time using the Single-Phase Method 90

4.7 Estimation of the Computation Time using the Three-Phase Method 93

4.8 Estimator for the Optimality Gap Using the Single-Phase Method . . 95

4.9 Performance of the two solution methods 97

5.1 The Resource Consumption Limit for the Campaign in Terms of Time 102

5.2 Collective Constraint Characteristics (Standard Exploration Strategy) 102

5.3 Energy Density for $NaBH_4/H_2O_2$ Fuel Cell 103

5.4 Power Requirement for Mars Surface Exploration Using Camper/UPV 105

5.5 Reference Exploration Route and Corresponding Fuel Budget 106

5.6 Fuel Mass Rate for Mars Surface Exploration Using Camper/UPV . . 106

5.7 Single Route Constraint Characteristics (Standard Tactic) 107
5.8 Mass Delivered on Mars Surface (Standard Strategy) 107
5.9 Orbit Period and Worst Case Cross Track Distance 110
5.10 Human Item Consumption Rates for the Reference Exploration Scenario 111
5.11 Human Item Mass Calculation for the Reference Exploration Scenario 111
5.12 Core Contents for the Single Supply Unit 112
5.13 Ballistic Coefficient Calculation for the MER Entry Vehicle 113
5.14 Design of an Entry Body for a Single Supply Unit 115
5.15 Strategies and Corresponding Tactics for the Mars Exploration Case . 116
5.16 GLRPP Instance Generation for Mars Surface Exploration Case Study 117
5.17 Result Summary - Mars Surface Exploration Case Study 118
5.18 Result Summary: Mars Surface Exploration Without Orbiting Depot Technology 123
5.19 Result Summary - Mars Surface Exploration Without ISRU Technology, With Orbiting Depot Technology 124
5.20 Result Summary - Mars Surface Exploration Without ISRU Technology, without Orbiting Depot Technology 124

6.1 Schools Selected for Recruiting 131
6.2 Airports Potentially Used for Recruiting 132
6.3 Campaign Characteristics of the Recruiting Case Study 132

Nomenclature

Abbreviations

ESS	Error Square Sum
GLRPP	Generalized Location Routing Problem with Profits
IP	Integer Program
JPL	Jet Propulsion Laboratory
LB	Lower Bound
LP	Linear Program
LRP	Location Routing Problem
MDVRP	Multi-Depot Vehicle Routing Problem
MDVRPP	Multi-Depot Vehicle Routing Problem with Profits
NASA	National Aeronautics and Space Administration
PDF	probability density function
TSP	Traveling Salesman Problem
UB	Upper Bound
VRP	Vehicle Routing Problem
VRPP	Vehicle Routing Problem with Profits

Symbols

J	profit sum for the GLRPP obtained using the single-phase method
K	profit sum for the GLRPP obtained using the three-phase method

V	profit sum for the MDVRPP
$\mathbf{A}$	matrix composed of vectors representing site subsets
$\mathbf{B}$	index set of potential bases
$\mathbf{C}$	distance matrix
$\mathbf{E}$	index set of potential exploration sites
$\mathbf{H}$	collective resource consumption matrix
$\mathbf{J}$	collection of indices for all subsets of $\mathbf{E}$
$\mathbf{J}_f^{b,s,k}$	index set of subsets which, combined with a base b, make feasible routes with respect to tactic k of strategy s
$\mathbf{R}_j$	subset of potential exploration sites with an index j
$\mathbf{S}$	index set of mission strategies
$\mathbf{T}^s$	index set of routing tactics for a mission strategy s
$\mathcal{A}$	index set of arcs
$\mathcal{C}$	index set of clusters
$\mathcal{G}$	complete graph ($= [\mathcal{N}, \mathcal{A}]$)
$\mathcal{N}$	set of nodes ($=\mathbf{B} \cup \mathbf{E}$)
$b \in \mathbf{B}$	potential base index
$i \in \mathbf{E}$	site index
$j \in \mathbf{J}$	site subset index
$k \in \mathbf{T}^s$	routing tactic index
$l \in \mathcal{C}$	cluster index
r_j	profit sum for subset of potential exploration sites j, ($=\sum_{i \in \mathbf{R}_j} v_i$)
$s \in \mathbf{S}$	mission strategy index
t_i	time required to obtain profits at potential exploration site i
v_i	profit value assigned to potential exploration site i
$\mathbf{c}_0^{s,k}$	single-route constraint per-route resource consumption coefficient vector (tactic k and strategy s)
$\mathbf{c}_d^{s,k}$	single-route constraint on-arc resource consumption coefficient vector (tactic k and strategy s)
$\mathbf{c}_\tau^{s,k}$	single-route constraint on-site resource consumption coefficient vector

	(tactic k and strategy s)
$\mathbf{d}_0^s$	collective constraint per-route resource consumption coefficient vector (strategy s)
$\mathbf{d}_d^s$	collective constraint on-arc resource consumption coefficient vector (strategy s)
$\mathbf{d}_\tau^s$	collective constraint on-site resource consumption coefficient vector (strategy s)
$\mathbf{l}_r^{s,k}$	single-route constraint resource consumption limit vector (tactic k and strategy s)
$\mathbf{l}_c^s$	collective constraint resource consumption limit vector (strategy s)
$\mathbf{n}^s$	maximum number of routes vector (strategy s)

Subscripts and Superscripts

$(\cdot)^*$	optimal value
$(\cdot)'$	transpose

Definitions

Base

A base is a node at which a route starts and ends.

Site

A site is a node which has an attribute called. An agent can visit the site and obtain the profit. The agent can be human and/or robotic.

Profit

Profit for a site is the benefit that can be achieved by visiting and exploring the site. It is expressed as a scalar value determined outside of the problem (e.g. by site surveys).

Route

A route is a directed cycle defined over the graph. The route contains exactly one base and starts and ends at the base. Each route selects a routing tactic which determines characteristics of the route.

Resource

A resource is something that is used to carry out a campaign. Consumption of the resource is expressed as the sum of three resource consumption elements: on-route, on-arc, and on-site resource consumption.

Single-Route Constraint

A single-route constraint is a constraint imposed on a route such that the total amount of resource consumed over the route cannot exceed a certain limit value.

Routing Tactic

A routing tactic is the set of characteristics for a route type which is composed of single-route constraints and a maximum number of routes.

Feasibility of a Route

A route is feasible if the route satisfies single-route constraints defined by the routing tactic selected by the route.

Mission

A mission is the collection of feasible routes which have a base in common. We assume that a single site should not be visited more than once by the routes for the mission. Each mission selects a mission strategy which determines characteristics of the mission. All routing tactics selected by the routes should be available in the mission strategy selected by the mission.

Collective Constraint

A collective constraint is a constraint imposed on a mission such that the total amount of resource consumed over all routes for the mission cannot exceed a certain limit value.

Maximum Route Number Constraint

A maximum route number constraint is a constraint imposed on a mission such that the total number of routes which select a certain routing tactic cannot exceed a certain limit value.

Mission Strategy

A mission strategy is the set of characteristics for the mission which is composed of collective constrains, available routing tactics, and mission cost.

Feasibility of a Mission

A mission is feasible if routes for the mission satisfy collective constraints defined by a mission strategy selected by the mission and the maximum route number constraints defined by routing tactics for the mission strategy.

Campaign

A campaign is a collection of missions which have a common objective and are carried out under the same budget source. We assume that a single site should not be visited more than once by the routes for all the missions in the campaign.

Budget Constraint

A budget constraint is a constraint imposed on a campaign that cost sum of missions for the campaign should be less than a certain limit value (budget).

Reachability

A site is reachable from a base if there exists an available routing tactic such that a route using the tactic which represents a round trip between the base and the site is feasible.

Cluster

A cluster is a subset of nodes which has the following properties: (1) For every site in the cluster there exist at least one base from which the site is reachable; (2) For every site in the cluster there exists no base outside the cluster from which the site is reachable; (3) For every base in the cluster there exist at least one site in the cluster that is reachable from the base; and (4) For every base in the cluster there exists no site outside the cluster that is reachable from the base.

Chapter 1

Introduction: Global Planetary Surface Exploration

The *Global Exploration Strategy* initiated by NASA in April 2006 is a strategy for exploring the solar system that encompasses the interests of many stakeholders including space agencies, academia, and commercial investors.[1] Intended to address two issues - *why we are returning to the moon*, and *what we are planning to do when we get there* - the strategy emphasizes a worldwide, or *global* coordination through which nations can collaborate both on individual projects and on the collective effort [22].

Initial elements of the strategy were announced in December 2006 [2], and a report with the title "The Global Exploration Strategy: The Framework for Coordination" was published in May 2007. An answer to the question "why global?" can be found in the following quote from the report [3].

> *One of the most fundamental human characteristics is a relentless curiosity that drives us to investigate the unknown. Throughout our history, we have looked beyond our apparent boundaries to the mysteries that lie beyond. Compelled to explore, to understand and to use the world in which*

[1] It has been discussed among experts from NASA and thirteen other space agencies (Australian, Canadian, Chinese, European, French, German, British, Indian, Italian, Japanese, Russian, South Korean and Ukrainian space agencies), non-governmental organizations, and commercial interests.

we find ourselves, we have spread across continents and oceans. We have probed the farthest reaches of the planet-the frozen poles, the deep oceans, the high atmosphere. With increasing intent and determination, we are resolved to explore our nearest companions-the Moon, Mars and some nearby asteroids. Our goal is not a few quick visits, but rather a sustained and ultimately self-sufficient human presence beyond Earth supported by robotic pathfinders. Sustainable space exploration is a challenge that no one nation can do on its own. This is why fourteen space agencies have developed The Global Exploration Strategy: The Framework for Coordination, which presents a vision for robotic and human space exploration, focusing on destinations within the solar system where we may one day live and work. It elaborates an action plan to share the strategies and efforts of individual nations so that all can achieve their exploration goals more effectively and safely.

Considering that future space exploration with a strong emphasis on sustainability will include large-scale and broad-scope activities, a *global* exploration strategy - an exploration by international operators - is the only feasible and reasonable way to successfully achieve the ambitious exploration goals.

On the other hand, global exploration also means exploration for planetary-wide objects. Here, "global" refers to the entire surface of a planetary body. Various stakeholder groups are interested in a planetary body and each group wants to obtain information from distinct regions of the body. To satisfy these diversified groups, we need to explore multiple locations distributed over the surface of the whole planet. In other words, *global planetary surface exploration* is highly desirable.

A number of space missions have explored surfaces of planetary bodies. Some of these past efforts were successfully completed as planned. Some missions - the Spirit and Opportunity Mars rover missions for example - even survived much longer than their original design lives. Table 1.1 summarizes space missions for surface exploration and Figure 1-1 shows a superimposed view of the traverses for the missions [1].

The amount of on-surface activities for the missions was limited for several reasons.

Table 1.1: Historical Planetary Surface Explorations

Mission	*Rover*	*Destination*	*Type*	*Length* [km]	*Year Landed*
U.S. Lunar					
Apollo 11	N/A	Moon	Human	0.3	1969
Apollo 12	N/A	Moon	Human	2.0	1969
Apollo 14	N/A	Moon	Human	3.3	1971
Apollo 15	Lunar Rover	Moon	Human	27.9	1971
Apollo 16	Lunar Rover	Moon	Human	27.0	1972
Apollo 17	Lunar Rover	Moon	Human	35.0	1972
Soviet Lunar					
Luna 17	Lunokhod 1	Moon	Robotic	10.5	1970
Luna 21	Lunokhod 2	Moon	Robotic	37.0	1973
U.S. Mars					
Pathfinder	Sojourner	Mars	Robotic	0.1	1997
MER-A	Spirit	Mars	Robotic	6.9	2005
MER-B	Opportunity	Mars	Robotic	11.4	2005

First, a much higher priority was given to successful in-space transportation, which was far more challenging for early missions than it is now. Second, capabilities of the surface mobility systems for early missions were limited. Last, especially for the case of the human lunar missions, surface stay times were relatively short.

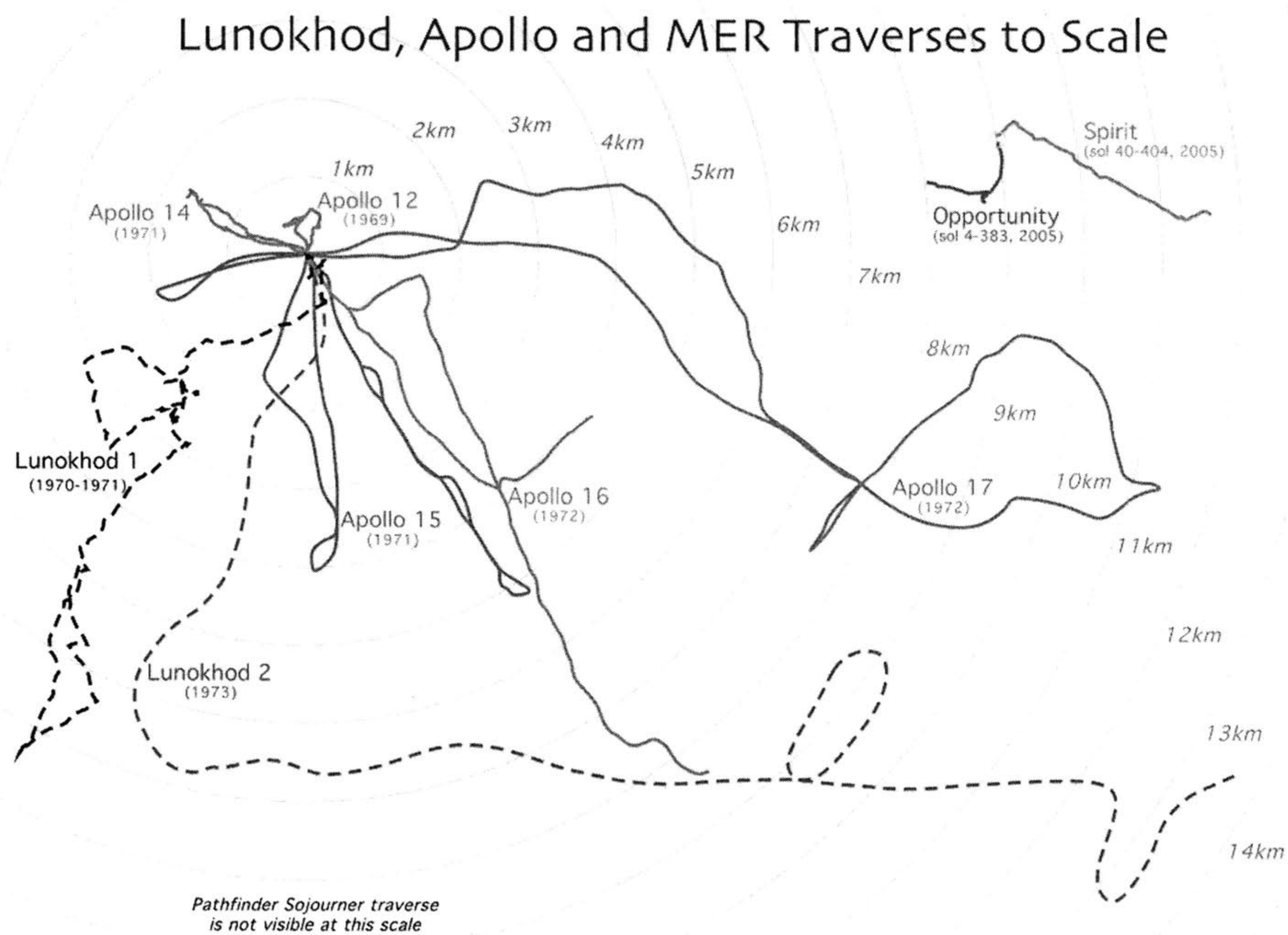


Figure 1-1: Superimposed Plot for Historical Planetary Surface Explorations [1]

The situation will be different for future planetary surface exploration missions. Accumulated experiences from past missions will make in-space transportation more reliable and easier to carry out; benefits from exploration will be considered more important than they used to be. In addition, improved capabilities of planetary surface mobility systems will be available[2] and surface stay times for future human missions will be much longer than those for previous missions. These factors make it very important to carefully design surface activities of the missions. Stakeholder groups will identify globally-distributed locations which they are interested in. The amount of resources that can be used for the missions will be limited, which prohibits visiting all the identified locations. Therefore, we need to design the surface exploration missions such that total benefits from the missions can be maximized.

The problem of optimizing a global planetary surface exploration *campaign* (a series of missions sharing an objective) is dealt with in this thesis. The problem is formulated as a routing problem class referred to as the *Generalized Location Routing Problem with Profits* (GLRPP).

This thesis develops a mathematical formulation and solution method for the GLRPP. Figure 1-2 shows a thesis roadmap. In Chapter 2, the GLRPP is developed using a global planetary surface exploration campaign optimization problem and mathematically formulated as an Integer Program (IP). Two solution methods to solve the GLRPP are presented in Chapter 3 (single-phase method) and Chapter 4 (three-phase method). Chapter 5 is dedicated to a case study for a global Mars surface exploration campaign optimization as a space application of a GLRPP. Chapter 6 deals with a case study for the *College Football Recruiting Problem* as a terrestrial application of the GLRPP. Chapter 7 summarizes this thesis and proposes future work as follow-up research.

[2]One recent study of a planetary surface vehicle reports the excursion capability of the vehicle of 500 km [42].

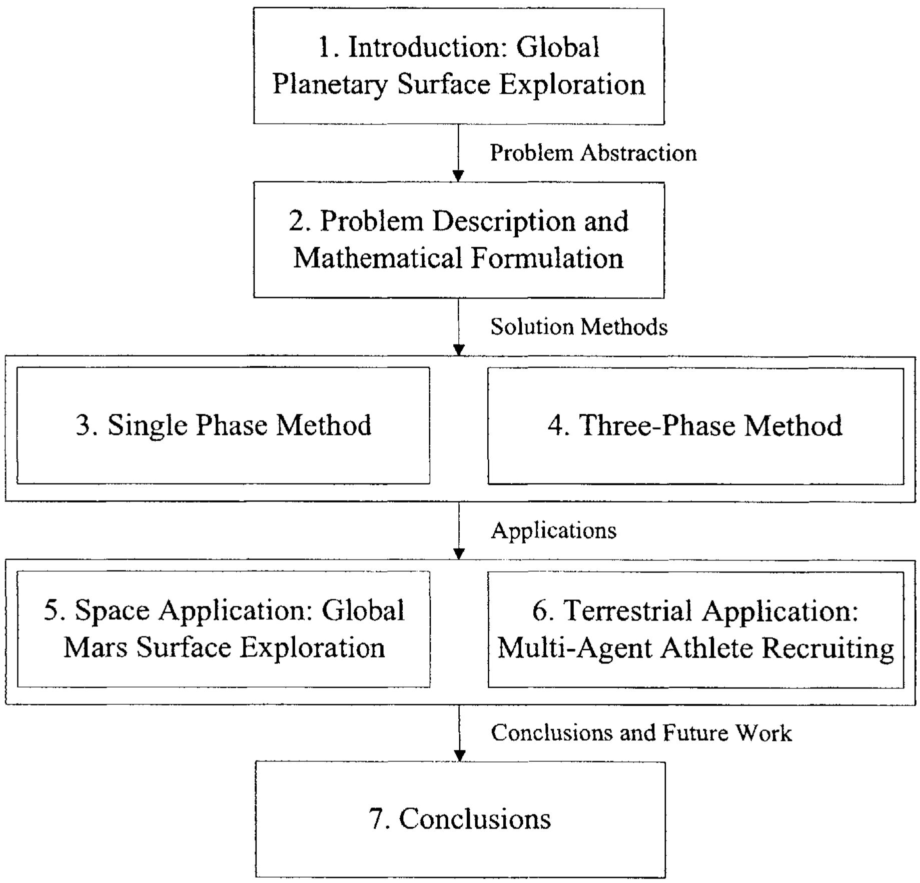


Figure 1-2: Thesis Roadmap

Chapter 2

Problem Description and Mathematical Formulation

This chapter describes an optimization problem for a planetary surface exploration campaign and develops the Generalized Location Routing Problem with Profits (GLRPP) framework. Concepts used throughout this chapter are introduced in Section 2.1. Section 2.2 describes an optimization problem for a planetary surface exploration campaign. The problem belongs to the class of routing problems which have been previously studied in Operations Research (OR). Literature review for the routing problems is provided in Section 2.3. Then, unique features of the GLRPP that have not been addressed in existing routing problems are introduced in Section 2.4. An abstraction and a mathematical formulation for the GLRPP are presented in Section 2.5 and Section 2.6, respectively. Finally a sample problem is provided that will be used to explain solution methods in Chapter 3 and 4.

2.1 Concepts used for Problem Description

2.1.1 Agent, Site and Profit

An *agent* visits locations and collects profits. A *site* is a location at which the agent can obtain profit[1]. The *profit* for each site is expressed as a scalar number. The amount of time to collect the profit is also assigned to each site. Candidate sites and profits / required time associated with the sites are determined from outside of the problem.[2] The agent can be either a human or robotic agent carrying out the exploration and obtaining profits. It is assumed that stakeholder groups for the exploration determine these candidate sites (potential exploration sites) and estimated profit values associated with the sites.[3] Figure 2-1 shows candidate exploration sites on Mars compiled from mission studies and individual contributions from scientists in fields like Geochemistry, Geology, Seismology, Meteorology, and Exobiology.[4]

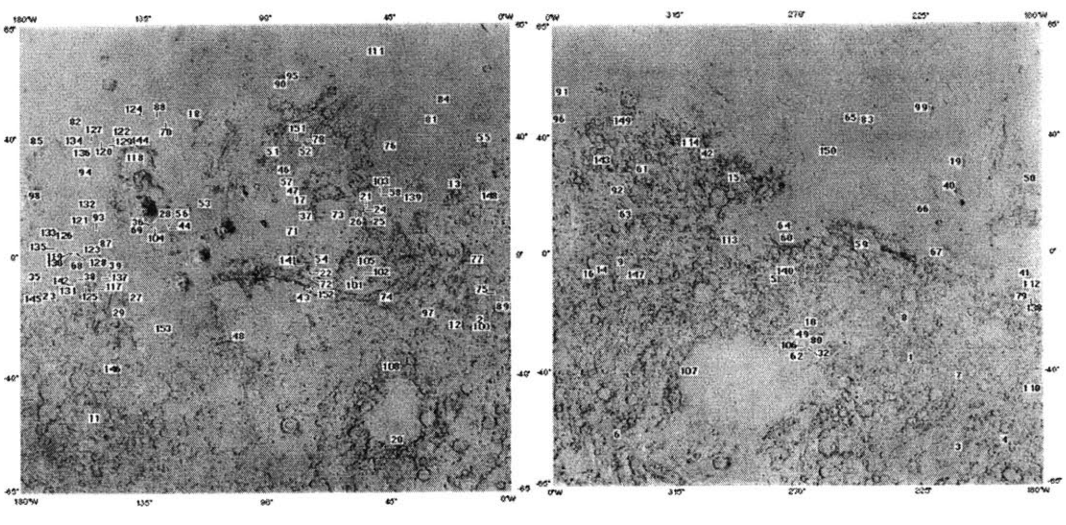


Figure 2-1: Mars Candidate Exploration Sites [40]

[1]Usually *profit* means {*(revenue) - (cost)*} in an economic sense (expressed as a dollar value). In this situation because there is no explicit revenue generated for exploring a site we don't generate financial profit but rather scientific value. Nevertheless the terminology *profit* is used throughout the thesis because it is a generally used terminology in operations research. *Benefit* or *reward* would be alternative terms.

[2]The term *candidate* is used because in the GLRPP all the sites need not to be visited.

[3]A discussion on whether this assumption is reasonable or not is dealt with in Chapter 5 and Chapter 7

[4]Originally these locations are referred to as *potential Landing Sites.* But they are selected because of scientific significance and can be regarded as *potential exploration sites* using the terminology in this thesis.

2.1.2 Base, Route, Routing Tactic, and Single-Route Constraint

A *base* is a location at which an agent starts and ends a *route*. A *route* is a sequence of locations (sites or bases) that represents movement of an agent over time. Only one base is included in a route, and the base is both a starting and an end location of the route. An agent collects profits by visiting sites belonging to the route. It is assumed that a site cannot be visited twice. Candidate base locations are externally determined. A *routing tactic* characterizes a route by specifying *single-route constraints* and the *maximum number of routes*. A single-route constraint is expressed by a constraining resource type, consumption coefficients (per route, on-arc, and on-site), and a consumption limit. A route is *feasible* if the route satisfies single-route constraints defined by the routing tactic selected by the route. The maximum number of routes limits the number of routes that can use a specific routing tactic in a mission.

For planetary surface exploration the potential bases are potential landing locations identified by mission planners. A spacecraft lands at one of the potential landing locations and establishes a base there. Potential landing locations should be determined so that spacecraft can land safely and bases can be easily established and maintained at the landing locations. Figure 2-2 shows examples of routing tactics for surface exploration. *Walking*, *rover*, and *depot-assisted rover* are tactics presented in this example. A single-route constraint for the walking tactic is that the exploring agent should return to the base within 8 hours of the route's start. A constraining resource type in this tactic is time. An agent is assumed to walk with a speed of 2 [km/hr], and an on-arc resource consumption coefficient (time spent per unit distance during transportation between locations) is 0.5 [hr/km]. Time spent at a site is directly counted towards the resource consumption and the on-site resource consumption coefficient is 1 [hr/hr]. Similarly, a single-route constraint for the *rover* tactic is that the exploring agent should return to the base before the fuel consumption exceeds the rover capacity (600 [kg]). It is assumed that the per-distance fuel consumption is 1 [kg/km] (= an on-arc resource consumption coefficient) and that the rover is traveling with an average speed of 3 [km/hr], which leads to the on-site resource consumption coefficient of 3 [kg/hr]. The *depot-assisted rover* has a larger fuel consumption limit than the standard *rover*, but the number of routes using this tactic cannot exceed 4.[5] Figure 2-3 shows how resource consumption is calculated for a feasibility check of a route with respect to a single-route constraint. The upper figure shows an infeasible route using the walking tactic. Time consumed on the route is 10 [hr], which is larger than the consumption limit and violates the single-route constraint (= 8 [hr]). The lower figure shows a feasible route using the rover tactic. Fuel consumption on the route is 460 [kg], which does not exceed the consumption limit and satisfies the single-route constraint.

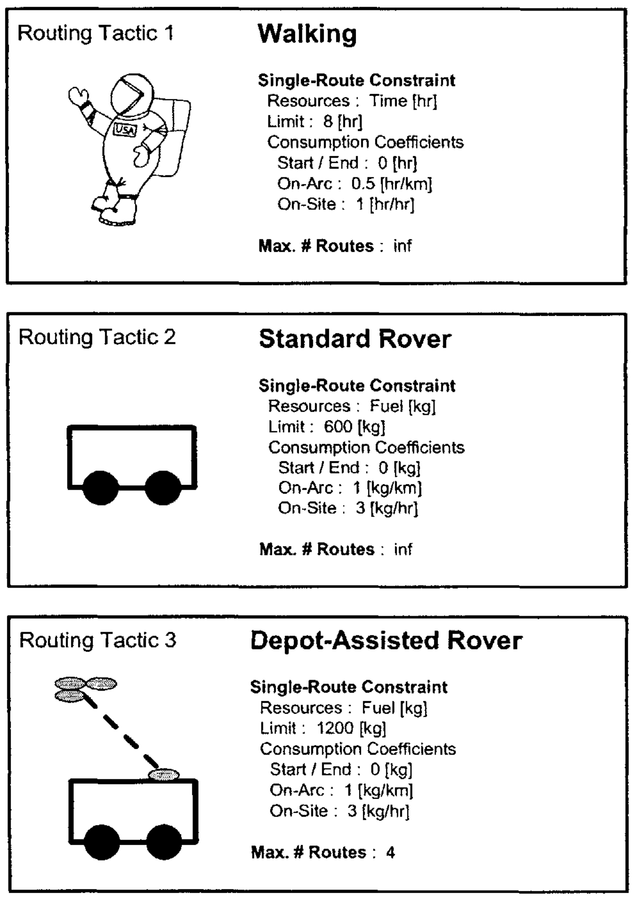
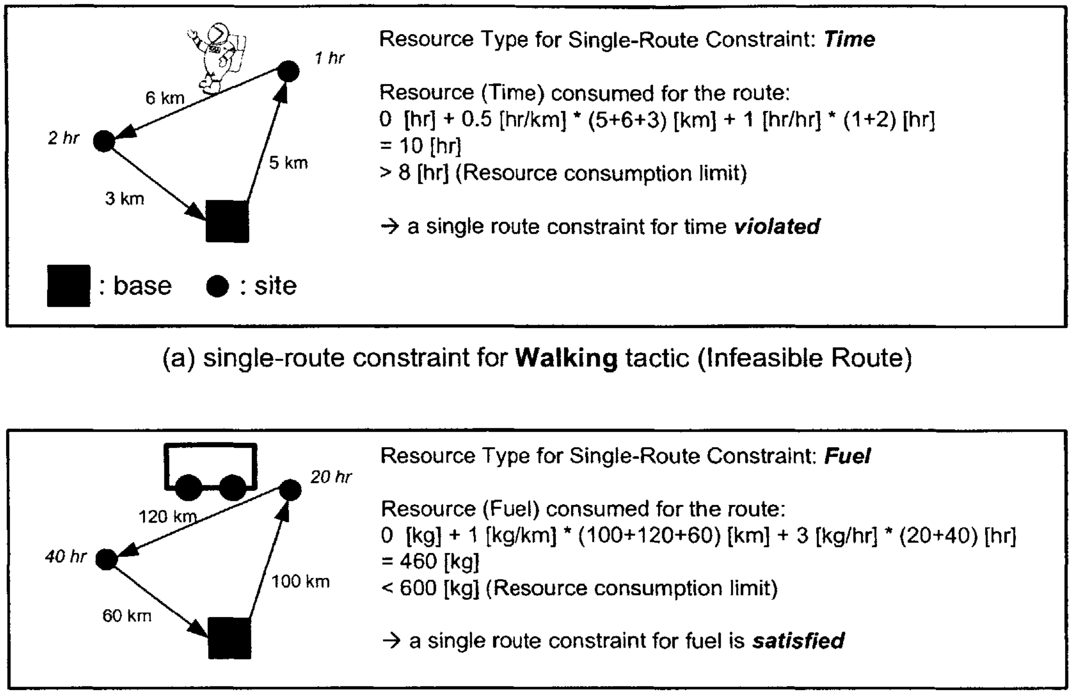


Figure 2-2: Examples of Routing Tactics

[5] Details on the orbiting depot are provided in Chapter 5.



(a) single-route constraint for **Walking** tactic (Infeasible Route)

(b) single-route constraint for **Rover** tactic (Feasible Route)

Figure 2-3: Resource Consumption and Feasibility Check for a Route

2.1.3 Mission and Mission Strategy

A *mission* is a collection of feasible routes sharing a base. We assume that there exist multiple technologies, each of which represents a *mission strategy*. A *mission strategy* characterizes the mission by specifying collective constraints, available routing tactics for the mission, and mission cost. Similar to a single-route constraint, a collective constraint is expressed by a constraining resource type, consumption coefficients, and a consumption limit. Routes belonging to the mission should use available routing tactics specified by the mission strategy. Mission cost is used to calculate the total cost of a campaign, on which a budget constraint is imposed.

A mission for a planetary surface exploration is a set of routes associated with a common base. Figure 2-4 shows examples of mission strategies which use routing tactics presented in Figure 2-2. Figure 2-5 shows how resource consumption constraints for a mission are calculated. The *rover* strategy presented in Figure 2-4 is considered. Resource (time) consumption for all routes included in each mission is

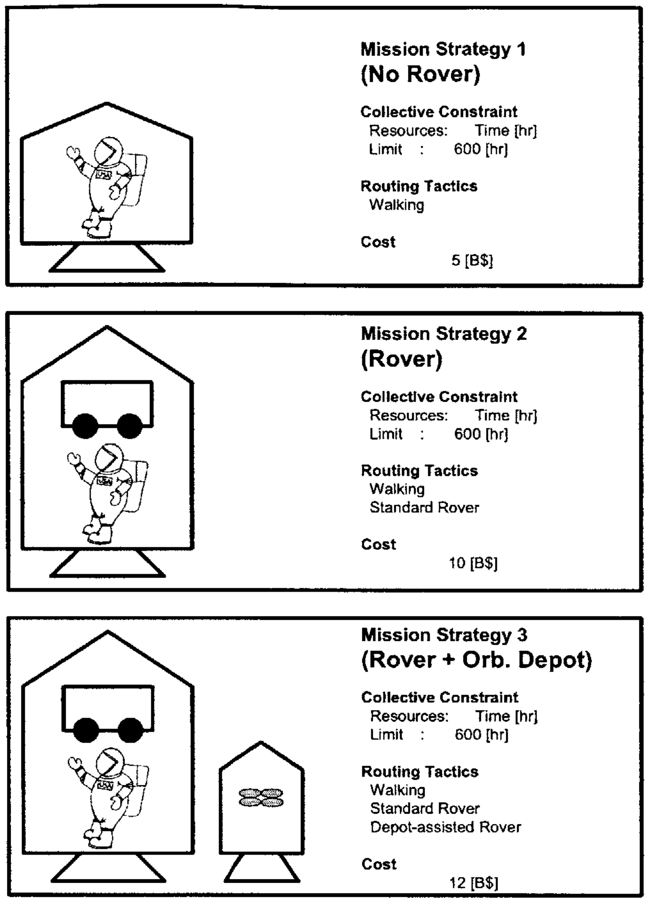
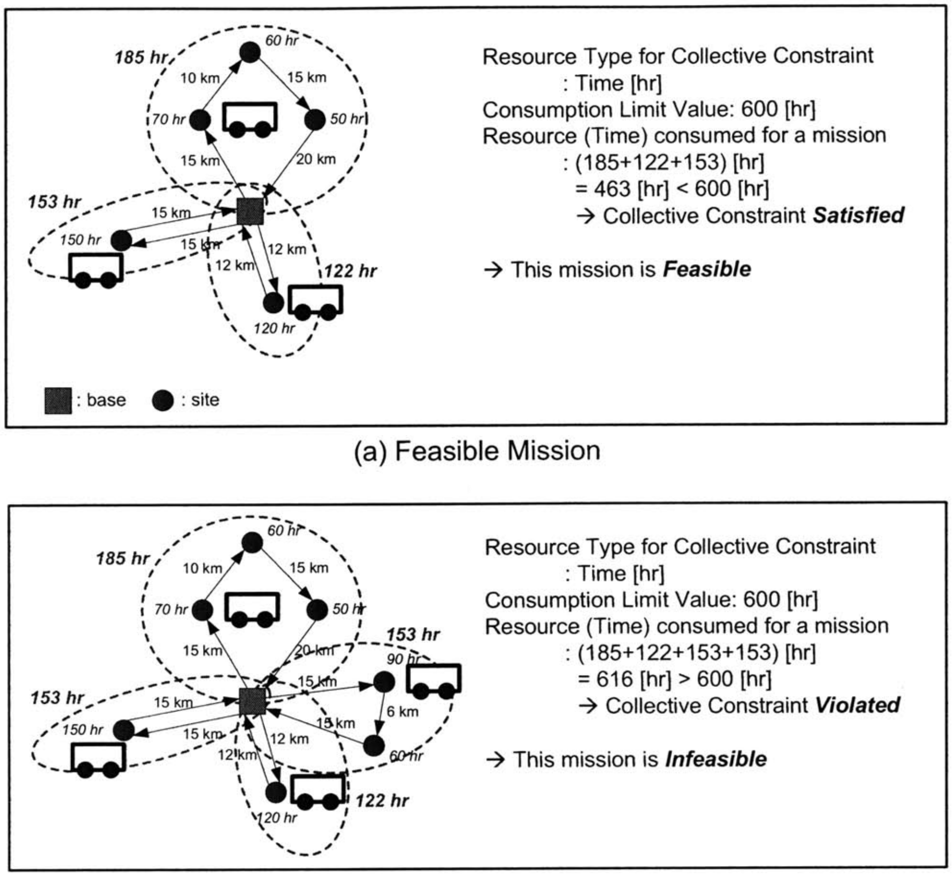


Figure 2-4: Mission Strategy Example

aggregated and compared with a limit value as a feasibility check. For the mission in Figure 2-5-(a), the sum of time consumed over all routes is less than the consumption limit and the mission is therefore feasible. We assume that the routes are executed sequentially. The total time for the mission in Figure 2-5-(b) is, however, larger than the consumption limit and the mission is infeasible.

2.1.4 Campaign and Budget Constraint

A *campaign* is composed of missions sharing an objective and budget. A subset of potential bases are selected for missions included in the campaign. A *budget constraint* is imposed on a campaign such that the total cost for missions comprising the campaign cannot exceed the budget. Figure 2-6 illustrates an example of a planetary surface exploration campaign. Three out of five potential landing locations are used



(a) Feasible Mission

(b) Infeasible Mission: Violates Collective Constraint

Figure 2-5: Resource Consumption for a Mission and Collective Constraint

as bases. Mission strategies presented in Figure 2-4 are used for missions comprising the campaign. Total cost for this campaign is (5+10+12) = 27 [B$].

2.2 Optimization of a Global Planetary Surface Exploration Campaign

Consider a surface exploration campaign for a planetary body such as Mars. Geographical information on candidate landing locations and exploration sites is given. Distance between any of two positions (potential landing locations or exploration sites) can be generated out of this information - with or without terrain information. A profit value assigned for each exploration site and the amount of time required to obtain the profit at the site is also assumed to be determined from outside the

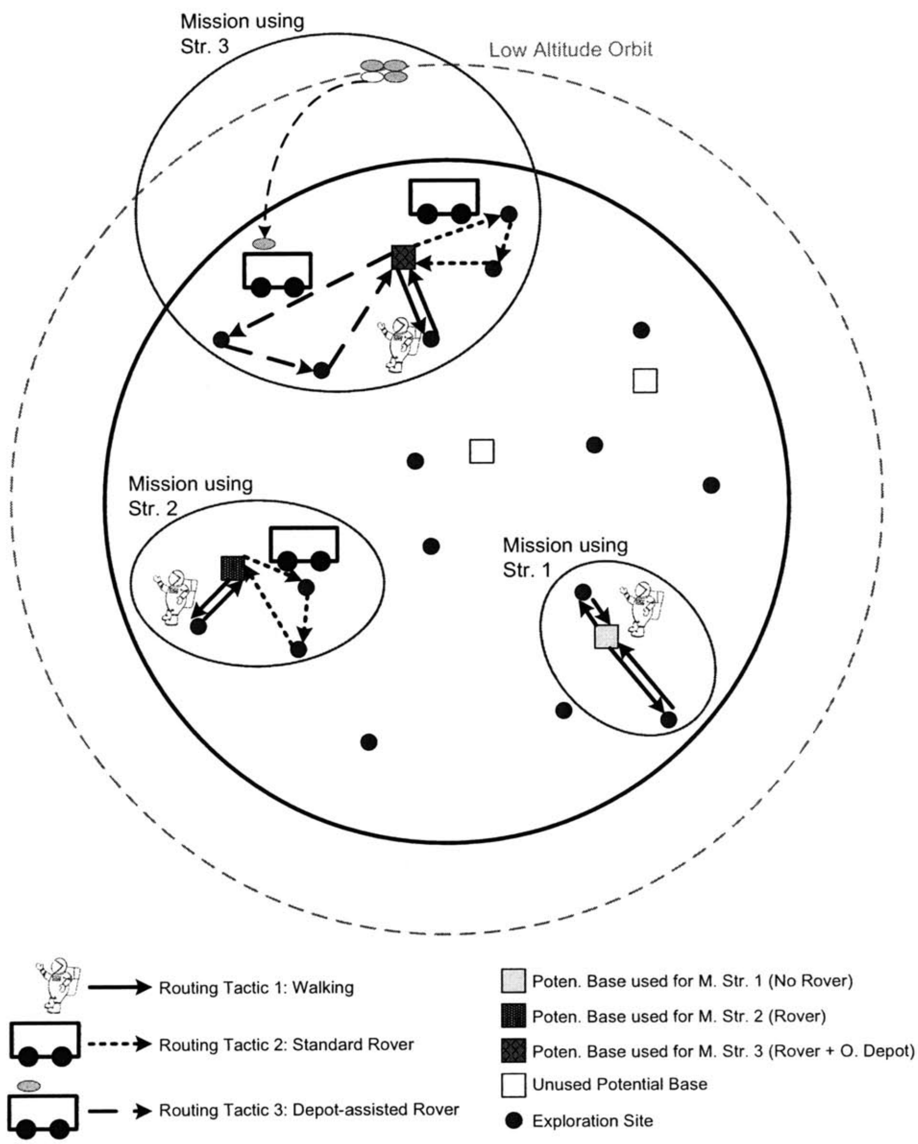


Figure 2-6: Planetary Surface Exploration Campaign Example

problem.

The amount of budget allowed for the campaign is given. Also the mission strategies that can be used for the campaign are pre-determined. Each mission strategy characterizes a mission by specifying a mission cost, collective constraints, and available routing tactics. A routing tactic characterizes a route by specifying single-route constraints and the maximum number of routes.

The overall objective is to maximize the sum of profits obtained from all visited exploration sites. Decisions for this problems are: (1) landing locations (bases) used

Table 2.1: Parameters, Decisions, Constraints, and an Objective for the Problem

Objective	Maximization of the profit sum obtained from explored sites
Decisions	1. Selection of landing locations (bases) 2. A mission strategy for each base used in 1 3. Routes for each of the base used in 1 4. Routing tactics for the routes in 3
Constraints	1. Feasibility of routes (single route constraints) 2. Feasibility of missions (collective constraints) 3. Feasibility of campaign (budget constraint) 4. Each site cannot be explored more than once
Parameters	1. Geographical information on destination body 2. Potential landing locations 3. Potential exploration sites and associated profit 4. Set of possible mission strategies 5. Set of possible routing tactics for each mission strategy 6. Budget

for exploration missions, (2) mission strategies chosen by the missions, (3) routes corresponding to each base, and (4) routing tactics for selected routes.

The following constraints are imposed: (1) Every route should satisfy single-route constraints (single route feasibility); (2) Routes which belong to a mission should satisfy collective constraints and maximum route constraints specified by the strategy selected by the mission and routing tactics chosen by the routes, respectively (mission feasibility); and (3) The cost total of missions comprising a campaign should not exceed the campaign budget (campaign feasibility).

2.3 Literature Review for Routing Problems

Route selection is a part of the decisions that should be made to optimize the global planetary surface exploration campaign. This problem can be placed in the *routing problem family.* Routing problems have attracted researchers in diverse fields for a long time. Many studies on the routing problems for various applications have been published to date. In this section we review publications on routing problems which are related to our optimization problem.

The classical routing problem is the *Traveling Salesman Problem* (TSP), which has been studied extensively in supply chain management [24, 34, 9, 50]. Dantzig et al. described the TSP in their seminal paper as follows: *Find the shortest route (tour) for a salesman starting from a given city, visiting each of a specified group of cities, and then returning to the original point of departure.* There are two main restrictions in this problem: (1) Only a single route and base (depot) for the route are allowed; and (2) All the cities (sites) in the given group must be visited. In extensions of the TSP, these two restrictions have been generalized and/or modified so that the problem can handle a broader scope of situations. Figure 2-7 presents a *lineage tree* of the routing problem family, in which generalization of the two restrictions is expressed.

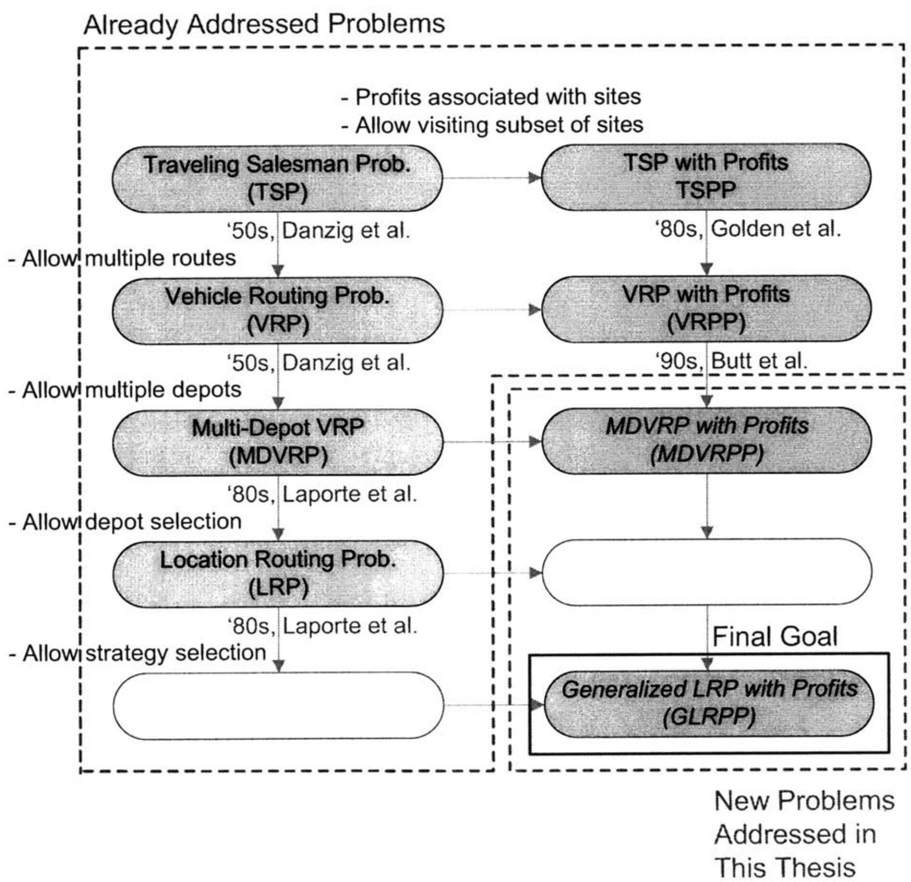


Figure 2-7: Lineage Tree of Routing Problem Family

Vertical evolution in the lineage tree represents generalization of the first restriction. The most straightforward extension of the TSP related with the first restriction (single route and base) is to allow multiple routes for visiting sites. The TSP becomes the *Vehicle Routing Problem* (VRP) by this extension. The VRP determines feasible routes which all start and end at a given depot to minimize total travel distance. Each site must be visited exactly once and some side constraints such as capacity of a vehicle (CVRP), travel distance of a route (DVRP), and time windows for visiting each site (VRPTW) are imposed if necessary [25, 21, 19, 20, 48, 45, 13, 14, 15, 28, 6, 35].

The next extension is to allow multiple depots as starting and end points of routes. The VRP becomes the *Multi-Depot Vehicle Routing Problem* (MDVRP) by this extension. Given multiple depots and sites, the MDVRP decides sets of routes to minimize total travel distance of all routes. The routes start and end at a depot included in the given depot set. Like the VRP, each site must be visited exactly once and side constraints are typically imposed on the MDVRP [37, 27, 23, 66, 56, 38].

The latest extension related to the first restriction is to allow selection of a subset of depots out of candidate depots. This problem is known as the *Location Routing Problem* (LRP). The LRP determines selection of depots and sets of feasible routes associated with the selected depots to minimize the sum of cost related with depot selection and/or total travel distance. Each site must be visited exactly once and some side constraints are typically imposed. The LRP is a very complex problem and relatively few studies related to this problem have been published so far [49, 46, 59, 64, 10].

Horizontal evolution in the lineage tree represents another extension by relaxing the restriction that each site must be visited exactly once. There are situations where all sites cannot be visited because of routing constraints. The objective of routing problems in this situation is qualitatively different from conventional routing problems in which a constraint that *every* site must be visited is imposed.

Assume that it is possible to quantify the profit obtained by visiting each site. There are three different problem classes that are defined by extension from the TSP [33]. The first problem class finds a route which maximizes the {(collected profit) - (travel costs)}. The *Profitable Tour Problem* (PTP) defined by Dell'Amico et al. belongs to this problem class [26]. The second problem class finds a route which maximizes collected profit such that travel costs does not exceed a certain value. The *Orienteering Problem* (OP), the *Selective Traveling Salesman Problem* (STSP), and the *Maximum Collection Problem* (MCP) belong to the second problem class [63, 47, 43]. The last problem class finds a route which minimizes travel costs such that collected profit is not smaller than a certain value. The *Prize-Collecting Traveling Salesman Problem* (PCTSP) is included in this problem class [7]. Similarly, the *Team Orienteering Problem* (TOP), the *Multiple Tour Maximum Collection Problem* (MTMCP), and the *Multivehicle Routing Problem with Profits* (MVRPP) are extensions from the VRP [18, 61, 16, 17, 41]. For consistency we refer to problems which are generated by extension from the TSP as the *Traveling Salesman Problem with Profits* (TSPP) and those by extension of the VRP as the *Vehicle Routing Problem with Profits* (VRPP).

Figure 2-8 shows problems obtained from these vertical/horizontal extensions in the lineage tree of the routing problem family.

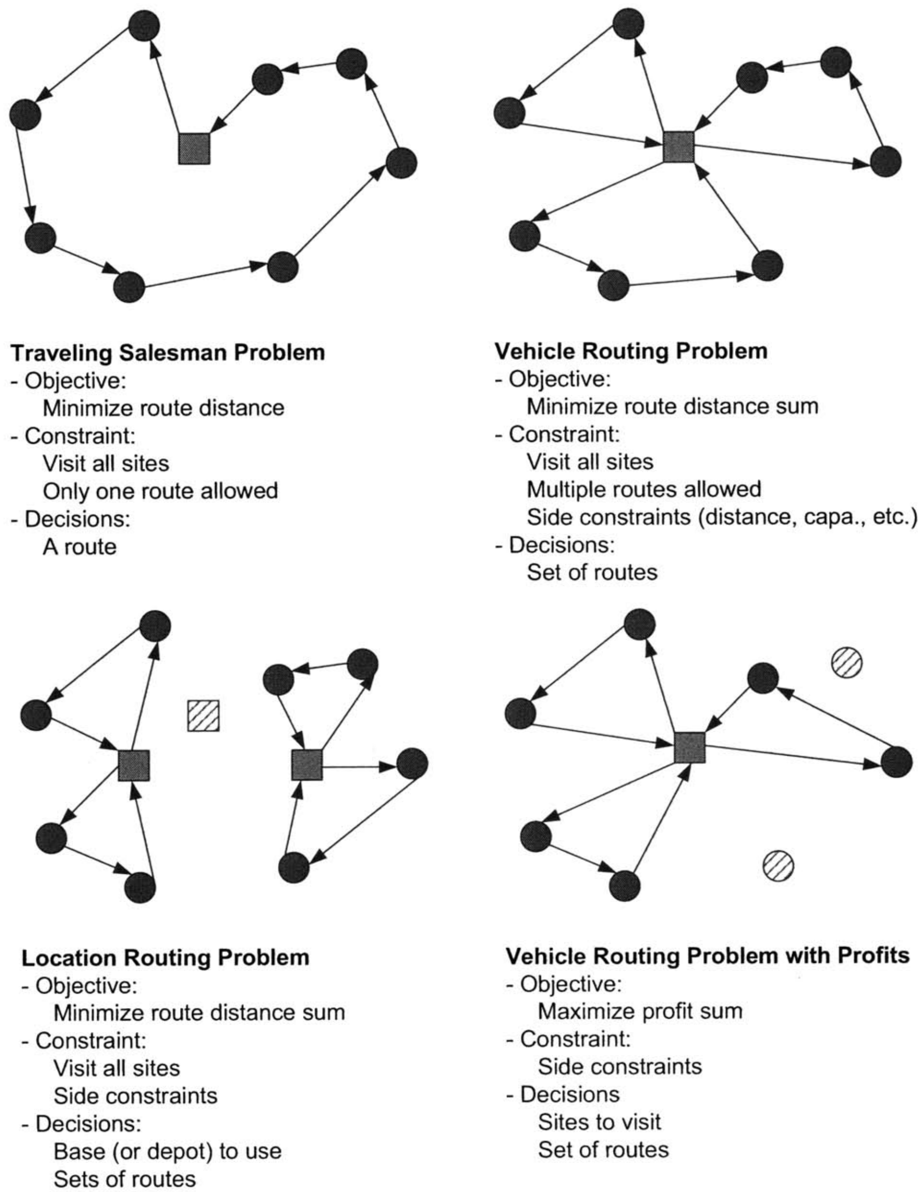


Figure 2-8: Selected Problems in Routing Problem Family

2.4 Unique Features of the Global Planetary Surface Exploration Campaign Optimization Problem

The global planetary surface exploration campaign optimization problem introduced in this chapter has some characteristics that have not been addressed in other routing problems published to date. This section provides a discussion of the new features that make this problem unique and challenging.

We first identify an existing routing problem class that is most similar to the campaign optimization problem. Selection of landing locations out of a candidate set is part our problem hence it can be classified as the *Location Routing Problem* (LRP). Two unique features of our problem are presented by comparison with the conventional LRP. The first is consideration of *profit* and the second is generalization of base selection from *whether to use or not to use each base* to *how to use each base.*

First, the objective of our problem is maximization of profit sum obtained from the visited sites. The conventional LRP minimizes the total travel cost and does not consider the *profit.* As pointed out in the literature review section, consideration of the profits requires determination of which sites are actually visited, which makes the problem more complex. So far no study on the *LRP with Profits* has been reported and our problem addresses this issue first in this thesis.

Second, our problem determines *how to use each potential base* while the classical LRP determines *whether each base is used or not.* Suppose we have three different exploration mission strategies presented in Figure 2-4. In this case, a decision related to a potential base is one out of the following four cases: (1) Use the potential base for a mission with a mission strategy 1; (2) Use the potential base for a mission with a mission strategy 2; (3) Use the potential base for a mission with a mission strategy 3; and (4) Do not use the potential base. In the conventional LRP, the decision is one out of two cases: (1) Use the potential base; and (2) Do not use the potential base. The former decision is a "multiple choice," which is a generalization of the latter, "binary choice." Therefore, a term "Generalized" is added at the beginning of our problem's title.

In summary, our problem can be regarded as the LRP with two unique features - consideration of the profit and generalization of the base selection. The problem is thus referred to as the *Generalized Location Routing Problem with Profits* (GLRPP). The GLRPP is placed in the lower right corner of the lineage tree presented in Figure 2-7.

2.5 Abstract Description of the GLRPP

As the first step in the mathematical formulation of the GLRPP, an abstract description of the problem is provided in this section. Consider a complete graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N} = (\mathbf{B} \cup \mathbf{E})$ is an index set for nodes and $\mathcal{A} = \{(i,j)|i,j \in \mathcal{N}\}$ is an index pair set for arcs. $\mathbf{B} = \{1, \ldots, n_B\}$ is an index set for potential bases (n_B: number of potential bases) and $\mathbf{E} = \{n_B + 1, \ldots, n_B + n_E\}$ is an index set for exploration sites (n_E: number of the exploration sites). Typically n_E is larger than n_B. For each potential site $i \in \mathbf{E}$, two real values v_i and t_i are assigned. v_i represents the profit that can be obtained by visiting site i, and t_i is the time required to collect the profit. Cost (distance) associated with each arc is expressed as a matrix $\mathbf{C} = [c_{i_1 i_2}]$ where $c_{i_1 i_2}$ represents the length of an arc from i_1 to i_2 ($i_1, i_2 \in \mathcal{N}$).

A *route* related to a potential base $b \in \mathbf{B}$ is a sequence of nodes representing the transportation between nodes. Potential base b starts the route, is followed by sites belonging to $\mathbf{E}$, and ends the route. A *mission* associated with base b is a set of routes related to b. A *campaign* is a set of all missions which are defined over graph $\mathcal{G}$ and are collectively pursuing an objective - profit sum maximization.

Resource consumption occurs in order to carry out a campaign; hence the resource consumption is required to carry out a mission and all associated routes. The resource consumption is expressed using three consumption classes: per-route, on-arc, and on-site consumption classes. A per-route class is a routing tactic specific constant representing the resource consumed when the route starts and/or ends. An on-arc class represents the resource consumed during transportation between nodes, and is proportional to the total distance of the route. An on-site class represents the resource consumed during activities to obtain profits, and is proportional to the sum of time spent on sites included in the route.

Given a base and a set of sites, only a minimum distance path that can be obtained from solving the TSP is considered as a viable route for the problem; we do not consider deliberately inefficient routes. Thus the total number of routes related to a base, regardless of the feasibility of routes, equals the total number of subsets of the site set $\mathbf{E}$ (2^{n_E}). We define $\mathbf{J} = \{0, \ldots, 2^{n_E} - 1\}$ as an index set of subsets of the site set $\mathbf{E}$.[6]

Routing tactic $k \in \mathbf{T}^s$, where $\mathbf{T}^s$ is a set of routing tactics available for mission strategy s, characterizes a route by specifying single-route constraints and a maximum number of routes. Single-route constraints are expressed using constraining resource types, resource consumption coefficient vectors ($\mathbf{c}_0^{s,k}$, $\mathbf{c}_d^{s,k}$, $\mathbf{c}_\tau^{s,k}$) and a resource consumption limit vector for the constraining resources ($\mathbf{l}_\tau^{s,k}$). The maximum number of routes $n^{s,k}$ limits the number of routes using the tactic $k \in \mathbf{T}^s$ included in a mission.

$J_f^{b,s,k} \subset J$ is an index set of feasible routes related to the base b with respect to a routing tactic k of a mission strategy s and is defined as the following:

$$J_f^{b,s,k} = \{j \in J | \underbrace{\mathbf{c}_0^{s,k}}_{per-route} + \underbrace{\mathrm{TSP}_j^b \cdot \mathbf{c}_d^{s,k}}_{on-arc} + \underbrace{(\sum_{i \in \mathbf{R}_j} t_i) \cdot \mathbf{c}_\tau^{s,k}}_{on-site} \leq \mathbf{l}_\tau^{s,k}\}, \tag{2.1}$$

where $\mathbf{R}_j$ is the exploration site subset with an index j and TSP_j^b is travel distance of the TSP solution for the nodes ($\{b\} \cup \mathbf{R}_j$).

Mission strategy $s \in \mathbf{S}$, where $\mathbf{S}$ is a set of available strategies for the campaign, characterizes a mission by specifying a set of available routing tactics ($\mathbf{T}^s = \{1, \ldots, t^s\}$), collective constraints, and a mission cost (C^s). Collective constraints are expressed using constraining resource types, resource consumption coefficient vectors ($\mathbf{d}_0^s$, $\mathbf{d}_d^s$, $\mathbf{d}_\tau^s$), and a resource consumption limit vector for the constraining resources

[6]Combined with a base index $b \in \mathbf{B}$, index $j \in \mathbf{J}$ can represent a route.

($\mathbf{l}_c^s$).

We assume all missions in a single campaign are funded under the same budget source. There is a budget constraint on the cost sum for all missions in a campaign; The cost sum cannot exceed a pre-determined budget M. Figure 2-9 shows the hierarchical structure of decisions and parameters for the GLRPP.

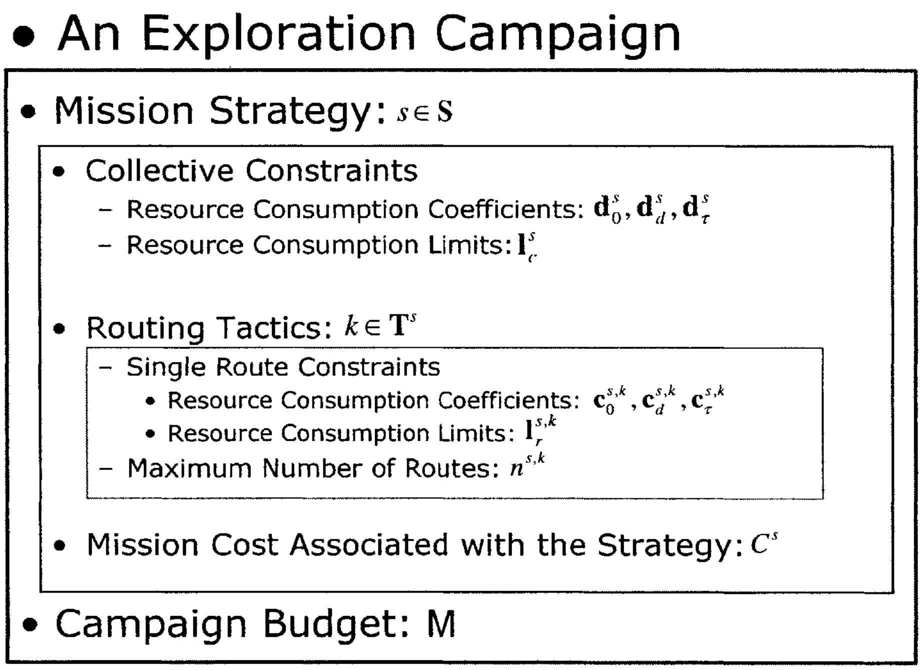


Figure 2-9: Decision and Parameter Hierarchy for the GLRPP

The objective of the campaign is to maximize the sum of profits that can be obtained during the campaign while not exceeding the campaign budget. Decision variables for this problem are selection of bases that are used, mission strategies for the used bases, selection of routes to visit sites from each of the selected bases, and routing tactics for the selected routes. The solution of the problem should satisfy single-route constraints, maximum route number constraints, collective constraints, and the campaign budget constraint. A constraint that a single site cannot be visited multiple times should also be satisfied.

The GLRPP is mathematically formulated as an integer programming (IP) problem in the next section.

2.6 Mathematical Formulation of the GLRPP

The origin of the mathematical formulation introduced in this section is the set-covering formulation for routing problems, which was first suggested by Balinski and Quandt [8] and has been successfully used to solve the *Vehicle Routing Problem* (VRP) and its variants [4, 28, 29, 55].

For the GLRPP, a set-packing formulation, which is a modified version of the original set-covering formulation, is used [17]. First a scalar version of the GLRPP formulation is presented as follows:

(GLRPPs) Generalized Location Routing Problem with Profits (Scalar Version)[7]

$$\max_{x_j^{b,s,k}, y^{b,s}} \sum_{b\in\mathbf{B}} \sum_{s\in\mathbf{S}} \sum_{k\in\mathbf{T}^s} \sum_{j\in J_f^{b,s,k}} \left(r_j \cdot x_j^{b,s,k} \right) \tag{2.2}$$

subject to

$$\sum_{b\in\mathbf{B}} \sum_{s\in\mathbf{S}} \sum_{k\in\mathbf{T}^s} \sum_{j\in J_f^{b,s,k}} \left(\mathbf{A}_j \cdot x_j^{b,s,k} \right) \leq \mathbf{1}_{n_E}, \tag{2.3}$$

$$\sum_{k\in\mathbf{T}^s} \sum_{j\in J^{b,s,k}} \left(\mathbf{h}_j^{b,s} \cdot x_j^{b,s,k} \right) \leq \mathbf{l}_c^s \cdot y^{b,s} \quad (\forall\, b\in\mathbf{B},\ \ \forall\, s\in\mathbf{S}), \tag{2.4}$$

$$\sum_{j\in J^{b,s,k}} x_j^{b,s,k} \leq n^{s,k} y^{b,s} \quad (\forall\, b\in\mathbf{B},\ \ \forall\, s\in\mathbf{S},\ \ \forall\, k\in\mathbf{T}^s), \tag{2.5}$$

$$\sum_{s\in\mathbf{S}} y^{b,s} \leq 1 \quad (\forall\, b\in\mathbf{B}), \tag{2.6}$$

$$\sum_{b\in\mathbf{B}} \sum_{s\in\mathbf{S}} C^s y^{b,s} \leq M, \tag{2.7}$$

$$x_j^{b,s,k} \in \{0,1\} \quad (\forall\, b\in\mathbf{B},\ \ \forall\, s\in\mathbf{S},\ \ \forall\, t\in\mathbf{T}^s), \tag{2.8}$$

$$y^{b,s} \in \{0,1\} \quad (\forall\, b\in\mathbf{B},\ \ \forall\, s\in\mathbf{S}). \tag{2.9}$$

$x_j^{b,s,k}$ and $y^{b,s}$ are binary decision variables for this problem. $x_j^{b,s,k}$ takes value

[7]To discriminate from the generic problem name, a mathematical formulation is printed using a Sans-Serif font. So the GLRPP refers to the *Generalized Location Routing Problem with Profits* generically, but GLRPPs represents a specific formulation for the problem.

1 if a route determined by base b and site subset $\mathbf{R}_j$ using routing tactic k and mission strategy s is included in the solution and takes value 0 otherwise. $y^{b,s}$ takes value 1 if mission strategy s is selected by base b and takes value 0 otherwise. r_j is the sum of profits for all sites belonging to site subset $\mathbf{R}_j$ and equals $\sum_{i \in \mathbf{R}_j} v_i$. Constraint (2.3) requires that each site be visited no more than once. Collective constraint (2.4) imposes that the resource sum for routes in a mission be bounded. $\mathbf{h}_j^{b,s}$ is a vector representing consumption of constraining resource types for collective constraints specified by mission strategy s on the route determined by base b and site subset $\mathbf{R}_j$ and is defined as follows:

$$\mathbf{h}_j^{b,s} = \underbrace{\mathbf{d}_0^s}_{per-route} + \underbrace{\mathrm{TSP}_\mathrm{j}^\mathrm{b} \cdot \mathbf{d}_d^s}_{on-arc} + \underbrace{(\sum_{i \in R_j} t_i) \cdot \mathbf{d}_\tau^s}_{on-site}. \tag{2.10}$$

Constraint (2.5) requires that total number of routes using a routing tactic (routing tactic k of mission strategy s) in a mission be at most a certain value ($n^{s,k}$). Constraint (2.6) imposes that no more than one strategy be selected by a base. If no strategy is selected for a base b ($y^{b,s} = 0$, $\forall s \in \mathbf{S}$), then base b is not used. Constraint (2.7) imposes that the sum of costs for all missions in a campaign not exceed budget M.

Next a matrix version formulation of the GLRPP is presented.

(GLRPP) Generalized Location Routing Problem with Profits (Matrix Version)

$$\min_{\mathbf{x},\mathbf{y}} (-\mathbf{r}'\mathbf{x}) \tag{2.11}$$

subject to

$$\mathbf{Ax} \leq \mathbf{1}_{n_1}, \tag{2.12}$$

$$\mathbf{Hx} - \mathbf{Ly} \leq \mathbf{0}_{n_2}, \tag{2.13}$$

$$\mathbf{E}_1\mathbf{x} - \mathbf{Ny} \leq \mathbf{0}_{n_3}, \tag{2.14}$$

$$\mathbf{E}_2\mathbf{y} \leq \mathbf{1}_{n_4}, \tag{2.15}$$

$$\mathbf{c}'\mathbf{y} \le \mathrm{M}, \tag{2.16}$$

$$\mathbf{0} \le \mathbf{x} \le 1, \quad \mathbf{0} \le \mathbf{y} \le 1, \quad \mathbf{x} \text{ and } \mathbf{y} \text{ are integers.} \tag{2.17}$$

$\mathbf{x}$ and $\mathbf{y}$ are decision variable vectors for GLRPP, which are composed of decision variables in the equations (2.8) and (2.9). Matrices $\mathbf{r}$, $\mathbf{A}$, $\mathbf{H}$, $\mathbf{L}$, $\mathbf{E}_1$, $\mathbf{N}$, $\mathbf{E}_2$, and $\mathbf{c}$ are created out of coefficients in constraints (2.3)-(2.9). n_1, n_2, n_3, and n_4 represent the numbers of rows for constraints (2.12), (2.13), (2.14), and (2.15). n_1 $(= n_E)$ equals the number of exploration sites. n_2 $(= n_B \cdot \sum_{s \in \mathbf{S}} |\mathbf{l}_c^s|)$ represents the total number of collective constraints for all potential bases using all possible mission strategies. n_3 $(= n_B \cdot \sum_{s \in \mathbf{B}} |\mathbf{T}^s|)$ is the total number of routing tactics for all potential bases using all possible mission strategies. n_4 $(= n_B \cdot n_S)$ is the total number of mission strategies for all bases.

Two different solution methods to solve the GLRPP are proposed in the next two chapters. Chapter 3 introduces a *single phase method* and applies a column generation procedure to the GLRPP formulation to obtain a solution for the GLRPP. Chapter 4 introduces a *three-phase method*. Each phase breaks the problem down into subproblems (*Divide* phase), solves each subproblem (*Conquer* phase), and synthesizes the results (*Synthesize* phase) to obtain a solution for the GLRPP.

2.7 Sample Problem

A sample GLRPP instance will be used to clarify the problem and to explain the two different solution methods. Figure 2-10 and Table 2.2 present the information on potential bases and sites. Note that the size of the exploration sites shown is proportional to the potential profit to be obtained when the site is visited. Table 2.3 shows the campaign budget, mission characteristics, and route characteristics of the problem.

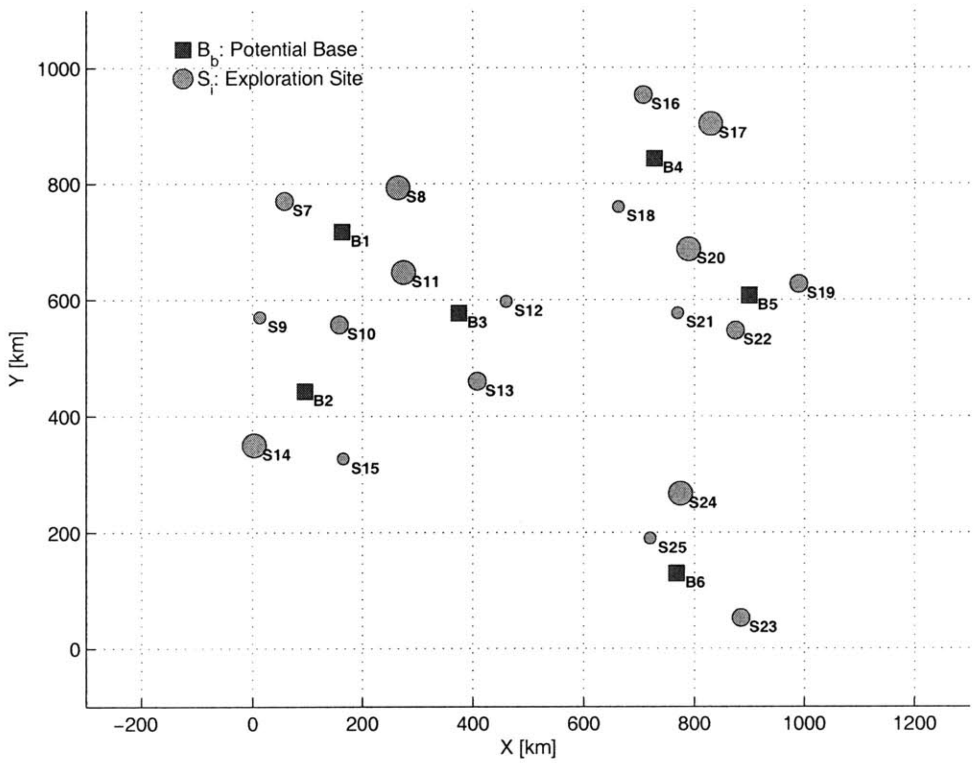


Figure 2-10: Sample GLRPP Instance

Table 2.2: Sample GLRPP Instance

Base No.	X [km]	Y [km]	Site No.	X [km]	Y [km]	Profit (v_i)	Time (t_i) [hr]
1	163	717	7	58	770	2	15
2	95	443	8	265	793	3	25
3	375	577	9	13	570	1	22
4	728	843	10	158	557	2	30
5	900	607	11	275	647	3	17
6	768	130	12	460	597	1	25
			13	408	460	2	30
			14	3	350	3	22
			15	165	327	1	28
			16	708	953	2	26
			17	830	903	3	17
			18	663	760	1	25
			19	990	627	2	13
			20	790	687	3	20
			21	770	577	1	14
			22	875	547	2	26
			23	885	53	2	23
			24	775	267	3	13
			25	720	190	1	11

Table 2.3: Strategies, Tactics and Budget for the Sample Problem

	Strategy I		Strategy II	
Collective Constraint				
Constraining Resource	Exploration time		Exploration time	
Coefficients	$d_0^1 = 0$	$[hr]$	$d_0^2 = 0$	$[hr]$
	$d_d^1 = 1/12$	$[hr/km]$	$d_d^2 = 1/12$	$[hr/km]$
	$d_\tau^1 = 1$	$[hr/hr]$	$d_\tau^2 = 1$	$[hr/hr]$
Consumption Limit	$l_c^1 = 300$	$[hr]$	$l_c^2 = 300$	$[hr]$
Routing Tactics				
Tactic 1	Extended		Standard	
Single-Route Constraint				
Resource	Fuel		Fuel	
Coefficients	$c_0^{1,1} = 0$	$[kg]$	$c_0^{2,1} = 0$	$[kg]$
	$c_d^{1,1} = 1$	$[kg/km]$	$c_d^{2,1} = 1$	$[kg/km]$
	$c_\tau^{1,1} = 3$	$[kg/hr]$	$c_\tau^{2,1} = 3$	$[kg/hr]$
Limit	$l_r^{1,1} = 600$	$[kg]$	$l_r^{2,1} = 300$	$[kg]$
Max. Number of Routes	$n^{1,1} = 1$	$[-]$	$n^{2,1} = 3$	$[-]$
Tactic 2	Standard			
Single-Route Constraint				
Resource	Fuel			
Coefficients	$c_0^{1,2} = 0$	$[kg]$		
	$c_d^{1,2} = 1$	$[kg/km]$		
	$c_\tau^{1,2} = 3$	$[kg/hr]$		
Limit	$l_r^{1,2} = 300$	$[kg]$		
Max. Number of Routes	$n^{1,1} = 1$	$[-]$		
Cost	$C^1 = 5$	$[-]$	$C^2 = 3$	$[-]$
Campaign Budget		18 $[-]$		

2.8 Chapter Summary

This chapter introduces the Generalized Location Routing Problem with Profits (GLRPP) by abstraction of an optimization problem for a global planetary surface exploration campaign. A literature review for routing problems related with the GLRPP is provided and unique features of the GLRPP are identified. The GLRPP is formulated as an IP problem and a sample problem is provided that will be used for the explanation of the solution methods in next two chapters.

Chapter 3

Single-Phase Method

3.1 Single-Phase Solution Method: Introduction

The single-phase method maintains the structure of the GLRPP formulation presented in Equations (2.11)-(2.17) throughout the solution procedure. First, a linear program relaxation (LP-relaxation) of the GLRPP is implemented. Then, an optimal solution for the relaxed problem is obtained using column generation. Finally, a near-optimal feasible solution to the GLRPP is obtained by solving an integer program (IP) problem created by columns identified during the column generation procedure for the relaxed problem. The LP relaxation solution provides an upper bound of the true optimum. So we obtain a near-optimal feasible solution to the GLRPP with the upper bound of an optimum.

The following sections in this chapter present details of the single-phase method such as formulating an LP-relaxation and its dual problem, solving the relaxed LP using column generation, and obtaining a near-optimal solution and an optimality gap. In the last section of this chapter the sample problem from Section 2.7 is solved using the single-phase method.

3.2 Formulating LP Relaxation of the GLRPP and Its Dual Problem

First an LP relaxation of the GLRPP is presented as follows:

(GLRPPLR) LP Relaxation of the GLRPP

$$\min_{\mathbf{x},\mathbf{y}} (-\mathbf{r}'\mathbf{x}) \tag{3.1}$$

subject to

$$\mathbf{A}\mathbf{x} \leq \mathbf{1}_{n_1}, \tag{3.2}$$

$$\mathbf{E}_1\mathbf{x} - \mathbf{N}\mathbf{y} \leq \mathbf{0}_{n_2}, \tag{3.3}$$

$$\mathbf{H}\mathbf{x} - \mathbf{L}\mathbf{y} \leq \mathbf{0}_{n_3}, \tag{3.4}$$

$$\mathbf{E}_2\mathbf{y} \leq \mathbf{1}_{n_4}, \tag{3.5}$$

$$\mathbf{c}'\mathbf{y} \leq \mathrm{M}, \tag{3.6}$$

$$\mathbf{x} \geq 0, \quad \mathbf{y} \geq 0. \tag{3.7}$$

Binary constraints imposed on the decision variables $\mathbf{x}$ and $\mathbf{y}$ in the GLRPP are now relaxed as nonnegative real constraints. Note that an upper bound for each decision variable is automatically set to be 1 by constraints (3.2) and (3.5), which is consistent with the original binary constraints.

To apply column generation to the GLRPPLR, dual variables associated with constraints of the GLRPPLR are required. A dual problem of the GLRPPLR is formulated as follows:

(GLRPPLRD) Dual Problem of the GLRPPLR

$$\max_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4,\mathrm{p}_5} (\mathbf{p}_1'\mathbf{1}_{n_1} + \mathbf{p}_4'\mathbf{1}_{n_4} + \mathrm{p}_5 \cdot \mathrm{M}) \tag{3.8}$$

subject to

$$\mathbf{p}_1'\mathbf{A} + \mathbf{p}_2'\mathbf{E}_1 + \mathbf{p}_3'\mathbf{H} \leq -\mathbf{r}', \tag{3.9}$$

$$-\mathbf{p}_2'\mathbf{N} - \mathbf{p}_3'\mathbf{L} + \mathbf{p}_4'\mathbf{E}_2 + \mathrm{p}_5\mathbf{c}' \leq \mathbf{0}', \tag{3.10}$$

$$\mathbf{p}_1 \leq \mathbf{0}, \quad \mathbf{p}_2 \leq \mathbf{0}, \quad \mathbf{p}_3 \leq \mathbf{0}, \quad \mathbf{p}_4 \leq \mathbf{0}, \quad \mathrm{p}_5 \leq 0. \tag{3.11}$$

$\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, and p_5 are dual variables associated with constraints (3.2), (3.3), (3.4), (3.5), and (3.6), respectively. Equations (3.9) and (3.10) are related to $\mathbf{x}$ and $\mathbf{y}$, respectively. Next, we discuss the column generation method to solve the GLRPPLR.

3.3 Solving the GLRPPLR Using Column Generation

Although binary constraints imposed on the decision variables for the GLRPP are relaxed in the GLRPPLR, it is still prohibitively difficult to solve because of the large problem size. The column generation method is used to solve the large size problem without enumerating all columns [11]. The method has been widely used to get optimal or near-optimal solutions for various types of routing problems for about 20 years [58, 62].

Usually only a small portion of routes out of all possible routes are used in the final solution. Restated in a physical sense, only a small portion of routes out of all possible routes are used in final exploration design. Thus, we can obtain an optimal solution for the GLRPPLR without solving the full-size problem as follows: (1) Efficiently select columns so that the columns include routes for the optimal solution; (2) Create an LP using selected columns; and (3) Solve the created problem to optimality.

For every iteration of the column generation procedure, we create an LP using columns that have been identified so far (the GLRPPLRc: a sub-problem of the GLRPPLR), and get optimal primary variables and optimal dual variables for the GLRPPLRc. (From now on, the columns that have been identified so far are referred to as the *current set of columns.*) Using these optimal variables, we can identify new columns that should be added to create a new GLRPPLRc. After the new GLRPPLRc is created, we go back and repeat the procedure for the next iteration. We stop iterating when we cannot identify any column to be added to the current set of columns.

Details of column generation are explained in the rest of this section. First, the formulation for the sub-problem of the GLRPPLR created using the current set of columns is presented.

(GLRPPLRc) Sub-Problem of the GLRPPLR Using the Current Set of Columns

$$\min_{\mathbf{x}_c, \mathbf{y}} (-\mathbf{r}_c' \mathbf{x}_c) \tag{3.12}$$

subject to

$$\mathbf{A}_c \mathbf{x}_c \leq \mathbf{1}_{n_1}, \tag{3.13}$$

$$\mathbf{E}_{1c} \mathbf{x}_c - \mathbf{N} \mathbf{y} \leq \mathbf{0}_{n_2}, \tag{3.14}$$

$$\mathbf{H}_c \mathbf{x}_c - \mathbf{L} \mathbf{y} \leq \mathbf{0}_{n_3}, \tag{3.15}$$

$$\mathbf{E}_2 \mathbf{y} \leq \mathbf{1}_{n_4}, \tag{3.16}$$

$$\mathbf{c}' \mathbf{y} \leq \mathrm{M}, \tag{3.17}$$

$$\mathbf{x}_c \geq 0, \quad \mathbf{y} \geq 0. \tag{3.18}$$

Columns related to $\mathbf{x}$ are obtained through column generation. All columns related to $\mathbf{y}$ are included in the *current* set of columns from the beginning.

Let $\mathbf{p}_{1c}^*$, $\mathbf{p}_{2c}^*$, $\mathbf{p}_{3c}^*$, $\mathbf{p}_{4c}^*$, and p_{5c}^* be optimal dual variables of the GLRPPLRc. These variables satisfy the dual feasibility conditions for the GLRPPLRc and the following relations hold:

$$\mathbf{p}_{1c}^{*}{}' \mathbf{A}_c + \mathbf{p}_{2c}^{*}{}' \mathbf{E}_{1c} + \mathbf{p}_{3c}^{*}{}' \mathbf{H}_c \leq -\mathbf{r}_c{}', \tag{3.19}$$

$$-\mathbf{p}_{2c}^{*}{}' \mathbf{N} - \mathbf{p}_{3c}^{*}{}' \mathbf{L} + \mathbf{p}_{4c}^{*}{}' \mathbf{E}_2 + \mathrm{p}_{5c}^{*} \mathbf{c}' \leq \mathbf{0}', \tag{3.20}$$

$$\mathbf{p}_{1c}^* \leq \mathbf{0}, \quad \mathbf{p}_{2c}^* \leq \mathbf{0}, \quad \mathbf{p}_{3c}^* \leq \mathbf{0}, \quad \mathbf{p}_{4c}^* \leq \mathbf{0}, \quad \mathrm{p}_{5c}^* \leq 0. \tag{3.21}$$

If $\mathbf{x}_c^*$ and $\mathbf{y}_c^*$ are also optimal for the GLRPPLR, $\mathbf{p}_{1c}^*$, $\mathbf{p}_{2c}^*$, $\mathbf{p}_{3c}^*$, $\mathbf{p}_{4c}^*$, and p_{5c}^* should satisfy (3.9), (3.10), and (3.11) as well. Because equations (3.20) and (3.21) are identical to (3.10) and (3.11), respectively, the only condition that should be checked for optimality of the GLRPPLRc is (3.9).

Inversely, if we can find a column that is not contained in the current sub-problem and optimal dual variables of the current GLRPPLRc do not satisfy (3.9) related to the column, then we can conclude that the column should be added to the next GLRPPLRc.

This procedure of identifying columns for which the dual feasibility condition related to the current optimal variables is not satisfied and adding those columns to the current problem to create a new GLRPPLRc is referred to as *column generation*. Efficient column generation is very important to increase the calculation speed of the overall optimization; it should be carefully designed with consideration of the problem characteristics.

A column generation procedure to solve the GLRPPLR is now introduced. Table 3.1 summarizes the column generation procedure introduced in this section. Let $x_j^{b,s,k}$ be a decision variable indicating whether a route for exploration subset $j \in \mathbf{J}_f^{b,s,k}$ corresponding to base $b \in \mathbf{B}$ using mission strategy $s \in \mathbf{S}$ and routing tactic $k \in \mathbf{T}^s$ is included in the solution. Also let $\mathbf{A}_j$ $(= \mathbf{A}_j^{b,s,k})$, $\mathbf{E}_{1j}^{b,s,k}$, and $\mathbf{H}_j^{b,s}$ $(= \mathbf{H}_j^{b,s,k})$ be column vectors of $\mathbf{A}$, $\mathbf{E}_1$, and $\mathbf{H}$ which are associated with decision variable $x_j^{b,s,k}$. An element of $\mathbf{r}$ which is associated with $x_j^{b,s,k}$ is r_j $(=\sum_{i \in \mathbf{R}_j} v_i)$. The ith element of $\mathbf{A}_j$ is equal to 1 if site i belongs to site subset $\mathbf{R}_j$ and 0 otherwise. An element of $\mathbf{E}_{1j}^{b,s,k}$ corresponding to base b, mission strategy s, and routing tactics k is equal to 1, and all other column elements are 0s. $\mathbf{H}_j^{b,s}$ is a collective resource consumption vector related to a route. Its elements related to base b and strategy s are $\mathbf{h}_j^{b,s}$ as defined in Equation (2.10) and all other column elements are 0s. Structures of the column vectors associated with decision variable $x_j^{b,s,k}$ are presented in Figure 3-1.

Table 3.1: Column Generation Procedure for Solving the GLRPPLR

STEP 1:	Calculate q_i's for all nodes $i \in \mathbf{E}$ using profits and dual solution.
STEP 2:	Sort q_i's in a descending order. $(q_{a_1} \geq \ldots \geq q_{a_p} > 0 > \ldots \geq q_{a_{n_E}})$
STEP 3:	Base b, mission strategy s, and routing tactic k are selected
STEP 4:	$Z = \{1\},\ P = \{1, \ldots, p\}$.
STEP 5:	$\mathbf{R}_Z \equiv \{a_i \mid i \in Z\}$. Let j be an index such that $\mathbf{R}_j = \mathbf{R}_Z$.
STEP 6:	If $j \notin J_f^{b,s,k}$, go to **STEP 8**.
STEP 7:	If $\sum_{i \in \mathbf{R}_j} q_i + p_2^{b,s,k} + (\mathbf{p}_3^{b,s})'\mathbf{h}_j^{b,s} > 0$, a column corresponding to b, s, k, and j is generated.
STEP 8:	If $j \in J_f^{b,s,k}$ in **STEP 6**, $Z \leftarrow \bar{Z}$. Otherwise, $Z \leftarrow \tilde{Z}$. $\bar{Z}$: P's subset. Lexicographically the next of Z. $\tilde{Z}$: P's subset. Lexicographically the next of Z where $\lvert\tilde{Z}\rvert \leq \lvert Z\rvert$.
STEP 9:	If no more Z's can be identified in **STEP 7** or the number of generated columns exceeds the maximum number of iterations, go to **STEP 3** and try with a different base, mission strategy, and routing tactics combination. Once the column generation has been carried out for all combinations of bases, mission strategies, and routing tactics, stop column generation. Otherwise, go to **STEP 5**.

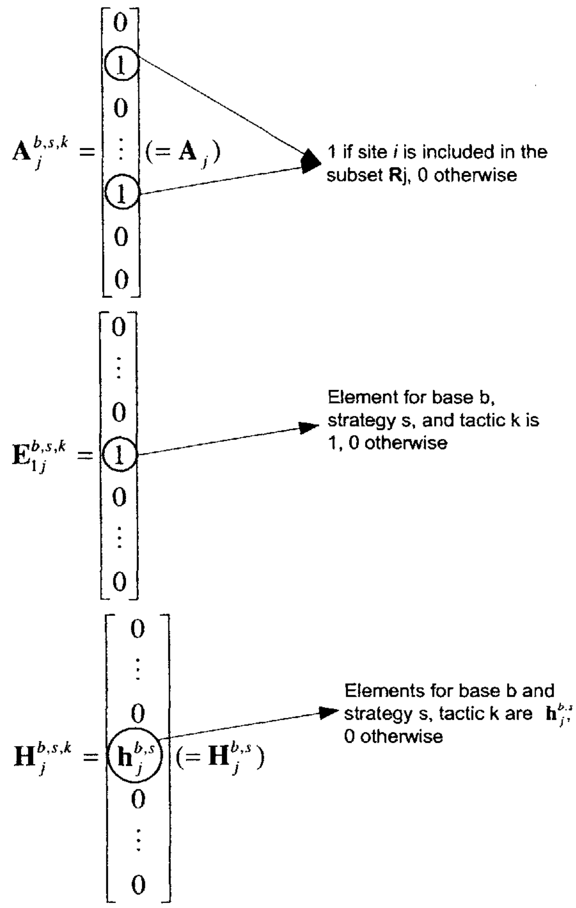


Figure 3-1: Structures of the columns associated with $x_j^{b,s,k}$

For every base, strategy, and tactic, we want to identify routes that are feasible (satisfying route constraints) and violate the dual feasibility condition expressed as Equation (3.9); these routes should be added as columns to make a new GLRPPLRc. That is, for each b, s, and k, we want to find out subsets of exploration sites, or j's, which satisfy the following condition:

$$\mathbf{p}_{1c}^{*}{}'\mathbf{A}_j + \mathbf{p}_{2c}^{*}{}'\mathbf{E}_{1j}^{b,s,k} + \mathbf{p}_{3c}^{*}{}'\mathbf{H}_j^{b,s} > -r_j. \tag{3.22}$$

Recall that the ith element of $\mathbf{A}_j$ equals 1 if site i belongs to $\mathbf{R}_j$, which is a subset of exploration sites with an index j, and 0 otherwise. So the following relation holds:

$$\mathbf{p}_{1c}^{*}{}'\mathbf{A}_j = \sum_{i \in \mathbf{R}_j} p_{1i}, \tag{3.23}$$

where p_{1i} is the ith element of $\mathbf{p}_{1c}^*$. Similarly, we can express $\mathbf{p}_{2c}^*{}'\mathbf{E}_{1j}^{b,s,k}$ and $\mathbf{p}_{3c}^*{}'\mathbf{H}_j^{b,s}$ using $\mathbf{E}_{1j}^{b,s,k}$ and $\mathbf{H}_j^{b,s}$ as follows:

$$\mathbf{p}_{2c}^*{}'\mathbf{E}_{1j}^{b,s,k} = \mathrm{p}_2^{b,s,k}, \tag{3.24}$$

$$\mathbf{p}_{3c}^*{}'\mathbf{H}_j^{b,s} = \left(\mathbf{p}_3^{b,s}\right)' \mathbf{h}_j^{b,s}, \tag{3.25}$$

where $p_2^{b,s,r}$ is an element of $\mathbf{p}_{2c}^*$ which is related to base b, strategy s, and tactic k, and $\mathbf{p}_3^{b,s}$ is a vector comprising elements of $\mathbf{p}_{3c}^*$ that are related to base b and strategy s.

Using the relation that $r_j = \sum_{i\in\mathbf{R}_j} v_i$ and Equations (3.24) and (3.25), we can rewrite Equation (3.22) as follows:

$$\sum_{i\in\mathbf{R}_j} q_i + p_2^{b,s,k} + \left(\mathbf{p}_3^{b,s}\right)' \mathbf{h}_j^{b,s} > 0, \tag{3.26}$$

where a new variable q_i is defined as $q_i \equiv (p_{1i} + v_i)$.

We know that $\mathbf{p}_{2c}^*$ and $\mathbf{p}_{3c}^*$ are non-positive from constraint (3.11). $\mathbf{h}_j^{b,s}$ is the resource consumption vector and should be non-negative. This leads to the conclusion that if the condition expressed as (3.26) holds, the summation $\sum_{i\in\mathbf{R}_j} q_i$ should be positive.

For each iteration, q_i's are sorted by a descending order and only i's with positive q_i's are used to create candidate routes. Each positive q_i is assigned a number identical to its rank, which is used for lexicographical indexing. For given b, s, and k, we start creating candidate routes using the site with the greatest q_i value. Suppose that we have a feasible candidate route at one point of the iteration. Then we add a site to the current feasible candidate site set to create a new route. A site set which is lexicographically the next of the current set is selected to represent the new route [57]. We first check the feasibility of the new route in terms of single-route constraints. If the new route is feasible, it also becomes a candidate route, and we check if it satisfies equation (3.26). If so, we add the new candidate route into the problem as a new column, otherwise we do not. In any case the new candidate route becomes a new *current* candidate route for the column generation iteration. If the new route is infeasible, then we try a site set that is lexicographically the next of the current site set whose size is no greater than the current site set. When we cannot find any new candidate site set which is lexicographically larger than the current one, we stop column generation for the specific base, mission strategy, and routing tactic combination and move on to a column generation procedure for the next combination. This procedure will be done exhaustively for all possible combinations of b, s, and k.

Note that this column generation procedure requires an initial set of columns at the beginning of the procedure. For every combination of base, mission strategy, and routing tactic, all feasible round trips between the base and sites are used for the initial columns.

When the column generation procedure is completed we have final constraint and objective matrices ($\mathbf{A}_f$, $\mathbf{E}_{1f}$, $\mathbf{H}_f$, and $\mathbf{r}_f$, where $(\cdot)_f$ represents the matrix created by columns at the final iteration of column generation). Also, we have the optimal solution for the GLRPPLR ($\mathbf{x}^*_{LP}$ and $\mathbf{y}^*_{LP}$) and the optimal profit sum for the GLRPPLR ($J^*_{LP} = \mathbf{r}'_f \mathbf{x}^*_{LP}$[1]). The GLRPPLR is a relaxed problem of the GLRPP and J^*_{LP} provides an upper bound of the optimal profit sum for the GLRPP.

3.4 Obtaining a Near-Optimal Solution for the GLRPP and an Optimality Gap

A solution of the GLRPPLR is likely to be fractional and not feasible for the GLRPP; we still have to find an integer solution. There are a number of ways to use the current set of columns to generate an optimal or a near-optimal integer solution. We use a cutting plane method to obtain a near-optimal integer solution. The cutting plane method does not generate any additional columns after the column generation for LP relaxation is completed. An IP is created using the current set of columns (GLRPPf),

[1]Although our problem is conceptually a maximization problem, the formulation of the GLRPP is expressed as a minimization problem for the convenience of handling. So an (optimal profit sum) equals the (-optimal objective function).

and an optimal solution of the GLRPPf is obtained ($\mathbf{x}^*_{IP}$, $\mathbf{y}^*_{IP}$, and J^*_{IP}). The optimal solution and profit sum for the GLRPPf is the near-optimal solution and profit sum for the GLRPP, respectively. Note that the optimal solution for the GLRPPf is not guaranteed to be optimal for GLRPP, which is the original IP formulation. There may be columns in the optimal solution of GLRPP that are not included in the GLRPPf. But an upper bound of the profit sum is provided by J^*_{LP}, and we can obtain an optimality gap which represents a worst-case bound on its relative error [62].

Let J^*, J^*_{LP}, and J^*_{IP} be optimal profit sums for the GLRPP, the GLRPPLR, and GLRPPf, respectively. The GLRPPLR is the relaxation of the GLRPP, therefore we can claim that J^*_{LP} is no smaller than J^*. Also the columns comprising the GLRPPf belong to the column set for the GLRPP and we can claim that J^* is no smaller than J^*_{IP}. Thus, the following relation holds:

$$J^*_{IP} \leq J^* \leq J^*_{LP}. \tag{3.27}$$

In addition, an optimality gap using the single-phase method can be calculated by the following equation:

$$G^s_{opt} = \frac{J^*_{LP} - J^*_{IP}}{J^*_{IP}} \cdot 100[\%]. \tag{3.28}$$

Figure 3-2 shows a flow chart for the single-phase solution method.

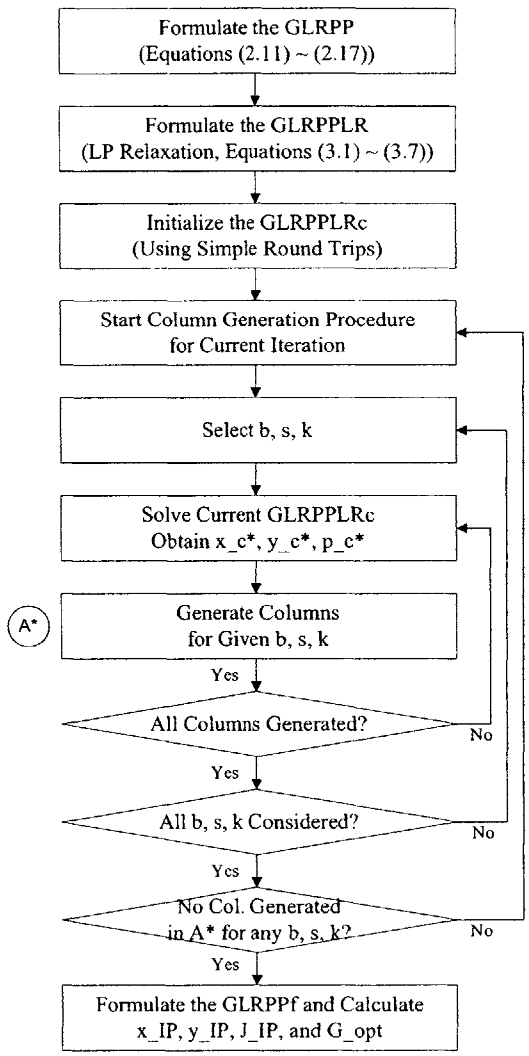


Figure 3-2: Flow Chart for the Single-Phase Solution Method

3.5 Sample Problem Using the Single-Phase Method

The sample problem described in Section 2.7 is solved using the single-phase solution method introduced in this chapter. Table 3.2 shows the summary of the sample problem solution obtained by the single-phase method. A total of 65 columns are created during the procedure to solve the GLRPPLR. An optimal solution to the GLRPPLR contained 23 nonzero decision variables. 15 nonzeros are route-related decisions ($\mathbf{x}$) and 8 nonzeros are base/strategy/tactics decisions ($\mathbf{y}$). 10 out of 15 nonzero $\mathbf{x}$ elements and 7 out of 8 nonzero $\mathbf{y}$ elements are fractional values, meaning that this optimal solution of GLRPPLR is not feasible for the original IP (GLRPP). The optimal profit sum for the GLRPPLR is 25.7, which is an upper bound of the optimal profit sum for the GLRPP. The GLRPPf is created using the set of generated columns. In the solution of the GLRPPf, eight routes are selected related to four bases. Three bases use the mission strategy 1 and one base uses the mission strategy 2. The total cost of the whole campaign is 18, which equals the budget constraint amount. The optimal profit sum for the GLRPPf is 25.0 [-] and the optimality gap for the sample problem is 3 [%]. Actual columns associated with the base 1 of the solution (columns for $y^{1,1}$, $x_{19}^{1,1,1}$, and $x_{1}^{1,1,2}$) are presented in Figure 3-4. Figure 3-3 graphically shows the sample problem solution.

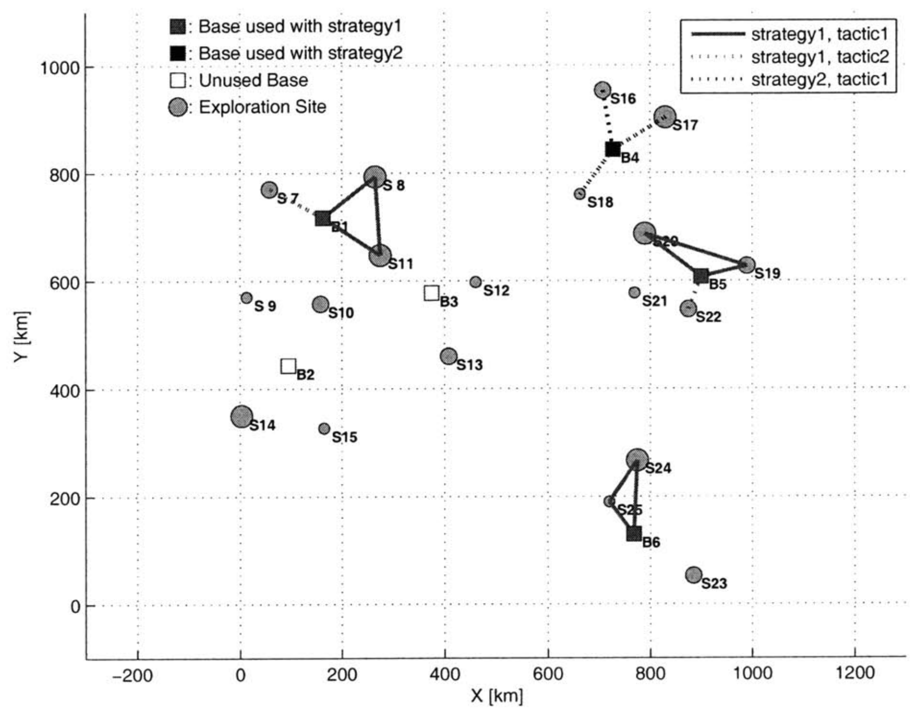


Figure 3-3: Sample Problem Solution using a Single-Phase Method

Table 3.2: Sample Problem Solution Summary - The Single-Phase Method

	GLRPPLR	GLRPPf
No. Gen. Col.	65	-
Nonzero Solutions (Fractional Values)		
y	$y^{1,1} = 1.00$ $(y^{2,1} = 0.30)$ $(y^{3,2} = 0.33)$ $(y^{4,1} = 0.33)$ $(y^{4,2} = 0.67)$ $(y^{5,1} = 0.67)$ $(y^{5,2} = 0.33)$ $(y^{6,1} = 0.50)$	$y^{1,1} = 1$ $y^{4,2} = 1$ $y^{5,1} = 1$ $y^{6,1} = 1$
x	$x_1^{1,1,2} = 1.00$ $x_{32}^{3,2,1} = 1.00$ $(x_{2048}^{4,1,2} = 0.33)$ $x_{512}^{4,2,1} = 1.00$ $x_{1024}^{4,2,1} = 1.00$ $(x_{32768}^{5,1,2} = 0.67)$ $(x_{4096}^{5,2,1} = 0.67)$ $(x_{32768}^{5,2,1} = 0.33)$ $(x_{262144}^{6,1,2} = 0.50)$ $x_{18}^{1,1,1} = 1.00$ $(x_{384}^{2,1,1} = 0.30)$ $(x_{10240}^{4,1,1} = 0.33)$ $(x_{24576}^{5,1,1} = 0.33)$ $(x_{12288}^{5,1,1} = 0.33)$ $(x_{393216}^{6,1,1} = 0.50)$	$x_1^{1,1,2} = 1$ $x_{512}^{4,2,1} = 1$ $x_{1024}^{4,2,1} = 1$ $x_{2048}^{4,2,1} = 1$ $x_{32768}^{5,1,2} = 1$ $x_{18}^{1,1,1} = 1$ $x_{12288}^{5,1,1} = 1$ $x_{393216}^{6,1,1} = 1$
Profit Sum	25.7	25.0
Opt. Gap	3 [PCT]	
Total Cost	18	18

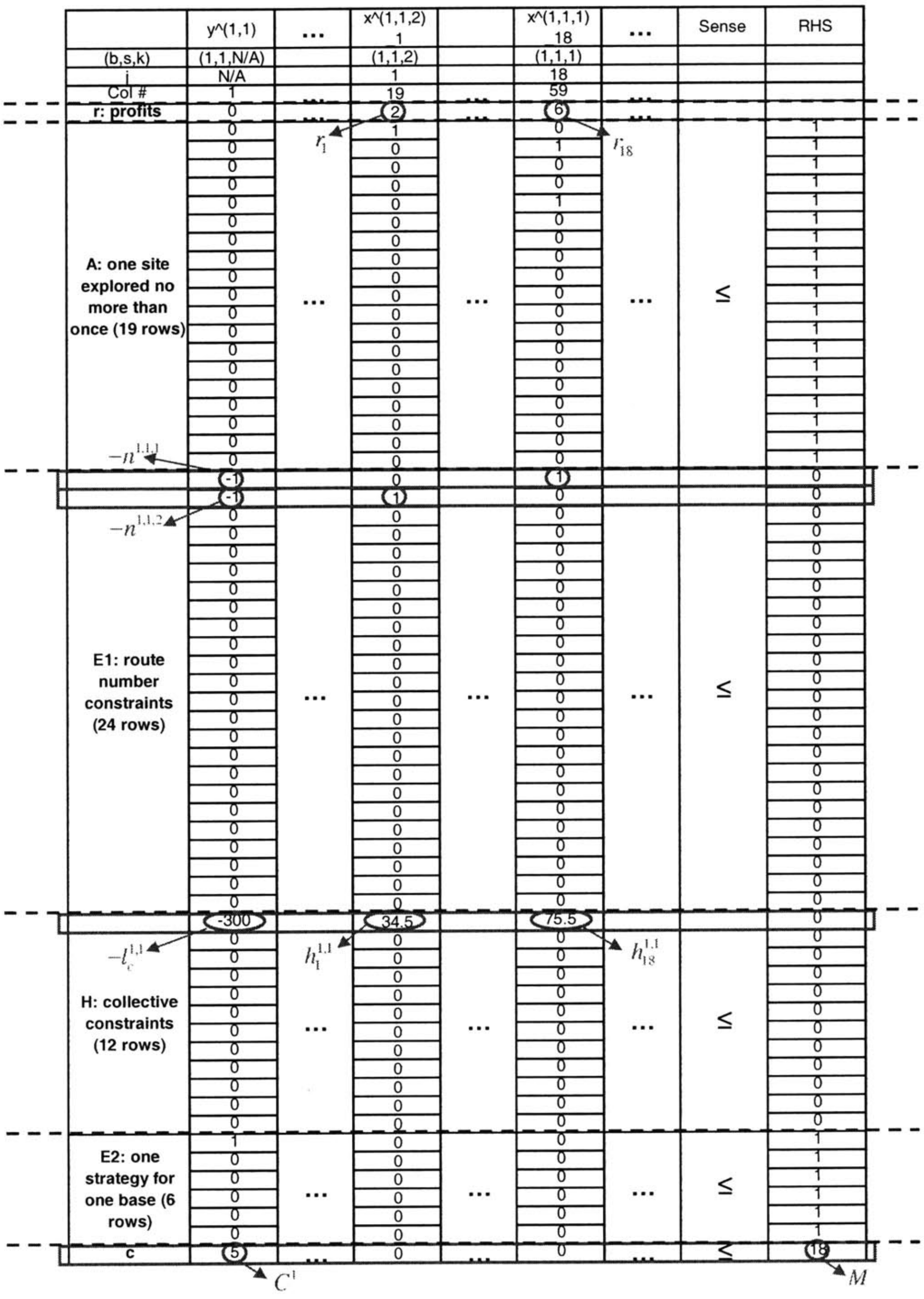


Figure 3-4: Columns associated with $y^{1,1}$, $x_1^{1,1,2}$, and $x_{18}^{1,1,1}$

3.6 Single-Phase Method Summary

The *Single-Phase Method* to solve the GLRPP is developed in this chapter. An original IP formulation of the GLRPP is relaxed to an LP (GLRPPLR). The relaxed LP is solved using a column generation method. Using columns generated during the procedure to solve the relaxed LP, a sub-IP of the original formulation (GLRPPf) is created and solved to obtain a near-optimal solution. We can also get an upper bound of the optimal profit sum from the solution of the GLRPPLR, and calculate an optimality gap from the upper bound and the near-optimal solution.

Chapter 4

Three-Phase Method

4.1 Three-Phase Solution Method: Introduction

The three-phase method is presented in this chapter as the second method to solve the GLRPP. The scope of decisions that should be made in the GLRPP is broad. The problem decides bases selected for missions, mission strategies for the selected bases, routes for each selected base, and routing tactics for the routes. The decision variable vector $\mathbf{x}$ of the GLRPP determines routes and routing tactics and the decision vector $\mathbf{y}$ determines bases and mission strategies for the bases. In the single-phase method, the two decision vectors are found simultaneously using one single IP problem. Here we consider a different approach.

Consider the case when decision vector $\mathbf{y}$ for the GLRPP is already determined. In this case, bases and associated mission strategies are selected by $\mathbf{y}$, and the GLRPP becomes a new problem. The new problem decides routes and routing tactics corresponding to each selected base to maximize profit sum subject to the constraints on resource consumption for routes and missions. We refer to this problem as the *Multi-Depot Vehicle Routing Problem with Profits* (MDVRPP). The MDVRPP is easier to solve than the GLRPP because $\mathbf{y}$ is already given and the only decision vector is $\mathbf{x}$.

Constraints (2.15) and (2.16) of the GLRPP formulation are related with only $\mathbf{y}$. Suppose we enumerate all $\mathbf{y}$'s satisfying the two constraints and solve the MDVRPP for each of them. We can find the optimal solution of the GLRPP by selecting the case which has the maximum MDVRPP solution out of the enumeration. Let us refer to this approach as the *Full Factorial Strategy Method.*

To use this *Full Factorial Strategy Method* for solving the GLRPP, we have to solve the MDVRPP for every $\mathbf{y}$ which satisfies constraints (2.15) and (2.16). The number of $\mathbf{y}$'s which satisfy the constraint (2.6) equals $(1 + n_S)^{n_B}$ because there are $(1 + n_S)$ choices for each base (either use the base using one of n_S mission strategies or do not use the base) and we have n_B bases. Some of these $\mathbf{y}$'s may not satisfy the constraint (2.7) and $(1 + n_S)^{n_B}$ is an upper bound for the number of feasible $\mathbf{y}$'s.

Consider a GRLPP instance presented in Figure 4-1-(a). The instance has two mission strategies and three potential bases, and the maximum number of feasible $\mathbf{y}$'s equals $(1 + 2)^3 = 27$. Figure 4-2 shows the GLRPP solution for the instance using the *Full Factorial Strategy Method.* A table enumerating all combinations of strategy selection for bases is created. Each row of the table is associated with a specific decision vector $\mathbf{y}$. Total cost for every row is calculated for the feasibility check of the campaign budget constraint. For every feasible row, the MDVRPP is solved and the profit sum for the solution is obtained. After the table is fully populated, we select the row which has the maximum profit sum value. Decision vectors $\mathbf{x}$ and $\mathbf{y}$ for the row are taken as the solution of the GLRPP instance.

One big hurdle inherent in this method is the large number of combinations for $\mathbf{y}$ when we have many potential bases - the number grows exponentially with the number of potential bases. When there are 2 mission strategies and the number of potential bases equals 10, the maximum number of the MDVRPP we have to solve is $(1 + 2)^{10} = 59,049$. For 100 potential bases, the number increases to $(1 + 2)^{100} \approx 5.2 \cdot 10^{47}$. Even though the MDVRPP is less complex than the GLRPP, the benefit of reducing the problem complexity is lost in this situation. An observation presents a hint to resolve this difficulty: *all bases do not have to be handled together to solve the MDVRPP.* The node set for the GLRPP is divided into two groups in Figure 4-1-(b). This group is referred to as a *cluster.* (A formal definition of a *cluster* is provided in the next section.) Each cluster is also a node set composed of bases and sites. Assume that no route can have nodes from multiple clusters.

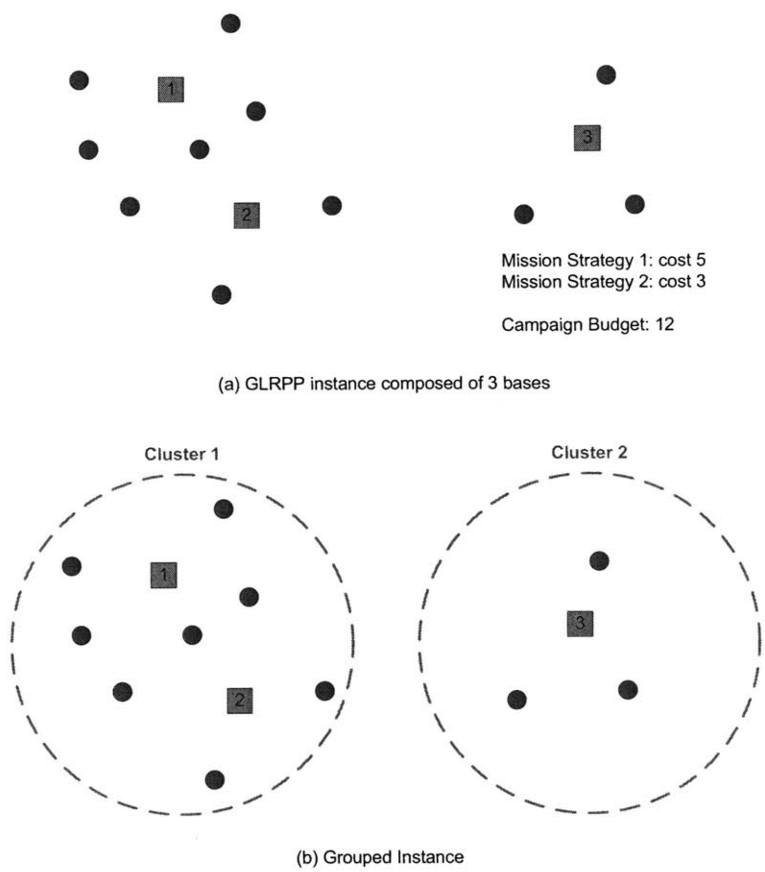


Figure 4-1: GLRPP Instance and Two Clusters for the Instance

Consider a feasible **y** and the associated MDVRPP. At the same time, we can also consider the MDVRPP for each cluster because a cluster is also a node set and decisions on **y** also fix mission strategies for bases in each cluster. Let optimal profit sums of the MDVRPPs for the original node set, cluster 1, and cluster 2 be V, V_1, and V_2, respectively. Then a relation among these three values that $V = V_1 + V_2$ holds because no route can have nodes from multiple clusters; hence, it is impossible that the solutions of MDVRPPs for cluster 1 and cluster 2 have any site in common.

Using this property, we can create any row of the table in Figure 4-2 using the two tables presented in Figure 4-3. For example, suppose we want to know the maximum profit sum for the MDVRPP when mission strategy 1 is used for base 1 and mission strategy 2 is used for base 2 and base 3. We refer to Table 1 (for cluster 1) of Figure 4-3 to find out the optimal profit sum for the MDVRPP when base 1 is used with strategy 1 and base 2 is used with strategy 2 (Case 7) and obtain the profit sum

	Base 1	Base 2	Base 3	Cost	Feasibility	Prof. Sum
Case 00	not used	not used	not used	0	Yes	0
Case 01	str. 1	not used	not used	5	Yes	8
Case 02	str. 2	not used	not used	3	Yes	5
...	...	...	..	...	...	...
Case 13	str. 1	str. 1	str. 1	15	No	N/A
...	...	...	...	...	...	...
Case 25	str. 1	str. 2	str. 2	11	Yes	15
Case 26	str. 2	str. 2	str. 2	9	Yes	12

Problem to solve:
Find a row with maximum profit sum.

Result:
Case 25

Totla # of the MDVRPP to solve:
3^3=27

Figure 4-2: Solving the Instance Using the *Full Factorial Strategy Method*

of 12. Similarly we refer to Table 2 (for cluster 2) to find out the profit sum for the MDVRPP when base 3 is used with strategy 2 (Case 2) and obtain the value of 3. Then the profit sum we want to know is calculated by summation of the two values: $12 + 3 = 15$, which is consistent with the information for the Case 25 of Figure 4-2. Now we can obtain the solution of the GLRPP for the instance. We solve an optimization problem to select one row from each table to maximize total profit sum subject to the constraint that cost sum for the selected rows do not exceed the campaign budget. Actual routes can be also taken from the routes related to the two rows from Table 1 and Table 2.

Let us compare the amount of effort made for the two methods presented in Figure 4-2 and Figure 4-3. In the *Full Factorial Method*, we solve the MDVRPP 27 times. The number is reduced to 12 (=9+3) times in the modified method.[1] We can reduce the number of the MDVRPP calculation typically more than 50 percent by dividing the whole node set to multiple clusters, solving the MDVRPPs for each cluster, and synthesizing the results from all clusters to obtain the final GLRPP solution. Reduction in calculation effort is more dramatic when a large node set is divided into many clusters. Consider the situation when there are 100 bases in the instance. We would have to solve the MDVRPP $5.2 \cdot 10^{47}$ times if we apply the *Full*

[1]The modified method requires additional efforts related to making clusters and finding the solution of the GLRPP out of multiple *result tables* but time consumed on these efforts are negligible.

Table 1

	Base 1	Base 2	Cost	Feasibility	Prof. Sum
Case 00	not used	not used	0	Yes	0
Case 01	str. 1	not used	5	Yes	8
Case 02	str. 2	not used	3	Yes	5
...	...	...	...	...	...
Case 7	str. 1	str. 2	8	Yes	12
Case 8	str. 2	str. 2	6	Yes	9

Table 2

	Base 3		Cost	Feasibility	Prof. Sum
Case 00	not used		0	Yes	0
Case 01	str. 1		5	Yes	4
Case 02	str. 2		3	Yes	3

Problem to solve:

Find a row from each table which maximizes total prof. sum subject to the condition that cost sum is no more than budget.

Result:

Case 7 for Table 1 and Case 2 for Table 2

Totla # of the MDVRPP to solve:

3^2+3^1=12

Figure 4-3: Solving the Instance Using the *Modified Method*

Factorial Strategy Method. But if it is possible to divide this instance into 20 clusters such that each cluster has 5 bases and the *Modified Method* is applied, the number of the MDVRPP we have to solve is reduced to $20 \cdot (1+2)^5 = 4,860$. This number is extremely small compared with the case of the *Full Factorial Strategy Method.*

In this chapter, the *Modified Method* is refined and proposed as the second solution method to solve the GLRPP. We refer to this method as the *Three-Phase Method*; The whole procedure of the method is composed of three phases, which are referred to as the *Divide Phase*, the *Conquer Phase*, and the *Synthesize Phase*, respectively. In the *Divide Phase*, the whole node set is divided into multiple groups, each of which is referred to as a *Cluster.* For each cluster we consider base / mission strategy combinations, and each combination is referred to as a *Cluster Strategy.* In the *Conquer Phase*, we solve the *Multi-Depot Vehicle Routing Problem with Profits* (MDVRPP) for every cluster strategy. The MDVRPP can be solved using a method very similar to the Single-Phase Method to solve the GRLPP. Finally the *Synthesize Phase* collects the MDVRPP results (near-optimal solution, near-optimal profit sum, and upper-bound profit sum) for all cluster strategies and solves an IP that selects one cluster strategy for each cluster, maximizing the total profit sum for the campaign subject to the budget constraint. Figure 4-4 shows procedures and information flows for the three-phase method.

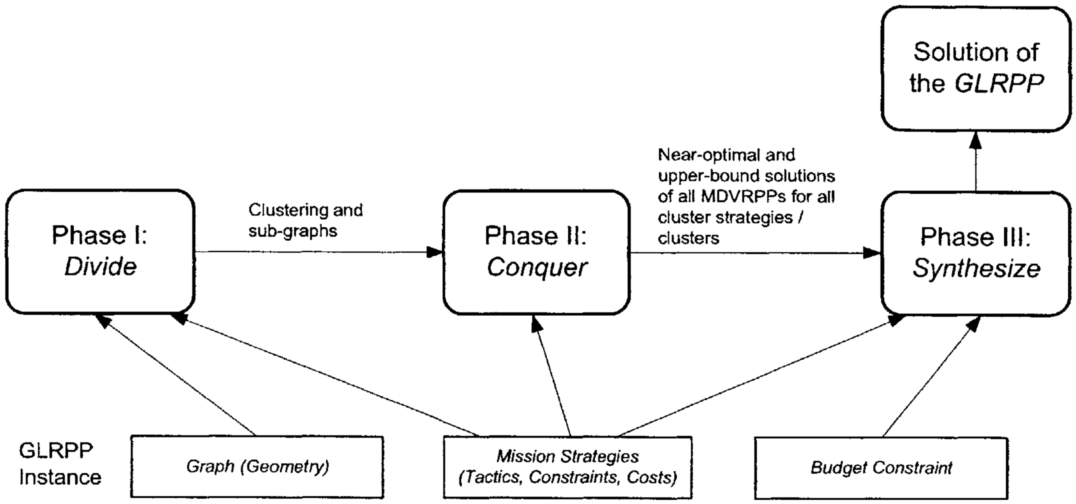


Figure 4-4: Three-Phase Method for Solving the GLRPP

The next three sections provide explanations on the *Divide Phase*, *Conquer Phase*, and *Synthesize Phase*, respectively. Then, a solution for the sample problem described in Section 2.7 is presented. Finally a comparison of numerical experiment results between the single-phase method and the three-phase method is provided in the last section.

4.2 Phase I - Divide

The objective of the *Divide Phase* is to partition the node set $\mathcal{N}$ into clusters so that any route composed of nodes belonging to different clusters cannot be feasible. We start the *Divide Phase* with identification of sites that can be visited from each base. First the term *Reachable* is defined.

Definition 1 (Reachable). *For base $b \in \mathbf{B}$ and site $i \in \mathbf{E}$, i* ***is reachable from*** *b if there exists mission strategy $s \in \mathbf{S}$ and a routing tactic $k \in \mathbf{T}^s$ such that a round trip between b and i is feasible for the routing tactic k. It can be also described using* *the following mathematical expression:*

$$\exists s \in \mathbf{S},\ \exists k \in \mathbf{T}^s,\quad s.t.\ \ \mathbf{c}_0^{s,k} + (c_{bi} + c_{ib}) \cdot \mathbf{c}_d^{s,k} + t_i \cdot \mathbf{c}_\tau^{s,k} \leq \mathbf{l}_r^{s,k}.$$

In this definition, $\mathbf{c}_0^{s,k}$, $(c_{bi} + c_{ib}) \cdot \mathbf{c}_d^{s,k}$, and $t_i \cdot \mathbf{c}_\tau^{s,k}$ are *per-route class*, *on-arc class*, and *on-site class* resource consumptions, respectively. Also $\mathbf{l}_r^{s,k}$ is the consumption limit vector of the constraining resources. Note that a base from which no site is reachable or a site which is not reachable from any base is not relevant to the problem and can be deleted from the problem instance. Throughout the discussion in this chapter, it is assumed that every considered base has at least one site which is reachable from the base and every potential site is reachable from at least one base. The definition of a *Proximity Set* follows.

Definition 2 (Proximity Set). *For base $b \in \mathbf{B}$, proximity set $\mathbf{P}_b$ is defined as a union of $\{b\}$ and the set of sites reachable from b.*

The *Divide Phase* generates a partition of the whole node set $\mathcal{N}$. This partition is referred to as *clustering* and a member of the partition is referred to as a *cluster.* The set of clusters should have the following properties.

1. Each cluster has at least one base and site.
2. For each site in a cluster, there exists at least one base in the same cluster such that the site is reachable from that base.
3. For each base in a cluster, there exists at least one site in the same cluster which is reachable from the base.
4. A cluster cannot be further partitioned into subsets which satisfy the properties described in 1-3.

Clustering and *cluster* are mathematically defined as follows.

Definition 3 (Clustering / Cluster). $\Gamma = \{\mathbf{C}_1, \cdots, \mathbf{C}_l, \cdots, \mathbf{C}_{n_C}\}$ *is a clustering of* $\mathbf{N}$ *and* $\mathbf{C}_l$ $(l = 1, \ldots, n_C)$ *is a cluster of* $\mathcal{N}$ *if all of the following conditions are satisfied.*

1. *Γ is a partition of $\mathcal{N}$.*

2. *For all $b \in \mathbf{B}$ and $l \in \{1, \ldots, n_C\}$, if $b \in \mathbf{C}_l$, then $\mathbf{P}_b \subset \mathbf{C}_l$*

3. *For all $b_1 \in \mathbf{B}$ and $l \in \{1, \ldots, n_C\}$, if $b_1 \in \mathbf{C}_l$ and $|\mathbf{C}_l \cap \mathbf{B}| \geq 2$, then there exists $b_2 \in \mathbf{B}$ such that $b_2 \in \mathbf{C}_l$ and $\mathbf{P}_{b_1} \cap \mathbf{P}_{b_2} \neq \{\}$.*

4. *For all $b_1, b_2 \in \mathbf{B}$ and $l_1, l_2 \in \{1, \ldots, n_C\}$, if $b_1 \in \mathbf{C}_{l_1}$, $b_2 \in \mathbf{C}_{l_2}$, and $l_1 \neq l_2$, then $\mathbf{P}_{b_1} \cap \mathbf{P}_{b_2} = \{\}$.*

Finally, *cluster base set* and *cluster site set* are defined.

Definition 4 (Cluster Base Set / Cluster Site Set). *Corresponding with a cluster $\mathbf{C}_l$ defined in* **Definition 3**, ***cluster base set*** $\mathbf{B}_l$ *is defined as follows:*

$$\mathbf{B}_l = \mathbf{C}_l \cap \mathbf{B}.$$

Also ***cluster site set*** $\mathbf{E}_l$ *is defined as follows:*

$$\mathbf{E}_l = \mathbf{C}_l \cap \mathbf{E}.$$

Note that generation of a clustering and corresponding clusters out of a node set $\mathcal{N}$ is analogous to identification of a set of disjoint *connected subgraphs* out of a given graph [12]. Each base in $\mathbf{B}$ can be seen as an analogy for the node in the graph and a non-emptiness of an intersection between two proximity sets related to two bases is an analogy for the existence of an (undirected) arc between two nodes.

Now we describe the *Divide Phase* using the concepts introduced in this section.

Phase 1: Divide

1. Given a GLRPP instance, construct proximity sets $\mathbf{P}_1, \ldots, \mathbf{P}_{n_B}$.

2. Using the proximity sets, construct a clustering $\Gamma = \{\mathbf{C}_1, \ldots, \mathbf{C}_{n_C}\}$ and corresponding base sets $(\mathbf{B}_1, \ldots, \mathbf{B}_{n_C})$ and site sets $(\mathbf{E}_1, \ldots, \mathbf{E}_{n_C})$.

Table 4.1: Pseudocode - Clustering

```
FOR i = 1:n_B,
    C_i = P_i;
END
i = 1;
n_C = n_B;
WHILE (1),
    WHILE (1),
        isMerged = 0;
        j = i+1;
        WHILE ( j <= n_C ),
            IF ( intersection(C_i, C_j) != NULL ),
                C_i = union( C_i, C_j );
                k = j;
                WHILE ( k <= n_C-1 ),
                    C_k = C_(k+1);
                END
                n_C = n_C -1;
                isMerged = 1;
                BREAK
            END
        j = j + 1;
        END
        IF (isMerged == 0),
            BREAK
        END
    END
    i = i+1;
    IF ( i == n_C ),
        BREAK
    END
END
```

Proximity set identification is straightforward and no additional explanation on this procedure is provided here. Cluster identification using the proximity sets needs some systematic effort. Table 4.1 shows the pseudo-code for an algorithm to identify clusters out of proximity sets.

4.3 Phase II - Conquer

In the *Conquer Phase*, we identify all possible mission strategy combinations for every cluster. Then, the MDVRPP is solved for each mission strategy combination. Both a near-optimal solution and the upper bound of the maximum profit are recorded to be used in the next phase. A combination of mission strategies for bases in one cluster is referred to as a *Cluster Strategy*. The *Cluster Strategy* is defined as follows:

Definition 5 (Cluster Strategy). *Let* $\mathbf{C}_l$ *be a cluster and* $\mathbf{B}_l = \{b_{l,1}, \ldots, b_{l,n_{B_l}}\}$ *be a cluster base set. Suppose that each base* $b_{l,m}$ *has a specific strategy* $s_{l,m}$. $(m = 1, \ldots, n_{B_l})$ *A cluster strategy* $\mathbf{s}_l = [s_{l,1}, \ldots, s_{l,n_{B_l}}]'$ *is defined as a strategy vector whose size is* n_{B_l} *and represents a combination of mission strategies selected by all bases in* $\mathbf{B}_l$.

For every base in a cluster, we have $(1 + n_S)$ choices for the use of the base: (1) using the base with strategy $s \in \mathbf{S}$ and (2) not using the base. Note that $s_l^i = 0$ if base i of the cluster is not used. Then the total number of cluster strategies for cluster l is calculated as follows:

$$n_l = (1 + n_S)^{n_{B_l}}. \tag{4.1}$$

Index g $(0 \leq g \leq n_l - 1)$ represents a cluster strategy $\mathbf{s}_l$ and is expressed as the following equation:

$$g = \sum_{m=1}^{n_{B_l}} (1 + n_B)^{m-1} \cdot s_{l,m}. \tag{4.2}$$

Now we discuss a method to solve the *Multi-Depot Vehicle Routing Problem with Profits* (MDVRPP), which is a sub-problem that should be solved for each cluster / cluster strategy. The MDVRPP can be seen as a special case of the GLRPP in which decision vector $\mathbf{y}$ is given. Therefore, the whole procedure of the single-phase method introduced in Chapter 3 is applicable with only minor modifications: (1) Mission strategy selection for bases ($\mathbf{y}$) is not a decision for the MDVRPP and the budget constraint is irrelevant; and (2) Column generation of the MDVRPP is composed of two loops - base b and routing tactic k.[2]

Chapter 3 is referred to for a detailed explanation of the solution procedure for the MDVRPP. Figure 4-5 shows the flowchart for the procedure to solve the MDVRPP.

[2]column generation of the GLRPP has three loops - base b, mission strategy s, and routing tactic k

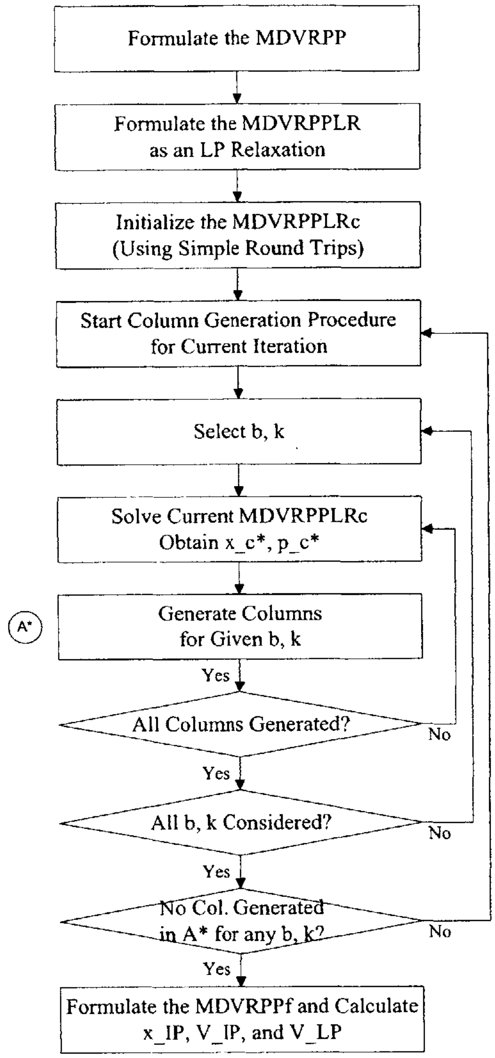


Figure 4-5: Flow Chart for the MDVRPP Solution Procedure

For cluster strategy g of cluster l, we record the near-optimal feasible solution $(\mathbf{x}^*_{IP})^{l,g}$, near-optimal feasible profit sum $(V^*_{IP})^{l,g}$, and upper bound of optimal profit sum $(V^*_{LP})^{l,g}$ of the MDVRPP. This information is used in the *Synthesize Phase*.

4.4 Phase III - Synthesize

The *Synthesize Phase* solves the problem of assigning one cluster strategy to each cluster to maximize the total profit sum for the whole campaign while ensuring that the selected set of strategies satisfies the budget constraint. A cluster strategy represents a combination of mission strategies used by bases included in the cluster. For each cluster strategy g of cluster l, we can get the near-optimal and upper-bound profit sums of the MDVRPP ($(V_{IP}^*)^{l,g}$ and $(V_{LP}^*)^{l,g}$) from the *Conquer Phase* and calculate the cost by summing up the costs associated with selected mission strategies for bases in cluster l. This problem is a relatively simple IP[3]. We first formulate a generic cluster strategy selection problem.

(CSSP) Cluster Strategy Selection Problem

$$\max_{z^{l,g}} \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left(w^{l,g} \cdot z^{l,g} \right) \tag{4.3}$$

subject to

$$\sum_{g=1}^{n_l} z^{l,g} = 1 \quad (\forall l \in \{1, \ldots, n_C\}), \tag{4.4}$$

$$\sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left(c^{l,g} \cdot z^{l,g} \right) \leq M, \tag{4.5}$$

$$z^{l,g} \in \{0,1\} \quad (\forall l \in \{1, \ldots, n_C\}, \quad \forall g \in \{1, \ldots, n_l\}). \tag{4.6}$$

Equation (4.3) indicates that we want to maximize the sum of profits that can be obtained over the whole campaign. Index l represents a cluster and index g represents a cluster strategy (defined in Equation (4.2)). $c^{l,g}$ is the cost sum of missions determined by the cluster strategy indexed as g and expressed as follows:

$$c^{l,g} = \sum_{m=1}^{n_{B_l}} C^{s_{l,m}}, \tag{4.7}$$

where g represents cluster strategy $\mathbf{s}_l = [s_{l,1}, \ldots, s_{l,n_{B_l}}]'$.

We define three different objective coefficients $(w^*)^{l,g}$, $(w_{IP}^*)^{l,g}$, and $(w_{LP}^*)^{l,g}$ as the following equations:

$$(w^*)^{l,g} = (V^*)^{l,g}, \tag{4.8}$$

$$(w_{LP}^*)^{l,g} = (V_{LP}^*)^{l,g}, \tag{4.9}$$

[3]This problem is an *Assignment Problem with a Side Constraint*

$$(w^*_{IP})^{l,g} = (V^*_{IP})^{l,g}, \tag{4.10}$$

where $(V^*)^{l,g}$, $(V^*_{LP})^{l,g}$, and $(V^*_{IP})^{l,g}$ represent a true optimal profit sum, an upper bound of the optimal profit sum, and a near-optimal profit sum of the MDVRPP for cluster l and cluster strategy g, respectively. Consider the following three problems.

(CSSPOpt) CSSP with Optimal MDVRPP Solution

$$\max_{z^{l,g}} \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*)^{l,g} \cdot z^{l,g}\right) \tag{4.11}$$

subject to Equations (4.4), (4.5), and (4.6).

(CSSPLP) CSSP with Upper Bound LP MDVRPP Solution

$$\max_{z^{l,g}} \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*_{LP})^{l,g} \cdot z^{l,g}\right) \tag{4.12}$$

subject to Equations (4.4), (4.5), and (4.6).

(CSSPIP) CSSP with Lower Bound IP MDVRPP Solution

$$\max_{z^{l,g}} \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*_{IP})^{l,g} \cdot z^{l,g}\right) \tag{4.13}$$

subject to Equations (4.4), (4.5), and (4.6).

Let K^*, K^*_{LP}, and K^*_{IP} be optimal solutions for CSSPOpt, CSSPLP, and CSSPIP, respectively. Note that K^* is the actual optimum for the GLRPP ($K^* = J^*$). We show that K^*_{LP} and K^*_{IP} provide an upper bound and a lower bound of J^*, respectively.

Theorem 1 (Lower and Upper Bounds of the GLRPP Solution Using The Three--Phase Method). *K^*_{IP} is a lower bound of J^*, and K_{LP} is an upper bound of J^*. That* *is, the following relations are satisfied:*

$$K^*_{IP} \leq J^* \leq K^*_{LP}. \tag{4.14}$$

Proof.

1. $$K^*_{IP} = \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*_{IP})^{l,g} \cdot (z^*_{IP})^{l,g} \right) \leq \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*)^{l,g} \cdot (z^*_{IP})^{l,g} \right)$$

$$\leq \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*)^{l,g} \cdot (z^*)^{l,g} \right) = K^* = J^*.$$

The first inequality holds because $(w^*_{IP})^{l,g} \leq (w^*)^{l,g}$, and the second inequality holds because $(z^*)^{l,g}$ is an optimal solution of CSSPOpt.

2. $$K^*_{LP} = \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*_{LP})^{l,g} \cdot (z^*_{LP})^{l,g} \right) \geq \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*_{LP})^{l,g} \cdot (z^*)^{l,g} \right)$$

$$\geq \sum_{l=1}^{n_C} \sum_{g=1}^{n_l} \left((w^*)^{l,g} \cdot (z^*)^{l,g} \right) = K^* = J^*.$$

The first inequality holds because $(z^*_{LP})^{l,g}$ is an optimal solution of CSSPLP, and the second inequality holds because $(w^*_{LP})^{l,g} \geq (w^*)^{l,g}$. □

Finally, an optimality gap for the overall GLRPP using the three-phase method is expressed as follows:

$$G^t_{opt} = \frac{K^*_{LP} - K^*_{IP}}{K^*_{IP}} \cdot 100[\%]. \tag{4.15}$$

4.5 Sample Problem Using the Three-Phase Method

In this section, the sample problem presented in Section 2.7 is solved using the three-phase method.

4.5.1 *Divide*

First we create proximity sets for the sample problem. Figure 4-6 graphically exhibits the proximity sets generated for the sample problem. Then a clustering procedure using the algorithm explained in the Table 4.1 is carried out. Iterations for making clusters out of the proximity sets are presented in Table 4.2.

Figure 4-7 shows a clustering result for the sample problem and Table 4.3 presents the final clusters and cluster-derived sets. After three iterations no change in the (temporary) clusters happens and the procedure terminates.

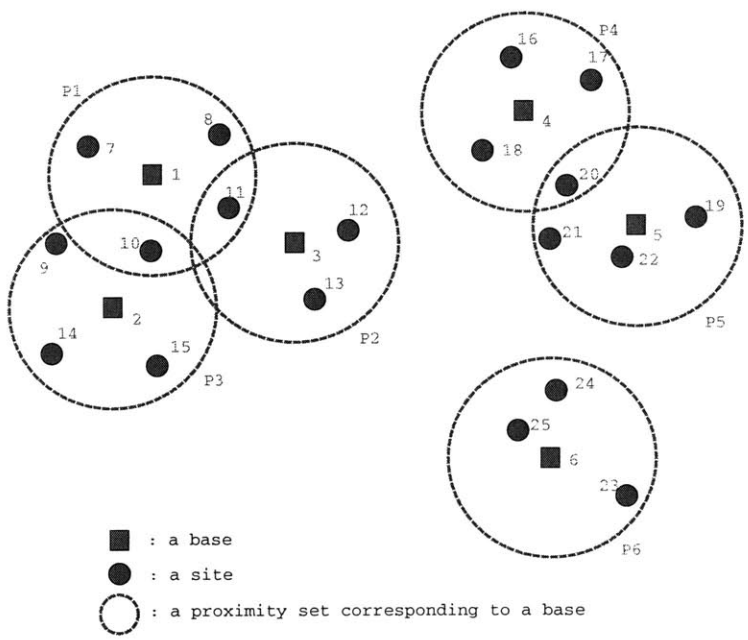


Figure 4-6: Proximity Sets for the Sample Problem

4.5.2 *Conquer*

In the *Conquer Phase*, the solution of the MDVRPP for every cluster / cluster strategy is obtained. As Table 4.3 presents, three clusters for the sample problem have been created as a result of the *Divide Phase.* Cluster 1 has three bases and the total number of strategies for the cluster equals $(1 + n_S)^{n_{B_1}} = (1 + 2)^3 = 27$. Similarly, cluster 2 has 9 cluster strategies and Cluster 3 has 3. Figure 4-8 shows all the cluster strategies for all clusters. The objective of the *Conquer Phase* is to fill in two columns

Table 4.2: Clustering Iteration for Figure 4-6

Initial	$C_1 \leftarrow P_1$
	$C_2 \leftarrow P_2$
	$C_3 \leftarrow P_3$
	$C_4 \leftarrow P_4$
	$C_5 \leftarrow P_5$
	$C_6 \leftarrow P_6$
Iteration 1	$C_1 \leftarrow C_1 \cup C_2 = P_1 \cup P_2$
	$C_2 \leftarrow C_3 = P_3$
	$C_3 \leftarrow C_4 = P_4$
	$C_4 \leftarrow C_5 = P_5$
	$C_5 \leftarrow C_6 = P_6$
Iteration 2	$C_1 \leftarrow C_1 \cup C_2 = P_1 \cup P_2 \cup P_3$
	$C_2 \leftarrow C_3 = P_4$
	$C_3 \leftarrow C_4 = P_5$
	$C_4 \leftarrow C_5 = P_6$
Iteration 3	$C_1 = P_1 \cup P_2 \cup P_3$ (unchanged)
	$C_2 \leftarrow C_2 \cup C_3 = P_4 \cup P_5$
	$C_3 \leftarrow C_4 = P_6$
Final	$C_1 = P_1 \cup P_2 \cup P_3$
	$= \{1, 2, 3, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$
	$C_2 = P_4 \cup P_5$
	$= \{4, 5, 16, 17, 18, 19, 20, 21, 22\}$
	$C_3 = P_6 = \{6, 23, 24, 25\}$

labeled as V_{IP}^* and V_{LP}^*, which represent the lower bound and the upper bound of the maximum profit sum, respectively, for the MDVRPP defined by each cluster strategy corresponding to each row.

The MDVRPPs corresponding to all the rows in Figure 4-8 are solved using the procedure introduced in Section 4.3. The solutions are summarized in Figure 4-9. For this sample problem the V_{IP}^* column is exactly the same as V_{LP}^* column, which means that the solutions for all MDVRPPs are true optima. This observation indicates that the final solution of the GLRPP, which is the result of the *Synthesize Phase*, is guaranteed to be a true optimum.

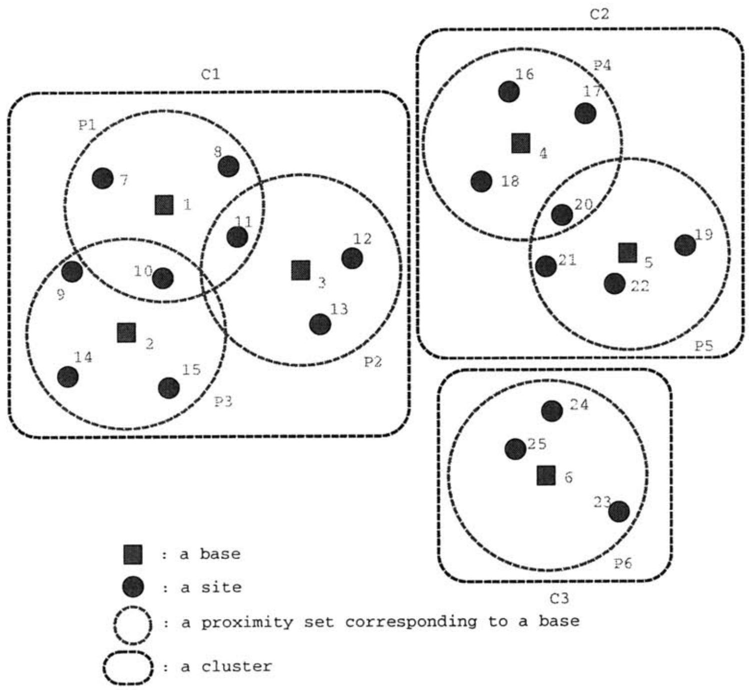


Figure 4-7: Clustering Result for Sample Problem

Table 4.3: Clustering-derived Sets in Figure 4-7

Proximity Sets	$P_1 = \{7, 8, 10, 11\}$
	$P_2 = \{11, 12, 13\}$
	$P_3 = \{9, 10, 14, 15\}$
	$P_4 = \{16, 17, 18, 20\}$
	$P_5 = \{19, 20, 21, 22\}$
	$P_6 = \{23, 24, 25\}$
Clusters	$C_1 = P_1 \cup P_2 \cup P_3$
	$C_2 = P_4 \cup P_5$
	$C_3 = P_6$
Cluster Base Sets	$B_1 = \{1, 2, 3\}$
	$B_2 = \{4, 5\}$
	$B_3 = \{6\}$
Cluster Site Sets	$E_1 = \{7, 8, 9, 10, 11, 12, 13, 14, 15\}$
	$E_2 = \{16, 17, 18, 19, 20, 21, 22\}$
	$E_3 = \{23, 24, 25\}$

Cluster	Cluster Strat.	Mission Strategy						V_IP*	V_LP*	Cost
		Base 1	Base 2	Base 3	Base 4	Base 5	Base 6			
	0	0	0	0						0
	1	1	0	0						5
	2	2	0	0						3
	3	0	1	0						5
	4	0	2	0						3
	5	0	0	1						5
	6	0	0	2						3
	7	1	1	0						10
	8	1	2	0						8
	9	2	1	0						8
	10	2	2	0						6
	11	0	1	1						10
#1 Bases: 1,2,3	12	0	1	2						8
	13	0	2	1						8
	14	0	2	2						6
	15	1	0	1						10
	16	1	0	2						8
	17	2	0	1						8
	18	2	0	2				Should be filled in during Conquer Phase		6
	19	1	1	1						15
	20	1	1	2						13
	21	1	2	1						13
	22	2	1	1						13
	23	1	2	2						11
	24	2	1	2						11
	25	2	2	1						11
	26	2	2	2						9
	0				0	0				0
	1				1	0				5
	2				2	0				3
#2 Bases: 4,5	3				0	1				5
	4				0	2				3
	5				1	1				10
	6				1	2				8
	7				2	1				8
	8				2	2				6
#3 Base: 6	0						0			0
	1						1			5
	2						2			3

Figure 4-8: Clusters and Cluster Strategies for the Sample Problem (Divide Phase)

Cluster	Cluster Strat.	Mission Strategy						V_IP*	V_LP*	Cost
		Base 1	Base 2	Base 3	Base 4	Base 5	Base 6			
	0	0	0	0				0	0.00	0
	1	1	0	0				8	8.00	5
	2	2	0	0				2	2.00	3
	3	0	1	0				4	4.00	5
	4	0	2	0				0	0.00	3
	5	0	0	1				6	6.00	5
	6	0	0	2				4	4.00	3
	7	1	1	0				12	12.00	10
	8	1	2	0				8	8.00	8
	9	2	1	0				6	6.00	8
	10	2	2	0				2	2.00	6
	11	0	1	1				10	10.00	10
#1 Bases: 1,2,3	12	0	1	2				8	8.00	8
	13	0	2	1				6	6.00	8
	14	0	2	2				4	4.00	6
	15	1	0	1				11	11.00	10
	16	1	0	2				9	9.00	8
	17	2	0	1				8	8.00	8
	18	2	0	2				6	6.00	6
	19	1	1	1				15	15.00	15
	20	1	1	2				13	13.00	13
	21	1	2	1				11	11.00	13
	22	2	1	1				12	12.00	13
	23	1	2	2				9	9.00	11
	24	2	1	2				10	10.00	11
	25	2	2	1				8	8.00	11
	26	2	2	2				6	6.00	9
	0				0	0		0	0.00	0
	1				1	0		7	7.00	5
	2				2	0		6	6.00	3
#2 Bases: 4,5	3				0	1		7	7.00	5
	4				0	2		4	4.00	3
	5				1	1		13	13.00	10
	6				1	2		11	11.00	8
	7				2	1		13	13.00	8
	8				2	2		10	10.00	6
#3 Base: 6	0						0	0	0.00	0
	1						1	4	4.00	5
	2						2	1	1.00	3

Figure 4-9: Result Summary of the Conquer Phase for the Sample Problem

4.5.3 *Synthesize*

In the *Synthesize Phase*, two IP problems (CSSPIP and CSSPLP) are created using the two columns (V_{IP}^* and V_{LP}^*) in Figure 4-9 and solved. The two columns from this sample problem are identical, and the lower bound of the optimal profit sum (K_{IP}^*) equals the upper bound of the optimal profit sum (K_{LP}^*). And the profit sum is also identical to the true optimum of profit sum (J^*) of the overall GLRPP - the optimality gap for this sample problem is 0 [%]. Figure 4-10 presents the resultant solution for the sample problem using the three-phase method. The profit sum of the sample problem using three-phase method is the same as that using the single-phase method (25 [-]). However actual routes to obtain the profits are different, which we can notice by comparing Figure 3-3 and Figure 4-10[4].

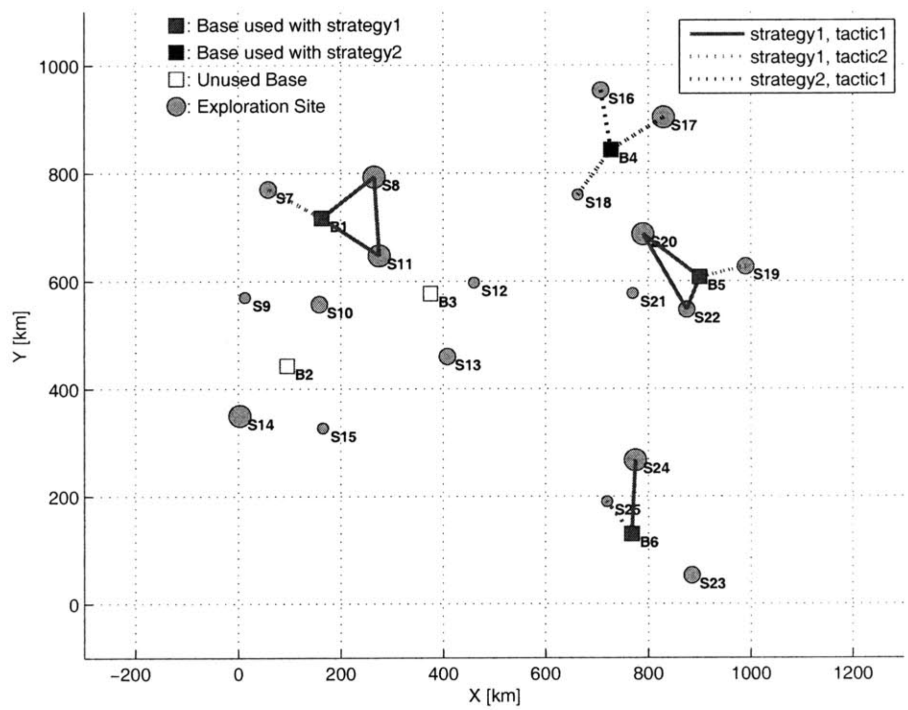


Figure 4-10: Sample Problem Solution Using a Three-Phase Method

Comparison of the two solution methods using numerical experiments is further addressed in the following section.

[4]We can know that there exist multiple optimal solutions in this sample problem.

4.6 Numerical Experiments

Numerical experiments for solving GLRPP instances using two solution methods introduced in this thesis are carried out. Two metrics representing the performance of the solution methods are obtained and compared. The first metric is the computation time (T_{cal}) representing the *efficiency* of a solution method and the second metric is the optimality gap (G_{opt})[5] representing the *effectiveness* of a solution method.

First, a brief discussion on generating GLRPP instances used in the numerical experiments is provided. Then, analysis results for the two performance metrics are presented. Estimators of the metrics for the solution methods are designed and parameters for the estimators are identified. Finally, discussion on the performance comparison of the two solution methods is presented, which can help selection of an appropriate solution method for a GLRPP instance.

4.6.1 GLRPP Instance Generation and Basic Statistics for Numerical Experiments

A GLRPP instance can be defined using information on potential base and site locations, mission strategies (including collective constraints and routing tactics), and campaign budget. Potential bases and sites are randomly generated inside a circular region with a uniform spatial distribution. The number of bases (n_B) and sites (n_E) are also randomly selected with a uniform distribution. Mission strategies and corresponding routing tactics / collective constraints for the sample problem in Table 2.3 are modified and used for the numerical experiments. We introduce a parameter referred to as the budget tightness (f_B), which is defined as the ratio of the campaign budget (M) to the maximum possible budget ($n_B \cdot \max_{s \in \mathbf{S}} C^s$). Hence the campaign

[5]For fair comparison of the optimality gaps from the two solution methods, the same basis should be used for optimality gap calculation. Thus, a *proxy* of true optimal profit sum (J^*_{pr}) is introduced as $J^*_{pr} = \max(J^*_{IP}, K^*_{IP})$. Using the value, an optimality gap for the single-phase method is calculated as $G^s_{opt} = \frac{J^*_{LP} - J^*_{pr}}{J^*_{pr}} \cdot 100$ [%] and an optimality gap for the three-phase method is calculated as $G^t_{opt} = \frac{K^*_{LP} - J^*_{pr}}{J^*_{pr}} \cdot 100$ [%].

Table 4.4: Characteristics of Instances for Numerical Experiments

Potential Bases and Sites	
Radius of an Expl. Region (R)	3000 [km]
Number of Potential Bases (n_B)	$\sim U(80, 180)$ [-]
Number of Potential Sites (n_E)	$\sim U(1000, 2000)$ [-]
Aerial Distribution of Locations	Uniform over all region
Mission Strategies (S)	Same as Table 2.3 except $l_c^1 = l_c^2 = 150$ [hr].
Routing Tactics ($\mathbf{T}^s$)	Same as Table 2.3 except $l_r^{1,1} = 400$ [kg], $l_r^{1,2} = l_r^{2,1} = 300$ [kg].
Budget Constraint	
Budget (M)	$n_B \cdot \max_{s \in \mathbf{S}} C^s$ [-]
Budget Tightness Parameter (f_B)	$2^u/2^5$ $(u \sim U(0,5))$
Size of Experiments	
No. of Instances (M)	100 [-]

budget for each problem instance is expressed as follows:

$$M = f_B \cdot n_B \cdot \max_{s \in \mathbf{S}} C^s. \tag{4.16}$$

Table 4.4 summarizes the characteristics of the random GLRPP instances generated for the numerical experiments.[6] Note that actual ranges of n_B and n_E for the GLRPP instances are usually smaller than the ranges presented in Table 4.4 because some sites are not reachable from any potential base and some potential bases have no site in their proximity sets, in which case the bases/sites are deleted from the problem instance.

All experiments are carried out on a PC Pentium-IV 3.8 GHz with 4GB of RAM under Windows XP. Both solution methods are implemented in C and compiled with the Microsoft Visual C++ 6.0 compiler. The CPLEX 10.0 callable library is used to solve the LPs and IPs that are encountered during the solution procedure.

Basic statistics for the numerical experiments are summarized in Table 4.5. The

[6] $U(a, b)$ represents a uniform distribution between a and b.

Table 4.5: Basic Statistics for the Numerical Experiment Results

Computation Time (T_{cal} [sec])				
	mean	std	min	max
Single-Phase	84	24	3	2417
Three-Phase	411	1562	4	15192
Optimality Gap (G_{opt} [%])				
	mean	std	min	max
Single-Phase	10.8	9.30	0.36	50.00
Three-Phase	0.40	0.53	0.00	3.57

mean values of the two performance metrics indicate that the single-phase method is generally better in terms of the computation time and the three-phase method is generally better in terms of the optimality gap.

4.6.2 Estimating Performance Metrics

Estimators for the computation time and the optimality gap using the two solution methods are proposed and parameters for the estimators are identified using the least-squares method [60]. Estimators for the computation time are first introduced.

Computation Time

For estimation of the single-phase method computation time, two parameters representing the size of the instance (n_B, n_E) and a parameter representing the geometric complexity of the instance ($n_{c,\mathrm{max}}$) are used. n_B is the number of potential bases and n_E is the number of potential sites. To define $n_{c,\mathrm{max}}$, we first define the geometric complexity of base b ($n_{c,b}$) as follows:

$$n_{c,b} = \sum_{i \in \mathbf{P}_b,\ i \neq b} \left(\frac{l_r}{h_i} \right), \tag{4.17}$$

where $\mathbf{P}_b$ is a proximity set for base b, l_r is the consumption limit of the constraining resource for the single-route constraint, and h_i is the resource consumption of the route represented by the round trip between b and i $(c_0 + (c_{bi} + c_{ib}) \cdot c_d + t_i \cdot c_T)$.[7] As resource consumption for the route between b and i gets smaller, the problem complexity related to the base increases because it is possible to visit more sites with limited resource (h_i in the denominator). Also when more sites are reachable from the base, the complexity increases (l_r/h_i values summed up for all i's in $\mathbf{P}_b$). Then we define the geometric complexity parameter for a GLRPP instance ($n_{c,\max}$) as follows:

$$n_{c,\max} = \max_{b \in \mathbf{B}} n_{c,b}. \tag{4.18}$$

An estimator for the computation time of instance m (T^s_m), using the number of potential bases (n^m_B), the number of sites (n^m_E), and the complexity parameter ($n^m_{c,\max}$) is proposed as follows:

$$\hat{T}^s_m = K_s \cdot (\alpha_s)^{n^m_B} \cdot (\beta_s)^{n^m_E} \cdot (\gamma_s)^{n^m_{c,\max}}, \tag{4.19}$$

where $\hat{T}^s_m$ is the estimator of T^s_m and K_s, α_s, β_s, γ_s are estimator parameters to be identified. We take the log of both sides of equation (4.19) and express $\log_{10} T^s_m$ as follows:

$$\begin{aligned} \log_{10} T^s_m &= \log_{10} \hat{T}^s_m + \epsilon^s_m \qquad (4.20) \\ &= \log_{10} K_s + (\log_{10} \alpha_s) n^m_B + (\log_{10} \beta_s) n^m_E + (\log_{10} \gamma_s) n^m_{c,b} + \epsilon^s_m, \end{aligned}$$

where ϵ^s_m is the estimation residual which is the difference between $\log_{10} T^s_m$ and $\log_{10} \hat{T}^s_m$. The Error Square Sum (ESS) is defined as follows:

$$\text{ESS} = \sum_{m=1}^{M} (\epsilon^s_m)^2. \tag{4.21}$$

Parameters K_s, α_s, β_s, and γ_s are identified so that the ESS is minimized. Identified parameters and R^2 value of the estimator for the single-phase method computation

[7] Generally there are multiple constraining resource types for the single-route constraints. In this case we take the most constraining resource type to calculate the geometric complexity index.

Table 4.6: Estimation of the Computation Time using the Single-Phase Method

Formula	$\hat{T}^s = K_s \cdot (\alpha_s)^{n_B} \cdot (\beta_s)^{n_E} \cdot (\gamma_s)^{n_{c,\max}}$	
Parameters	K_s	0.7418
	α_s	1.0072
	β_s	1.0029
	γ_s	1.0144
Goodness of Fit	R_s^2	0.7529

time are presented in Table 4.6.[8]

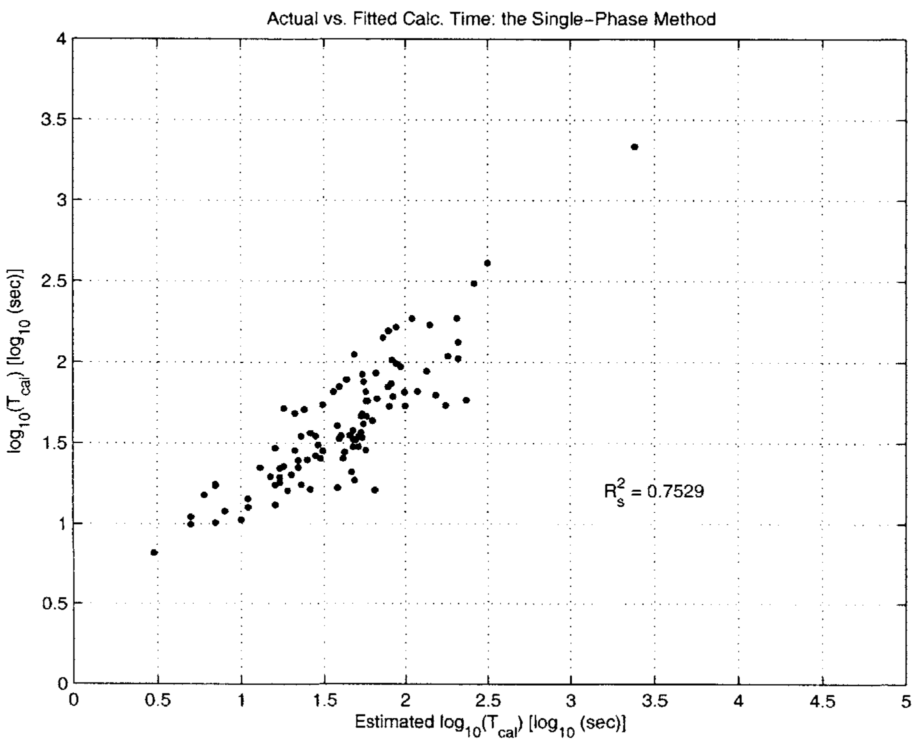


Figure 4-11: Single-Phase Method Computation Time - Actual vs. Fitted

The total time required to solve a GLRPP instance using the three-phase method is determined by the total number of the MDVRPPs that should be solved for the instance and the time required to solve each MDVRPP. The total number of the MDVRPPs that should be solved for the instance is $\sum_{l=1}^{n_C}(1+n_S)^{n_{B,l}}$ where $n_{B,l}$ is the number of bases in cluster l. $n_{B,l}$ for every cluster contributes to the total number, but

[8] R^2 value represents the fraction of the variation that can be explained by the fitting to the total variance of true values.

the number is dominated by the maximum value of $n_{B,l}$ out of all clusters. Therefore, the maximum number of bases in a cluster ($n_{B,max}$) is used as a parameter for the estimator. The time required to solve a single MDVRPP is dependent on the size of each cluster and geometric complexity. We take the maximum value for $n_{E,l}$ out of all clusters and use it as the second parameter ($n_{E,max}$). Mathematical expressions for $n_{B,max}$ and $n_{E,max}$ are presented as follows:

$$n_{B,max} = \max_{1 \leq l \leq n_C} n_{B,l}, \quad n_{E,max} = \max_{1 \leq l \leq n_C} n_{E,l}. \tag{4.22}$$

Also we use $n_{c,\max}$ as the third parameter for the three-phase method computation time estimator representing the geometric complexity of the instance.

An estimator of the computation time for solving instance m using the three-phase method ($\hat{T}^t_m$) is presented as follows:

$$\hat{T}^t_m = K_t \cdot (\alpha_t)^{n^m_{B,\max}} \cdot (\beta_t)^{n^m_{E,\max}} \cdot (\gamma_s)^{n^m_{c,\max}}. \tag{4.23}$$

$n^m_{B,\max}$ is the maximum number of bases in a cluster and $n^m_{E,\max}$ is the maximum number of sites in a cluster for instance m. $n^m_{c,\max}$ is the geometric complexity of instance m. And K_t, α_t, β_t, and γ_t are parameters for the estimator. The parameters are identified using the least-squares method. The R^2 value representing the goodness of the fit is 0.83. Figure 4-12 shows a plot for (actual three-phase method computation time) vs. (estimated three-phase method calculation time) using Equation (4.23). Table 4.7 summarizes the identified parameters for the estimator of the three-phase method computation time.

Now, we find a relation among the GLRPP instance characteristics which makes $\hat{T}^s$ and $\hat{T}^t$ identical. The relation is expressed as a boundary in the instance characteristics space. For problem instances whose characteristics are located on one side of the boundary, the single-phase method is advantageous, and for the instances with characteristics on the other side, the three-phase method is advantageous. We refer to this boundary as the *Performance Advantage Boundary*.

Three different attributes of a problem instance affect computation time of the

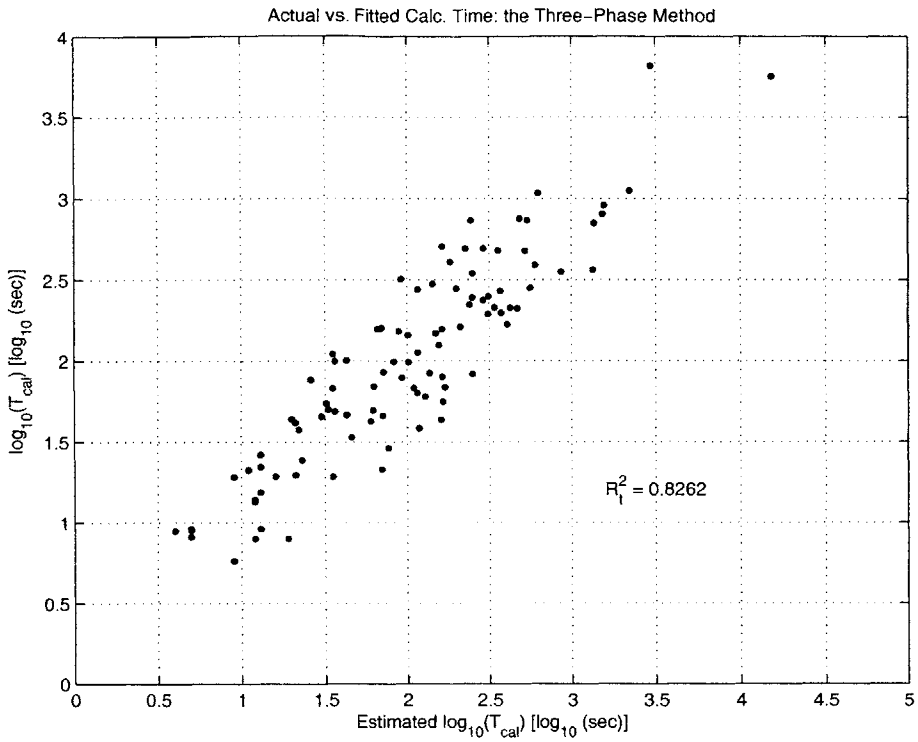


Figure 4-12: Three-Phase Method Computation Time - Actual vs. Fitted

instance: (1) overall problem size (n_B and n_E), (2) clustering complexity ($n_{B,\max}$ and $n_{E,\max}$), and (3) geometric complexity ($n_{c,\max}$). We first express the boundary in the plane whose axes represent $n_{c,\max}$ and $n_{B,\max}$ (expressed as a function of cluster complexity and geometric complexity), then we show how the performance boundary shifts as the overall problem size changes.

For simplicity we assume that the ratio of $n_{B,\max}$ to n_B equals the ratio of $n_{E,\max}$ to n_E. Then $n_{E,\max}$ can be expressed as follows:

$$n_{E,\max} = \left(\frac{n_E}{n_B}\right) \cdot n_{B,\max}. \tag{4.24}$$

Equation (4.23) can be re-written using this relation as follows:

$$\log_{10} \hat{T}^t = \log_{10} K_t + (\log_{10} \alpha_t + (n_E/n_B) \log_{10} \beta_t) \cdot n_{B,\max} + \log_{10} \gamma_t \cdot n_{c,\max}. \tag{4.25}$$

The boundary can be obtained by letting $\log_{10} \hat{T}^t = \log_{10} \hat{T}^s$. Given n_B and n_E, we

Table 4.7: Estimation of the Computation Time using the Three-Phase Method

Formula	$\hat{T}^t = K_t \cdot (\alpha_t)^{n_{B,\max}} \cdot (\beta_t)^{n_{E,\max}} \cdot (\gamma_t)^{n_{c,\max}}$	
Parameters	K_t	0.0268
	α_t	2.8441
	β_t	1.0312
	γ_t	1.0091
Goodness of Fit	R_t^2	0.8262

can derive the expression for the boundary in the $n_{c,\max}$-$n_{B,\max}$ plane as follows:

$$C_1 \cdot n_{c,\max} + C_2 \cdot n_{B,\max} + C_3 = 0 \tag{4.26}$$

where,

$$C_1 = -(\log_{10} \gamma_s - \log_{10} \gamma_t),$$
$$C_2 = \log_{10} \alpha_t + (n_E/n_B) \cdot \log_{10} \beta_t,$$
$$C_3 = \log_{10} K_t - \log_{10} K_s - \log_{10} \alpha_s \cdot n_B - \log_{10} \beta_s \cdot n_E.$$

Figure 4-13 shows the performance boundary (for calculation time) for the case of $n_B = 50$ and $n_E = 500$. Problem instances whose characteristics are located in the lower right part of the boundary can be solved more quickly using the single-phase method, and instances with characteristics in the upper left part of the boundary can be solve more quickly using the three-phase method. The single-phase method is more advantageous for problem instances with higher clustering complexity and lower geometric complexity. The clustering complexity increases the number of the MDVRPPs that the three-phase method should solve, and the effect of the geometric complexity on the computation time for the single-phase method is larger than that for the three-phase method.

Equation (4.26) indicates that the performance boundary is also dependent on the problem size (n_B and n_E). Figure 4-14 shows how the boundary moves as the problem gets less and more complex. As the problem gets bigger (large n_B and n_E)

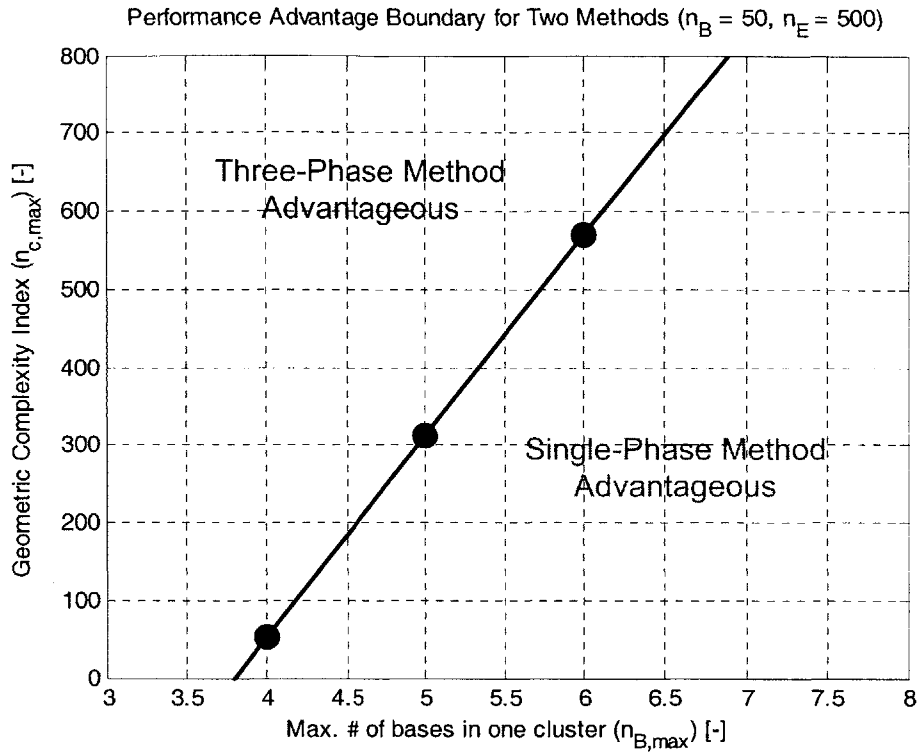


Figure 4-13: Performance Boundary

the three-phase method becomes more and more advantageous and the boundary shifts to the right.

Optimality Gap

An estimator for the optimality gap using the single-phase method is proposed considering an observation that the optimality gap for the single-phase method (G^s_{opt}) gets larger when the budget tightness factor (f_B) gets smaller. Also, it is observed that the change in G^s_{opt} is larger for small f_B. Hence, estimator $\hat{G}^s_{opt}$ is proposed as follows:

$$\hat{G}^s_{opt,m} = K^s_0 + K^s_1 \cdot f_{B,m} + K^s_2 \cdot \frac{1}{f_{B,m}}, \tag{4.27}$$

where $f_{B,m}$ is the budget tightness for instance m and K^s_0, K^s_1, and K^s_2 are estimator parameters. The parameters are identified using the least-squares method. The R^2 value for the fit is 0.76. Identified parameters are presented in Table 4.8. The same estimator is used to estimate the optimality gap using the three-phase method, but

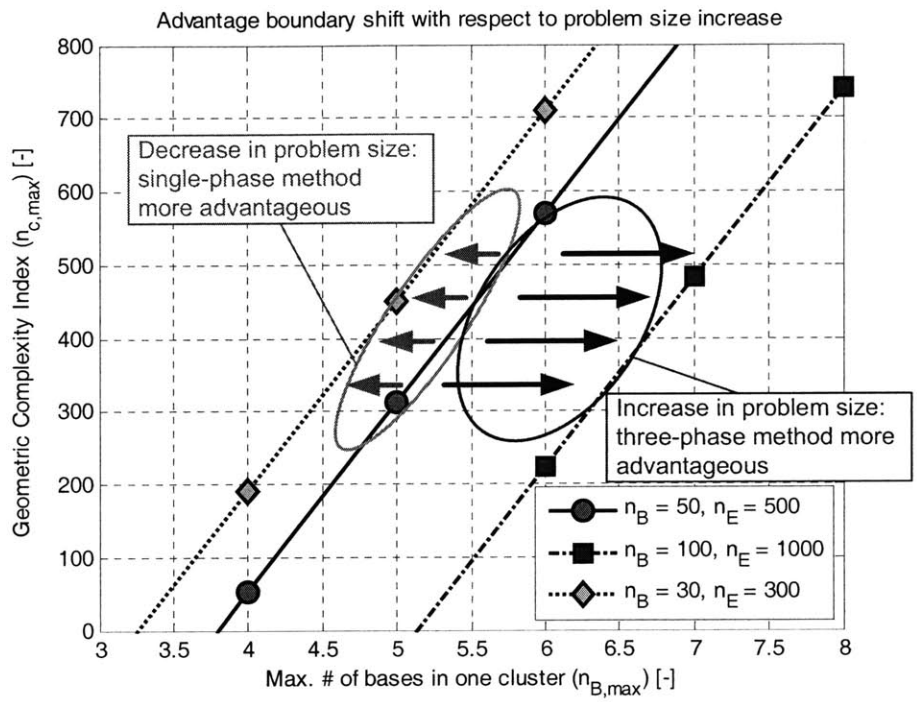


Figure 4-14: Performance Boundary Shift with Problem Size Change

Table 4.8: Estimator for the Optimality Gap Using the Single-Phase Method

Formula	$\hat{G}^s_{opt} = K^s_0 + K^s_1 \cdot f_B + K^s_2 \cdot \frac{1}{f_B}$	
Parameters	K^s_0	6.23
	K^s_1	-7.88
	K^s_2	0.88
Goodness of Fit	R^2	0.76

the R^2 value for this fit was only 0.09. Other estimators do not give meaningful fitting results, either.

Figure 4-15 shows the optimality gap for the single-phase method (G^s_{opt}) and the three-phase method (G^t_{opt}) plotted together. We can see the gaps for the three-phase method are smaller than those for the single-phase method most of the time. This indicates that the effectiveness of the solutions obtained from the three-phase method is generally better. As f_B decreases, or as the budget constraint gets tighter, G^s_{opt} gets larger, that is the single-phase method provides the less effective solution. When f_B approaches 0, the gap increases very rapidly, which is why a $1/f_B$ term is required for the estimator. When the f_B value is close to 1, the gaps for the single phase method are close to those for the three-phase method. For most space exploration problems there are many more sites than the budget allows to visit and f_B is very small. In those cases the three-phase method should be used to avoid large optimality gaps.

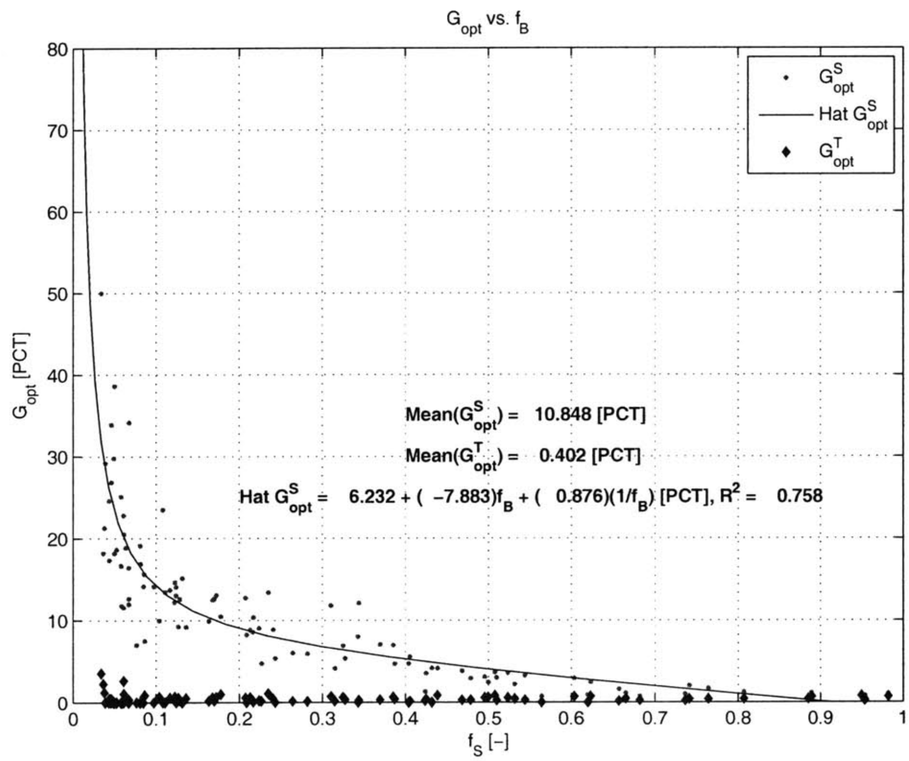


Figure 4-15: Optimality Gap, Single-Phase Method and Three Phase Method

4.6.3 Summary of Numerical Experiments

Numerical experiments for solving GLRPP instances using the single-phase method and the three-phase method have been discussed in this section. Two performance metrics are analyzed for both methods. The calculation time (T_{cal}) indicates the efficiency of each method and the optimality gap (G_{opt}) indicates the effectiveness of each method. Analysis results for the two methods are used to identify the *advantage boundary* which separates the instance characteristics space into two regions based on the performance of the solution methods.

For the generated instances, the single-phase method is superior in terms of the

Table 4.9: Performance of the two solution methods

	Single-Phase Method	**Three-Phase Method**
Calculation Time (T_{cal})	Advantageous for instances with: 1. Smaller problem size 2. Higher clustering complexity 3. Lower geometric complexity	Advantageous for instances with: 1. Larger problem size 2. Lower clustering complexity 3. Higher geometric complexity
Optimality Gap (G_{opt})	Large for a tight budget constraint	Advantageous in general

calculation time and the three-phase method is superior in terms of the optimality gap, in general.

We proposed estimators for the computation time and optimality gap using two solution methods. The estimator for the single-phase method computation time uses problem size (n_B, n_E) and geometric complexity ($n_{c,\max}$). The three-phase method computation time is estimated using clustering complexity ($n_{B,\max}$, $n_{E,\max}$) and geometric complexity ($n_{c,\max}$) as inputs. We found an expression for the *advantage boundary* and visually illustrated the boundary. The three-phase method has an advantage over the single-phase method when problem size is large, clustering complexity is low, and geometric complexity is high.

For the optimality gap, performance of the three-phase method is generally better than that of the single-phase method. The gap for the single-phase method gets larger as the budget constraint gets tighter (small f_B). A numerically significant estimator for the three-phase method optimality gap could not be found, since it is generally very small (¡ 1%).

Note that the estimator parameters presented in this section are only valid for the instances whose characteristics lie within range of the GLRPP instances generated for the experiments. However, the performance patterns for the two methods will be valid for GLRPP instances with characteristics in different ranges. Table 4.9 summarizes the performance of the solution methods regarding the two metrics that are considered in this thesis.

4.7 Three-Phase Method Summary

The three-phase solution method for solving the GLRPP is presented in this chapter. Each phase of the method (1) divides the problem into multiple sub-problems referred to as the *Multi-Depot Vehicle Routing Problem with Profits* (MDVRPP), (2) solves each subproblem using a procedure similar to the single-phase method presented in Chapter 3, and (3) synthesizes the results of the second phase to obtain a near-optimal solution and an upper bound / optimality gap for the whole GLRPP. The performance analysis for the two solution methods in terms of two metrics - the computation time and the optimality gap - is also carried out.

Chapter 5

Space Application: Global Mars Surface Exploration

5.1 Introduction

In this chapter, global Mars surface exploration is formulated as the Generalized Location Routing Problem with Profits (GLRPP), and the problem is solved using the three-phase method. Recent results for the Mars surface exploration vehicle study are used to make the case up to date and realistic [42].

An orbiting depot technology and an in-situ resource utilization (ISRU) technology are introduced as mission strategies. We demonstrate a methodology to evaluate a technology by comparing the solutions of the GLRPP with and without the technology.

5.2 Global Mars Surface Exploration Campaign Optimization

Consider an optimization problem for the global Mars surface exploration campaign. We assume that potential landing locations and exploration sites on Mars surface are pre-determined and given to the problem (**B**: potential bases, **E**: sites). We also assume the profit that can be obtained by exploring each site and the amount of time required to obtain the profit are assigned to each site as real numbers (v_i: profit, t_i: time).[1]

There are multiple exploration technology options ($\mathbf{S}$: strategies). Characteristics for each technology are constraints imposed on resource consumption for the overall mission using the technology (collective constraints), methods to carry out exploration on Mars surface (routing tactics), and the cost associated with the technology. Each routing tactic to carry out exploration on Mars surface ($k \in \mathbf{T}^s,\ s \in \mathbf{S}$) is characterized by constraints imposed on resource consumption for the route using the tactic (single-route constraints) and the number of routes that can use the method ($n^{s,k}$).

The objective of this Mars exploration problem is to find missions, bases and technology choices for the missions, routes included in each mission, and routing tactics for the routes to maximize the sum of profits obtained by exploring sites subject to constraints on a single route, a mission, and the whole campaign.

5.3 Selection of Potential Bases and Exploration Sites

We first assume that candidate bases and exploration sites have been externally determined. For previous surface exploration missions with very limited mobility, practically a landing location is not different from an exploration site. Any selected landing location should be safe for descent and ascent and be scientifically interesting. The location selection procedure for the Mars Exploration Rover (MER) mission, in consideration of both landing system topology requirements and scientific interests, is well summarized in papers written by Golombek et al. in 2003 and by Anderson et al. in 2003 [39, 5]. The surface exploration system that is used for this case study has a range of 500 [km]. This advanced capability enables agents to carry out exploration

[1]We refer to Chapter 2 for the detailed explanation on concepts such as campaign, mission, route, strategy, tactic, budget constraint, collective constraint, and single-route constraint.

far from the landing location, which makes it possible to select landing locations and exploration sites separately.

We also assume that the profit value assigned to each site is externally determined (e.g. by a committee of scientists and mission planners). Multiple stakeholder groups are interested in Mars surface exploration. For a specific exploration site, different stake-holder groups can show different degrees of interest. For the GLRPP, a profit value for a site should be determined as a scalar number. It would be best if an "agreement" on the profit value for the site could be obtained through discussion. In a more realistic situation, it is very difficult to reach the agreement on the valuation of visiting the site and group decision procedures such as the Delphi Method should be used to determine the value [51].

5.4 Mission Strategies and Routing Tactics

Two mission strategies are considered for this case ($\mathbf{S} = \{1, 2\}$). One is the standard strategy ($s = 1$), and the other is the orbiting depot strategy ($s = 2$).

We first describe the standard strategy. There is one collective resource constraint for total exploration time. To calculate the maximum exploration time for the mission (l_c^1), we assume that agents are staying on Mars' surface for 600 [days]. 150 [days] of set-up time and 100 [days] of wrap-up time are assumed. Using an additional assumption that a quarter of the working time is dedicated to exploration ($f_{exp} = 0.25$), the consumption limit for the mission is 2100 [hr] ($=l_c^1$). Table 5.1 exhibits the procedure to calculate the l_c^1 value for the standard strategy.

Now we calculate the resource consumption coefficients for the collective constraint. First we assume that no amount of time is required for starting/terminating a route ($d_0^1 = 0$). The on-arc resource consumption coefficient (d_d^1), and on-site resource consumption coefficient (d_τ^1) are determined by operational characteristics of the surface exploration vehicle. An unpressurized Vehicle (UPV)/Camper concept by Hong is used for this case study [42, 31]. Figure 5-1 and 5-2 show the conceptual diagram of the camper and the UPV-camper assembly, respectively.

Table 5.1: The Resource Consumption Limit for the Campaign in Terms of Time

Stay Time (T_{stay})	600	[day]
Set-up Time (T_{setup})	150	[day]
Wrap-up Time (T_{wrapup})	100	[day]
Work Time (T_{work})	350	[day]
Exploration Fraction (f_{exp})	0.25	[-]
Exploration Time (T_{exp})	87.5	[day]
Exp. Time Limit (l_c^1)	2100	[hr]

Table 5.2: Collective Constraint Characteristics (Standard Exploration Strategy)

Resource Type	Exploration Time		[hr]
Coefficients	d_0^1	0.0	[hr]
	d_d^1	$3/V_V = 0.2$	[hr/km]
	d_τ^1	3.0	[hr/hr]
Constraint	l_c	2100	[hr]

Figure 5-4 summarizes the reference scenario for planetary surface exploration [42]. Based on this scenario, we assume that for every one-hour activity time there is corresponding two-hour inactivity time, and the actual time consumption is the triple of the activity time.

Assuming that the agents are driving with a speed of V_V using the surface exploration vehicle, the activity time spent per unit distance equals $1/V_V$. Considering the inactivity time, d_d^1 is set to be $3/V_V$, and considering a vehicle speed of 15 [km/hr] the coefficient is 0.2 [hr/km]. Using the same logic, d_τ is set to be 3 [hr/hr]. Table 5.2 summarizes the characteristics of the collective constraint associated with the standard exploration strategy.

For the standard strategy, there is only one routing tactic - the standard tactic. No restriction for the number of routes using the standard tactic is imposed and $n^{1,1}$ is set to ∞. There is one single-route constraint for the standard tactic - fuel consumption. We first calculate the consumption limit. The power source for this case study is a $NaBH_4/H_2O_2$ fuel cell [52, 53]. The energy density of this fuel cell is summarized in

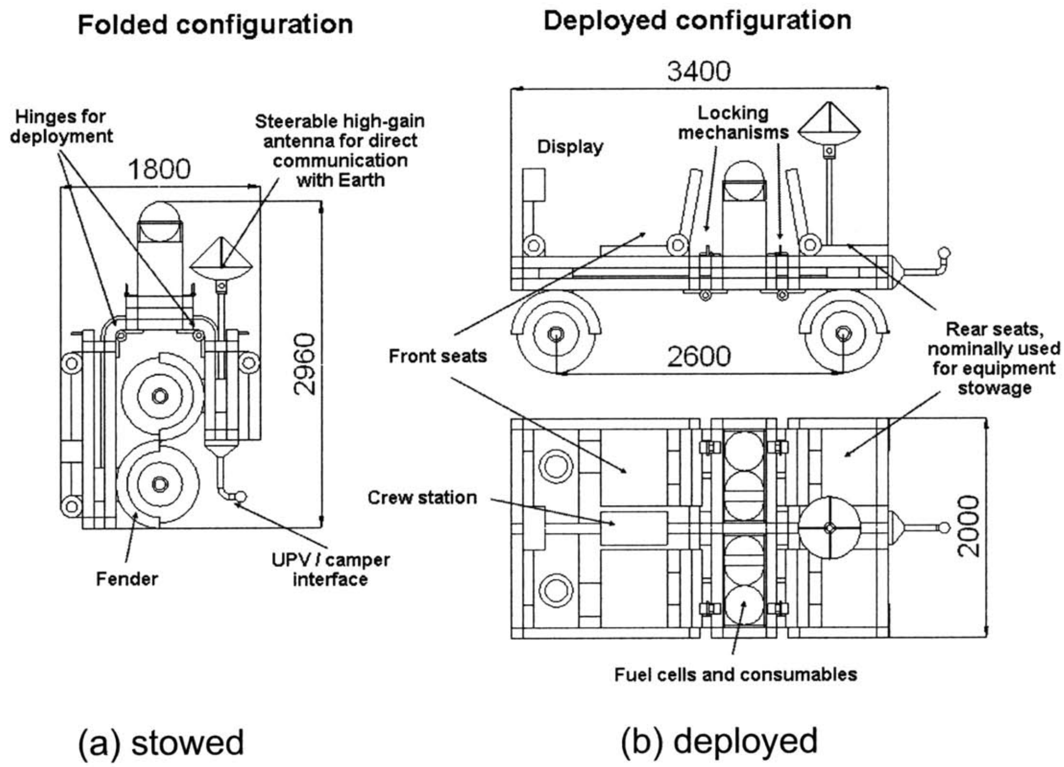


Figure 5-1: Surface Exploration Vehicle - UPV

Table 5.3: Energy Density for $NaBH_4/H_2O_2$ Fuel Cell

Ideal Energy Density (ρ_I)	2580	[WH/kg]
Overall Efficiency (η_o)	0.2653	[-]
Effective Energy Density ($\rho_E = \eta_o \rho_I$)	684.5	[WH/kg]

Table 5.3 and the activity-based power requirement for Mars surface exploration is presented in Table 5.4. We assume that 85 [%] of the power requirement is provided by the fuel cell and the remaining 15 [%] is provided by the solar panel. Table 5.5 exhibits the total amount of fuel required to complete the exploration scenario. Using the result, we conclude that fuel capacity of the surface exploration vehicle is 677 [kg] ($=m_f^c$), which is $l_r^{1,1}$ (single-route fuel consumption limit).

Then the consumption coefficients are calculated. The activity-based fuel consumption rates are presented in Table 5.6. We assume that there is no starting/terminating fuel consumption and $c_0^{1,1}$ is set to be 0 [kg]. For every active hour, the vehicle stays inactive for two hours. $c_d^{1,1}$, which is the fuel consumption per unit distance, is ex-

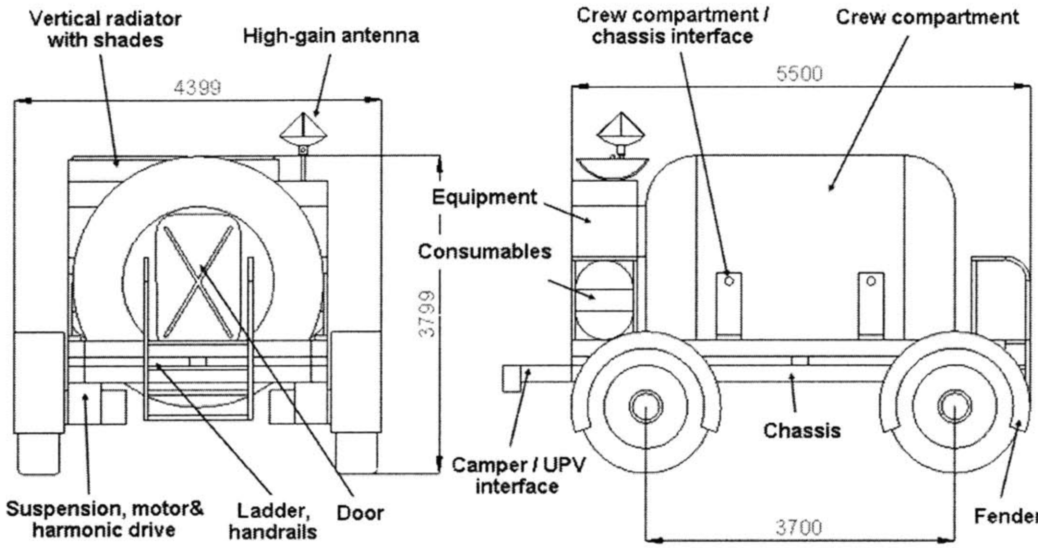


Figure 5-2: Surface Exploration Vehicle - Camper

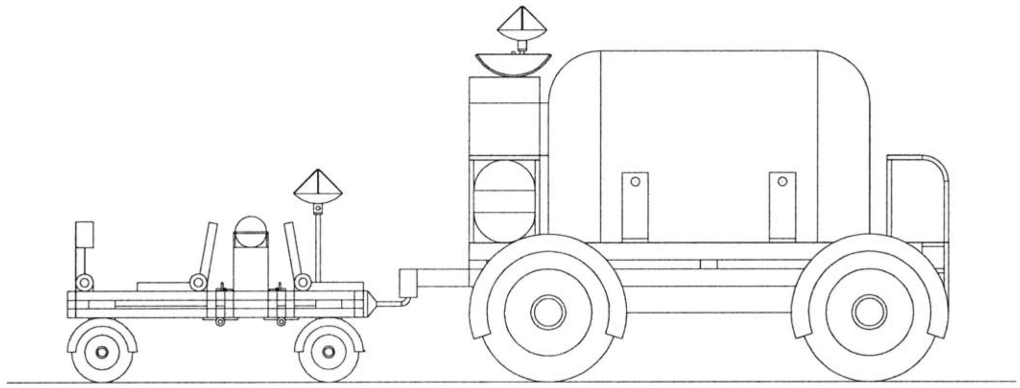


Figure 5-3: Surface Exploration Vehicle - UPV/Camper Assembly

pressed as $(\dot{m}_D/V_V + 2\dot{m}_I/V_V)$. Also $c_\tau^{1,1}$, which is the fuel consumption per unit active time on site, can be expressed as $((\dot{m}_S+2\dot{m}_E)/3+2\dot{m}_I)$. Table 5.7 summarizes the single-route constraint characteristics.

Finally cost associated with the strategy is calculated. It is very difficult to obtain the cost of a mission as a monetary value. Based on the assumption that it costs more to deliver a larger amount of mass to the planetary surface, the "Mass Delivered on Planetary Surface [MT]" is used as a proxy metric for the cost.

Zubrin's *Mars Direct* mission plan is modified and used to calculate the mass delivered to the surface of Mars [68]. The *Mars Direct* mission plan is composed

Day 1	I, 8[hr]	D, 8[hr]		I, 8[hr]
Day 2	I, 8[hr]	D, 8[hr]		I, 8[hr]
Day 3	I, 8[hr]	E, 6[hr]	S, 2[hr]	I, 8[hr]
Day 4	I, 8[hr]	E, 6[hr]	S, 2[hr]	I, 8[hr]
Day 5	I, 8[hr]	E, 6[hr]	S, 2[hr]	I, 8[hr]
Day 6	I, 8[hr]	D, 8[hr]		I, 8[hr]
Day 7	I, 8[hr]	D, 8[hr]		I, 8[hr]

I : Inactivity
D : Driving
E : Exploration
S : Science

Figure 5-4: Reference Scenario for a 7-day Excursion

Table 5.4: Power Requirement for Mars Surface Exploration Using Camper/UPV

Activity	Camper [W]	UPV [W]	Total [W]
Driving	10718	1108	11826
Science	1993	0	1993
Inactivity	1418	0	1418
Exploration	547	1108	1655

of two flights. The first flight delivers the Earth Return Vehicle (ERV) and the propellant production facility used to generate propellant for returning to the Earth. The second flight delivers the habitat, the crew, and the rovers. We do not change the contents of the first flight. The mass of rovers included in the second flight is replaced by the mass of the camper-UPV surface mobility system proposed by Hong [42]. Zubrin assumed that the rovers can use the propellant generated for the ERV (CH_4/O_2), which is not applicable for this case study. Thus 0.5 [MT] of a fuel cell regeneration system, which is the same amount of mass as the propellant production plant in Zubrin's original mission plan, is added to the second flight. The calculated total mass delivered to Mars surface is 54 [MT], which C^1. Table 5.8 summarizes the mass allocation for each flight for this case study and the total mass delivered to Mars' surface.

Now the orbiting depot strategy is introduced. An orbiting depot is a cluster of supply units circling in low Mars orbit. Each supply unit in the depot is capable of landing at a predetermined location on Mars' surface. Exploring agents can reach the

Table 5.5: Reference Exploration Route and Corresponding Fuel Budget

Activity	Time [hr]	FC Power [W]	Energy [Wh]	Fuel Mass [kg]
Driving	32	9461	302743	442
Science	8	1594	9566	14
Inactivity	112	1134	127048	186
Exploration	16	1324	23830	35
Total	168		463188	677

Table 5.6: Fuel Mass Rate for Mars Surface Exploration Using Camper/UPV

Activity	Tot. Power [W]	Fuel Cell Power [W]	Fuel Mass Rate [kg/hr]
Driving	11826	10052	14.69 ($\dot{m}_D$)
Science	1993	1694	2.47 ($\dot{m}_S$)
Inactivity	1418	1205	1.76 ($\dot{m}_I$)
Exploration	1655	1407	2.06 ($\dot{m}_E$)

landed supply unit and increase the exploration range using the additional fuel and consumables included in the supply unit.

The collective constraint for the orbiting depot strategy is the same as that of the standard strategy. So, coefficients d_0^2, d_d^2, and d_τ^2 are identical to d_0^1, d_d^1, and d_τ^1, and resource consumption limit l_c^2 equals l_c^1.

The orbiting depot strategy has two routing tactics. The first routing tactic is the standard routing used for the standard strategy. The second routing tactic is the depot-assisted tactic. Both tactics have one single-route constraint imposed on the total amount of fuel consumption for each route.

We assume that the surface exploration vehicle used for the standard strategy is also used for the orbiting depot strategy. Thus resource consumption coefficients for the two tactics of the orbiting depot strategy are identical to those for the standard strategy ($c_0^{2,1}{=}c_0^{2,2}{=}c_0^{1,1}$, $c_d^{2,1}{=}c_d^{2,2}{=}c_d^{1,1}$, $c_\tau^{2,1}{=}c_\tau^{2,2}{=}c_\tau^{1,1}$). The consumption limit for the single-route constraint and the maximum number of routes for tactic 1 of the orbiting depot strategy are also identical to those for the standard strategy ($l_\tau^{2,1}{=}l_\tau^{1,1}$

Table 5.7: Single Route Constraint Characteristics (Standard Tactic)

Constraining Resource	Fuel Consumption		[kg]
Coefficients	$c_0^{1,1}$	0.0	[kg]
	$c_d^{1,1}$	$\dot{m}_D/V_V + 2\dot{m}_I/V_V = 1.14$	[kg/km]
	$c_\tau^{1,1}$	$(\dot{m}_S + 3\dot{m}_E)/4 + 2\dot{m}_I = 5.68$	[kg/hr]
Limit	$l_\tau^{1,1}$	677	[kg]

Table 5.8: Mass Delivered on Mars Surface (Standard Strategy)

	Flight 1	**Flight 2**	**Total**
Contents	ERV, ISRU Facility	Habitat, Crew, Surf. Exp. Vehicle	
Mass Delivered	28 [MT]	26 [MT]	54 [MT]

and $n^{2,1}$=$n^{1,1}$).

We present the details of the Mars orbiting depot in the next section and identify the cost associated with the orbiting depot strategy and the characteristics of the depot-assisted routing tactic.

5.5 Design of the Mars Orbiting Depot

5.5.1 Concept of the Orbiting Depot

An orbiting depot is a cluster of supply units deployed to a Martian orbit. Each prepackaged supply unit contains supply items (propellant and food) that can extend the Mars surface exploration range.

A conceptual diagram of the orbiting depot is presented in Figure 5-5. Without the support of the orbiting depot, only the standard routing tactic is available. A vehicle starts exploration from a base and returns to the base before it runs out of fuel or consumables (food, water, and oxygen). In this case the amount of fuel and consumables that can be used during a single route is determined by the capacity of the surface exploration vehicle. Now suppose we have an orbiting depot in Martian

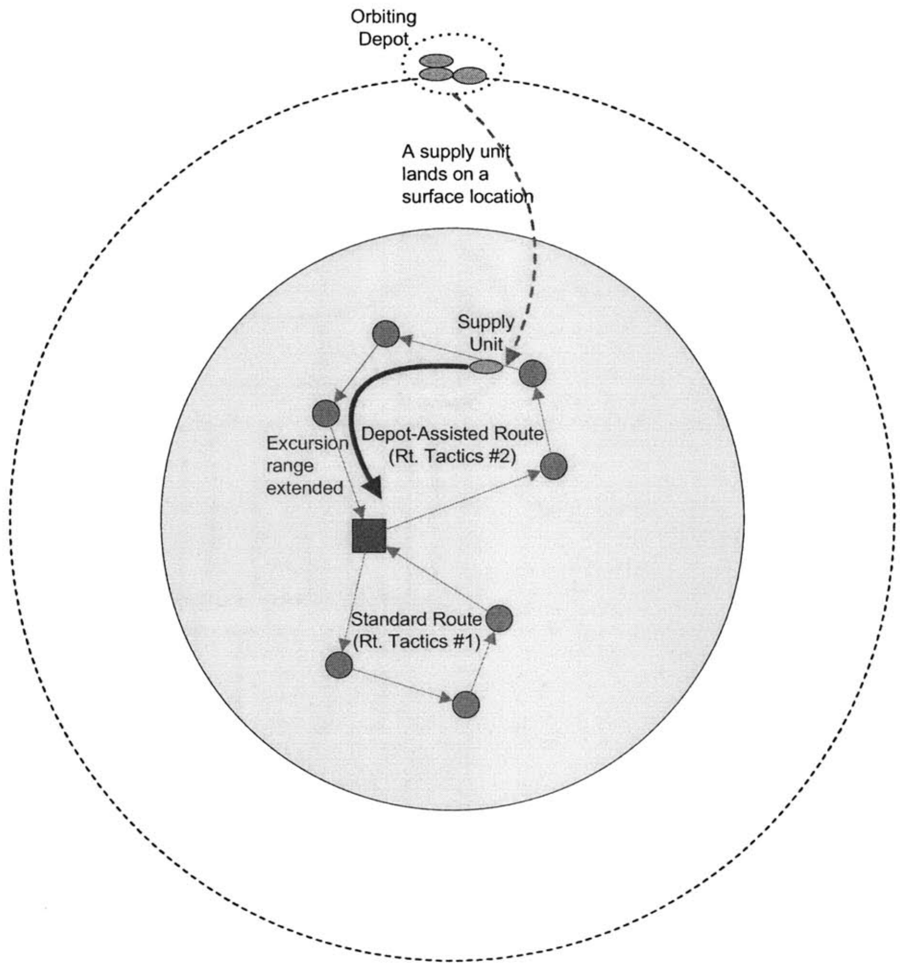


Figure 5-5: Conceptual Diagram for the Orbiting Depot and Its Functionality

orbit and it is possible to command the depot to drop a supply unit filled with fuel and consumables. If the supply unit lands in such a location that the surface exploration vehicle can get the unit before it runs out of fuel or consumables, the vehicle can then travel from the supply unit's landing location for almost the same distance as it has already gone through. In effect, the additional supply from the orbiting depot almost doubles the capability of the surface exploration vehicle. This section deals with the design of the Mars orbiting depot. Discussions on orbit selection, the single supply unit design, and the orbiting depot assembly design are presented in the following subsections.

5.5.2 Orbit Selection

As the first step of the orbiting depot design, we select relevant elements of the orbit in which the depot is deployed. We assume that the orbit is circular and the eccentricity (e) of the orbit is 0. We also assume that the inclination of the orbit is 90 degree so that the supply unit can access the full latitude range.

Then we determine the radius of the orbit. If the orbit altitude is too low, large deceleration by the atmospheric drag increases the amount of propellant required to maintain the orbit. It is known that the atmospheric density on the surface of Mars is less than 1/100 of the sea-level atmospheric density of the Earth, which is equivalent to the density at about 30 [km] altitude above the Earth. Keating et al. predicted, based on Mars Global Surveyor experimental data, that Mars atmospheric density at the altitude of 160 [km] ranges between $1.0 \cdot 10^{-10}[kg/km^3] \sim 0.45 \cdot 10^{-10}[kg/km^3]$ [32]. This level of atmospheric density is small enough for orbit maintenance, and if the radius is larger than this value the atmospheric drag is not problematic. Thus we require that the altitude of the orbit should be larger than 160 [km].

We consider one more thing for selection of the orbit radius. For each orbit, the ground track of the depot sweeps the full latitude and longitude, but the ground track may not exactly pass over the point in which we are interested. To access a specific target point on the surface, the orbiting depot waits until its ground track is closest to the target point and releases the supply at the right time with proper initial cross track velocity. The release time controls the along-track landing position and the initial cross-track velocity controls the cross-track landing position.

The worst case cross-track position error takes place at the equator. It is known that the shift in the crossover longitude value of the ground track at the equator can be expressed as follows [65]:

$$\Delta\phi = T_{orb}/T_{sol} \cdot 360 \text{ [deg]}, \tag{5.1}$$

where T_{orb} is the orbital period, T_{sol} is one Mars day, and $\Delta\phi$ is the longitude shift

Table 5.9: Orbit Period and Worst Case Cross Track Distance

k [−]	T_{orb} [sec]	R_{orb} [km]	h_{orb} [km]	$\Delta\phi$ [deg]	$D_{crs,eq}$ [km]
3	14796	6198	2788	60	893
4	11097	5116	1706	45	670
5	8878	4409	999	36	536
6	7398	3904	494	30	446
7	6341	3523	113	26	383

for consecutive ground tracks. The worst case error is expressed as follows:

$$D_{crs,eq} = 0.5 \cdot R_{Mars} \cdot \Delta\phi \cdot (\pi/180), \tag{5.2}$$

Actually we have two chances per day - once ascending from south to north, once descending from north to south. If the descending ground track crosses exactly at the center of the two consecutive ascending ground track, we can reduce the cross track distance as follows:

$$D_{crs,eq} = 0.25 \cdot R_{Mars} \cdot \Delta\phi \cdot (\pi/180). \tag{5.3}$$

We can achieve this goal by selecting an orbit that has a period of the following form:

$$T_{orbit} = T_{sol}/(2k), \;\; (k = 1, 2, \ldots). \tag{5.4}$$

Table 5.9 shows the family of orbits whose periods satisfy Equation (5.4). We choose the orbit with k value of 6 so that its orbital altitude is higher than 160 [km]. In this case the radius of the orbit is 3904 [km] and the altitude of the orbit is 494 [km]. Other orbit elements such as the longitude of node (Ω) or the argument of pericenter (ω) are irrelevant to the operation of the orbiting depot and are not discussed in this thesis.

Table 5.10: Human Item Consumption Rates for the Reference Exploration Scenario

Item	**Density**	**Unit**	**Rate**	**Unit**
Food	500	$[kg/m^3]$	2	[kg/day/person]
Water	998	$[kg/m^3]$	5	[kg/day/person]
Oxygen			0.63	[kg/day/person]
EVA Water	998	$[kg/m^3]$	5	[kg/EVA day/person]

Table 5.11: Human Item Mass Calculation for the Reference Exploration Scenario

Item	**Rate**	**Unit**	**Person**	**Time** [day]	**Mass** [kg]
Food	2	[kg/day/person]	2	7	28
Water	5	[kg/day/person]	2	7	70
Oxygen	0.63	[kg/day/person]	2	7	9
EVA Water	5	[kg/EVA day/person]	2	1	10
Total					117

5.5.3 Individual Supply Unit Design

An individual supply unit is designed so that the unit can double the exploration range. Table 5.10 and 5.11 present the amount of food, water, and oxygen required for the reference scenario shown in Figure 5-4. Mass and volume of the supply unit core are presented in Table 5.12.

The unit should be designed so that the contents can be protected from the deceleration and heating during entry. We select the entry vehicle of the Mars Exploration Rover (MER) mission as the reference [44]. It is reported that the descent velocity of the MER entry vehicle is 440 [m/sec] at the altitude of 9.1 [km], where a supersonic parachute is deployed. Our objective is to design a "lander" which satisfies the parachute deployment condition. Governing equations for the entry vehicle are given as follows:

$$\dot{\mathbf{v}} = -(\mu/|\mathbf{r}|^3) \cdot \mathbf{r} - (a_D/|\mathbf{v}|)\mathbf{v}, \tag{5.5}$$

$$\dot{\mathbf{r}} = \mathbf{v}, \tag{5.6}$$

Table 5.12: Core Contents for the Single Supply Unit

	Mass [kg]	Density [kg/m^3]	Volume [m^3]
Food	28	500	0.01
Water	80	998	0.14
Oxygen	9		
Fuel	677	1313	0.52
Food Container	3		0.01
Water Container	8		0.15
Oxygen Container	2		0.05
Fuel Container	135		0.65
Total	942		0.87

where

$$a_D = (1/2) \cdot \frac{\rho |\mathbf{v}|^2}{m/(S \cdot c_D)}, \tag{5.7}$$

$$\rho = f(h) = f(|r| - r_M). \tag{5.8}$$

We can find three parameters that affect the descent trajectory in these equations: initial velocity, initial radius, and the $m/(S \cdot c_D)$ value. The $m/(S \cdot c_D)$ value, often expressed as β, is referred to as the "ballistic coefficient." β can be interpreted as the relative magnitude of the inertial force compared to the drag. An entry body with low β is easily decelerated by the atmospheric drag and has smaller terminal velocity [67, 54].

The ballistic coefficient for the MER entry vehicle (β_{MER}) is calculated in Table 5.13. The MER entry vehicle entered Martian atmosphere directly from the interplanetary trajectory and the initial condition of the entry vehicle is more severe than that of our individual supply unit. If we can design the supply unit assembly so that the β value of the assembly is smaller than that of the MER entry vehicle, we can ensure that the parachute deployment condition is met.

We start the design with an assumption that the mass of the whole supply unit assembly is 2,500 [kg]. The lander mass of the MER entry vehicle was 190 [%] of its core content (rover). Considering that the entry condition of our supply unit is less

Table 5.13: Ballistic Coefficient Calculation for the MER Entry Vehicle

Mass Calculation (m)			
Rover	185		
Lander	348		
Backshell / Parachute	209		
Heat Shield	78		
Propellant	50		
Total		870	[kg]
Reference Area (S)		4.15	$[m^2]$
Drag Coefficient (C_D, Assumed)		1.0	[-]
Ballistic Coefficient (β_{MER})		209	$[kg/m^2]$

severe and the core contents are not high-tech instruments like those for the MER, we reduce the ratio from 190 [%] to 85 [%]. The mass allocation results are presented in Table 5.14.

Using the assumption of the total mass of the supply unit and the mass allocation results, we can calculate the propellant mass as 230 [kg]. This amount of propellant should be able to provide sufficient ΔV required to carry out cross-track maneuvers. Assuming that the I_{sp} value of the engine for the unit is 350 [sec], the available ΔV is $(g_0 \cdot I_sp \cdot \ln \frac{m_0}{m_f}) = 330$ [m/sec]. Considering 100 [m/sec] of ΔV is required for de-orbiting and additional 230 [m/sec] of ΔV can be used for deceleration and targeting of the entry body, the amount of fuel is reasonable. This completes the mass allocation for the supply unit assembly as shown in Table 5.14.

We assume that the drag coefficient of the entry body is 1 [-]. To achieve the target ballistic coefficient value, the surface area of the entry body should be 12 $[m^2]$. For the circular cross section of the body, we can conclude that the maximum diameter of the entry surface is 3.9 [m]. We set the outer volume of the assembly as three times the core volume and design the body as a tapered shape whose minimum diameter equals a quarter of the maximum diameter. Figure 5-6 shows the shape of the supply unit.

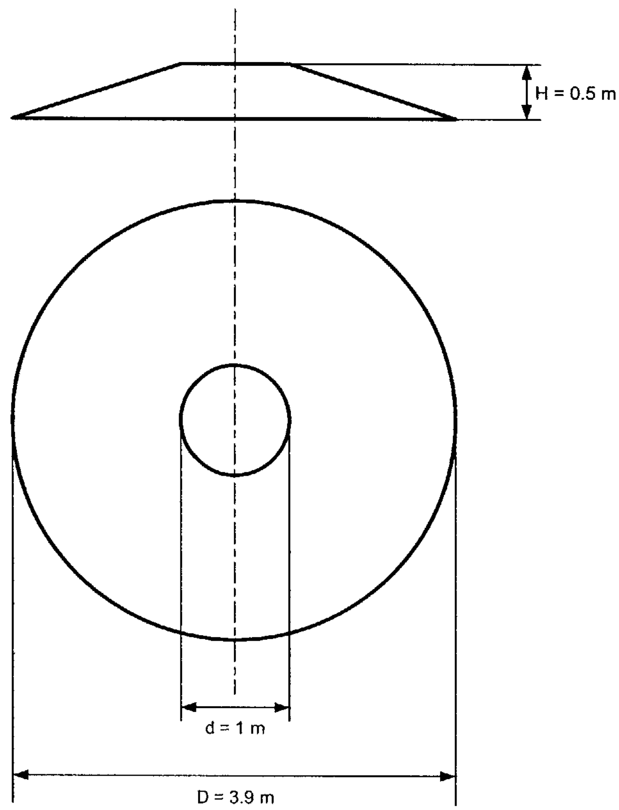


Figure 5-6: Shape of the Supply Unit

Table 5.14: Design of an Entry Body for a Single Supply Unit

Characteristics	Design Value		Note
Mass Allocation			
Core Contents	942	[kg]	
Outer Structure	1041	[kg]	Lander + Heat Shield + Backshell
Parachute	287	[kg]	
Propellant	230	[kg]	Available ΔV = 330 [m/sec]
Total Mass	2500	[kg]	
Dimension			
Max Diameter	3.9	[m]	
Min Diameter	1.0	[m]	
Height	0.5	[m]	
Ballistic Coeff.	209	$[kg/m^2]$	

5.5.4 Characteristics of the Orbiting Depot Strategy

We assume that four individual supply units support one mission using the orbiting depot strategy ($n^{2,2} = 4$). The amount of fuel and consumables contained in an individual supply unit equals the capacity of the surface exploration vehicle. We consider 10 [%] margin that agents may need to travel more due to the landing position error of the individual supply unit, and set the resource limit for the single-route constraint as 1.9 times the surface vehicle fuel capacity ($l_r^{2,2} = 1.9 \cdot 677 = 1286$ [kg]). Finally the cost of the orbiting depot strategy is expressed as $C^2 = (C^1 + 4{\cdot}2.5)$ [MT] = 64 [MT]. Strategies and tactics used in the Mars exploration case study are summarized in Table 5.15.

Table 5.15: Strategies and Corresponding Tactics for the Mars Exploration Case

	Standard Strategy		Orbiting Depot Strategy	
Collective Constraint				
Resource	Exploration time		Exploration time	
Coefficients	$d_0^1 = 0$	$[hr]$	$d_0^2 = 0$	$[hr]$
	$d_d^1 = 0.2$	$[hr/km]$	$d_d^2 = 0.2$	$[hr/km]$
	$d_\tau^1 = 3$	$[hr/hr]$	$d_\tau^2 = 3$	$[hr/hr]$
Limit	$l_c^1 = 2100$	$[hr]$	$l_c^2 = 2100$	$[hr]$
Routing Tactics				
Tactic I	Standard		Standard	
Single-Route Constraint				
Resource	Fuel		Fuel	
Coefficients	$c_0^{1,1} = 0$	$[kg]$	$c_0^{2,1} = 0$	$[kg]$
	$c_d^{1,1} = 1.14$	$[kg/km]$	$c_d^{2,1} = 1.14$	$[kg/km]$
	$c_\tau^{1,1} = 5.68$	$[kg/hr]$	$c_\tau^{2,1} = 5.68$	$[kg/hr]$
Limit	$l_r^{1,1} = 677$	$[kg]$	$l_r^{2,1} = 677$	$[kg]$
Max. Number of Routes	$n^{1,1} = \infty$	$[-]$	$n^{2,1} = \infty$	$[-]$
Tactic II			Depot-Assisted	
Single-Route Constraint				
Resource			Fuel	$[kg]$
Coefficients			$c_0^{2,2} = 0$	$[kg]$
			$c_d^{2,2} = 1.14$	$[kg/km]$
			$c_\tau^{2,2} = 5.68$	$[kg/hr]$
Limit			$l_r^{2,2} = 1286$	$[kg]$
Max. Number of Routes			$n^{2,2} = 4$	$[-]$
Cost	$C^1 = 54$	$[MT]$	$C^2 = 64$	$[MT]$

Table 5.16: GLRPP Instance Generation for Mars Surface Exploration Case Study

Potential Bases and Sites	
Region on Mars Surface	Global
Number of Potential Bases (n_B)	100 [-]
Number of Potential Sites (n_E)	1000 [-]
Aerial Distribution of Locations	Uniform over all regions
Mission Strategies (S)	Presented in Table 5.15
Routing Tactics (T^s)	Presented in Table 5.15
Budget Constraint	
Budget (M)	700 [MT]
Size of Experiments	
No. of Instances (N)	10 [-]

5.6 Numerical Examples

5.6.1 Problem Instances

Ten GLRPP instances for the Mars surface exploration are created. For each instance 100 potential bases and 1000 sites[2] are randomly generated on the surface of Mars. The total campaign budget is set as 700 [MT] delivered to the surface. Table 5.16 summarizes the parameters used in the instance generation. The three-phase solution method is used to obtain the solutions.

Figure 5-7 shows the result of the "Divide" phase for one of the created instances. We can see four big clusters (out of the total 38 clusters).

5.6.2 Results

Table 5.17 summarizes the results of the numerical experiments (average of 10 instances). Each campaign in the ten problem instances is composed of eleven missions - ten out of them use the orbiting depot strategy and only one mission uses the standard strategy.

Figure 5-8 shows the solution for the instance presented in Figure 5-7. Ellipses

[2]154 actual Mars candidate exploration sites shown in Figure 2-1 [40] are included in these sites.

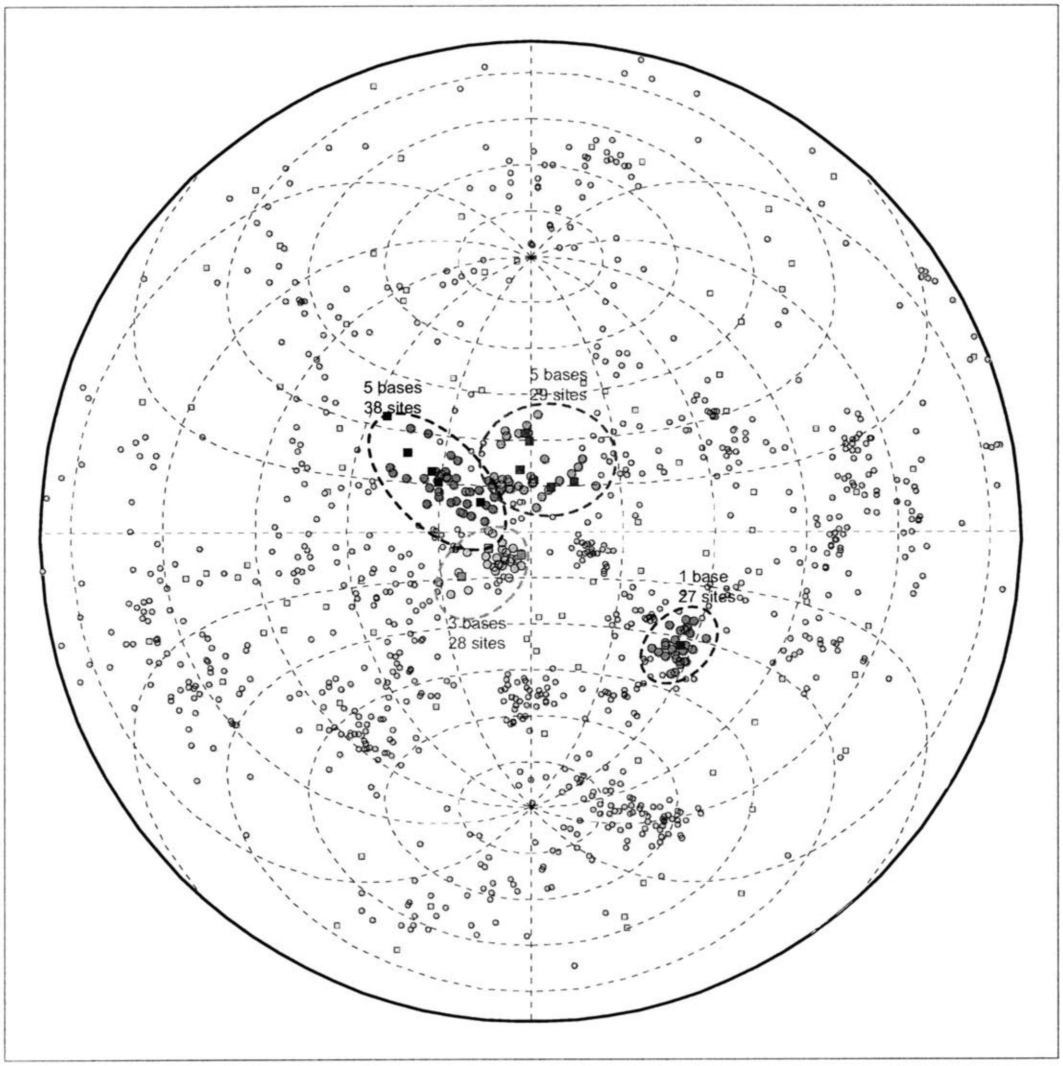


Figure 5-7: Instance 4 - Clustering Result

Table 5.17: Result Summary - Mars Surface Exploration Case Study

	Profits [–]	No. of Missions [–]	Prof. per Mission [–]	Prof. per Cost [–/MT]
Standard Strategy	11.3	1	11.3	0.21
O.Depot Strategy	276.2	10	27.6	0.43
Overall Campaign	287.5	11	26.1	0.41

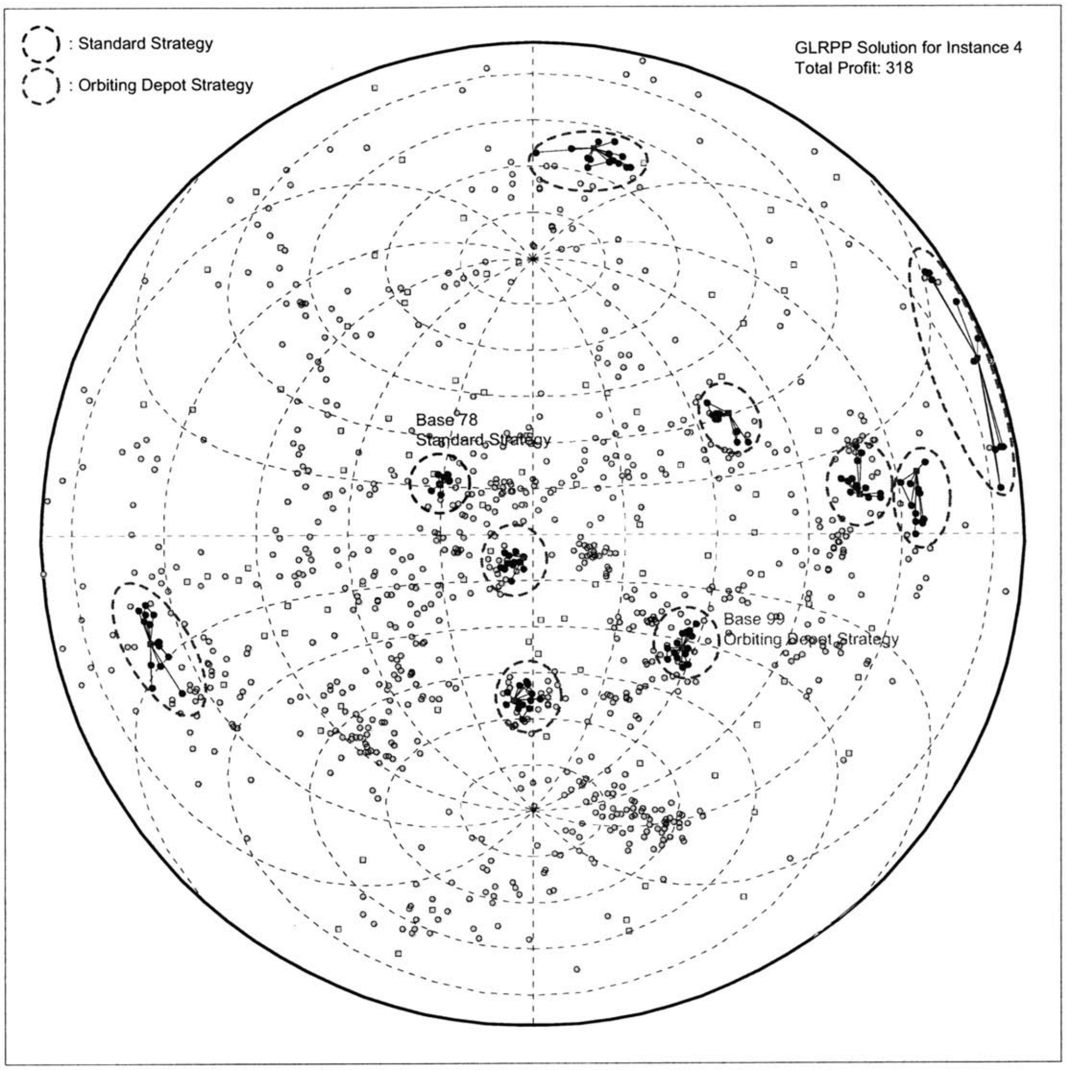


Figure 5-8: Instance 4 - Solution

with dashed blue lines and dashed red lines represent the mission using the standard strategy and orbiting depot strategy, respectively. Figure 5-9 shows routes for the mission using the standard strategy (base 78). The mission is composed of 8 standard routes and the sum of profits obtained during the mission is 16 [-]. Figure 5-10 shows the routes for the mission using the orbiting depot strategy (base 99). Out of 10 routes comprising the mission, 6 routes use the standard routing tactic and 4 routes use the depot-assisted routing tactic. The profit sum for this mission is 36 [-] - more than twice the profit sum for the mission using base 78.

The last two columns of Table 5.17 exhibit the profit sum per mission and profit

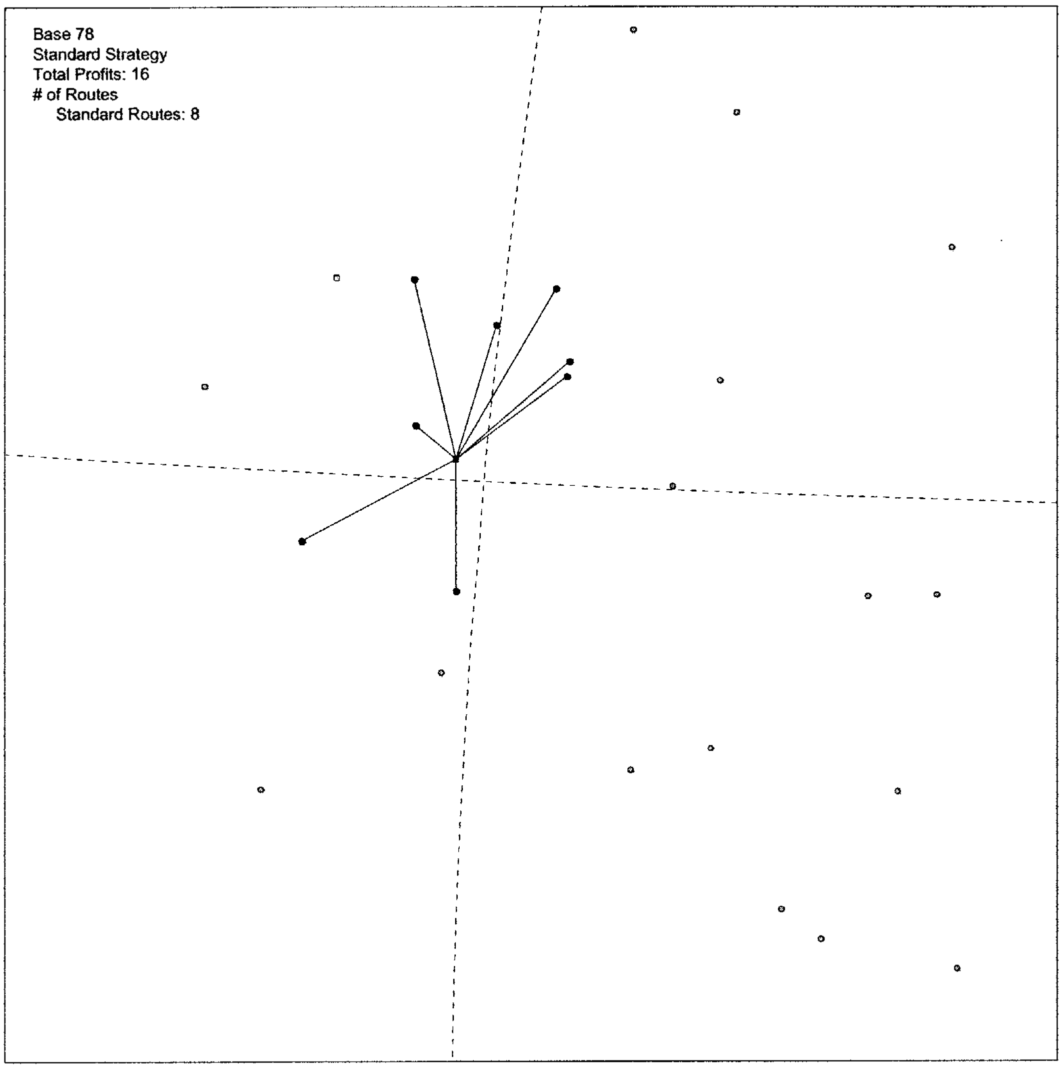


Figure 5-9: Routes for the Mission Using Base 78 (Standard Strategy)

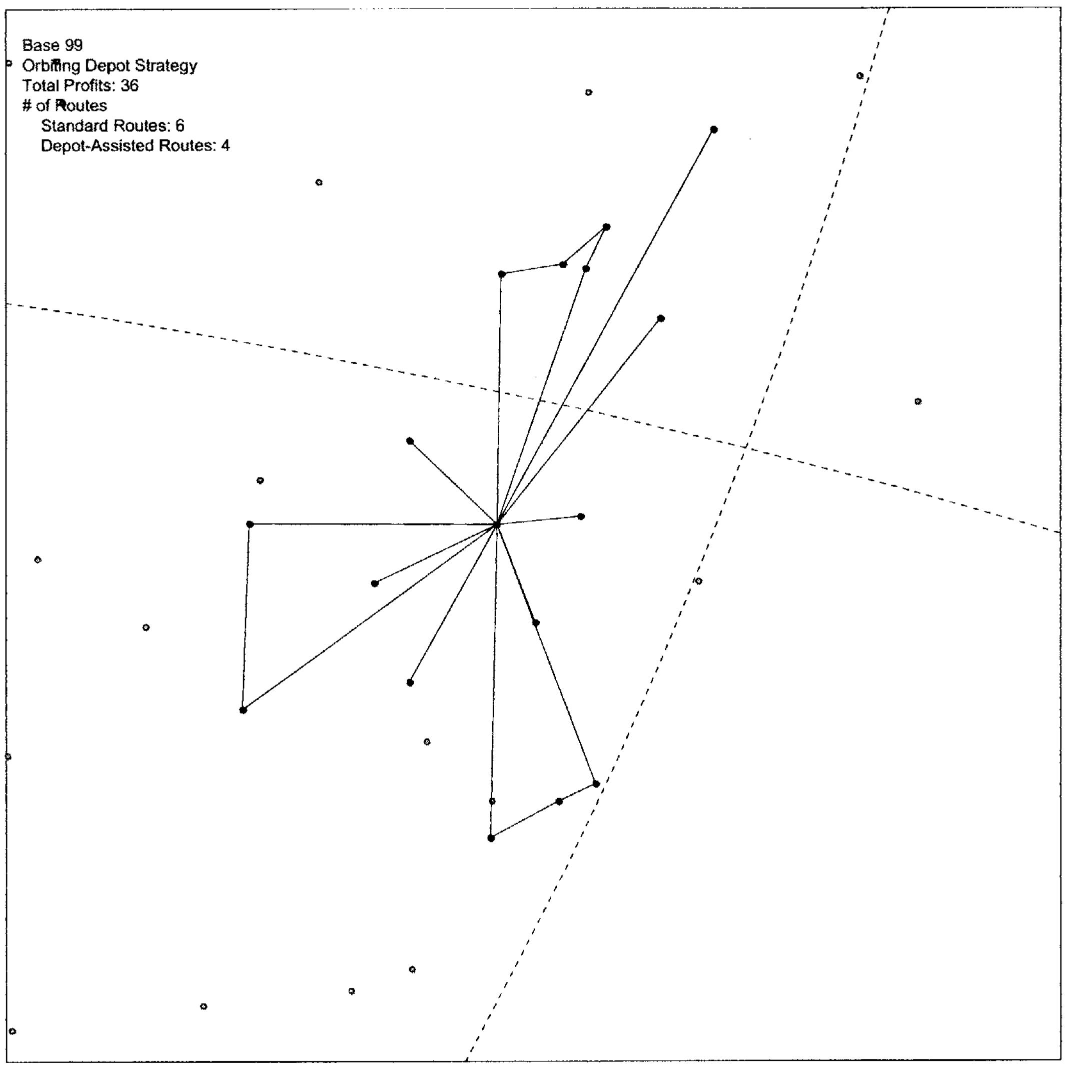


Figure 5-10: Routes for the Mission Using Base 99 (Orbiting Depot Strategy)

sum per cost for the two strategies. The mission cost associated with the orbiting depot strategy is 64 [MT], which is 19 [%] higher than the cost associated with the standard strategy (54 [MT]). But the profit sum per mission for the orbiting depot strategy is 27.6 [-/-], which is about 144 [%] higher than the profit sum per mission for the standard strategy (11.3 [-/-]). This discrepancy leads to the large gap in the profit sum per cost, which is 0.43 [-/MT] for the orbiting depot strategy and 0.21 [-/MT] for the standard strategy. The optimizer chooses the orbiting depot strategy primarily, and if there is some budget left for the standard strategy but not for the orbiting depot strategy, it chooses the standard strategy.

So far we have analyzed the solutions for the Mars surface exploration instances. Under assumptions made for the case study, the orbiting depot strategy shows a larger profit per cost than the standard strategy, and is selected with higher frequency.

This result is based on an assumption that the required technology (orbiting depot technology in this case) already exists because the cost imposed by the selection of the orbiting depot strategy is the variable cost. A methodology to assess the value of the technology itself is proposed in the next subsection.

5.6.3 Value of a Technology

We propose a finite difference technique to calculate the value of a technology using the GLRPP framework. In the problem instance presented in the previous subsection, it has been implicitly assumed that we already have the orbiting depot technology - to manufacture the supply units and orbiting depot assembly, to send them to the Martian orbit, and to operate the depot for supporting the Mars surface exploration mission.

Suppose that we do not have the orbiting depot technology. In this case, only the standard strategy is available for the missions. The amount of profit from the global Mars surface exploration campaign in this case can be obtained by solving the GLRPP using only the standard strategy. Table 5.18 summarizes the result of the GLRPP solutions using only the standard strategy for the instances generated in Subsection 5.6.1. The average profit sum for this case is 131.7 [-], which is 155.8 [-] less

Table 5.18: Result Summary: Mars Surface Exploration Without Orbiting Depot Technology

	Profits [–]	No. of Missions [–]	Prof. per Mission [–]	Prof. per Cost [–/MT]
Standard Strategy	131.7	12	11.0	0.20
With Orbiting Depot	287.5			
Value of Orbiting Depot	155.8			

than the profit sum for the result presented in Table 5.17 (287.5 [-]). This difference (155.8 [-]) is the marginal profit that can be obtained by having the orbiting depot technology. It can be interpreted as the **value of the orbiting depot technology** expressed as exploration profits.

Similarly the value of the In-Situ Resource Utilization(ISRU) technology is also calculated. The standard strategy and the orbit-depot strategy assume that the propellant used for the Earth Return Vehicle (ERV) and the surface exploration vehicle is produced on Mars surface using the ISRU technology. Without the ISRU technology, we have to bring more propellant from the Earth.

The amount of propellant used for ERV in the case study is 82 [MT] [68]. The fuel capacity 677 [kg] is based on a 7-day exploration scenario. Total time allowed for the surface exploration is 2100 [hr] and the maximum amount of fuel for the surface transportation is $0.677 \cdot \frac{2100}{24 \cdot 7} = 8$ [MT]. Assuming that the ISRU plant mass is relatively small, the increase in mass that has to be delivered from the Earth is 90 [MT].

This situation can be interpreted as an increase in cost. So the mission cost of the standard strategy becomes $(54 + 90) = 144$ [MT] and that for the orbiting depot strategy becomes $(64 + 90) = 154$ [MT]. Table 5.19 summarizes the GLRPP solution results. Average profit sum is 132 [-]. This value is less than the average profit sum with the ISRU technology by 155.5 [-] So the value of the ISRU technology is 155.5 [-].

The value of the ISRU technology when it is used without the orbiting depot

Table 5.19: Result Summary - Mars Surface Exploration Without ISRU Technology, With Orbiting Depot Technology

	Profits [-]	No. of Missions [-]	Prof. per Mission [-]	Prof. per Cost [-/MT]
Standard Strategy	0	0	N/A	N/A
O.Depot Strategy	132.0	4	33.0	0.21
Overall Campaign	132.0	4	33.0	0.21

Table 5.20: Result Summary - Mars Surface Exploration Without ISRU Technology, without Orbiting Depot Technology

	Profits [-]	No. of Missions [-]	Prof. per Mission [-]	Prof. per Cost [-/MT]
Standard Strategy	65.2	4	16.3	0.11

technology is also calculated. Table 5.20 summarizes the GLRPP solution results without ISRU technology and without the orbiting depot technology. The average profit sum in this case is 65.2 [-] and the difference from the results for (with ISRU / without orbiting depot) is (131.7 - 65.2) = 66.5 [-]. So the value of the ISRU technology without the orbiting depot technology is 66.5 [-]. Figure 5-11 shows the summary for the value of the orbiting depot technology and the ISRU technology expressed as the additional profit generated by the exploration campaign.

5.7 Mars Surface Exploration Summary

A global Mars surface exploration campaign optimization problem is formulated as the Generalized Location Routing Problem with Profits (GLRPP) and is solved using the three-phase solution method. Two mission strategies - the standard strategy and the orbiting depot strategy - are considered.

Problem instances with 100 potential bases and 1000 exploration sites are generated and solved. A methodology to assess the value of a technology through finite

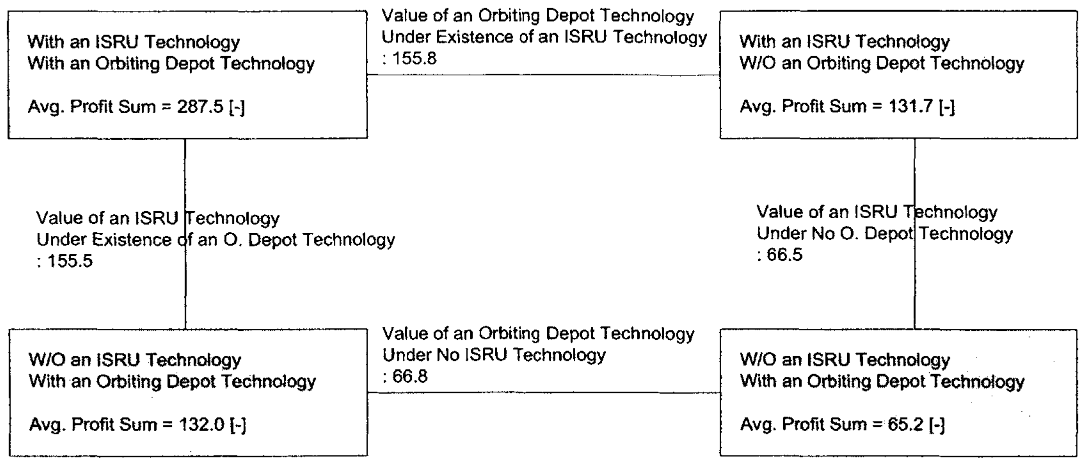


Figure 5-11: Value of Technology - Orbiting Depot and ISRU

differencing is proposed. Solutions with different technology options - with and without the orbiting depot technology and the ISRU technology - are obtained and the value of the orbiting depot technology and the ISRU technology are calculated by comparing the results.

Chapter 6

Terrestrial Application: College Football Recruiting

6.1 Introduction

Athlete recruiting was first studied by Butt and Cavalier in 1994 as an application of the Vehicle Routing Problem with Profits (VRPP), in which only one base was used for recruiting.[1] They proposed a heuristic procedure called MAXIMP which sequentially builds routes using a weighting scheme based on profits and travel time. An exact method for the same class of problem was proposed by Butt and Ryan in 1999 [17]. They used a column generation procedure to solve an LP relaxation problem of the original IP formulation for the VRPP. A near-optimal solution with a relatively small optimality gap was found by the method described in the paper.

This chapter deals with recruiting of college football athletes by agents representing a National Football League (NFL) team. The agents carry out recruiting by visiting schools to gather information. It is assumed that the value of information that can be gathered from each school is pre-determined. The objective of the problem is to maximize the total sum of information gathered from all visited schools.

This problem can be viewed as a terrestrial application of the Generalized Location

[1]This problem is referred to as the Multiple Tour Maximum Collection Problem (MTMCP) in their original publication [16].

Routing Problem with Profits (GLRPP). In the following sections, we formulate this problem as the GLRPP, create a GLRPP instance using the data from 2006-season NCAA football league result, and solve the problem using the three-phase method introduced in Chapter 4.

6.2 Problem Description

We assume that multiple recruiting agents are traveling together as a group. Once the group of recruiting agents arrive at an airport, each agent separately visits schools by car located in the area around the airport. After the agents complete visits to the schools within the area, they reunite at the airport and fly to the next airport. To simplify the problem it is assumed that flights between airports in the problem instance are available when they are required.

Each school is assigned a real scalar value representing the potential of the athletes of the school's football team. The objective of the problem is to maximize the sum of the potential values for visited schools. The whole recruiting campaign should be finished within a pre-determined period.

Using the terminology of the GLRPP, the whole activity of visiting schools in this recruiting application is the "campaign" and the collection of visiting routes associated with an airport is the "mission." The airports and schools are potential bases (**B**) and sites (**E**), respectively. There is a "single-route constraint" that each agent who leaves an airport should return to it within the time duration specified by the routing tactics for the problem. There is no collective constraint. Finally, the sum of man-days for all missions should be less than a pre-determined limit value, which is a "budget constraint" for the campaign in the GLRPP terminology.

6.3 School and Airport Selection and Campaign Characteristics

Schools included in the National Collegiate Athletic Association (NCAA) football division I-A and airports that are close to the schools are selected as sites and candidate bases, respectively. Some schools share airports and the number of the airports is smaller than the number of the schools. We assume that the aggregated potential of the school's football team has high correlation with the final 2006 NCAA football league ranking and assigned profit value of 1 [-] $\sim$ 5 [-] based on the ranking [30].

A total of 118 schools and 55 airports are selected for the case study. Table 6.1 exhibits the schools selected for the case study and profit values assigned to the schools. Table 6.2 is the list of the airports used for the case study.

There is only one mission strategy in this case study: two-day recruiting. The type of "cost" in this recruiting problem is "man-days spent on recruiting" and the cost of the mission is 4 [man-days] (2 agents travel for 2 days). The campaign budget is 80 [man-days]. There is no collective constraint for the strategy. There is only one routing tactic in the mission strategy: driving. We assume that two agents are traveling together and once they arrive at an airport each of them carries out recruiting separately for two business days and returns to the airport. The constraining resource for the single route constraint is time. Considering traffic jams in urban areas, we assume the traveling speed is 30 [km/hr]. The resource consumption limit is 16 [hr] = 960 [min]. The characteristics of the whole campaign for this case study are summarized in Table 6.3.

Finally, we assign stay-time at school i (t_i). Usually the recruiting agents spend more time at a school with higher potential. So we assign the stay time for each school using the following equation:

$$t_i = t_f + t_v \cdot v_i, \tag{6.1}$$

where t_i is the stay time at school i (minutes), t_f is the fixed stay time (= 120 [min]),

v_i is the profit value for school i, and t_v is the variable stay time (= 24 [min/-]).

Table 6.1: Schools Selected for Recruiting

Profit	**Schools**
5	Ohio State, Florida, Michigan, LSU, Louisville, Wisconsin (6 schools)
4	Oklahoma, USC, Boise State, Auburn, Notre Dame, Arkansas (6 schools)
3	West Virginia, Virginia Tech, Wake Forest, Rutgers, Tennessee, Texas (6 schools)
2	Brigham Young, California, Texas A&M, Nebraska, Boston College, Oregon State, TCU (7 schools)
1	New Mexico, New Mexico State, Baylor, Southern Methodist, North Texas, Georgia, Connecticut, Vanderbilt, Middle Tennessee, Idaho, Tulane, Army, Buffalo, Clemson, South Carolina, Kent State, Akron, Ohio, Miami (OH), Illinois, Northern Illinois, Marshall, Cincinnati, Colorado, Air Force, Colorado State, Wyoming, Iowa State, Michigan State, Central Michigan, Western Michigan, Eastern Michigan, UTEP, Oregon State, Oregon, Fresno State, Hawaii, UAB, Auburn, Houston, Rice, Ball State, Purdue, Indiana, UNLV, UCLA, Texas Tech, Kentucky, Arkansas State, Kansas State, Kansas, Missouri, UCF, Mississippi, Mississippi State, Memphis, Miami (FL), Florida Atlantic, Florida International, Minnesota, Louisiana-Lafayette, Louisiana-Monroe, Louisiana Tech, Southern Miss, Oklahoma State, Northwestern, Penn State, Temple, Arizona State, Arizona, Pittsburgh, Washington State, North Carolina State, North Carolina, Duke, East Carolina, Virginia, Nevada, South Florida, San Diego State, Washington, Stanford, San Jose State, Utah, Utah State, Syracuse, Alabama, Troy, Florida State, Bowling Green, Toledo, Tulsa (93 schools)

Table 6.2: Airports Potentially Used for Recruiting

HIK, MIA, RSW, MCO, IAH, MSY, AUS, TLH, BTR, ELP, SAN, AFW, TCL, PHX, ATL, LBB, LAX, CAE, HSV, ABQ, MEM, OKC, RDU, GSO, LAS, BNA, TUL, XNA, FAT, SFO, CRW, ADW, CVG, MCI, RNO, IND, CMH, LAF, PIT, EWR, SLC, LNK, DSM, ORD, BDL, DTW, BOS, DEN, SYR, MSN, BOI, EUG, MSP, PUW, SEA

Table 6.3: Campaign Characteristics of the Recruiting Case Study

Mission Strategy	Two-Agent Two-Day Recruiting	
Collective Constraint	N/A	
Routing Tactic	Driving	
Single-route constaint		
Constraining resource	time	[min]
Consumption coefficients	$c_0 = 0$	[min]
	$c_d = 2$	[min/km]
	$c_\tau = 1$	[min/min]
Consumption limit	960	[min]
Maximum route number	2	[–]
Cost	$C = 4$	[man-day]
Campaign Budget	80	[man-day]

6.4 Numerical Result

We use the three-phase method for the case study. The *Divide Phase* generates clusters for the problem instance. Reachability of a school from an airport is determined by checking the feasibility of the round-trip route between the airport and the school (with respect to the constraint on time consumption). The largest cluster has 11 bases (airports) and 33 sites (schools), which contains more than a quarter of the total schools. Figure 6-1 shows three big clusters from the results of the *Divide Phase.*

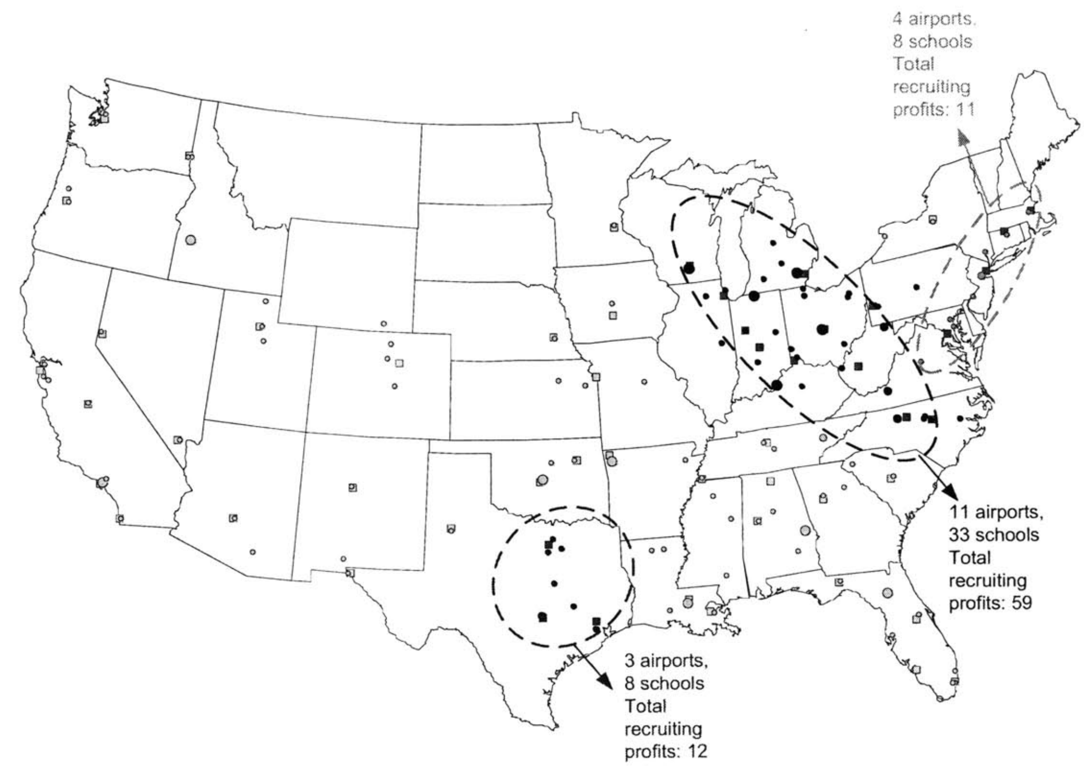


Figure 6-1: College Football Recruiting - Clustering Result

The solution of the GLRPP instance for the case study is presented in Figure 6-2. All schools with profit value of 5 [-] (2006 NCAA ranking 1-6) are included in the solution. Up to three schools are included in one route. The LP-relaxation profit sum (K_{LP}) for the problem instance is 116.50 [-], which is an upper bound of the optimal profit sum. The near-optimal feasible profit sum (K_{IP}) is 116 [-] and the optimality gap for the solution is 0.43 [%]. The total profit available from all the candidate schools is 179 [-], and we obtain about 68 [%] of the total profit from the GLRPP solution.

Figure 6-3 shows schools visited by the solution, routes to visit the schools, airports used as bases for the routes, profits obtained from the routes, and cumulative man-days. Note that in the GLRPP framework we do not consider the optimization of the mission sequence, and the order to carry out the missions can be arbitrary. The "mission" for this application is the set of recruiting activities based on a specific airport, and the order may be also important. To optimize the mission order, we can use the result of the Traveling Salesman Problem (TSP) for the graph composed of airports included in the solution. This task is not within the scope of this thesis.

Figure 6-2: College Football Recruiting - the GLRPP Solution

6.5 College Football Recruiting Summary

A problem to optimize the recruiting of college football athletes is dealt with in this chapter as a terrestrial application of the GLRPP. We generated a GLRPP instance using the schools in the NCAA football division I-A and airports which can reach the schools. We used the three-phase solution method and successfully obtained a solution with a very low optimality gap (0.43 [%]).

Cumulative Man-Day	Airport	School 1	School 2	School 3	Profits
	MCO	Florida			5
4	MCO	South Florida	UCF		2
	AUS	Texas			3
8	AUS	Texas A&M			2
	AFW	Baylor			1
12	AFW	North Texas	Southern Methodist	TCU	4
	MSY	LA-Lafayette			1
16	MSY	LSU	Tulane		6
	LAX	San Diego State			1
20	LAX	USC	UCLA		5
	ATL	Auburn			4
24	ATL	Georgia Tech	Georgia		2
	TUL	Oklahoma			4
28	TUL	Arkansas			4
	RDU	North Carolina	Wake Forest		4
32	RDU	NC State	Duke		2
	CRW	Virginia Tech			3
36	CRW	Marshall	Ohio		2
	CVG	Ohio State			5
40	CVG	Kentucky	Cincinnati		2
	IND	Louisville			5
44	IND	Ball State	Miami (OH)		2
	LAF	Notre Dame			4
48	LAF	Indiana	Purdue		2
	PIT	Akron	Kent State		2
52	PIT	Pittsburgh	West Virginia		4
	DTW	Michigan	Michigan State		6
56	DTW	Bowling Green	Toledo	Eastern Michigan	3
	MSN	Northern Illinois			1
60	MSN	Wisconsin			5
	SFO	Stanford			1
64	SFO	San Jose State	California		3
	EWR	Temple			1
68	EWR	Army	Rutgers		4
	SLC	Utah State			1
72	SLC	Utah	Brigham Young		3
	DEN	Air Force			1
76	DEN	Colorado	Colorado State		2
80	BOI	Boise State			4
				Profit Total	**116**

Figure 6-3: Used Airports and Visited Schools

Chapter 7

Conclusions

7.1 Thesis Summary

Given locations and profits for potential bases and sites, a problem to maximize the sum of profit obtained over a campaign by making decisions on selection of bases to use, selection of strategies for the missions using the bases, selection of routes to visit sites, and selection of routing tactics the routes use, with constraints imposed on each route, each mission, and the overall campaign is formulated and solved in this thesis.

The problem, referred to as the *Generalized Location Routing Problem with Profits* (GLRPP), is originally based on an optimal design problem for a global planetary surface exploration campaign. Chapter 1 explains global planetary surface exploration as the background of this thesis. Chapter 2 describes the GRLPP and provides the Integer Program (IP) formulation.

Two solution methods for solving the problem are proposed in Chapters 3 and 4. The single-phase method presented in Chapter 3 directly uses the original formulation presented in Chapter 2 to create a linear program relaxation (LP relaxation) to obtain a near-optimal solution of the GLRPP and an upper bound / optimality gap of the solution. The three-phase method presented in Chapter 4 divides the problem into multiple sub-problems referred to as the *Multi-Depot Vehicle Routing Problem with Profits* (MDVRPP), solves each subproblem using a procedure similar to the single-phase method presented in Chapter 3, and synthesizes the MDVRPP solutions to obtain a near-optimal solution and an upper bound / optimality gap for the whole GLRPP. A performance analysis for the two solution methods is also provided.

A global Mars surface exploration campaign optimization problem is presented in Chapter 5 as a space application of the GLRPP. A realistic design of a planetary surface exploration vehicle is used and an orbiting depot is considered as a technology option. Problem instances with 100 potential bases and 1000 exploration sites are generated and successfully solved using the three-phase method. A methodology to calculate the value of a technology is proposed and used for valuation of an orbiting depot technology and an In-Situ Resource Utilization technology used in the case study.

In Chapter 6 a college football recruiting problem is introduced as a terrestrial application of the GLRPP. A problem instance is created out of the NCAA football division I-A schools and airports from which agents can reach the schools. The problem is successfully solved using the three-phase solution method within a very small optimality gap.

7.2 Contributions

As has been pointed out previously, this research was originally inspired by the problem to optimize a "Global Planetary Surface Exploration Campaign," and the development of a framework for the campaign optimization is the first contribution of this thesis.

A case study for Mars surface exploration campaign optimization using the framework is provided. Using realistic information for the space and surface elements of the exploration campaign, problems for many potential landing locations and exploration sites have been successfully solved.

An "orbiting depot" concept has been proposed as one of the mission strategies for Mars surface exploration for the case study. A preliminary design of the orbiting depot has been presented. Under the problem settings for this thesis, the orbiting depot technology improves the profit-obtaining capability of a mission significantly with relatively small cost increase and is found to be a very attractive option.

This thesis also contributes to the development of a problem class referred to as the Generalized Location Routing Problem with Profits (GLRPP). The GLRPP can be seen as an extension of the "Location Routing Problem" in two different directions. A mathematical formulation and two solution methods for the GLRPP were developed and performance analysis for the solution methods was also carried out.

A systematic expression of resource consumption using three resource consumption classes - the per-route, the on-arc, and the on-site consumption classes - is another contribution of this thesis. We use this method to express the consumption of many resource types in the context of the routing problem.

7.3 Suggested Future Work

7.3.1 Improved Solution Procedure

The performance analysis result for the two solution methods proposed in the thesis indicates that each solution method has a weakness for solving a problem instance with certain characteristics. The single phase method has been found to be less attractive in terms of providing an upper bound when the budget constraint is tight. The three-phase method is expensive for instances whose number of bases in one cluster is large.

There may be some room for improvement for both solution methods. The single phase method in this thesis is using a simple LP relaxation of the original GLRPP to get an upper bound. For the future work, the Lagrangian relaxation technique could be used to provide a better upper bound by putting the constraints (2.16) and (2.17) into the objective [36].

A methodology to accelerate the conquer phase of the three-phase method is suggested. A column represents a route associated with a base, a mission strategy for the base, and a routing tactic. A "good" column - or a "good" route - for one cluster strategy may be a "good" column for another cluster strategy with high probability.

We propose a future study for a smart way to reuse generated columns in previously solved MDVRPPs so that we can accelerate the column generation procedure of the *Conquer Phase*, which takes a lot of calculation time.

7.3.2 Balanced Optimization Scheme - Consideration of Profit Vectors

One major limitation of the GLRPP framework is the way it deals with "profits." Under the current framework, profit value to visit a site is set as a *scalar* value. There are multiple stakeholder groups interested in visiting sites and obtaining profits. And valuations of a site from different stakeholder groups will likely be different. Assignment of a single profit value to each site implies that the values associated with a site from different stakeholder groups are somehow aggregated into one single number. Once they are merged into a number, the optimization procedure for this thesis cannot determine the original evaluation of a site from each group. A profit value 1 obtained from site A is identical to the profit value 1 obtained from site B in terms of the objective function. How the interests of different groups are aggregated to yield the profit value 1 is irrelevant in the current framework. In this case, even when the overall profit sum obtained by visiting sites is maximized, if we track back to the original evaluation of sites by different groups, there may be some groups which could not obtain much value in terms of their own interests and are not happy about the optimization result.

To prevent this situation, an optimization scheme referred to as a "balanced optimization" is proposed as future work. In the balanced optimization, evaluation of a site i by m different groups is expressed as a vector $\mathbf{v}_i$ as follows:

$$\mathbf{v}_i = [v_1^i, \ldots, v_m^i]'. \tag{7.1}$$

When we visit sites, we can consider m different types of profit sums, each type represents a profit sum from the perspective of a specific stakeholder group. Assume that each type of profit is normalized and one type of profit for all sites add up to

1. Then the sum of type μ profits for visited sites can be interpreted as the "degree of satisfaction of group μ obtained by the visits." We propose to optimize the degree of satisfaction of the "least" happy group. So we make decisions to maximize the minimum profit sum out of all types of profit sums. A mathematical formulation for the proposed optimization scheme is expressed as follows.

(GLRPVP) Generalized Location Routing Problem with Vector Profits

$$\min_{\mathbf{x},\mathbf{y},\mathrm{s}}(-\mathrm{s}) \tag{7.2}$$

subject to

$$-\mathfrak{R}\mathbf{x} + \mathbf{1}_m \mathrm{s} \leq \mathbf{0}_m, \tag{7.3}$$

$$\mathbf{A}\mathbf{x} \leq \mathbf{1}_{n_1}, \tag{7.4}$$

$$\mathbf{E}_1\mathbf{x} - \mathbf{N}\mathbf{y} \leq \mathbf{0}_{n_2}, \tag{7.5}$$

$$\mathbf{H}\mathbf{x} - \mathbf{L}\mathbf{y} \leq \mathbf{0}_{n_3}, \tag{7.6}$$

$$\mathbf{E}_2\mathbf{y} \leq \mathbf{1}_{n_4}, \tag{7.7}$$

$$\mathbf{c}'\mathbf{y} \leq \mathrm{M}, \tag{7.8}$$

$$\mathbf{0} \leq \mathbf{x}, \mathbf{y} \leq 1, \quad \mathbf{x} \text{ and } \mathbf{y} \text{ are integers.} \tag{7.9}$$

This GLRPVP formulation is very similar to the GLRPP formulation presented as equations (2.11)~(2.17). A new constraint expressed as equation (7.3) is added and the objective function is changed to maximization of s, which is also a new variable appearing in equation (7.3). $\mathfrak{R}$ is defined as a matrix whose μ^{th} row is $\mathbf{r}'_\mu$ where the j^{th} element of $\mathbf{r}'_\mu$ is the profit sum related with a route expressed as the j^{th} column of the constraint matrices. $\mathfrak{R}$ is expressed as follows:

$$\mathfrak{R} = \begin{bmatrix} \mathrm{r}'_1 \\ \vdots \\ \mathrm{r}'_m \end{bmatrix}. \tag{7.10}$$

The μ^{th} element of vector $\mathfrak{R}\mathbf{x}$ is the total profit sum for the campaign expressed by route selection $\mathbf{x}$ based on the values assigned by a stakeholder group μ. If the profit values are normalized, the profit sum can be interpreted as the degree of satisfaction for group μ. s is lower than each element in vector $\mathfrak{R}$; By maximizing s in the objective function (7.3), we maximize the degree of satisfaction of the least satisfied group.

7.3.3 Potential Applications of the GLRPP

The GLRPP is a general framework that can handle the most complex routing problems. It supports simultaneous decisions on selection of bases, associated exploration strategies, and routes. Global recruiting of an international company, drilling of exploration wells in the oil/gas industry, and military operations over multiple locations are all potential applications of the GLRPP. In the case of the exploration well drilling, for example, seaports where drilling ships depart and return are potential bases ($\mathbf{B}$), candidate drilling locations are sites ($\mathbf{E}$), and estimated amount of oil/gas for each site (i) is the profit value (v_i). Studies for these applications using the GLRPP framework presented in this thesis are also suggested as future work.

Bibliography

[1] Introduction to space system architecture, February 2006. Lecture Notes for Space Systems Engineering(16.89), MIT Dept. Aero/Astro.

[2] Nasa unveils global exploration strategy and lunar architecture, December 2006. NASA NEWS, RELEASE: 06-361.

[3] The global exploration strategy: The framework for coordination. Technical report, NASA and 13 space agencies from around the world, 2007.

[4] Y. Agarwal, K. Mathur, and H. Salkin. A set-partitioning-based exact algorithm for the vehicle routing problem. *Networks*, 19:731–749, 1989.

[5] F. Anderson, A. Haldemann, N. Bridges, M. Golombek, and T. Parker. Analysis of mola data for the mars exploration rover landing sites. *Journal of Geophysical Research*, 108:ROV25 1–27, 2003.

[6] P. Badeau, M. Gendreau, F. Guertin, J.-Y. Potvin, and E. Taillard. A parellel tabu search heuristics for the vehicle routing problem with time windows. *Transportation Research C*, 5:109–122, 1997.

[7] E. Balas. The prize collecting traveling salesman problem. *Networks*, 19:621–636, 1989.

[8] M. Balinski and R. Quandt. On an integer program for a delivery problem. *Operations Research*, 12:300–304, 1964.

[9] M. Bellmore and J. Malone. Pathology of traveling salesman problem subtour-elimination algorithms. *Operations Research*, 19:278–307, 1971.

[10] R. Berger, C. Coullard, and S. Daskin. Location-routing problems with distance constraints. *Transportation Science*, 41:29–43, 2007.

[11] D. Bertsimas and J. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, 1997.

[12] B. Bollobás. *Modern Graph Theory*. Springer, 2001.

[13] J. Bramel and D. Simchi-Levi. A location based heuristics for general routing problems. *Operations Research*, 43:649–660, 1995.

[14] J. Bramel and D. Simchi-Levi. Probablistic analyses and practical algorithms for the vehicle routing problem with time windows. *Operations Research*, 44:501–509, 1996.

[15] J. Bramel and D. Simchi-Levi. On the effectivenss of set covering formulation for the vehicle routing problem with time windows. *Operations Research*, 45:295–301, 1997.

[16] S. E. Butt and T. M. Cavalier. A heuristic for the multiple tour maximum collection problem. *Computer and Operations Research*, 21:101–111, 1994.

[17] S. E. Butt and D. M. Ryan. An optimal solution procedure for the multiple tour maximum collection problem using column generation. *Computer and Operations Research*, 26:427–441, 1999.

[18] I. Chao, B. Golden, and E. Wasil. The team orienteering problem. *European Journal of Operations Research*, 88:475–489, 1996.

[19] N. Christofides and S. Eilon. An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, 20:309–318, 1969.

[20] N. Christofides, A. Mingozzi, and P. Toth. Exact algorithms for the vehicle routing problem based on the spanning tree and shortest path relaxations. *Mathematical Programming*, 20:255–282, 1981.

[21] G. Clarke and J. Wright. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12:568–581, 1964.

[22] D. Cooke and D. Beaty. Dawn of a new space age: the global exploration strategy, July 2006. Briefing Material at 2006 July NAC Planery Science Subcommittees Meeting.

[23] J-F. Cordeau, M. Gendreau, and G. Laporte. A tabu search heuristic for periodic and multi-depot vehicle routing problems. *Networks*, 30:105–119, 1998.

[24] G. Dantzig, R. Fulkerson, and S. Johnson. Solution of a large-scale traveling salesman problem. *Journal of the Operations Research Society of America*, 2:393–410, 1954.

[25] G. Dantzig and R. Ramser. The truck dispatching problem. *Management Science*, 6:80, 1959.

[26] M. Dell'Amico, F. Maffioli, and P. Värbrand. On prize-collecting tours and the asymmetric traveling salesman problem. *International Transactions in Operational Research*, 31:515–530, 1988.

[27] G. Desaulniers, J. Lavigne, and F. Soumis. Multi-depot vehicle scheduling with time windows and waiting costs. *European Journal of Operational Research*, 111:479–494, 1998.

[28] M. Desrochers, J. Desrosiers, and M. Solomon. A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research*, 40:342–354, 1992.

[29] J. Desrosiers, F. Soumis, and M. Desrochers. Routing with time windows by column generation. *Networks*, 14:545–565, 1984.

[30] ESPN. College football. http://sports-ak.espn.go.com/ncf/index.

[31] B. Bairstow et al. Extensible planetary surface mobility systems. Final report, MIT 16.89/ESD 352 Spring 2006 Class, 2006.

[32] G. M. Keating et al. The structure of the upper atmosphere of mars: In situ accelerometer measurements from mars global surveyor. *Science*, 279:1672–1676, 1998.

[33] D. Feillet, P. Dejax, and M. Gendreau. Traveling saleman problem with profits. *Transportation Science*, 39:188–205, 2005.

[34] M. Gendreau, A. Hertz, and G. Laporte. New insertion and postoptimization procedures for the traveling salesman problem. *Operations Research*, 40:1086–1094, 1992.

[35] M. Gendreau, A. Hertz, and G. Laporte. A tabu search heuristic for the vehicle routing problem. *Management Science*, 40:1276–1290, 1994.

[36] A. Geoffrion. Lagrangian relaxation and its uses in integer programming. *Mathematical Programming Study*, 2:82–114, 1974.

[37] B. Gillett and J. Johnson. Multi-terminal vehicle dispatch algorithm. *Omega*, 4:711–718, 1976.

[38] I. Giosa, I. Tansini, and I. Vierra. A cluster insertion heuristic for single and multiple depot vehicle routing problems with backhauling. *Journal of the Operational Research Society*, 53:977–984, 2002.

[39] M. Golombek, J. Grant, T. Parker, D. Kass, J. Crisp, S. Squires, A. Haldemann, M. Adler, W. Lee, N. Bridges, R. Arvidson, M. Carr, R. Kirk, P. Knocke, R. Roncoli, C. Weitz, J. Schofield, W. Zurek, P. Christensen, R. Fergason, F. Anderson, and J. Rice Jr. Selection of mars exploration rover landing sites. *Journal of Geophysical Research*, 108:ROV13 1–48, 2003.

[40] Ronald Greeley and Peggy Ellen Thomas. Mars landing site catalog: The electronic version. http://cmex.ihmc.us/marstools/mars_cat/mars_cat.html.

[41] C. Gueguen. *Méthodes de résolution exacte pour les problèmes de tournées de véhicules*. Unpublished doctoral dissertation, École Centrale Paris, 1999.

[42] S. Hong. Design of power systems for extensible surface mobility systems on the moon and mars. Master's thesis, Massachusetts Institute of Technology, 2007.

[43] S. Kataoka and S. Morito. An algorithm for the single constraint maximum collection problem. *Journal of Operational Research Society Japan*, 31:515–530, 1988.

[44] Jet Propulsion Laboratory. Mars exploration rover mission: The mission. http://marsrover.nasa.gov/mission.

[45] G. Laporte. The vehicle routing problem: an overview of exact and approximate algorithms. *European Journal of Operational Research*, 59:345–358, 1992.

[46] G. Laporte, F. Louveaux, and H. Mercure. Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, 39:71–78, 1999.

[47] G. Laporte and S. Martello. The selective traveling salesman problem. *Discrete Applied Mathematics*, 26:193–207, 1990.

[48] G. Laporte, Y. Nobert, and M. Desrochers. Optimal routing under capacity and distance restrictions. *Operations Research*, 33:1050–1073, 1985.

[49] G. Laporte, Y. Nobert, and S. Taillefer. Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation Science*, 22:161–172, 1988.

[50] E. Lawler, J. Lenstra, A. Kan, and D. Shmoys. *The Traveling Salesman Problem.* John Wiley and Sons, 1985.

[51] H. Linstone and M. Turoff. *The Delphi method : techniques and applications.* Addison-Wesley, 1975.

[52] N. Luo, G. Miley, P. Shrestha, R. Gimlin, R. Burton, J. Rusek, and F. Holcomb. H_2O_2-based fuel cells for space power systems. In *Proceedings of 3rd International*

Energy Conversion Engineering Conference, San Francisco, California, August 2005.

[53] K. Raman, N. Choudhury, and A. Shukla. A high output voltage direct borohydride fuel cell. *Electrochemical and Solid-State Letters*, 7:A488–A491, 2004.

[54] F. Regan. *Re-Entry Vehicle Dynamics.* AIAA, 1984.

[55] C. Ribeiro and F. Soumis. A column generation approach to the multi-depot vehicle scheduling problem. *Operations Research*, 42:41–52, 1994.

[56] S. Salhi and G. Nagy. A cluster insertion heuristic for single and multiple depot vehicle routing problems with backhauling. *Journal of the Operational Research Society*, 50:1034–1042, 1999.

[57] R. Séroul. *Programming for Mathematicians.* Berlin: Springer-Verlag, 2000.

[58] D. Simchi-Levi, X. Chen, and J. Bramel. *The Logic of Logistics - Theory, Algorithms, and Applications for Logistics and Supply Chain Management.* Springer, 2004.

[59] R. Srivastava and W. Benton. The location-routing problem: considerations in physical distribution system design. *Computers and Operations Research*, 17:427–435, 1990.

[60] H. Stark and J. Woods. *Probability, Random Process, and Estimation Theory for Engineers.* Prentice-Hall, 1994.

[61] H. Tang and E. Miller-Hooks. A tabu search heuristic for the team orienteering problem. *Computer and Operations Research*, 32:1379–1407, 2005.

[62] P. Toth and D. Vigo. *The Vehicle Routing Problem.* Society for Industrial and Applied Mathematics, Philadelpha, PA, 2001.

[63] T. Tsiligirides. Heuristic methods applied to orienteering. *The Journal of the Operational Research Society*, 35:797–809, 1984.

[64] D. Tuzun and L. Burke. A two-phase tabu search approach to the location routing problem. *European Journal of Operational Research*, 116:87–99, 1999.

[65] J. Wertz. *Space Mission Analysis and Design.* Microcosm Press, 1999.

[66] T. Wu, C. Low, and J-W. Bai. Heuristic solutions to multi-depot location-routing problems. *Computers and Operations Research*, 29:1393–1415, 2002.

[67] P. Zarchan. *Tactical and Strategic Missile Guidance.* AIAA, 2002.

[68] R. Zubrin. *The Case for Mars.* Simon & Schuster New York, 1996.