航空 航天

CHINA AEROSPACE
STUDIES INSTITUTE

In Their Own Words:

Research on Intelligent Planning of On-Orbit Servicing Tasks

Translations from Chinese source documents

The "In Their Own Words" series is dedicated to translations of Chinese documents in order to help non-Mandarin speaking audiences access and understand Chinese thinking. CASI would like to thank all of those involved in this effort.

In the "In Their Own Words" series, CASI and its collaborators aim to provide Chinese texts that illustrate thoughtful, clearly articulated, authoritative foreign perspectives on approaches to warfare at the strategic, operational, and tactical levels.

Printed in the United States of America
by the China Aerospace Studies Institute

To request additional copies, please direct inquiries to
Director, China Aerospace Studies Institute,
Air University, 55 Lemay Plaza, Montgomery, AL 36112

E-mail: Director@CASI-Research.ORG
Web: http://www.airuniversity.af.mil/CASI
https://twitter.com/CASI_Research @CASI_Research
https://www.facebook.com/CASI.Research.Org
https://www.linkedin.com/company/11049011

CASI would like to acknowledge the work and effort of the team at BluePath Labs for their help with this translation. This translation was done through an automated translation engine. While we have tried to ensure that this is an accurate and readable translation, we expect there to be inconsistencies and errors. In addition, the page numbers of this translation correspond to the page numbers of the original (i.e., the page numbered page ten in this document has the content of the page numbered ten in the original).

China Aerospace Studies Institute

The mission of the China Aerospace Studies Institute (CASI) is to advance the understanding of the strategy, doctrine, operating concepts, capabilities, personnel, training, and organization of China's aerospace forces and the civilian and commercial infrastructure that supports them.

CASI supports the Secretary, Chief of Staff of the Air Force, the Chief of Space Operations, and other senior Air and Space leaders. CASI provides expert research and analysis supporting decision and policy makers in the Department of Defense (DoD) and across the U.S. government. CASI can support the full range of units and organizations across the United States Air Force (USAF), U.S. Space Force (USSF), and the DoD. CASI accomplishes its mission through conducting the following activities:

- CASI primarily conducts open-source native-language research supporting its five main topic areas.
- CASI conducts conferences, workshops, roundtables, subject matter expert panels, and senior leader discussions to further its mission. CASI personnel attend such events, government, academic, and public, in support of its research and outreach efforts.
- CASI publishes research findings and papers, journal articles, monographs, and edited volumes for both public and government-only distribution as appropriate.
- CASI establishes and maintains institutional relationships with organizations and institutions in the PLA, the PRC writ large, and with partners and allies involved in the region.
- CASI maintains the ability to support senior leaders and policy decision makers across the full spectrum of topics and projects at all levels, related to Chinese aerospace.

CASI supports the U.S. Defense Department and the China research community writ-large by providing high quality, unclassified research on Chinese aerospace developments in the context of U.S. strategic imperatives in the Asia-Pacific region. Primarily focused on China's military air, space, and missile forces, CASI capitalizes on publicly available native language resources to gain insights as to how the Chinese speak to and among one another on these topics.

AUTHOR BIOGRAPHIES[1]

Liu Bingyan is currently engaged in postdoctoral research at the Institute of War Studies of the Academy of Military Sciences and is a doctor of military operations research and an assistant researcher. He has been engaged in the research of spacecraft on-orbit servicing mission planning for a long time, has participated in the development of a number of key model projects in the direction of on-orbit servicing and the research of military scientific research topics, has published more than 30 academic papers in core journals at home and abroad such as SCI and EI, has won the first prize of the military modeling competition two consecutive times, obtained 15 authorized invention patents, and published three academic monographs in the Military Science Press, Guangming Daily Publishing House and Northwestern Polytechnical University Press.

Ye Xiongbing is currently the deputy director and researcher of the Combat Experiment Center of the Institute of War Studies of the Academy of Military Sciences, a doctor of military operations research, and a doctoral supervisor. He is a representative of the 19th National Congress of the Communist Party of China, a young and middle-aged expert of the whole army, an expert in the direction of command and control of the Academy of Military Sciences, and a leader in military operations research. He has been engaged in research work in the field of military operations research and command and control for a long time, and has made important achievements in command and control modeling and simulation engineering technology, research on experimental theory and method of military tasks, and experimental demonstration of key practical problems, and has published more than 60 academic papers, presided over and participated in more than 50 key research projects, and successfully completed the experimental demonstration of more than 20 major practical topics assigned by the headquarters of the Central Military Commission. He has won 11 first and second prizes for the army's scientific and technological progress, was selected into the national "Millions of Talents Project," and has been commended as "National Outstanding Scientific and Technological Worker," "National Young and Middle-aged Experts with Outstanding Contributions," and "Outstanding Communist Party Member of the Whole Army." He enjoys the special government allowance of the State Council and has been awarded the second-class merit once and the third-class merit three times.

Gao Yong, currently director and researcher of a certain department of the Strategic Support Force, Ph.D. from Tsinghua University, and a member of the military's "14th Five-Year Plan" aerospace field expert group. He has been engaged in research in the field of military aerospace technology and on-orbit services for a long time, presided over the completion of a number of key scientific research and engineering construction projects for the military and national defense, and led the troops to be at the leading level in China in terms of on-orbit servicing technology, successively published more than 50 academic papers, obtained more than 20 invention patents, won more than 10 second and third prizes for military scientific and technological progress, and published 3 academic monographs in the National Defense Industry Press and the Military Science Press.

[1] Translator's note: as provided in an online description of the book (http://item.kongfz.com/book/49290540.html) and not part of original book.

Research on Intelligent Planning of On-Orbit Servicing Tasks

[在轨服务任务智能规划研究]

By Liu Bingyan [刘冰雁], Ye Xiongbing [叶雄兵], Gao Yong [高勇], and Fang Shengliang [方胜良]

China Astronautic Publishing House [中国宇航出版社]

China Aerospace Studies Institute

Research on Intelligent Planning of On-Orbit Servicing Tasks

By Liu Bingyan, Ye Xiongbing, Gao Yong, and Fang Shengliang

China Astronautic Publishing House

Beijing

Book-in-edition cataloging (CIP) data
Research on Intelligent Planning of On-orbit Servicing Missions/Liu Bingyan et al.--Beijing: China Aerospace Press, 2021.12
ISBN978-7-5159-1949-2
I.(1) In... II.(1) Liu... III.(1) Spacecraft – Flight Control IV.(1) V448.2

CIP Data Verification (2021) No. 131869 of the Chinese Version Library

Editor in charge: Peng Chenguang Cover design: Yuxing Culture

Publisher: China Yuhang Publishing House (CAPN)
Address: No. 8, Fucheng Road, Beijing Zip Code: 100830
(010)60286808 (010)68768548
Website: www.caphbook.com
Distribution: Xinhua Bookstore
Issuance Department: (010) 60286888 (010) 60286887
(010) 68371900 (010) 60286804 (fax)
Retail Store: Reader Services
(010) 68371105
Printing: Tianjin Picture-in-Picture Printing Co., Ltd
Edition: December 2021 1st edition The first printing was held in December 2021
Specifications: 880x1230 Folio: 1/32
Sheet: 8 Color inserts: 12 sides Word count: 230 thousand words
Book number: ISBN978-7-5159-1949-2
Price: 68.00 yuan

If there is any problem with the printing quality of this book, you can contact the distribution department for exchange

PREFACE

With the continuous improvement of space research, development and application capabilities, countries have successively developed and launched a large number of spacecraft for various mission requirements. As an important means to ensure the long-lasting and stable operation of spacecraft in complex space environments, on-orbit servicing has become an important development direction in the field of space technology. The on-orbit servicing is built on the basis of the high development of aerospace science and technology, and has the characteristics of highly dispersed service objects, high-end and intensive service means, and accurate and efficient service methods. As the suddenness of various on-orbit servicing tasks becomes increasingly prominent, the requirements for the timeliness, resource constraints and autonomy of mission planning are also getting higher and higher. Automation and intelligence are the inevitable choices for the development of on-orbit servicing mission planning technology, and the operational practice of international space on-orbit services shows this. In order to further meet the practical needs of mission planning throughout the whole process of on-orbit servicing, and better deal with such a kind of mission planning problem with long duration and uneven type of situation, this book combines the development of on-orbit servicing technology and highlights the needs of risk disposal, explores the ways and means of applying advanced artificial intelligence technology to on-orbit servicing mission planning, and provides scientific planning means for on-orbit servicing through the combination of theoretical research and simulation experiments, and provides reference for solving mission planning problems in other fields.

This book is written on the basis of keeping up with the international research frontier and independent innovation achievements, and systematically expounds the basic concepts and principles, basic theories and methods, mathematical models and algorithms, case simulations and results of intelligent planning of on-orbit servicing tasks. The book is divided into 7 chapters. Chapter 1 is an introduction, which expounds the basic concepts, outlines the technological developments, and analyzes the current state of research. Chapter 2 analyzes the whole process of on-orbit servicing mission, analyzes the needs and clarifies the focus of the book, and introduces a research framework that uses intelligent methods to construct on-orbit servicing mission planning. Chapter 3 addresses the fact that the service users are large and scattered, and the service force in orbit is limited , and based on the research framework of intelligent planning, this paper meets the requirements of pre-planning and disposing of situations, satisfies the characteristics of the composite service mode, and gives full play to the advantages of forward transmission and reverse training of Deep Q Networks to solve the problem of spacecraft target allocation in orbit, and introduces an in-orbit target allocation method in the composite service mode. In Chapter 4, aiming at the prominent risk of untimely avoidance of space debris threat, this chapter gives full play to the advantages of rapid operation and real-time avoidance of artificial potential field method to solve the problem of spacecraft orbit temporary avoidance path planning, and introduces a spacecraft orbit temporary avoidance path planning method. Chapter 5 aims at the problem that the target is not cooperative and the equilibrium strategy is difficult to obtain, and the dynamic game is combined with optimal control to effectively solve the spacecraft orbit game problem, and a real-time planning method for spacecraft orbit game is introduced. Chapter 6 focuses on the practical application of intelligent planning for on-orbit servicing tasks, carries out the design of the mission planning system, and conducts the simulation analysis of the mission planning system to test the feasibility and effectiveness of the intelligent planning method in this book. Chapter 7 summarizes the main work of the book and the next research directions.

The research work of this book has received the concern of many leaders of the Academy of Military Sciences, the University of Aerospace Engineering, the China Academy of Space Technology, and other units, and has been guided by Si Guangya, Zhang Zhizhi, Yu Xiaohong, Xiong Wei, Zhang Yasheng, Yue Zhihong, Wang Xinbo, Liu Biliu, Liu Xiaohe, Zhou Chifei, Dong Xianzhou, Song Xumin, Jia Jun, and other expert leaders. At the same time, Ma Junwei, Liu Haiqiang, Liu Yuchen, Zhao Yongsheng, Cai Zongbao, Li Yue, Song Jiaqian, Li Yadong and other leaders gave strong support, Gang Jianxun, Yu Hongyuan, Ma Xinyi, and Shi Zhan, doctors of the Academy of Military Sciences, participated in the relevant research work, and Wang Wenwen, Xin Jiang, Zhang Qiyang, Dong Fang, Fang Ying, Zhang Song, Wang Tao, Wan Kang, Yin Wenhao, Zhang Kai, Wang Shusheng, etc., put forward valuable suggestions, and we would like to express our gratitude to them.

The publication of this book is fortunate to have the care and support of Academician Zhou Zhixin of the Chinese Academy of Sciences, and we would like to express our deep gratitude.

Due to the limited level of the authors, it is inevitable that there will be inappropriateness in the book, and readers are invited to criticize and correct.

The Authors
Beijing, December 2021

CONTENTS

Chapter 1: Introduction

1.1: Research Background of On-Orbit Servicing Mission Planning

"The clear-eyed guard against misfortune, and the wise plan for disasters the future may bring."[1] Forward-looking development of space technology is to grasp the way to win in space in the future, and seizing the frontier of space technology is to create a weapon to win space advantages. Looking at today's outer space,[2] with the circumnavigation of the "constellation" and the layout of the "eye in the sky," human spaceflight is entering the era of on-orbit servicing in an all-round way, and on-orbit servicing technology will become an important symbol of the development of aerospace in the 21st century, and will have a wide and far-reaching impact on space activities.[1]

Since the launch of the first spacecraft in 1957, various spacecraft for different purposes have been launched into orbit for more than 60 years, playing an irreplaceable role in navigation, meteorology, and communications. Looking at the vigorous development of the aerospace industry, although people's ability to enter and use space has made great progress so far, the long-standing problems of high cost, high risk, and "one-time life" have not been fundamentally solved, so that factors such as cost and risk have been affecting and restricting the long-term sustainable development of human space industry.[1] These spacecraft in orbit with an investment of hundreds of millions or even billions of yuan are limited by the aging of fuel and components, and their lives are extremely fragile, and perhaps an accident will end their service life. The emergence of these problems is mainly reflected in the following two aspects: First, the one-time characteristics of spacecraft in orbit are obvious. At present, most spacecraft are launched into orbit until the end of service, during which the hardware is difficult to upgrade and update, and if it needs to be updated, new spacecraft can only be launched into orbit.

[1] On January 18, 2016, General Secretary Xi Jinping quoted the "Three Kingdoms" quote "The clear-eyed guard against misfortune, and the wise plan for disasters the future may bring" at the seminar on the study and implementation of the spirit of the Fifth Plenary Session of the 18th Central Committee of the Communist Party of China, and pointed out: We must be proactive, take precautions, see the subtleties, prevent the gradual progress, play a good first move, fight the initiative battle, and be prepared to deal with any form of contradictions, risks and challenges.

[2] Outer space was first proposed by the 34th President of the United States, Dwight David Eisenhower (October 14, 1890 - March 28, 1969) on January 20, 1957. The concept of "outer space" was used for the first time. Subsequently, names such as "space" appeared and appeared as a legal term in international space legislation.

Second, the maintainability of spacecraft in orbit is poor. Spacecraft have high requirements for technical reliability, but the survivability is weak, and there is a lack of effective maintenance and guarantee means, and any failure may lead to the failure or scrapping of the entire spacecraft. In this context, on-orbit services have gradually been valued and have become the forefront and hot spot of the development of the aerospace field.

Early on-orbit servicing missions were mainly manned on-orbit services. In the early 80s of the 20th century, with the successful launch of the space shuttle, American astronauts carried out technical verification of on-orbit servicing based on this platform, which made the manned on-orbit servicing vigorously developed. The United States has successively carried out in-orbit maintenance of the Solar Maximum Mission satellite and the in-orbit maintenance of the Hubble Space Telescope, which reflects the tremendous economic benefits of on-orbit services. Manned services in orbit have their own unique advantages, but they require direct human involvement in the operation and are risky. Due to its advantages of low risk and low cost, unmanned autonomous on-orbit servicing has gradually become the research focus of various aerospace powers, especially the rapid development of mission planning related technologies in recent years, which provides effective technical support for autonomous on-orbit services.

Mission planning is the activity of using auxiliary tools to plan and calculate spacecraft on-orbit services, and it is an important scientific means in unmanned autonomous on-orbit services. In reality, the on-orbit servicing has the characteristics of insufficient precursor, high lack of information, and strong urgency, which determine that its mission planning is different from that of conventional mission planning. To this end, the on-orbit servicing task planning needs to focus on the overall situation, start from the details, think carefully, carefully plan, and plan scientifically, so as to achieve "planning in the middle of the battle, and winning thousands of miles away."

In recent years, the rapid development of artificial intelligence technology has provided new possibilities for the further realization of intelligent planning. Intelligent planning of on-orbit servicing tasks is based on orbit dynamics and orbit design, injecting operations research analysis, artificial intelligence theories and methods into it, and combining them with modern technical support means such as planning, modeling, and simulation.

1.2: CONCEPTS RELATED TO ON-ORBIT SERVICING MISSION PLANNING

Any theory must first clarify the messy and, if to say, confusing, concepts and concepts. Only when there is a common understanding of the name and the concept can it be possible to study the problem clearly and smoothly, and to stand on the same footing with the reader. Without a precise definition of their concepts, it is impossible to thoroughly understand their internal laws and interrelationships.[2] Therefore, in order to study the problem of on-orbit servicing mission planning, it is necessary to first clarify the relevant concepts of on-orbit servicing and mission planning, so as to unify the understanding and promote the in-depth research of methods and technologies.

1.2.1: On-Orbit Servicing

On-orbit servicing (OOS) refers to the space engineering activities that extend the life of in-orbit spacecraft, expand the functions of in-orbit spacecraft, and improve the performance of in-orbit spacecraft through humans, robots, or both.[1] As an effective means to enhance the performance of spacecraft, extend the life of spacecraft, reduce costs and risks, and reduce the number of obsolete spacecraft, on-orbit services mainly include on-orbit assembly, on-orbit maintenance and logistics support, covering a wide range of fields such as orbit transfer, on-orbit assembly, and space debris clearance.

Through the collaborative operation of robotic arms, space robots and relevant personnel, on-orbit services can complete the assembly, maintenance, capture and other space tasks of space targets, so as to reset the on-orbit system, extend the on-orbit life of the target, improve the target operation ability, and maintain the space environment. On-orbit services such as spacecraft on-orbit maintenance and upgrading, on-orbit assembly, on-orbit refueling, on-orbit observation, space rendezvous and docking, and on-orbit capture and removal of non-cooperative targets can provide new technologies and concepts for spacecraft on-orbit operation and mission execution, and have the following application prospects and advantages:[3]

(1) Extend the in-orbit life of spacecraft

By replenishing spacecraft fuel, batteries and other consumables, replacing and upgrading failed parts, and cleaning up space garbage, the in-orbit working life of the spacecraft can be effectively extended, and the safety of the spacecraft in orbit will be improved.

(2) Improve the availability of spacecraft payloads

When the spacecraft payload fails, the payload function can be restored by on-orbit maintenance of the failed parts, so as to improve the timeliness and flexibility of the space payload.

(3) Increase the scalability of spacecraft components

Due to the limitations of volume and weight, it was not possible to carry out a monolithic launch of a spacecraft with large equipment in one launch. Therefore, the complete configuration of the spacecraft can be realized through multiple launches of functional modules and the use of on-orbit assembly, so that the spacecraft has good scalability.

(4) The capture of non-cooperative goals

Due to the increase of current space missions, more and more space junk is formed by abandoned spacecraft, and the on-orbit servicing technology can realize the capture of non-cooperative targets such as space junk, and the captured non-cooperative targets can be used for on-orbit reorganization, return to the ground for reuse, or dismantle again as space junk.

(5) Reduce the cost of the task cycle

Through in-orbit replacement and upgrading, the redundant configuration of some satellite components is avoided, which reduces the design and production costs, and at the same time, the disassembly, launch and assembly of system components can also reduce the cost of the whole satellite launch. In addition, through the in-orbit maintenance of failed satellites in multi-satellite missions, satellite relaunch and replacement will be avoided, further reducing the cost of mission systems and improving reliability.

(6) Improve the flexibility and ability of on-orbit system tasks

Spacecraft are usually designed for specific tasks, and for spacecraft that have completed tasks or whose technical performance has been reduced, through on-orbit upgrades and maintenance, the spacecraft can have better working performance and can perform new space missions, so as to ensure the technological advancement and flexibility of the spacecraft.

With the continuous progress of space technology, the on-orbit servicing technology used in the maintenance of faulty or failed spacecraft and the cleaning of space junk has developed rapidly, and has shown a wide range of application prospects. The service objectives of these on-orbit services carried out in the close range of space targets can be divided into cooperative goals and non-cooperative goals. Among them, the development of on-orbit servicing technology for cooperation goals is relatively mature, and the on-orbit servicing technology for cooperation goals represented by space station application and rendezvous and docking, such as the International Space Station, China's Shenzhou series spacecraft and Tiangong docking, has become increasingly mature and has achieved huge economic and social benefits.

However, since the vast majority of spacecraft in orbit are non-cooperative targets, they cannot provide cooperation information, which will limit the scope of application of the on-orbit servicing technology of cooperative targets. The non-cooperative objectives are service objectives that do not have cooperation markers and cannot provide effective information, and their non-cooperation characteristics mainly include no pre-laid cooperation signs, no prior information, no control over them, and no communication with them, etc., and these factors are the key to achieving docking and capture, which are directly related to the success of space operation missions, which greatly increases the complexity of the problem.

Many future on-orbit servicing missions, such as repairing failed and tumbling satellites, actively removing space debris, and grabbing failed satellites, will have extremely non-cooperative service objectives or completely unknown target characteristics. The orbital maneuvering segment and the service implementation segment are the critical stages in the on-orbit servicing of spacecraft, and they are also the stages where the near-orbit operation of non-cooperative targets is very different from the traditional rendezvous and docking activities. The research on on-orbit servicing mission planning for non-cooperative goals is of great practical significance.

1.2.2: Task Planning and Intelligent Planning

The word mission planning originates from the English "mission planning," which is a foreign word, and some scholars interpret the word "mission planning" from the perspective of operations research as: mission planning is the key link in the process of mission planning that needs to use scientific planning methods to clarify specific matters, is the scientific allocation, optimization and selection of resources, space, tasks and task implementation process, and is to concretize what to do, how to do it, and who will do it.[4]

Mission planning is to plan a set of actions (or behaviors) that can be directly executed by various entities from the initial state to the target state under the condition of clear objectives and resources, taking into account various needs and constraints, and make the implementation results as good as possible to achieve the mission objectives.[5] Therefore, the on-orbit servicing mission planning is mainly to map the purpose of the on-orbit servicing, the force resources and the space environment to a unified mathematical space, use relevant theories, mathematical tools and computer technology, and scientifically calculate the force resources, orbital maneuver and service implementation according to certain target criteria under the constraints of resource conditions, operation rules and value orientation, and plan out the target allocation strategy, orbital maneuver path and service implementation plan.

The goal of on-orbit servicing mission planning is to make full use of on-orbit force resources through global operation research and scientific calculation, to achieve the maximum efficiency of on-orbit servicing, and to provide effective auxiliary decision support for the in-orbit decision-making of ground command centers and spacecraft.

The concept of intelligent planning originates from artificial intelligence and is the application of artificial intelligence technology in the field of task planning, and its main idea is: to recognize and analyze the surrounding environment, according to the expected goals, use intelligent methods to implement reasoning on a number of alternative actions and resource constraints, and comprehensively formulate a reasonable implementation strategy or scheme that can achieve the goal.[6] Intelligent planning is an important research field in the field of artificial intelligence, and it is also an interdisciplinary discipline covering multiple fields, including knowledge representation and reasoning, human-computer interaction, and cognitive science. This technology has shown great application prospects in the field of industrial scheduling, aviation, robot control and other fields. Accordingly, the intelligent planning involved in this book mainly refers to the use of relevant intelligent methods to solve several task planning problems, which is a new exploration of the transformation and application of the latest artificial intelligence technology to task planning problems.

Therefore, the intelligent planning of on-orbit servicing tasks will use intelligent methods to carry out autonomous reasoning on several alternative resource allocations, orbital maneuvering and service actions according to the service tasks and the performance of the in-orbit spacecraft platform, and form scientific, reasonable and feasible strategies and schemes accordingly.

1.3: Overview of the Development of On-Orbit Servicing Technology

Since the end of the 50s of the 20th century, the development of aerospace technology has profoundly changed human life and gradually developed into a new field closely related to human survival and development. As an important application field of advanced space technology, on-orbit servicing began with the Soviet Union's replenishment mission to the space station and then received special attention in space projects such as the Soviet spacecraft and the U.S. space shuttle.

The high-risk and high-investment characteristics of the aerospace field make the balance between cost and benefit pay special attention. The application prospects for on-orbit services in the aerospace field are mainly reflected in the following five aspects:

1) Directly repairing fault problems through on-orbit services can greatly reduce the cost of replacing whole satellites and improve the operational availability of spacecraft in orbit.

2) Component replacement or propellant replenishment through on-orbit services, update and upgrade in-orbit spacecraft, and maintain the flexibility and technological advancement of spacecraft in on-orbit servicing.

3) Realize on-orbit assembly and configuration through on-orbit services, improve the mission execution capability of on-orbit spacecraft, and improve the scalability of spacecraft.

4) Replace or upgrade the failed or obsolete components of the spacecraft in orbit through on-orbit services, replenish consumables, extend the working life of the spacecraft in orbit, and maintain the optimal operating state of the spacecraft in orbit, so as to improve the safety of the spacecraft.

5) On-orbit replacement and upgrade through on-orbit services, so as to reduce the difficulty, cycle and cost of pre-spacecraft design and manufacturing, and reduce the cost and expense of the whole life cycle.

As early as the mid-60s of the 20th century, the "on-orbit servicing" was marked by the activities of astronauts out of the capsule. In the 60s of the 20th century, the United States and the Soviet Union first carried out in-depth research on the concept and technology of "on-orbit servicing," in the early 80s, in the American "Challenger" mission, astronauts used robotic arms to successfully capture faulty satellites and carry out maintenance operations, demonstrating the feasibility of on-orbit services, in the 90s, the United States successively carried out in-orbit maintenance of the Sun Peak mission satellite and the Hubble Space Telescope, and promoted the development and research of related service technologies through specific projects, after the 90s, The United States, represented by the in-orbit assembly and maintenance of the International Space Station, has matured the on-orbit servicing technology. Looking back on the development of on-orbit services (see Figure 1-1), since the concept of on-orbit services was proposed, the research and development of on-orbit services in various countries has never stopped. After more than 50 years of continuous development and breakthroughs, the United States, Russia and other space powers have accumulated rich and mature experience in on-orbit servicing research and have made a series of achievements that have attracted worldwide attention.

In recent years, countries (regions) have continued to promote the research of on-orbit servicing projects, mainly involving auxiliary orbit change, debris removal, on-orbit fuel filling and life extension, on-orbit assembly, and on-orbit maintenance and upgrading. Among them, the United States focuses on on-orbit assembly, on-orbit fuel refueling tasks and related technologies, Europe focuses on low-Earth orbit (LEO) debris removal missions and related technologies, and Japan and Germany rely on their own advanced robotic arm technologies to carry out on-orbit servicing projects.

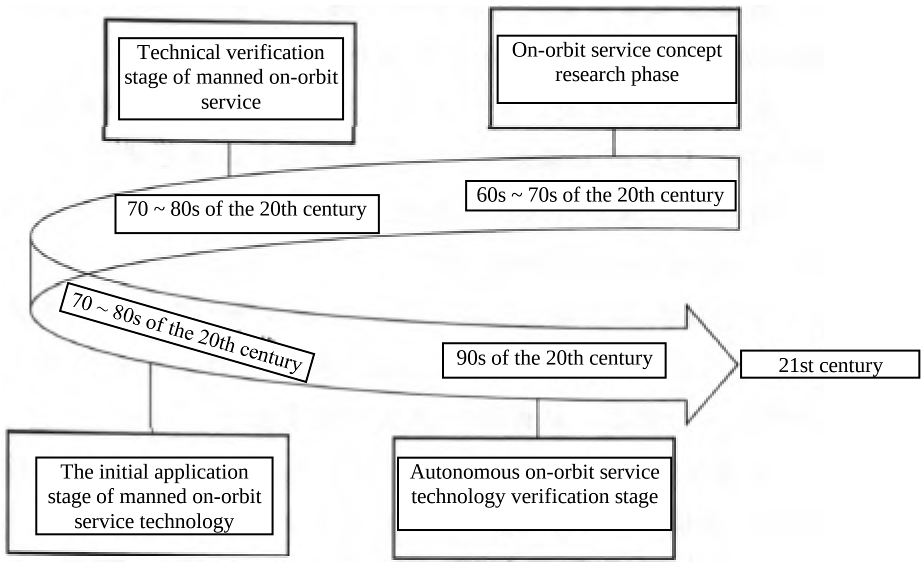


Figure 1-1: The Main Development History of On-Orbit Services

1.3.1: Debris Removal

While the development and application of space technology has brought great benefits to global development, it has also brought about problems such as space junk and space debris. With the development of space activities, the crisis of a major explosion of space debris is also approaching.[7] Space debris, also known as space junk, is fully defined as space debris in the 1999 United Nations Space Debris Technical Report: Space debris means all man-made objects, including debris and components, located in Earth orbit or re-entering the dense atmosphere that are incapable of functioning and for which there is no reason to be expected to perform or continue to perform their intended function or any other function approved or likely to be authorized, whether or not their owner can be identified.

Since the Soviet Union launched its first artificial satellite in 1957, as of January 1, 2021, mankind has carried out more than 10,000 space launches, put more than 9,000 spacecraft into Earth orbit, and has a total of 3,372 satellites in orbit around the world. Of all the spacecraft that have been launched, only a little more than 2,000 are functional, and the vast majority of the others have become space junk due to their loss of functionality. By the end of 2020, more than 20,000 pieces of debris over 10 cm had been cataloged by the Space Surveillance Network (SSN), accounting for only 0.02% of the total amount of debris.

The number of microscopic wastes that cannot be catalogued has reached several thousand tons and more than 1 billion.[8] Human space activities have caused a large amount of space debris.[9] They operate at high speeds in space orbit for long periods of time, and if left unchecked, the threat to spacecraft and space activities will become increasingly serious.[10] How to mitigate the adverse effects of space debris on spacecraft in orbit has become a prominent issue facing all countries in the world, especially the major space powers.

In order to cope with the "Kessler effect" that may be brought about by the "haze explosion" of low-earth orbit debris that is accelerating, the space powers and the international community will pay more attention to the space debris problem and increase measures to deal with the space debris crisis. These measures include monitoring, early warning, protection, mitigation and removal of space debris, the most important of which is debris removal.[7] Debris removal is primarily a service that uses spacecraft to capture non-cooperative space debris in orbit and quickly move it to a target orbital position or to eliminate combustion in the Earth's atmosphere. In view of the special geometric and motion characteristics of space debris, a variety of technical means have been developed at home and abroad, which can be roughly divided into three categories: transition deorbiting, increasing drag deorbiting, and capturing deorbiting according to the different applied forces,[11] as shown in Figure 1-2. Figure 1-3 shows the capture and deorbit methods, and the various methods are summarized in Table 1-1 through the inductive analysis of each method.

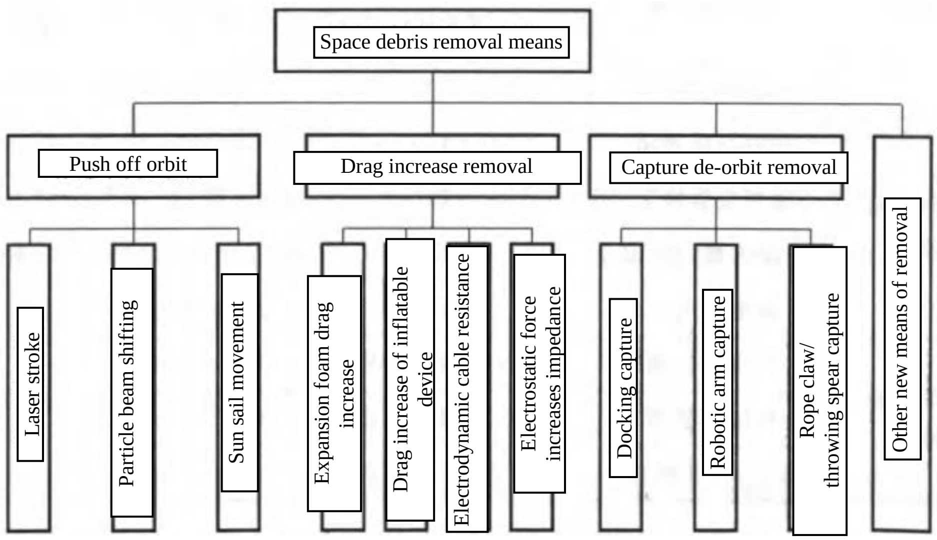


Figure 1-2: Classification of Space Debris Removal Means

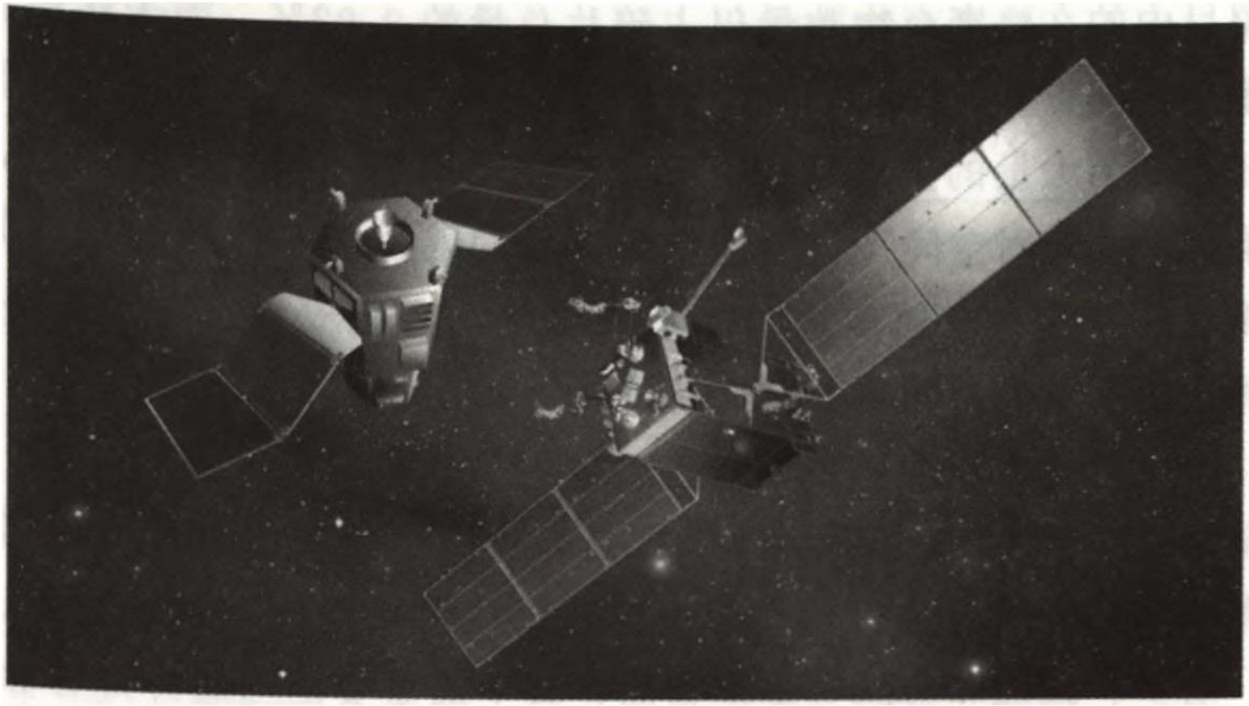


Figure 1-3: Capture and Deorbit Removal in Debris Removal

Table 1-1: Summary of various debris removal methods

Debris removal means		Orbital height range	Remove the target	Application prospects
Push off orbit to remove	Laser shift	LEO (Ground-based\Airborne)	1~10 cm	Higher
	Particle beam progression	LEO, MEO, GEO (Space-based)	>20 cm	Higher
	The sun sails on	LEO, MEO, GEO	>10 cm	Higher
Drag increase removal	Expansion foam increases resistance	LEO	>10 cm	Higher
	The inflator increases the resistance	LEO	>10 cm	Higher
	Electrodynamic cable resistance	LEO	1~10 nm	Average
	Electrostatic force increases impedance	LEO	<10 cm	Low
Capture removal	Docking capture	LEO, MEO, GEO	0.1~1 m	High
	Robotic arm capture	LEO, MEO, GEO	0.1~1 m	High
	Rope claw / throwing spear capture	LEO, MEO, GEO	0.1~1 m	High

With the increasing number of space debris, the probability of space collision is greatly increased, the space environment is more severe, and orbital resources are becoming more and more scarce.

At present, the world's space powers are vigorously developing debris removal projects, such as the European Space Agency's "European Deorbit" (e. Deorbit) project, Germany's "Deutsche Orbitale Servicing Mission" (DEOS) project, Switzerland's "CleanSpace One" (CSO-1) project, and Japan's "Space Debris Micro Remover" (SDMR) project.[12]

"European Deorbit" (e. Deorbit) is a 2012 European Space Agency (ESA) proposed engineering project to remove large space debris in low orbit to verify the technology of low orbit debris removal, and is an on-orbit servicing mission that uses in-orbit spacecraft to capture 800-1,000 km of non-cooperative large space debris belonging to the European Space Agency in low-Earth orbit/sun-synchronous orbit (SSO) and quickly move it to the Earth's atmosphere for combustion and elimination.

The " Deutsche Orbitale Servicing Mission " (DEOS) consists of two parts: service satellites and customer satellites, which can use a dedicated capture mechanism to capture rolling non-cooperative customer satellites, which is a follow-up autonomous satellite technology demonstration project proposed by the German Aerospace Center (DLR) in 2007 on the basis of the "Space System Demonstration and Verification Technology Satellite" (TECSAS) project.

The CleanSpace One (CSO-1) project is an in-orbit debris removal test project officially launched by Switzerland in 2012 to solve the problem of lack of non-cooperative target location information and difficult detection, aiming to demonstrate the active debris removal technology of small satellites by deorbiting the "Swiss Cube" (SwissCube).

The Space Debris Micro Remover (SDMR) is a microsatellite test project being developed by the Japan Aerospace Exploration Agency (JAXA) to demonstrate and validate active space debris removal technology, using a small satellite gripping robotic arm to capture space target debris and deorbit it using an electric tether.

1.3.2: Auxiliary Orbit Change

In order to provide better conditions for space exploration and at the same time save as much cost as possible, the United States and other space powers put forward the idea of realizing auxiliary orbit change in space in the 20th century. The main purpose of this concept is to reduce the costs associated with space launches. Once the auxiliary orbit change is achieved in space, the cost of space activities will be greatly reduced, so that humans can achieve more complex space missions in space.

Auxiliary orbit change is an on-orbit servicing that helps various target spacecraft to carry out orbit transfer, and its orbit change range can be transferred from low Earth orbit to geostationary orbit (GEO) and even deep space orbit.

Auxiliary orbit change (see Figure 1-4) is an important technology to ensure the stable, efficient and high-quality operation of spacecraft, which not only has important technical significance, but also has more intuitive economic benefits. At present, the design and launch of spacecraft are mostly one-time or designed to complete a certain task, so its life only ensures the completion of the mission, but with the expansion of the in-orbit mission, the spacecraft needs to perform the orbit transfer task, and the use of its own propellant is bound to cause a large loss of fuel, and the importance of auxiliary orbit change is highlighted. In recent years, the United States has mainly had auxiliary projects such as "Solar Electric Propulsion" Tug (SEP) and "Orbital Mobile Vehicle" (OMV).

Figure 1-4: Schematic Diagram of Auxiliary Orbit Change in On-Orbit Service

The "Solar Electric Propulsion" Tug (SEP) is an on-orbit servicing project proposed by the National Aeronautics and Space Administration (NASA) in 2011 for the purpose of assisting orbit change, mainly using advanced solar sail panels and high-efficiency Hall thrusters, and its thruster specific impulse is 10 times that of conventional chemical propulsion systems, which can reach 3,000 s, which can provide efficient thrust for large deep space probes, focusing on assisting large deep space probes to complete deep space orbit change tasks, and can also be used for low Earth orbit auxiliary orbit change tasks.

The Orbital Maneuvering Vehicle (OMV) is a low-cost launch device that uses an evolved disposable launch vehicle secondary payload adapter to provide small satellites with shared launch opportunities, so as to effectively complete a number of on-orbit servicing tasks such as optimizing the deployment of small satellite constellations.

1.3.3: On-Orbit Filling

After the traditional spacecraft is launched into orbit, the propellant and each component module are solidified, which will cause some problems: first, after the initial propellant is exhausted, even if other systems can still work normally, the service life of the spacecraft is about to end; second, in order to save propellant and prolong the time of use in orbit, the maneuvering ability of spacecraft is often limited, and it is necessary to have the strength to "take it easy"; third, as a large system, the failure of a key component of the spacecraft will cause the entire satellite to fail to work, which cannot help but make people feel embarrassed.[14] Table 1-2 shows the proportion of propellant required for spacecraft to enter low Earth orbit (LEO) under different missions, and it can be seen from the table that in order to complete these space missions, spacecraft need to consume more than half of their carrying capacity.[15]

Table 1-2: Proportion of propellant entering LEO for different missions

Target orbit	Propellant mass ratio
Geosynchronous Transfer Orbit (GTO)	42%
Lunar transfer orbit	50%
Mars transfer orbit	60%
Geostationary Orbit (GSO)	61%
The surface of the moon	75%

In-orbit refueling (see Figure 1-5) refers to the in-orbit operation of using service spacecraft to resupply propellant to the target spacecraft in space orbit, that is, to resupply gas and liquid for satellites, space stations and other spacecraft, which is similar to the role of "air tanker" and is an important part of the on-orbit servicing operation technology system.[14] The demand for human space exploration is increasing day by day, especially high-orbit missions and deep space missions will become more and more frequent, and the corresponding fuel consumption problem will become more and more prominent. If on-orbit refueling becomes in-orbit refueling for spacecraft like a ground car refueling station,[16] through the development of technologies such as in-orbit refueling and on-orbit module replacement, the spacecraft can be refueled and faulty parts replaced, which will provide a new power source for the spacecraft to re-enter the working state.

Therefore, the on-orbit servicing operation technology represented by on-orbit filling has a wide range of application prospects.

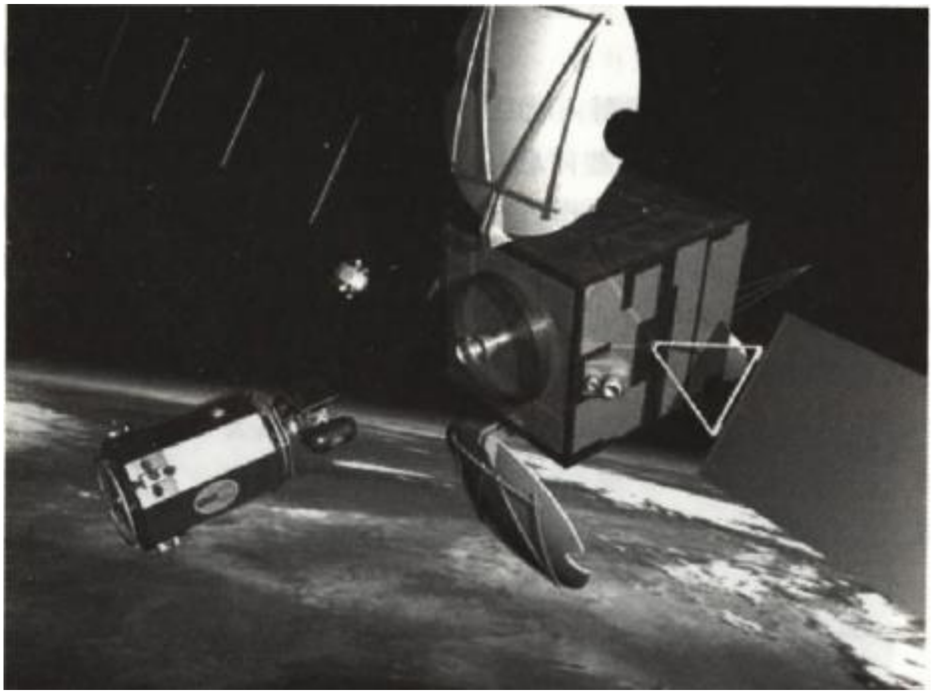


Figure1-5: Schematic Diagram of On-Orbit Filling in On-Orbit Service

As the number of in-orbit satellite assets continues to increase, and other components such as antennas and payloads can often continue to be used at the end of the satellite's fuel life, the life of the satellite can be extended and greater economic benefits can be obtained if the in-orbit server can be used for in-orbit refueling or to provide the power system for the target satellite. At present, the United States has "Restore-L" (Restore-L) in-orbit refill mission, "unmanned life extension vehicle" (MEV) and other research projects, which have the characteristics of high flexibility, good upgradeability and low risk, so as to further improve the sustainable use of space assets.[12]

The Restore-L in-orbit replenishment program is a pilot program initiated by NASA's Satellite Services Capabilities Office (SSCO) in 2014. The project aims to store and provide fuel for in-orbit spacecraft, because its unique system is a fuel transportation system, which can refuel target satellites for a long time and multiple tasks, thus effectively validating the on-orbit refueling technology for low-Earth orbit cooperative and non-cooperative targets.

The Unmanned Life Extension Vehicle (MEV) is a geostationary satellite life extension project proposed by US Space and Orbital ATK in 2011. The project uses the GEOStar3 platform and Cygnus rendezvous and docking technology, which has the characteristics of high flexibility, good upgradeability and low risk, and can be transferred to the next customer spacecraft for further service after separation from the target spacecraft, and its rendezvous and docking system can be reused to serve multiple targets, so as to ensure the sustainable use of space assets.

1.3.4: On-Orbit Assembly

At present, due to the limitation of space launch technology, only small-sized spacecraft or corresponding components can be launched, so it is necessary to launch spacecraft accessories into orbit in batches in orbit assembly technology and complete the on-orbit assembly task with the help of advanced robotic arm technology.

On-orbit assembly refers to the connection of different components in space to build a structure, subsystem, subsystem unit and other space facilities, or to separate one or more structures and reassemble them, including the on-orbit construction, replacement, connection, combination or reorganization of spacecraft, space systems and space structures, as small as module replacement, installation and deployment of battery arrays and antennas, as well as on-orbit docking of large independent modules, and the construction of larger-scale large-scale space structures. According to the scale of the on-orbit assembly task object, the spacecraft on-orbit assembly task can be divided into five levels from the top to the bottom layer,[17] as shown in Table 1-3.

Table 1-3: Spacecraft: In-Orbit Assembly Levels

Hierarchy name	Description	Example
Spacecraft combination	A combination of two or more spacecraft in orbit: spacecraft + spacecraft = (combined) spacecraft	Automatic Transfer Vehicle (ATV), H-2 Transfer Vehicle (HTV) and International Space Station (ISS) combination, etc.
Functional expansion	Add a functional module or module to the spacecraft: module (module) + spacecraft = (assembled) spacecraft	Autonomous replacement of replaceable modules (ORU): Engineering Test Satellite-7 (ETS-7) Orbital Express

Continued

Hierarchy name	Description	Example
Whole satellite assembly	Module (module) + module (module) = (assembled) spacecraft	"Phoenix" program, storage and deployment program (in-orbit manufacturing, assembly and deployment of CubeSat), "Dragonfly" project
Module assembly	Module (Components) + Modules (Components) = (Assembly) Modules	Aggregation of cellularized satlets in the "Phoenix" program, in-orbit assembly of large structures
In-orbit manufacturing	Raw material + raw material = (manufactured) parts	Space assembly of large structural systems; SpiderFab system, Storage & Deployment

There are many space assets, large roles, and high economic value, but failures are also common. At present, when the satellite fails or the satellite component technology is backward, the backup satellite can only be re-launched for replacement, but the replacement cycle is long and the cost is very high. If satellites can be repaired and assembled in orbit to reconstruct space systems, it can reduce economic losses and increase the resilience of space infrastructure.

The development history of on-orbit assembly technology (see Figure 1-6) can be divided into three stages by tracing the development process of space on-orbit assembly technology in the past 40 years: the concept formation and research stage of space on-orbit assembly technology, the research and exploration stage of manned on-orbit assembly technology, and the research and exploration stage of manned on-orbit assembly technology and independent assembly technology. On-orbit assembly has a wide range of functions, such as on-orbit docking and assembly of space structures, space vehicles or space systems. This in-orbit formation consists of antennas, solar arrays, and other attachments that can be operated by other equipment or astronauts.

In this field, the United States has also carried out a large number of project construction, such as "Phoenix," "Dragonfly," "SpiderFab," "Archinaut," "Robotic Assembly Modular Space Telescope" (RAMST), etc.[18]

The Phoenix program, which consists of the Payload Orbital Delivery System (PODS), the Tender and Satlets, is an on-orbit servicing demonstration and verification project launched by the U.S. Defense Advanced Research Projects Agency (DARPA) in 2012 to reuse valuable components from decommissioned geostationary orbit (GEO) satellites to build a new spacecraft. Demonstration verifies the technology of extracting and utilizing components (antennas, etc.) from end-of-life satellites, which is a sign of the evolution of on-orbit servicing technology from low to high orbit.

Figure 1-6: Schematic Diagram of an On-Orbit Assembly in an On-Orbit Service

“Dragonfly,” a spin-off of NASA's Langley Research Center and three commercial companies, is part of NASA's 'TippingPoint' ESC series of spin-offs from the Phoenix program in 2015. The project aims to assemble and reconstruct large solid-state RF reflectors in orbit, demonstrate the in-orbit assembly technology of autonomous antennas, and change the assembly mode of existing satellites.

"SpiderFab" is a conceptual research project proposed by the U.S. Tether Unlimited in 2012 and funded by the NASA Innovative Advanced Concept Project, which will study the concept of space in-orbit manufacturing system, and in the future, it will realize the use of 3D printing and other technologies to independently manufacture ultra-large space structures and multi-functional space system components in orbit, and at the same time use the "SpiderFab Bot" (SpiderFab Bot) to integrate large space structures in orbit to overcome the launch limitations of launch vehicles. Changing the way spacecraft are developed and deployed.

"Archinaut,” also known as "Multifunctional Space Robot Precision Manufacturing and Assembly System,” is a technology platform developed by commercial companies such as Northrop Grumman funded by NASA in 2015, and also belongs to the series of special topics of NASA's "Emerging Space Capability Turning Point.”

The project mainly includes a 3D printed additive fabricator, which is mainly used to manufacture and expand system structures, and robotic arms, which are mainly used for positioning and in-orbit assembly operations, can autonomously maneuver in space, can attach themselves to spacecraft, can be repaired and upgraded by adding or removing external components, can also remove and reuse parts from decommissioned spacecraft, and even clean up space debris.

"Robotic Assembly Modular Space Telescope" (RAMST) is a new concept of ultra-large space telescope design proposed by NASA Jet Propulsion Laboratory (JPL), NASA Goddard Space Flight Center (GSFC) and the University of California in July 2016 in the journal "Astronomical Telescope Instruments and Systems." Performing astronomical observation missions can lay the foundation for future on-orbit assembly tasks of large space structures.

1.4: Research Progress on Related Planning Technologies

Although the research on intelligent planning of on-orbit servicing tasks is an emerging field of intelligent planning application, it is also inseparable from the theoretical and practical basis of existing related problems and methods. This section will analyze the current status of research in related fields at home and abroad, compare and comprehensively review, and aim to explore the areas that need to be worked on in existing research and clarify the focus of this book.

1.4.1: Intelligent Planning of On-Orbit Servicing Tasks

1.4.1.1: Development status of intelligent planning

Intelligent planning of on-orbit servicing tasks is the application of advanced artificial intelligence technology to on-orbit servicing task planning, which is a new direction of intelligent development in the field of on-orbit servicing task planning. Represented by the United States, foreign countries have been vigorously developing the aerospace application of artificial intelligence technology, among which the direction of intelligent planning has attracted much attention. In April 2012, the United States announced the strategic roadmap for the development of robots, tele-operated robots and autonomous systems,[20, 21] which for the first time systematically proposed the development direction and key technologies of intelligent on-orbit servicing mission planning.

In October 2016, the U.S. National Science and Technology Commission released the U.S. National Strategy Plan for Artificial Intelligence Research and Development,[21, 22] which outlines seven development strategies for AI and points out the application direction of AI technology in on-orbit servicing mission planning. The U.S. Department of Defense and NASA are jointly conducting research on intelligent planning of space in-orbit missions to develop an autonomous planning system for in-orbit missions that can be constructed, evolved, and capable of autonomous planning for multiple missions.[23, 24]

At present, the X-37B orbiter, orbit express,[25] Space Maneuverable Vehicle (SMV),[26] Space Operation Vehicle (SOV), experimental satellite (eXperimental Satellite System (XSS)[27] and Demonstration of Autonomous Rendezvous Technology (DART).[28] The main purpose of which is to explore and develop new spacecraft with autonomous detection, tracking, identification and service capabilities, which can quickly maneuver to approach the target or rendezvous and dock with the target, and complete various on-orbit maneuvering or on-orbit services.[1] In addition, NASA's DS-1 has been equipped with the first spacecraft autonomous control system RA (Remote Agent), and with RAX-PS, the satellite can independently generate mission planning schemes according to high-level targets, and realize autonomous planning of missions such as observation and data transmission.[29]

The predecessor of China's research on intelligent planning of in-orbit missions is the intelligent autonomous control of spacecraft that began in the 80s of the 20th century. Mr. Yang Jiaqi[1] once pointed out that intelligent control and intelligent planning are the development direction of future aerospace technology, and explained in "The Development of Intelligent Autonomous Control Technology in China's Space Program" that "due to the problems of traditional control technology in spacecraft attitude and orbit control, various space countries have developed autonomous control technology more than 10 years ago, and it is more necessary for China to develop this technology."[30] It can be seen that with the continuous development of intelligent planning technology, intelligent planning of on-orbit servicing tasks will have the characteristics of high integration, self-organization, self-decision-making, and efficient cost ratio, and is bound to become the most effective planning method in the future aerospace field.

[1] Yang Jiaqi (1919-2006), a native of Wujiang, Jiangsu Province, an academician of the Chinese Academy of Sciences, a native of the International Academy of Astronautics, a well-known expert in automation and space technology in China, one of the main founders of automation and control technology and aerospace, and one of the advocates of the Outline of the High-tech Research and Development Plan (referred to as the "863 Program").

1.4.1.2: Current status of development in areas of concern

The "target allocation," "orbital avoidance path planning" and "orbital game strategy planning" that are the focus of this book have been developed in the corresponding fields abroad.

(1) Target allocation related technologies

Spacecraft target assignment in orbit is the application of Target Assignment (TA) technology in emerging fields, and has always been an important research direction in modern warfare and combat force application abroad.[31] The United States began relevant research as early as the 50~60s of the 20th century, and institutions such as the Massachusetts Institute of Technology and the Institute for Defense Analysis have been committed to the study of target allocation, and put forward key technologies such as force optimization and resource allocation, and resource allocation in the C4ISR environment.[32] The research on target allocation problems in relevant research institutions in the United States mainly originates from traditional theoretical techniques such as linear programming, dynamic programming, countermeasure theory and graph theory, but in recent years, with the development of environment and tasks, the goal allocation problem has gradually evolved to multi-parameter and multi-constraint optimization problems, and the traditional methods will not have the optimal solution in polynomial time, resulting in the emergence of a new generation of heuristic intelligent algorithms represented by genetic algorithm, ant colony algorithm and tabu search algorithm. Therefore, it provides a new solution for such target allocation problems with multi-objective, multi-constraint, non-differentiable, and nonlinear characteristics.[33]

(2) Orbit avoidance path planning related technologies

As an important technology for spacecraft active defense in orbit, orbital avoidance path planning is being developed by foreign aerospace organizations and institutions.[34] For example, after the crash of the space shuttle Challenger, NASA began to research on space debris avoidance technology, and successively provided technical support for the safe operation of spacecraft through regional early warning methods and avoidance technologies.[35] In 2009, eight spacecraft orbital avoidance maneuvers were conducted, of which two were spacecraft avoidance and the rest were space debris avoidance. Since 2004, the European Space Agency (ESA) has been developing orbital avoidance technologies for the Russian Cosmos-1486 and Cosmos-1424 satellites twice, and in 2010 it conducted five avoidance experiments on space debris and spacecraft.[36] The *Centre national d'études spatiales* (CNES) of France has been engaged in the research of space collision avoidance technology since the Cerise satellite collided with space debris on July 24, 1997, and carried out six evasive maneuvers in 2009 and 2010 for potentially dangerous events in orbit.[37]

In 2008, the German Space Control Center (GSOC) began to establish a collision warning and avoidance system, and in November 2009, the TerraSAR-X satellite used the system to avoid a collision with the debris of the Russian Cosmos-2251 satellite.[38] In recent years, the Japan Aerospace Exploration Agency (JAXA) has conducted research on satellite collision avoidance maneuver frameworks and technologies for its Advanced Land Observation Satellite (ALOS) using TLS cataloging data and radar reconnaissance data.[39]

(3) Techniques related to game strategy planning

As an important part of on-orbit servicing task planning, game strategy planning has irreplaceable practical significance. At present, the X-37B orbiter, Orbit Express,[25] Space Maneuverable Vehicle (SMV),[26] and Experimental Satellite System (XSS)[27] and Demonstration of Autonomous Rendezvous Technology (DART)[28] are demonstrating and verifying orbital maneuvering spacecraft with fast maneuvering and autonomous approach capabilities, and their main purpose is to explore and develop autonomous detection, tracking, and chasing capabilities. A new type of spacecraft capable of quickly chasing or rendezvous with targets and completing a variety of on-orbit servicing tasks.[1]

In recent years, universities and industrial departments in China, such as Tsinghua University, National University of Defense Technology, Harbin Institute of Technology, Northwestern Polytechnical University and Beijing University of Aeronautics and Astronautics, as well as China Aerospace Science and Industry Corporation, China Aerospace Science and Technology Corporation and China Electronics Technology Group, have continuously increased investment in on-orbit services in close connection with actual needs, and have a strong academic level in related fields. For example, Wang et al.[45] proposed a satellite attitude planning method based on deep reinforcement learning to obtain the optimal intelligent output of discrete behavior by using the attitude angular velocity to stabilize as a reward to obtain the optimal intelligent output of discrete behavior, so as to restore the satellite to a stable state. Luo et al.[46] analyzed the breakthroughs made by many scholars in recent years in introducing deep learning models into object tracking planning methods and looked forward to future research directions in the fields of sample shortage, multivariate classification problems, and real-time computing.

Liu Zhao et al.[47] designed a dynamic resource planning method that can avoid co-channel interference by combining reinforcement learning methods to reduce the probability of system congestion and further effectively utilize channel resources.

1.4.1.3: Analysis of the current situation

Intelligence is a new trend and direction in the development of the aerospace field after mechanization and informatization. From the relevant research work at home and abroad, it can be seen that artificial intelligence, represented by deep learning, reinforcement learning and other theories, is in the ascendant, and its promotion of the innovation of task planning methods is undoubted. The on-orbit servicing has the characteristics of highly dispersed deployment, high-end intensive means, fast and accurate action, accurate and efficient service, and its mission planning has certain particularities, which puts forward some new requirements for the application of artificial intelligence technology:

1) The on-orbit servicing mission planning needs to run through the whole process of on-orbit servicing, and the general planning method adopts a one-way solution process of problem description, problem modeling, method solving and result evaluation, with obvious progressive characteristics and relatively fragmented processes, which is not conducive to solving dynamic and reciprocating planning problems such as orbit games between spacecraft. Therefore, it is necessary to innovate the thinking concept, improve the methods and means, and explore the use of new methods to build a mission planning technical framework for the whole process of on-orbit servicing.

2) At present, the relevant research on on-orbit missions mainly focuses on orbit design and trajectory optimization, and the research on the uneven supply and demand quantity, space debris interference, and non-cooperative target services is still relatively weak. Due to the existence of uncertainties and unexpected factors, the on-orbit servicing scheme obtained under the deterministic assumption may not be applicable during the implementation process. In the composite service mode, target allocation, spacecraft orbit temporary avoidance and spacecraft orbit game are planning problems with obvious uncertainties and suddenness in orbit services, and more applicable planning methods can be explored with the help of the advantages of artificial intelligence technology in incomplete information processing, dynamic interaction and game update.

3) The research on task planning methods is usually based on modeling and solving the problem by classical optimization after appropriately simplifying the problem. For the solution of the on-orbit servicing mission planning problem, it is necessary to pay more attention to the orbital characteristics of spacecraft and consider the applicability of the solution algorithm to the composite service mode, the satisfaction of different avoidance needs and preferences, and the effectiveness of the bilateral control problem.

Trying to integrate the latest artificial intelligence technology to explore the intelligent algorithm for solving the problem of on-orbit servicing mission planning will help to further improve the ability of on-orbit servicing mission planning.

1.4.2: Target Allocation Method

1.4.2.1: Current status of methodological research

The target allocation method is a method to dispatch in-orbit spacecraft resources for multiple service targets to achieve the best allocation effect, which is an important condition to ensure the smooth progress of on-orbit services in the current informatization and even future intelligent process. At present, it is becoming more and more difficult for the artificial target allocation method to meet the task requirements under the condition of informatization and intelligence, and the target allocation method with a high degree of automation has become an indispensable key technology in various fields.

Objective allocation problems are essentially nonlinear combinatorial optimization problems, which can usually be solved by mathematical methods such as ergodic optimization method, enumeration method, branch delimitation method, and secant plane method. However, the solution speed of such mathematical methods decreases significantly with the increase of goals and the increase of scale, which makes it difficult for the solution accuracy of such traditional methods to adapt to the increasingly complex goal allocation problems. At present, the intelligent optimization algorithms represented by genetic algorithm and ant colony algorithm have better performance in solving target allocation problems because they have lower requirements for the continuity, differentiability and interpretability of the solution function, and also have strong adaptability to uncertainty problems.

(1) Target allocation method based on genetic algorithm

Genetic Algorithm (GA) is an optimization algorithm proposed by Professor Holland to simulate the survival of the fittest and survival of the fittest in nature, which has the advantages of high robustness, good parallelism, strong adaptability and global optimization, and has been widely used in backpack problems, scheduling problems and travel salesman problems, and has also been widely studied in target allocation problems. Wang Wei,[48] Tian Wen,[49] Yang Shanliang[50] and Chang Tianqing[51] and other scholars have studied the target allocation method based on genetic algorithm from the aspects of problem coding, population initialization, and mutation operation according to the realistic characteristics of the target allocation problem, and achieved good allocation results. In addition, scholars such as Wang Lei,[52] Chen Si,[53] Liu Zhen[54] and Yan Yuduo[55] have adapted the genetic algorithm according to the characteristics of the decision-making behavior modeling and optimization problem of target allocation. The global search ability of the algorithm is improved, which further improves the solving efficiency of the allocation model.

However, due to its unique solution steps such as population initialization, construction of feasible solution space, and selection of cross-variation, each link of genetic algorithm will have a great impact on the allocation results, resulting in limitations and shortcomings such as irregularities and uncertainties in gene coding, long iteration time, easy premature convergence, and falling into local optimal solutions.

(2) Target allocation method based on ant colony algorithm

Ant colony algorithm (ACO) is a swarm intelligence algorithm proposed by Italian scholars Colorni A and Dorigo M, which is composed of a group of unintelligent or slightly intelligent individuals who exhibit intelligent behaviors through mutual cooperation, which has the advantages of distribution, positive feedback and robustness, and is widely used in the shortest path problem, traveling salesman problem and sequence ordering problem. In terms of target allocation, scholars such as Huang Guorui,[56] Wu Congmeng,[57] and Yuan Mei[58] used the parallelism and positive feedback mechanism of ant colony algorithm to effectively solve the problem of weapon target allocation. In addition, scholars such as Xiong Yu,[59] Cui Lili,[60] and Su Miao[61] improved the ant colony algorithm, expanded the parallel search space, enhanced the ant local search experience and the global population search experience, improved the global search ability of the algorithm, and achieved better results in search quality and timeliness.

However, the search behavior of ant colony algorithm needs to be selected through pheromones, which has a strong dependence on the accumulation of pheromones on each path, resulting in a large dependence on parameter setting, slow convergence speed of the algorithm and the possibility of falling into the local optimal solution, and the obtained results also have some shortcomings in diversity and accuracy.

(3) Target allocation method based on particle swarm optimization

The particle swarm algorithm (PSO) was jointly proposed by Dr. Eberhart and Dr. Kennedy. It is an evolutionary algorithm that simulates the predation behavior of bird flocks and utilizes the information sharing of individuals in the group to make the movement of the entire group from disorder to order in the problem-solving space. It is widely used in many optimization fields due to its advantages such as few parameters and easy implementation. In terms of target allocation, scholars such as Gao Shang,[62] Yang Qisong,[63] and Chen H[64] combined particle swarm optimization with other optimization algorithms to provide a good idea for solving the weapon target allocation problem. In addition, scholars such as Li Xinran,[65] Zhang Jiao,[66] Fan Chengli[67] and Liu Shuangying[68] have improved the particle swarm optimization algorithm from the aspects of optimizing the constraint space and defining the particle change rate to balance the conflict between the global optimization ability of the algorithm and the convergence speed. This makes particle swarm optimization more suitable for large-scale weapon target allocation problems.

However, particle swarm optimization usually lacks dynamic adjustment of velocity, is easy to fall into local optimum, and cannot effectively solve discrete and combinatorial optimization problems, and needs to be used in combination with other optimization algorithms for nonlinear combinatorial optimization problems such as target allocation.

1.4.2.2: Application status in the aerospace field

The key to the goal allocation problem in the aerospace field is to solve the problem of non-deterministic (NP) difficulty in allocating spacecraft resources of different types and different orbits to targets with different orbits and priorities, which is also a polynomial complexity for the purpose of achieving maximum efficiency or minimum cost. In view of the characteristics of target allocation problems in the aerospace field, some scholars have studied the methods of integer programming, genetic algorithm and particle swarming, and some representative research results are shown in Table 1-4.

Table 1-4: Representative research results on spacecraft target allocation methods in orbit

Year	Scholar	What to study
2006	NG et al.[69]	Based on the formal description of the constraints of the data transmission resources, the general representation of the linear model of the problem and the scheduling objective function are given, and the satellite data transmission resource allocation model is established, which effectively reduces the resource conflict between the data transmission tasks
2010	Ouyang Qi et al.[29]	The research on the allocation of on-orbit servicing targets to high-orbit coplanar targets is carried out, and a linear programming model with fuel consumption as the optimization goal is established, and the optimal time allocation strategy is obtained
2017	Zhu Xiaoyu et al.[70]	Combining the space fuel station technology with the "one-to-many" on-orbit refueling problem, a round-trip in-orbit refueling allocation model based on the fuel station was constructed, and the genetic algorithm was used to solve the problem
2017	Xiao Hai et al.[71]	Considering the service efficiency and fuel consumption, a target allocation model of multi-orbit service vehicles was established, and a model solving method based on tabu discrete particle swarm optimization was proposed
2017	Zhou H et al.[72]	In this paper, the target allocation problem of spacecraft in-orbit refueling service in the "one-to-many" mode is studied, and the orbit transfer fuel consumption is taken as the optimization goal, and the genetic algorithm is used to solve the problem
2018	Tan Yinglong et al.[74]	
2018	Tao et al.[75]	Facing the development trend of unified station-network resource allocation for TT&C and data transmission tasks in the future, a station-network resource allocation method based on genetic algorithm is proposed, and the simulation results show that the method can effectively improve the success rate of resource scheduling and the satisfaction of task requirements in a variety of simulation scenarios

Continued

Year	Scholar	What to study
2019	Li Xiamiao et al.[76]	In order to improve the application efficiency of relay satellites and the completion rate of data transmission tasks, the relay satellite resource allocation model of resumable transmission is considered, and the relay satellite allocation model and heuristic algorithm are proposed based on conflict risk assessment, and the simulation results of various heuristic algorithms are compared, and it is proved that the algorithm can effectively improve the task completion rate and reduce resource loss

1.4.2.3: Analysis of the current situation

Target allocation methods based on genetic, ant colony or particle swarm algorithms have the advantages of strong robustness, good optimization effect and easy implementation, and have been widely used in nonlinear combinatorial optimization problems such as backpack, scheduling and traveling sales. Although spacecraft target allocation in orbit is a nonlinear combinatorial optimization problem, it needs to be more combined with the practical problems of on-orbit servicing, focusing on the current situation of many scattered objects and limited on-orbit servicing force, and considering the conditions such as assignment constraints, burnup constraints and timeliness constraints, which makes the application of conventional methods have certain shortcomings:

1) The conventional method seldom considers the influence of different allocation modes on the effect of target allocation, while in the actual target allocation, there are complex temporal characteristics such as pros and cons between different allocation modes and parallel serial execution relationships. Therefore, it is necessary to consider and select the most suitable allocation mode according to the actual situation in the target allocation.

2) The main purpose of conventional methods is to obtain the optimal feasible solution, and the comprehensive consideration of factors such as fuel consumption, timeliness and robustness that restrict the allocation strategy is relatively rare. However, in the allocation of spacecraft on-orbit targets, it is necessary to consider the assignment constraints, burn-up constraints and timeliness constraints as a whole to generate an allocation strategy with fuel saving, high timeliness and robustness, and minimize the possibility of strategy readjustment in the implementation process.

3) The polymorphism of constrained space is a significant feature of the spacecraft on-orbit target allocation process, and the conventional methods often weaken or simplify the constraints, and little consideration is given to the influence of different constraints. Therefore, it is necessary to fully consider the influence of various constraints on target allocation and study the target allocation method under multi-objective and multi-constraint conditions.

1.4.3: Circumvention of Path Planning Methods

(1) Current status of methodological research

The basis of orbital avoidance path planning is path planning. The basic methods of path planning research and implementation at home and abroad can be roughly summarized into global path planning methods and local path planning methods based on the control of surrounding environment information. Global path planning refers to the method of establishing a suitable environment model based on the known information and then planning a suitable path based on the complete knowledge of the global environment information without the need for real-time update of the environment. Typical global planning methods include the visual method[1], the grid method[2], and the free space method[3], among which A*[4][79, 80] and Dijkstra[5][81-84] are used. Algorithms are the most commonly used. Global path planning, because the environmental information is known in advance, no longer needs to collect a large amount of surrounding environment information, which can reduce the amount of calculation, but its planning effect is closely related to the division of environmental granularity. Local path planning is an online planning method[6], which refers to the method of collecting surrounding environment information in real time and dynamically updating the path without fully understanding the surrounding environment. Typical local planning methods include artificial potential field method[7] and,

[1] The visual method is to link the points of the starting point and obstacle with the target point to form a closed connection map, which is a method to convert the path search into the shortest distance problem from the starting point to the end point.

[2] The grid method is a method of dividing the spatial area into network units (grids) with binary information, and the units are connected to each other but do not overlap, and the obstacles are labeled

[3] The free-space method is a method that treats a moving object as a point, enlarges the obstacle according to a certain proportion, and expresses the space as a graph structure, and then conducts a graph search.

[4] A* algorithm is an algorithm that combines heuristics and formalization, which has the advantages of small time overhead, but has the disadvantages of heuristic rules that are not easy to obtain, large memory consumption, and large state dimensionality.

[5] Dijkstra algorithm is a path search method that starts from the starting point and finds the shortest distance of the next node to the target point, which is suitable for structured environments, but has the shortcomings of high complexity, blind search, and low search efficiency.[26]

[6] Online planning method: focus on relying on the sensor to obtain the current environment information, the information is constantly updated with the change of the environment, so it is necessary to continuously collect information and dynamically feedback the correction information.

[7] The artificial potential field method is a path planning method that sets the environment of the moving body as a virtual electric potential field and is subjected to positive and negative gravitational potential fields.

genetic algorithm [1], ant colony algorithm[2], Voronoi graph method[3], and fuzzy logic algorithm.[4][85-87] Local path planning is carried out when the environmental information is completely unknown or partially unknown, and has the advantages of strong real-time performance and fast response speed.[88, 89]

The Artificial Potential Field (APF) method is one of the local path planning methods, and the characteristics of other local path planning algorithms are shown in Table 1-5. In contrast, the artificial potential field method has the advantages of clear mathematical description, fast operation, small amount of calculation, low hardware requirements, and smooth planning path,[90] and is currently widely used in the path planning research of unmanned aerial vehicles, unmanned vehicles, bionic humans, etc., and some representative research results are shown in Table 1-6.

Table 1-5: Comparison of local path planning methods

Path planning methods	Advantages	Shortcomings	Scope of application
Artificial potential field method	The planning speed is fast, the amount of calculation is small, the real-time avoidance is strong, and a smooth path can be obtained	The target is unreachable, and it is easy to fall into a local minimum	Real-time online planning
Genetic algorithms	Strong flexibility, good robustness, and a wide range of applications	The calculation time of the model is long, the accuracy is insufficient, and the timeliness is insufficient	Offline planning
Ant colony algorithm	It is easy to implement, parallel distributed, and robust	There is the problem of early maturity and convergence, and it is easy to fall into the local minima trap	Parallel distributed planning
Voronoi diagram	Easy to implement, time-sensitive, and compatible	The resulting path is not smooth enough	Offline planning
Fuzzy logic algorithms	Strong adaptability to the unknown environment, and effective response to uncertain situations	The robustness is poor, the accuracy of the obtained path is not high, and the stability is poor	Online planning

[1] Genetic algorithm, by simulating the natural evolution process, according to the preferential search theory, using the selection, intersection, and mutation methods to screen the optimal path.

[2] Ant colony algorithm is a biomimetic algorithm that imitates ants foraging for food and finds the optimal path between them and food.

[3] The Voronoi diagram method, also known as the Thiessen polygon method, divides the spatial region into N and converts each region into a set of points, so as to find the shortest distance between points.

[4] Fuzzy logic algorithm, which combines intelligent methods and fuzzy control to simulate the thinking mode of the human brain.

Table 1-6: Representative research results on the artificial potential field method to solve the path planning problem

Year	Scholar	What to study
2016	He Renke et al.[91]	In order to solve the problem of route planning in complex environment and high-dimensional space, a UAV route planning method based on simulated situation field was proposed, which effectively shortened the path length and calculation time, and improved the safety of route planning
2018	Yang Lichun et al.[92]	A route planning method based on artificial potential field is proposed, which improves the ability of the aircraft to cope with dynamic environmental changes and solves the problem that the traditional artificial potential field method is easy to fall into the local minimum
2019	Shang Pu[93]	In order to further solve the obstacle avoidance problem in the flight path, the path planning and obstacle avoidance research were carried out based on the artificial potential field method, and the obstacle avoidance equation of the repulsion field function was constructed, which improved the shortcomings of the traditional method that are easy to fall into the local minimum
2019	Cheng Zhi et al.[94]	In view of the shortcomings of the traditional artificial potential field method in the path planning, an improved artificial potential field method was proposed, which further improved the planning efficiency of the obstacle avoidance path
2019	Jia Zhengrong et al.[95]	In view of the shortcomings of the artificial potential field method in the application of complex obstacle avoidance, a path planning method based on obstacle convex optimization was proposed, which improved the problem of falling into the local minimum
2020	Xu Xiaoqiang et al.[96]	In order to further improve the problems of the traditional artificial potential field method in the path planning of local minimum points and trap areas, an improved artificial potential field method was proposed, which enabled the robot to react before falling into the local minimum point or trap area, and improved the algorithm rate

In addition, in the practical application of the path planning method, the expression of spatial coordinates should also be considered, which is related to the complexity and computation of the entire path planning model. As a global coordinate system, the sampling timeliness and path fitting difficulty of Cartesian coordinate system are affected and restricted by the length of the trajectory, which makes the path planning model relatively complex. During the 2007 DAPRA Challenge, a coordinate system called Frenet was successfully applied,[97, 98] which solved the problem that the relative position of the driving carrier and the road in the path planning technology is not easy to represent, and has become the main spatial modeling method for intelligent driving in recent years. Table 1-7 lists some of the representative research results.

Table 1-7: Representative research results on the application of Frenet coordinates in path planning

Year	Scholar	What to study
2012	Werling M et al[99]	In order to solve the problems of autonomous lane change, lane merging and distance maintenance of autonomous vehicles driving along highways, an autonomous driving trajectory model based on Frenet coordinate system was proposed, and a path safety and comfort penalty function with speed change rate as the core was constructed, so that the vehicle movement path with smooth, comfortable and safer vehicle movement could be selected
2019	Wang Wei et al.[100]	Based on the Frenet coordinate system, the problems of path following and lane line detection of autonomous driving are further studied, and the adaptability to the complex and changeable actual traffic environment is improved
2019	Long Xiang et al.[101]	
2019	Wang Shajing[102]	Based on the Frenet coordinate system, the problem of vehicle motion trajectory description is studied, and it is concluded that Frenet coordinates are only related to the selection of reference lines, and the trajectory fitting calculation is simple, which greatly simplifies the motion description model and improves the calculation efficiency, which is a high-performance and low-overhead spatial modeling method
2015	Zhao Ningning et al.[103]	The Frenet coordinate system is applied to unmanned surface vehicles, spacecraft, aircraft carriers and other fields, which opens up new ideas for path planning research
2018	Su Fei et al.[13]	
2019	Yan Yongsuo[105]	

(2) Application status in the aerospace field

With the increasing demand for spacecraft orbital maneuvering and the continuous expansion of on-orbit services, the avoidance path planning technology, as the main active defense measure of spacecraft, will be an important guarantee for the safety of on-orbit services in the future. Table 1-8 lists the representative research results of some of the application of the avoidance path method in the aerospace field.

Table 1-8: Representative research results on the application of the avoidance path method in the aerospace field

Year	Scholar	What to study
2010 2011	Guo Yanning et al. Qian Yu and others	In order to solve the problem of spacecraft active defense in orbit, a pulsed optimal trajectory nonlinear programming model for the avoidance-return orbit planning problem was established, and the spacecraft avoidance path planning was studied by using potential function and other methods, revealing the influence of the avoidance maneuver of the spacecraft on fuel consumption, maneuvering distance, return time and other factors.[106, 107]
2012 2013	Yao Dangnai, et al. Sang Charl Lee et al. Zhang Jingrui et al	The relative motion model of spacecraft and space target is studied by using artificial potential field and other methods, and the maneuvering strategy of spacecraft active avoidance of space targets at close range is proposed, which effectively verifies the effectiveness of spacecraft proximity autonomous avoidance strategy.[108-110]

Continued

Year	Scholar	What to study
2014 2015	LEE U et al Miao Yuanming and others HU Q et al.	In this paper, the spacecraft orbit avoidance path planning problem under multi-constraint space is studied, and the avoidance path planning problem is transformed into an optimization problem, and the spacecraft orbit avoidance path planning method is proposed by using proportional differentiation and artificial potential field, which effectively improves the effectiveness of obstacle avoidance, obstacle area, or forbidden area.[111-113]
2017	Yu Dateng and others DONG H, et al.	In view of the increasingly serious threat of non-cooperative rendezvous faced by spacecraft, the linear relationship between the maneuvering pulse and relative motion of spacecraft is analyzed, and the effective avoidance strategies and methods of spacecraft in the face of non-cooperative rendezvous are studied, which is helpful to improve the space survivability of spacecraft.[114, 115]
2018	Su Fei and others FAN SHIPENG et al. SHEN Q et al.	In view of the dangerous intersection of spacecraft in orbit, the optimal avoidance method of spacecraft in-plane maneuver is studied by drawing on the idea of artificial potential field, and the optimal avoidance pulse is obtained under the premise of ensuring the lowest collision probability, and the effectiveness of the method is verified by two-dimensional plane trajectory simulation analysis.[13, 116-118]
2020	Li Haohao et al Yun Chaoming et al. Wang Guogang et al. Carmen Pardini et al.	In view of the active approach threat of spacecraft with active maneuvering capability to the target spacecraft, the collision between spacecraft and other space targets is analyzed, the spacecraft active avoidance methods and strategies are studied, and the effectiveness and feasibility of the collision avoidance strategy are verified.[119-122]

(3) Summary of the current situation

Path planning methods based on Dijkstra, RRT or Voronoi graphs have the advantages of strong search ability, simple structure and easy implementation, and have been widely used in unmanned aerial vehicles (UAVs) and unmanned vehicles. In contrast, there are few studies on the path planning method in spacecraft orbit temporary avoidance, but the dynamics and control strategies of orbit avoidance have been studied in depth at home and abroad,[109, 114, 161] which provides a solid technical foundation for the realization of avoidance path planning. In the planning of spacecraft orbital temporary avoidance path, it is necessary to pay more attention to the prominent risk of untimely space debris attack and avoidance, and the timeliness of orbit avoidance and the optimality of fuel consumption should be considered while successfully avoiding space debris,[112, 163, 164] which makes the direct application of conventional methods have certain limitations:

1) Combined with the temporary event characteristics of orbit avoidance, how to systematically analyze the avoidance strategies and behavioral decisions after the occurrence of space debris disturbance events; how to construct a relative motion coordinate system reasonably; how to use path planning methods for scientific planning; there are relatively few existing studies addressing these issues.

2) The artificial potential field method has the shortcomings of local minima and oscillation, coupled with the unique orbit characteristics of space and the different needs and preferences of spacecraft to avoid space targets, so that the artificial potential field method is not suitable for directly solving the problem of autonomous planning of orbit avoidance paths, but there are few improved studies related to it.

3) Space debris attack events have the characteristics of suddenness, coupling and destructiveness, and how to take into account the factors of avoidance safety, fuel consumption, minimum offset and braking time when planning the avoidance path; how to meet different preferences and multiple constraints; these are all questions that deserve further in-depth exploration.

1.4.4: Gambling Strategy Planning Methods

1.4.4.1: Current status of methodological research

In the study of game problems, although classical guidance control theory,[123] optimal guidance theory,[124] and modern nonlinear guidance theory can be used, they are mostly restricted by factors such as strong target maneuverability and changeable maneuvering laws.[125] As a result, the theory of differential countermeasures has become the mainstream method at present. The differential countermeasure theory can establish a game model between the two sides based on the optimal control theory of bilateral adversarial measures and can consider the incomplete information and symmetry degree in the game between the two sides, which is more suitable for this kind of continuous time dynamic game problem.

The earliest development of differential countermeasures originated from the study of the "pursuit and escape problem," which is a subfield of countermeasure theory under the constraints of differential equation dynamics, and is currently mainly focused on the military field or the study of confrontational factors.[127-131] Due to the high complexity and dimensionality of the differential countermeasure problem, numerical methods such as target method, coordination method and intelligent optimization algorithm are usually used to solve the problem.

(1) Target method

As a common method to solve the problem of differential countermeasures, the target method takes the necessity of finding the saddle point of the differential countermeasure problem as the starting point, and continuously modifies the initial conditions to make the differential equation satisfy the boundary value condition. For example, scholars such as Qu Xiangju,[132] Wang Hua,[133] Feng Haoyang[134] and Peng Kun[135] have studied the solution method based on the target method for the problem of flight differential strategy, which greatly improves the computational efficiency.

Zhang Qiuhua,[136] Peng Qiqing,[137] Liao Yihuan[138] and Zhang Zixiong[139] combined the target method with genetic algorithm or Gaussway Popularization Law to solve the problem of orbital game between spacecraft. A hybrid optimization strategy is proposed to improve the problem convergence, obtain the optimal control strategy for both adversarial sides, and ensure the accuracy and robustness of understanding.

By using the target method to solve the problem of differential countermeasures, the saddle point value can be obtained directly while obtaining the high-precision numerical solution. However, due to the complexity of the differential equations, the complexity of the constraint space and the nonlinearity of the boundary value conditions, it is difficult to solve the corresponding boundary value problem.[127] In addition, since the core of the target method is the improved Newtonian iterative method, the selection of the initial value requires a deep understanding of the kinetic model, otherwise it may lead to the difficulty of the algorithm convergence.

(2) Matching method

Different from the target method, the matching method transforms the differential countermeasure problem into an optimal control problem and seeks the zero point of the algebraic equation and boundary conditions through optimization methods such as quadratic programming or gradient recovery to obtain the optimal solution. For example, scholars such as Li Longyue,[140, 141] SHIMA T,[142] and Yu Jianglong[143] studied the modeling process and solution method of the pursuit and escape problem based on the theory of differential countermeasures, established an orbital game model, and proposed a solution method by using the coordination method. Scholars such as Zhao Jisong,[144] Xue Guohao[145] and Zhu Yingxuan[146] used the coordination method to solve the problem of differential countermeasures, and proposed corresponding planning methods, so as to ensure the accuracy and fast calculation speed, which can meet the needs of real-time game tracking optimization.

Although the coordination method is better than the target method in terms of convergence, its solution time and calculation accuracy are heavily dependent on the initial value, and it faces problems such as dimensional explosion and small convergence domain. In addition, in the differential countermeasures in the aerospace field, the solution process is different from the general maximum or minimum countermeasure problem, because the saddle point needs to consider the optimal control quantity of the two parties' strategy selection, which will involve the strategic equilibrium problem of the dynamic control of the two sides, which also makes the coordination method have shortcomings in convergence speed and accuracy.

(3) Intelligent optimization algorithm

Modern intelligent optimization algorithms, represented by genetic algorithms, ant colony algorithms, and neural networks, are derived from human simulations of natural phenomena, and have the advantages of not being easy to fall into local minima, and have a wider range of applications. For example, scholars such as Rauwolf G A,[147] Coverstone Carroll [148] and Larry D D [149] have used genetic algorithms to solve differential countermeasures for extra-atmospheric orbit transfers and have achieved many research results.

Later, scholars such as Chen G,[150] Kumar G N,[151] Zeng Guoqiang,[152] and Wang Jie[153] improved the genetic algorithm and applied it to solve other differential countermeasures such as satellite orbit change. In addition, scholars such as Qingzhen Z,[154] Xie Lei,[155] Ozan,[156], Wang Yin,[157] and Cheng Lin[158] have used ant colony, Simulated annealing and neural networks have been used to study the differential countermeasures of orbital transfer.

The modern intelligent optimization algorithm avoids the initial value guessing problem of the indirect method and has a broad application prospect. However, when solving the actual parameterization problem, intelligent optimization algorithms such as genetic algorithm often have the problem of precociousness, that is, individuals with large adaptability lose their diversity prematurely due to rapid diffusion, thus facing the problem of local optimality and unable to ensure the convergence of the objective function in a limited time.

1.4.4.2: Application status in the aerospace field

The on-orbit servicing of spacecraft to non-cooperative targets is a deep integration of optimal control and dynamic game, which can be described as an orbital game problem.[165] It is a typical sequential decision oriented to incomplete information, which is a dynamic game process in which the optimal behavior is taken under the condition of only knowing one's own state and the current finite state of the opponent, and not knowing the future behavior strategy of the other party. In view of the problem of continuous dynamic conflict between the two sides, some scholars at home and abroad have studied it through differential countermeasures. In contrast, the relevant research in foreign countries started earlier and achieved certain results,[167, 168] while the domestic research on spacecraft orbit game started late, and some scholars have begun to carry out related research in recent years, and the representative research results are shown in Table 1-9.

Table 1-9: Representative research results on the application of orbital game in the field of aerospace

Year	Scholar	What to study
2014	Zhang Qiuhua et al.	The strategy problem in the orbital game of two spacecraft is described by applying the differential countermeasure theory, and the countermeasure research is transformed into a high-dimensional time-varying nonlinear two-point boundary value problem for numerical solution.[136]
2016	Wang Qiang et al.	The interception rendezvous at the end of the satellite is regarded as tracking and escape, and transformed into a zero-sum differential countermeasure problem, and the interception miss-target amount and fuel consumption are used as the quadratically optimal objective functions to derive the satellite orbit suboptimal control strategy.[169]

Continued

Year	Scholar	What to study
2016	Chang Yan et al	The quantitative differential countermeasure method is used to analyze the differential countermeasure problem of space rendezvous chasing under continuous thrust, and a method to solve the differential countermeasure problem by nonlinear programming is proposed.[165]
2017	Zhu Hai et al	Considering the orbital game problem between spacecraft as a zero-sum differential countermeasure problem, the saddle point solution method of the differential countermeasure problem is studied, and the optimal control law of the chase game is proposed, so as to extend the solution method of the orbital chase game problem.[171, 172]
2019	Wu Qichang et al.	In order to solve the hot problem of spacecraft chasing game, the pursuit strategy was studied by using the survival countermeasure theory, and a numerical solution method of the chasing strategy based on ant colony algorithm was proposed.[173]
2019	Sun Songtao et al. Hao Zhiwei et al.	In order to solve the time-fixed orbit chase problem, a two-party optimal control strategy solution method is proposed based on the semi-direct point matching method, which improves the convergence of the two-point boundary value problem and brings a new idea for the differential countermeasure problem of chase game.[174-177]
2019	Zhao Lin et al.	In order to study the orbital game problem of spacecraft in three-dimensional space, combined with the theory of differential countermeasures, the description of the optimal control strategy of the tracker is obtained.[178]
2020 2021	Wang Chunbao et al. Feng Haoyang et al. Wang Yuqi et al.	The process of spacecraft rendezvous and terminal interception is regarded as a kind of differential game problem, and then the game strategy method is studied based differential game theory, which provides a simple, efficient and stable calculation method for strategy solution.[179, 134, 181]
2020	Luo Ya et al.	This paper reviews the research status of spacecraft orbit game based on differential countermeasures, sorts out the development context and current research hotspots of differential countermeasure theory, and looks forward to the follow-up research direction of spacecraft orbit game.[126]

1.4.4.3: Summary of the current situation

Strategic planning of orbital games has always been a difficult and thorny problem because it involves complex differential equations, nonlinear constraints, and incomplete information assumptions.[141, 183] The solution is based on the target method, the matching point method or the intelligent optimization algorithm, which has the advantages of simple model structure, mathematical analysis and calculation, and has a good application effect for the optimal control problem. Although the real-time planning of spacecraft orbit game strategy is an optimal control problem, it reflects more of the sequential game process of the two sides, and it is necessary to focus on the continuous dynamic interaction characteristics of the two sides, which makes the application of conventional methods have certain shortcomings:

1) Through the in-depth combing and analysis of the existing relevant literature, it can be seen that the real-time planning of spacecraft orbit game strategy is still a relatively new field, and the existing research is still in its infancy, and the existing results have not yet formed a complete and systematic theory and method.

2) In the application of existing differential countermeasures, most of them assume that both sides of the game know exactly the payment function of each other, while in the spacecraft orbit game, the two sides often cannot know the purpose of the opponent's game, which is an incomplete information differential game, and it is urgent to develop new research methods on the basis of the existing theories and methods.

3) With the rapid development of the new generation of artificial intelligence methods represented by deep reinforcement learning, according to its advantages in self-learning and self-optimization, decision-making control problems are not limited by task mode, and have been widely used in military, computer, transportation and other fields, and remarkable results have been achieved.[184-187] Although these studies have enabled the application of deep reinforcement learning algorithms in the field of control decision-making, they still face similar problems as tabular reinforcement learning in continuous space applications, that is, the number of operations that need to be explicitly represented increases exponentially with the increase of the number of operation dimensions, which needs to be further studied and improved.

1.5: THE CONTENT AND CHAPTER ARRANGEMENT OF THIS BOOK

1.5.1: Contents of the Book

In view of the outstanding risks of on-orbit services, such as uneven supply and demand, space debris disturbance, and non-cooperative target services, this book is supported by the most advanced artificial intelligence technology, and generally expounds an intelligent planning research framework for the whole process of on-orbit services in accordance with the discussion idea of "raising problems, analyzing problems, providing methods, and verifying and analyzing," and on this basis, it focuses on "on-orbit target allocation under the composite service mode," "spacecraft orbit temporary avoidance path planning" and "spacecraft orbit game strategy real-time planning" method.

(1) Research framework for on-orbit servicing mission planning

In view of the actual demand that mission planning should run through the whole process of on-orbit servicing, the relevant concepts of mission planning are defined, and the disposal of mission planning for the whole process of on-orbit servicing is explored. This paper analyzes the needs of on-orbit servicing task planning, focuses on difficult problems, and puts forward the main contents of this book: taking into account the general solution method of task planning, and introducing a research framework for task planning that uses intelligent methods to construct the whole process of on-orbit servicing.

(2) Research on on-orbit target allocation under composite service mode

In view of the problem of "nonlinear combinatorial optimization" presented by spacecraft on-orbit target allocation, this paper explores how to construct a target allocation model considering both execution benefit and energy consumption efficiency, and studies how to give full play to the autonomous computing advantages of forward transmission and reverse training of Deep Q Networks algorithm, which is suitable for spacecraft on-orbit target allocation. This paper introduces an optimization method that can allocate targets based on prior information such as target priority, service success probability, and fuel consumption estimation in the composite service mode.

(3) Research on the planning of spacecraft orbit temporary avoidance paths

In view of the problem of "multi-limited shortest path" presented by the temporary avoidance of spacecraft orbit, a space motion model that can not only express the relative motion of space debris but also take into account the absolute motion along the set transfer orbit is explored, and how to give full play to the advantages of simple description, fast operation and real-time avoidance of artificial potential field method to solve the problem of temporary avoidance path planning of spacecraft orbit. In this paper, a path optimization model that can comprehensively consider the factors of target avoidance, fuel saving, minimum offset and braking flexibility is studied, and a spacecraft orbit temporary avoidance path planning method that can meet different avoidance needs and preferences is introduced.

(4) Research on real-time planning of spacecraft orbit game strategy

In view of the problem of "continuous dynamic interaction" presented by the orbital game between spacecraft, this paper tries to combine the dynamic game with optimal control to effectively solve the problem of sequential game, explores a differential countermeasure model that can effectively describe the interactive motion between spacecraft, and studies a branched deep reinforcement learning architecture with multiple sets of parallel neural networks and shared decision-making modules. This paper introduces a real-time planning method of orbital game strategy that can cope with the dynamic interaction between the two sides, high autonomy and timely decision-making.

(5) Design and application of on-orbit servicing mission planning system

Based on the application requirements of on-orbit servicing experiments and ground simulation training, the foundation, requirements and key points of the construction of the mission planning system are studied, and the application requirements of the on-orbit servicing mission planning system are clarified.

Design the system architecture of the task planning system, build the functional architecture of the task planning system, and realize the systematization, modularization, easy operation, easy expansion and easy maintenance characteristics of the system; based on the case scenario, the on-orbit servicing task planning system is used to test the feasibility and effectiveness of the intelligent planning method for on-orbit servicing tasks, so as to better meet the needs of on-orbit servicing task planning, so as to provide strong support for on-orbit experiments or ground simulation training.

1.5.2: Chapter Arrangement

According to the research content, the book is divided into seven chapters, which are developed in the order of research question formulation (Chapter 1), problem analysis (Chapter 2), methodological research and case verification (Chapter 3 to 5), and system design and application (Chapter 6), and finally summary and outlook (Chapter 7), as shown in Figure 1-7.

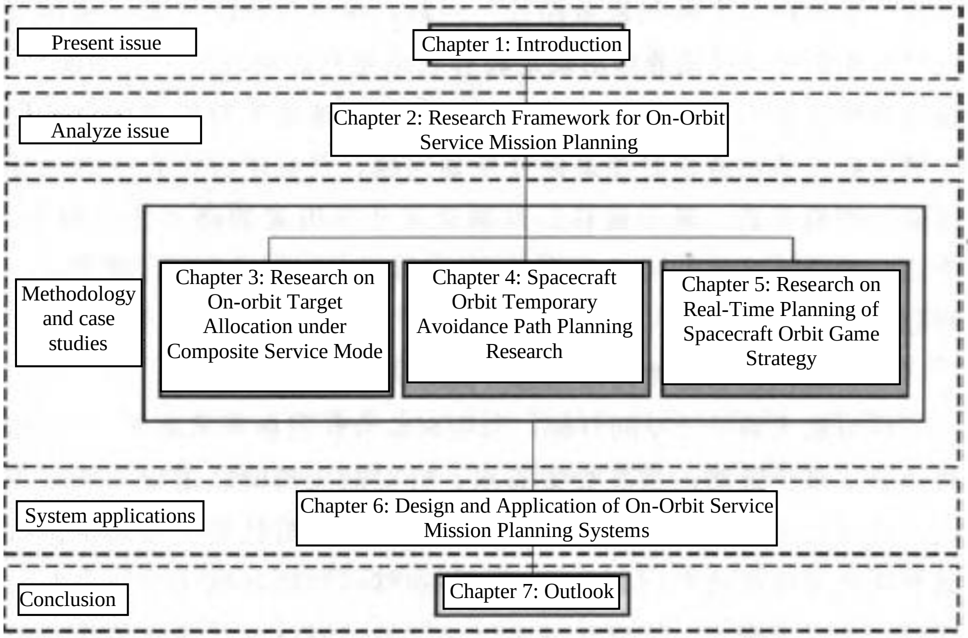


Figure 1-7: Chapters in This Book

Chapter 1 is the introduction. Raise questions, describe the research background, elaborate relevant concepts, analyze the development status and shortcomings of related task planning methods, and give the main content and chapter arrangement of this book.

Chapter 2 is a research framework for on-orbit servicing mission planning. Facing the whole process of on-orbit servicing, on the basis of the overview of on-orbit servicing task planning, the demand analysis is carried out and the research content is clarified, and the intelligent method is introduced to construct the research framework of on-orbit servicing task planning. Chapter 3 is a study on the allocation of on-orbit targets under the composite service mode. Based on the research framework of intelligent planning, facing the needs of pre-planning and disposing of situation, meeting the characteristics of composite service mode, and giving full play to the autonomous computing advantages of forward transmission and reverse training of Deep Q Networks to solve the problem of spacecraft target allocation in orbit, through the description and modeling of target assignment problems, an autonomous target allocation algorithm based on improved Deep Q Networks is introduced, and the comparative advantages and applicability of the algorithm are tested through case analysis.

Chapter 4 is a study on the planning of temporary avoidance paths in spacecraft orbits. Based on the research framework of intelligent planning, facing the needs of temporary planning situation disposal, coping with the problem of space debris attack, and giving full play to the advantages of rapid operation and real-time avoidance of artificial potential field method to solve the problem of spacecraft orbit temporary avoidance path planning, through the description and modeling of orbit avoidance problem and the description of orbit avoidance maneuver based on Frenet coordinate system, an orbit temporary avoidance path generation algorithm based on Frenet and improved artificial potential field is introduced. The comparative advantages and applicability of the algorithm are tested through case analysis.

Chapter 5 is a study on real-time planning of spacecraft orbit game strategy. Based on the research framework of intelligent planning, facing the needs of real-time planning situation disposal and coping with non-cooperative goals, this paper takes the combination of dynamic game and optimal control to effectively solve the spacecraft orbit game problem, carries out the analysis of the existence and consistency of orbit game problem description and modeling, and equilibrium strategy, introduces an orbital game strategy solving algorithm based on branched deep reinforcement learning, and tests the comparative advantages and applicability of the algorithm through case analysis.

Chapter 6 is the design and application of the system for the planning of on-orbit servicing missions. Based on the requirements of on-orbit servicing and ground simulation application, the foundation, requirements and key points of the construction of the mission planning system were introduced, the architecture of the mission planning system was designed, and the functional page of the mission planning system was constructed. Based on the case scenario, the feasibility and effectiveness of the intelligent planning method for on-orbit servicing tasks are comprehensively tested.

Chapter 7 is the outlook. This paper summarizes the main work of the research in this book and looks forward to the next research direction.

On-orbit servicing mission planning mainly refers to the scientific planning and planning of elements and activities such as target allocation, orbital maneuvering, and service implementation of in-orbit spacecraft. Task planning, characterized by dynamic interaction, is always closely related to its planning needs, and the pursuit of timeliness, controllability and autonomy constitutes the internal driving force for the development of intelligent planning. It is helpful to clarify the classification and characteristics of on-orbit servicing task planning, stage division and problem description, situation disposal and planning process, analyze the planning needs and difficult problems, and analyze the general solution methods and their solution difficulties, which will help to further grasp the core essence of mission planning and put forward a research framework based on intelligent planning.

2.1: Overview of On-Orbit Servicing Mission Planning

According to the implementation process, the on-orbit servicing task can be divided into three stages: target allocation, orbital maneuvering and service implementation. On-orbit servicing task planning can be divided into three steps: pre-planning, real-time planning and real-time re-planning, and can be divided into four steps according to the planning process: information elements, target decomposition, constraint analysis and decision-making calculation.

2.1.1: Overview of Phases and Planning

In the face of the complexity of the environment, the continuity of the mission, and the dynamic variability of the service objectives, the on-orbit servicing requires mission planning to run through the beginning and end of the on-orbit servicing. For such a long-duration mission planning problem, a phased classification planning strategy is adopted to divide an on-orbit servicing task into three stages: target allocation, orbital maneuver and service implementation,[188] and each stage corresponds to a specific mission planning problem. Figure 2-1 shows the phases of on-orbit servicing task planning.

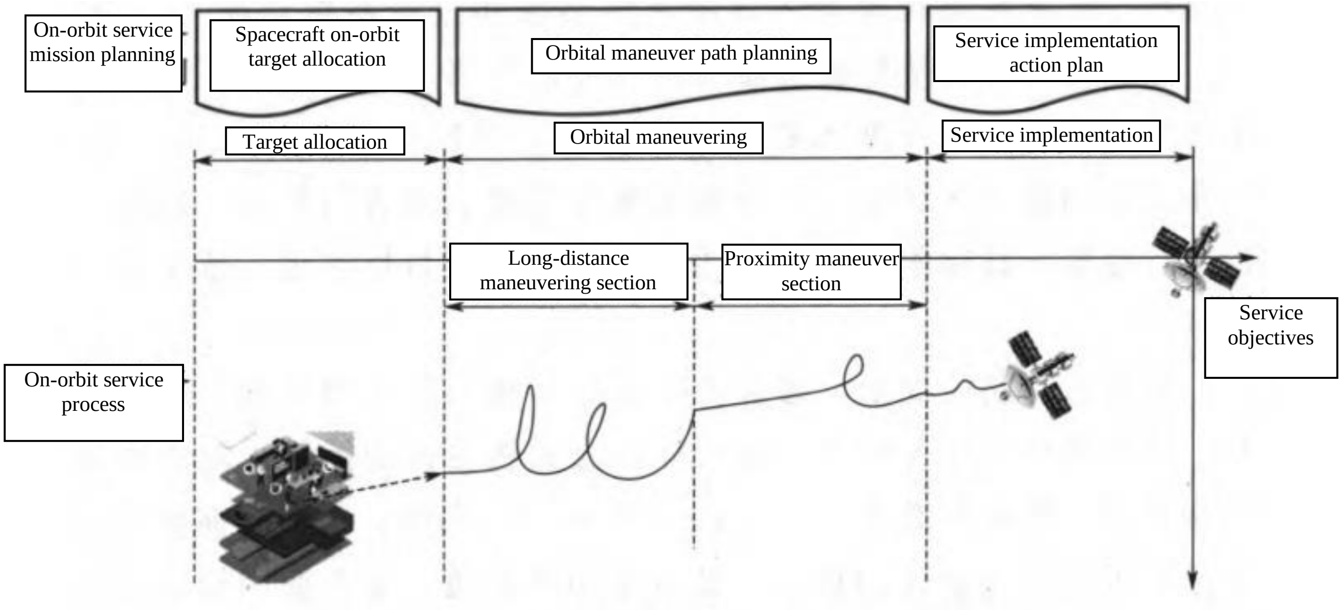


Figure 2-1: Phases of On-Orbit Service Task Planning

(1) Spacecraft target allocation in orbit

Normally, goal allocation is the initial link and an important part of task planning, which is a decision-making process that determines strength, goals, actions, and strategies on the basis of grasping the situation, based on factors such as environment, resources, and goals, and through scientific calculation, gradual deepening, and continuous improvement. Whether the target allocation is correct or not, and whether the decision can be made accurately, directly determines the process and effect of the on-orbit servicing.

Due to the uncertainty of the space environment, especially the orbital motion characteristics of the service targets, the service targets in certain periods and areas are suddenly relatively concentrated, and the temporary tasks are prominent, and it is necessary to temporarily increase forces or urgently assign tasks. Therefore, the allocation of spacecraft on-orbit targets needs to be based on the mission, according to the current orbit situation, combined with the priority of each service target, considering various on-orbit resource conditions, screening benefit indicators, and focusing on considering and estimating the implementation effect of the subsequent orbital maneuver stage and service implementation stage, and planning with this goal, so as to assign service objectives, determine service strategies, select orbit transfer methods, and formulate service initiation time and duration for in-orbit spacecraft.

(2) Orbital maneuver path planning

Orbital maneuvering is the process of orbital transfer carried out by a spacecraft to reach the vicinity of the orbital position of the service target after the target is assigned. According to the space distance between the spacecraft and the service target and the maneuver distance, the orbital maneuver can be divided into long-distance maneuver and close-range maneuver.[189]

Among them, the long-distance maneuver section refers to the process in which the spacecraft completes several orbital maneuvers after leaving the original orbit according to the assigned task until the sensor on board the spacecraft captures the service target; the proximity maneuver section refers to the time from the time the sensor on board the spacecraft captures the service target to the time when the spacecraft flies to a certain point near the service target, so as to capture the target orbit and reduce the approach speed.

Orbital maneuver path planning usually takes into account the relevant task requirements such as orbital height, velocity impulse, and multi-body environment, and then evaluates the availability of geosynchronous orbit, elliptical orbit or Lagrange point orbit that can be taken to complete orbital maneuvering, and then carries out a compromise of orbit design according to the early target allocation: whether to use circular orbit or whether to use eccentricity orbit, and then select orbit height and inclination, determine the cost of maneuver options, and finally form an orbit design scheme. In addition, the realization of orbital maneuvering of on-orbit servicing tasks also touches on the spatio-temporal system, task constraints, force analysis, orbit design and trajectory optimization, etc., and has clear requirements for time, space, energy, and even the necessary point, avoidance point, overflight point, and orbiting point, etc., which involves the shortest path problem of multi-constraint and multi-variable.

(3) Service implementation action plan

Service implementation usually refers to the process of providing services to the target when the spacecraft orbit maneuvers to the vicinity of the service target. Service implementation is the core link to complete the entire on-orbit servicing task, and the service objectives are accurately controlled to achieve the expected results.

Service implementation action planning is the process of selecting behaviors and determining strategies when the spacecraft approaches the service target in order to meet the necessary conditions such as position, velocity, relative attitude and angular rate required for service implementation. The service implementation action plan should focus on all possible responses to the service objectives, actively respond to all adverse effects caused by uncertain behaviors, make scientific decisions, plan accurately, and always grasp the initiative to ensure the smooth completion of the entire on-orbit servicing task.

2.1.2: Analysis of the On-Orbit Servicing Mission Planning Process

On-orbit servicing mission planning needs to have the ability to effectively deal with various situations and be able to effectively plan according to scientific processes.

2.1.2.1: Disposal of task planning

The on-orbit servicing mission planning is mainly oriented to the space environment and on-orbit servicing missions, and the results obtained from the planning should be refined to the implementable strategies or schemes. However, in practical applications, the space environment, mission requirements, and target states are often dynamically changing, and the actual situation is often inconsistent with the prior planning, the execution effect of the pre-sequence mission is inconsistent with the plan, and even unexpected situations such as on-orbit equipment failure, human error, and debris attack are encountered, resulting in the delay or cancellation of the mission. Therefore, on the basis of prior planning, it is necessary to consider the latest battlefield situation and mission requirements, carry out temporary planning adjustment or real-time planning, update the on-orbit servicing plan, and determine the next strategy and action in a timely manner.

For the whole process of on-orbit servicing, for such a long-lasting task planning problem, the phased classification planning strategy shown in Figure 2-2 is adopted, and three disposal methods are adopted, namely, pre-planning, temporary planning, and real-time planning.[190] In Figure 2-2, pre-planning covers the whole process of on-orbit servicing, focusing on pre-planning. Temporary planning to deal with temporary events in the process of on-orbit servicing operations, focusing on timely adjustment; real-time planning focuses on the service implementation process, with an emphasis on emergency response.

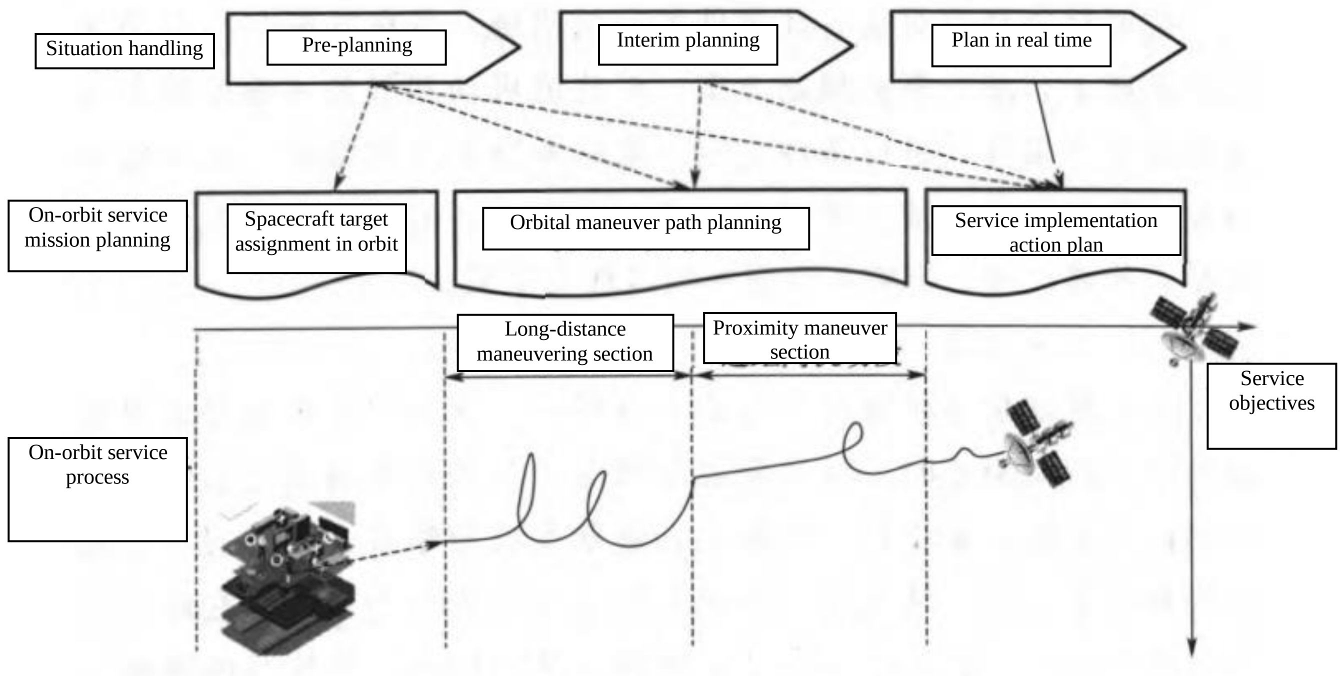


Figure 2-2: Disposal of On-Orbit Service Mission Planning

(1) Pre-planning

Prior planning is the process of determining the target allocation strategy, orbital maneuver path and service implementation plan according to the grasp of the space situation, service purpose and target situation, combined with the spacecraft's power attributes and in-orbit resources, before the spacecraft performs the on-orbit servicing mission. Pre-planning focuses on the completeness, informativeness and implementability of the planning strategy.

Ex-ante planning is the core link and important part of on-orbit servicing mission planning, which is a decision-making process that determines the objectives, paths and strategies on the basis of mastering the service objectives and based on factors such as environment, resources, and objectives. Whether the pre-planning is reasonable or not is directly related to the feasibility of the on-orbit servicing action plan and affects the whole process of on-orbit servicing.

(2) Interim planning

Interim planning is a process of temporary planning and adjustment in the face of temporary events such as space debris and space disturbance in orbit service operations, and it is necessary to carry out temporary planning and adjustment according to the current space situation, target state and changes in the event time sequence. The interim plan focuses on the continuity of the original policy and strives to make the original policy available with the least modification and the least cost.

Temporary planning is an effective measure to deal with temporary events, which is the process of setting conditions, adjusting strategies and modifying temporary events in the process of orbital maneuvering and service implementation based on the pre-planning plan on the basis of grasping the evolution of the spatial situation, so as to make the strategy and plan meet the actual situation. Whether the interim planning is effective is directly related to whether the mission can be sustained and affects the process and results of on-orbit services.

(3) Real-time planning

Real-time planning is a process of real-time planning and regeneration of the strategy or scheme based on the latest current situation when there is an unforeseen emergency situation or a large deviation from the original strategy in the on-orbit servicing operation. Real-time planning focuses on the ability to respond to emergencies or policy deviations, and strives to obtain the optimal strategy in the shortest time, so as to minimize the execution risk while ensuring that the task can continue.

As an effective way to deal with emergencies, real-time planning should actively respond to all adverse effects caused by emergencies on the basis of obtaining real-time situation, based on pre-planning and temporary planning schemes, scientific decision-making, rapid planning, and effectively deal with mission planning problems under the characteristics of unaware development trend, non-communication at the information level, and incomplete prior knowledge, so as to ensure the smooth completion of the entire on-orbit servicing task. Whether real-time planning is feasible is directly related to the results of on-orbit services and affects the completion and effect of task execution.

2.1.2.2: Task planning process

The on-orbit servicing task planning is mainly based on the task list and the current state of the target formulated according to the overall service intent, combined with the target expectation of the task, under the premise of satisfying the constraints of various conditions and ensuring the effectiveness and availability of resources, effectively dispatching various on-orbit resources, carrying out task planning for each stage, outputting the executable optimal strategy or scheme, and evaluating and optimizing the task completion and execution benefits of the on-orbit servicing, so as to test the rationality of the results obtained in the planning. Figure 2-3 shows the task planning process.

(1) Elements of mission information

The input of on-orbit servicing mission planning mainly includes five information elements: first, the resource element, that is, the in-orbit spacecraft directly used in the on-orbit servicing mission, which is characterized by the number of in-orbit, payload type and service capability; the second is the time element, which mainly includes service timing, action speed, duration, etc.; the third is the space element, which mainly includes objective conditions such as physical space, electromagnetic space, and orbital state; fourth, the elements of action, mainly including space maneuvering, orbit avoidance and orbit game; fifth, information elements, mainly including target information, environmental information, and accusation information.

(2) Task objective decomposition

In the on-orbit servicing task planning, it is necessary to integrate and decompose the mission information elements and transform them into tasks at each stage. Among them, the allocation of spacecraft targets in orbit mainly includes target selection, grade evaluation, resource allocation, force scheduling and program formulation. Orbital maneuver path planning mainly includes mechanical calculation, mode selection, orbit design and trajectory optimization. The service implementation action plan mainly includes target locking, orbit approach, action planning and strategy generation.

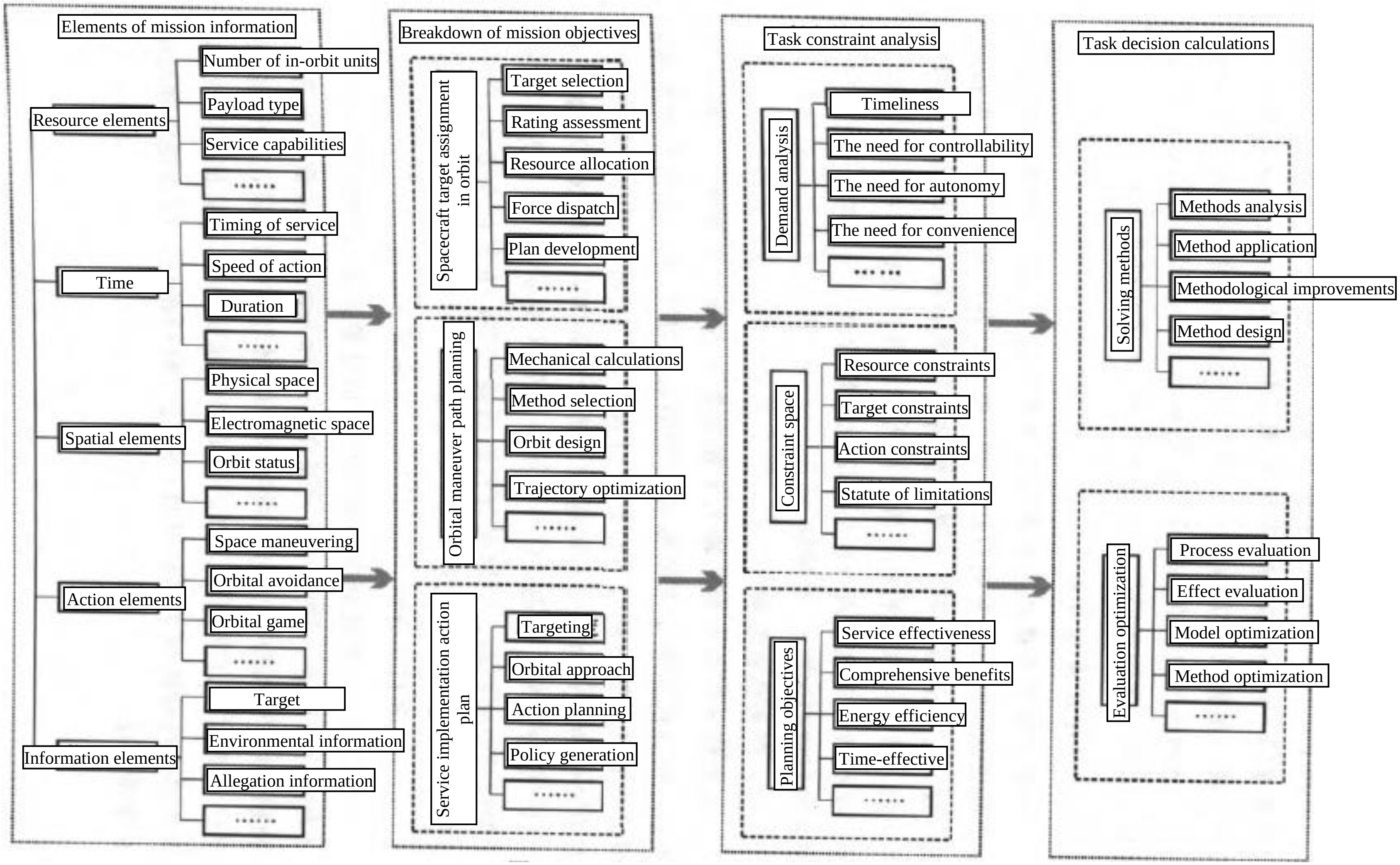


Figure 2-3: On-Orbit Service Task Planning Process

(3) Task constraint analysis

According to the decomposition tasks of each stage, the corresponding demand analysis and constraint combing are carried out, the constraint space in terms of resources, goals, actions and timeliness is clarified, the controllable scope is mastered, and the planning objectives are established from the aspects of service effect, comprehensive benefit, energy consumption efficiency and time benefit.

(4) Task decision calculation

Task decision calculation is the core step in the on-orbit servicing task planning process, which takes the planning goal as the traction, integrates various task constraints according to the mission information elements, and uses or designs the most appropriate planning method to solve and calculate the tasks at each stage. According to the general solution method, it can be divided into problem description, problem modeling, method solution and result evaluation, and finally obtain the executable event list and implementation plan, and provide a feasibility analysis report verified by simulation.

2.2: On-Orbit Servicing Task Planning Demand Analysis

With the continuous expansion of the scope and functions of on-orbit services, the uncertainty and unpredictability of on-orbit servicing tasks have increased, which has brought new problems and challenges to mission planning and given rise to new requirements for intelligent planning applications. Based on the analysis of the overall requirements of on-orbit servicing mission planning, this section focuses on the uneven supply and demand quantity, space debris disturbance, and the outstanding risks of non-cooperative targets, and puts forward three key issues in this book.

2.2.1: Overall Demand Analysis

On the basis of the traditional mission planning requirements, the on-orbit servicing task emphasizes the three requirements of timeliness, constraints and autonomy, which is the general focus of the research on the intelligent planning of on-orbit servicing tasks.

(1) Timeliness requirements

Generally speaking, spacecraft on-orbit servicing and conventional in-orbit movement have different timeliness requirements for their respective mission planning. The timeliness requirements of the conventional in-orbit movement of spacecraft, such as orbit transfer and attitude and orbit adjustment, are not obvious, on the contrary, more attention is paid to fuel saving. In contrast, on-orbit services are more time-sensitive.

First of all, the time, airspace and nature of the spacecraft participating in the service are different each time, the direct purpose, mission and use mode of the service are different, and the space environment is changing rapidly, which makes the service process flow forward rapidly, with different characteristics at different times and different orbit positions, resulting in significant dynamic differences in the on-orbit servicing in different periods. Secondly, the on-orbit servicing is the dynamic interaction between the spacecraft and the service target, which in turn causes the on-orbit servicing to have an obvious and rapid chain reaction. This remarkable dynamic difference and rapid chain reaction of spacecraft on-orbit servicing undoubtedly make its mission planning have obvious timeliness.

The objective requirement of on-orbit servicing, which cannot delay the timing and act relatively quickly, determines that its mission planning must be time-sensitive. With the rapid development of aerospace technology and its gradual application in command and control, the timeliness of mission planning will become more prominent and significant. On the one hand, the traditional mission planning method is formulated by the ground center in advance, and then the operation control system uploads the instructions to the spacecraft in orbit, during which the time span is large and the efficiency is low, which makes it difficult to respond to the current situation in a timely manner, and it is even more difficult to achieve real-time planning. On the other hand, as the speed at which spacecraft performance is improved is significantly faster, on-orbit services can be faster and more effective. According to reports, the US orbiter can maneuver in orbit at more than 20 times the speed of sound and has the ability to provide global service for 2 hours. This requires changing the planning model, improving planning techniques, shortening the time interval from situation discovery to program formulation, and improving the overall timeliness of task planning.

(2) The need for controllability

On-orbit servicing is a special form of conventional on-orbit movement, and mission planning with obvious intent, purpose and value orientation will play a leading role. Unlike conventional orbital transfers and their orbital adjustments, mission planning for on-orbit services always serves the needs of the mission and is implemented in close alignment with purpose and intent. The on-orbit servicing mission planning is manifested in the control and selection of service style, maneuver mode, relative position, fuel consumption and time efficiency in the constraint space. The control and selection of the constraint space of the task plan is determined or restricted by various internal and external factors. From the perspective of internal structure, the control and selection of confined space are mainly restricted by the attributes, performance and state of the spacecraft in orbit. From the perspective of external conditions, the control and selection of the constrained space are subject to the environment and the status of the service target, such as multi-body perturbation, orbital environment, target orbit position, target attributes and flight status. In addition, the form of on-orbit servicing should follow the characteristic laws of space science, orbit dynamics and cybernetics, so mission planning must be obviously controllable.

The controllability of mission planning does not mean that mission planning is a subjective form of on-orbit servicing rather than an objective form or attribute of existence. On the contrary, the controllability of mission planning reflects the dynamic interaction attribute of on-orbit services, so it can be said that the controllability of mission planning is the special manifestation of the objectivity of dynamic interaction on spacecraft. The controllability of mission planning is not arbitrarily chosen according to subjective will, and the objective space situation, space environment, spacecraft performance, target characteristics, etc., determine the control scope and boundaries of the constrained space. Although the control of the constraint space of mission planning cannot exceed the possibility range specified by the objective conditions, the constraint space still has great controllability or selectivity. The application of space forces, the design of maneuvering orbits, and the handling of emergency situations are all realized according to the specific environment and conditions, and the constraints can be transformed into feasible space. Therefore, the success of task planning depends to a considerable extent on the control and selection of constraint space.

(3) The need for autonomy

The autonomy requirement of on-orbit servicing task planning mainly refers to the automatic and autonomous task decision-making and calculation in the planning process. The on-orbit servicing is different from the conventional on-orbit movement, especially the orbital maneuvering and service implementation is more fluid and changing than any other on-orbit behavior, and its mission planning must be quickly adjusted with this change in order to accurately reflect the changing objective situation. Especially in the high-end intensive technical means of orbital service, the precise and fast service action of the behavior mode, its command control and action implementation rhythm is significantly accelerated, the transformation of various service styles and service forms is particularly frequent, the opportunity is fleeting, improve the autonomy of task planning from the task to the implementation of the program, improve the rate of policy generation, will be an effective measure to further improve the service effect.

On-orbit servicing is the product of the high development of aerospace technology, its process is complex, the degree of automation and intelligence is high, and mission planning needs to pay more attention to the requirements of accuracy, timeliness and emergency, not only in strict accordance with the procedures and steps of mission planning, but also from the continuous innovation of the planning system and mode, from the theory, methods and means provided by cutting-edge science and technology, to maximize the automation and autonomy of mission planning.

However, the current ground planning and manual operation methods are difficult to cope with the ever-changing situation of space, and it is difficult to meet the application needs of rapid response and flexible mobility. Therefore, it is necessary to gradually change from ground manual planning to space-based autonomous planning, and develop to intelligent planning, so as to alleviate ground pressure, save operating costs, reduce satellite-ground interaction, improve rapid response and independent decision-making capabilities, and further possess the intelligent planning capability of immediate implementation.

2.2.2: Main Research Questions

The on-orbit servicing mission planning runs through the whole process of on-orbit servicing, which not only covers routine tasks such as target allocation, orbital maneuvering and service implementation, but also needs to focus on some events that are complex, temporary, and difficult to deal with emergencies. In order to better solve all aspects of the on-orbit servicing mission planning problem, the following three difficult problems are analyzed to clarify the key points of this book.

(1) Target allocation in the composite service mode

The pre-planning covers the whole process of on-orbit servicing target allocation, orbital maneuvering and service implementation, and it is necessary to provide scientific and effective target allocation strategies, orbital maneuver paths and service implementation plans for on-orbit servicing actions, which involves many links, wide scope and complex content. Due to the orbital characteristics of the on-orbit servicing, once the in-orbit spacecraft is selected and the service target is determined, the velocity impulse can be calculated and the space transfer orbit can be designed in combination with the orbit change mode, and then the service implementation plan can be formulated according to the type of payload carried by the spacecraft. With the development of aerospace science and technology, the research on spacecraft orbit dynamics, orbital maneuver path design, transfer trajectory optimization and payload utilization has become relatively mature, which can provide effective support for advance planning in orbital maneuvering and service implementation. In contrast, spacecraft on-orbit target allocation, as an important part of prior planning and as a prerequisite for orbital maneuver path planning and service implementation strategy planning, is still in a state of little relevant research and lack of technical support.

At present, spacecraft target allocation in orbit mainly considers the mode of one spacecraft serving one space target (referred to as "one-to-one") or one spacecraft serving multiple space targets in turn (referred to as "one-to-many"). With the continuous expansion of the on-orbit servicing action mode, the spatial amplitude and scope will become larger and larger, the speed will become stronger and stronger, and the service process will become more and more complex. In order to improve the effect and success probability of on-orbit servicing, a composite service model combining one spacecraft serving multiple targets in turn (referred to as "one-to-many") and multiple spacecraft serving one target at the same time (referred to as "many-to-one") has also been gradually adopted.

Compared with the single service model, this composite service model needs to take into account the power input, resource allocation and service effect, and has high requirements for the comprehensive decision-making ability of pre-planning and currently needs to rely on manual assistance. Therefore, breaking through the traditional single allocation principle, adapting to the composite service model, and maximizing the minimum investment of resources and the maximization of service effect are the key issues that need to be considered in the current allocation of spacecraft in orbit.

The allocation of spacecraft on-orbit targets has to face multiple space targets, each of which has different orbital positions, different types, and different service priorities. There are many factors to consider, including resource factors, fuel factors, load factors and timeliness factors; it is necessary to construct multiple goals, namely, the implementation of efficiency goals and energy efficiency goals, etc., and assign one or a set of orderly target tasks to the spacecraft through scientific calculation, so as to optimize the overall efficiency and resource allocation of the on-orbit servicing system on the basis of maximizing the completion of tasks. This process mainly solves the problem of "who serves whom," which is essentially a nonlinear combinatorial optimization problem, which belongs to the non-deterministic (NP) problem of polynomial complexity. Commonly used planning methods include integer programming, genetic algorithm, ant colony optimization and particle swarm optimization, but in practical application, it is found that these methods need to deal with "one-to-one" and "one-to-many" decision-making problems respectively due to algorithm limitations, and the applicability to the target allocation problem in the composite service mode is obviously insufficient. Therefore, it is necessary to adapt to the composite service model presented by the development of on-orbit services, focus on the nonlinear combination optimization problem faced by target allocation, improve the methods and means, and propose an effective target allocation method, so as to effectively control the cost-effectiveness ratio of resource input while improving the comprehensive benefits of target allocation.

(2) Orbit temporary circumvention path planning problem

Interim planning is a strategy adjustment and plan modification to effectively respond to temporary events on the basis of prior planning. The interim plan covers the implementation process of orbital maneuvering and services, and when faced with temporary events such as space debris attacks and mechanical failures, the original strategy or scheme is temporarily adjusted and modified and strives to make the original strategy continuously available with the least modification and minimum cost.

Spacecraft perform on-orbit servicing missions, usually in accordance with predetermined orbital parameters such as orbital pattern, eccentricity, altitude and inclination.

However, the inherent characteristics of in-orbit spacecraft determine their vulnerability, and orbital maneuvering in the space environment is more susceptible to interference and threats. In a series of space regulatory documents, the United States specifically mentioned the incoming threat of spacecraft flight and the dangers of space debris. At present, the accusation method of ground planning and data transmission is time-consuming and delayed, and it is difficult to effectively deal with temporary situations such as space debris attacks. How to effectively improve countermeasures and enhance the active defense capability of spacecraft is a key issue that needs to be considered in the current interim plan.

The orbital maneuver path of a spacecraft is usually planned in advance according to the maneuver taken, and evasion defenses are considered based on the space debris database. However, the space environment is dynamic and unpredictable, and an average of 15% of new space debris per year will pose a huge threat to spacecraft in orbital maneuvering. The United States put forward the concept of collision avoidance in 2018, and pointed out that "collision avoidance refers to the movement of spacecraft to avoid collision with approaching dangerous objects to ensure the safety of astronauts and space payloads." In order to ensure the smooth execution of on-orbit services and prevent the spacecraft from being damaged, it is necessary to temporarily plan the orbital avoidance path and quickly realize the evasion maneuver in order to ensure the smooth execution of the on-orbit servicing and prevent the spacecraft from being damaged. This means that the orbit temporary avoidance path planning should not only avoid space debris but also consider the minimization of fuel consumption and the optimality of orbit offset, and is also limited by the common limitations of avoidance mode, braking time and other factors, which is a multi-restriction shortest path problem. At present, a large number of studies have been carried out at home and abroad on spacecraft rendezvous/interception, orbiting/companion flight, skimming/traversal patrol, etc., compared with it, the research on avoidance maneuver and its path planning also has important application value, and also has complex and interconnected mathematical and physical processes, but only less attention has been paid to it at present.[191] Therefore, as an important on-orbit active defense technology, the path planning method and technology of orbital temporary avoidance need to be further developed, so as to effectively ensure the safety and survival of spacecraft in orbit.

(3) Real-time planning of orbital game strategy

In the process of on-orbit servicing implementation, the spacecraft is very close to the target, but the target behavior is difficult to predict. If it is a non-cooperative target, the time to give the spacecraft a response is very short, and it is difficult to deal with the pre-planning and temporary planning. To ensure that on-orbit servicing tasks can be completed, real-time planning is required to obtain the optimal strategy in the shortest possible time.

At present, most of the on-orbit servicing mission planning is based on the assumption of complete information, that is, assuming that the service target is a cooperative goal, the spacecraft can not only accurately obtain all the target information used for planning but also grasp the in-orbit flight status and behavior of the target.

However, a large number of demonstrative studies and data show that the spacecraft is close to conventional targets in orbit, and such targets are very likely to be non-cooperative targets. As a result, the spacecraft will not be able to continue to perform the original on-orbit servicing mission, and it will even be difficult to accurately obtain the next flight direction of the target, and it will no longer be able to make completely rational decisions, resulting in the need to consider the orbit game with the target in the process of operation. How to use incomplete situation information and limited knowledge and experience to obtain the most reasonable behavior strategy in the current state is a difficult problem to be considered in real-time planning.

From the perspective of spacecraft, the orbit game problem is a dynamic game process in which the optimal behavior is adopted and the space approximation is finally completed under the condition that only the self-state and the current finite state of the target are known, and the future behavior strategy of the target is unknown. In addition to operating in a continuous and dynamically changing space environment, the space target in the orbital game also has the characteristics of typical non-cooperation, that is, no communication at the information level, no cooperation in maneuvering behavior, and incomplete prior knowledge. The spacecraft orbit game is a deep integration of orbit optimal control and dynamic game, which is a typical sequential decision-making process oriented to incomplete information, and its essence is a continuous dynamic interaction problem of bilateral control. Different from the general maximum or minimum control problem, the spacecraft orbit game involves the balance of the behavior strategies of both parties, the problem dimension is doubled, and the information conditions of both parties are considered in real time, the sequential decision game model is complex, and the pre-planning and temporary planning methods are obviously insufficient. At present, the research work related to spacecraft orbit game is still in the exploratory stage, and it is necessary to strengthen the research on the real-time planning direction of orbit game strategy to improve the emergency response capability of spacecraft in orbit service.

2.3: Research Framework of Intelligent Planning for the Whole Process of On-Orbit Servicing

In view of the obvious heterogeneity, multiple constraints, complex space and dynamic variability of on-orbit servicing task planning, this paper explores the use of intelligent methods to construct a mission planning research framework for the whole process of on-orbit servicing, so as to perceive the dynamic changes of tasks and environment more keenly, improve the solution direction more quickly, and better reflect the artificial intelligence ideas of autonomous interaction, self-learning and automatic feedback in the task planning process.

2.3.1: Intelligent Programming Solution

The general solution methods of mission planning problems are usually divided into problem description, problem modeling, method solving and result evaluation, but there may be difficulties in heterogeneity, constraints, spatiality and dynamics when applied to on-orbit servicing mission planning. Therefore, on the basis of giving full play to the advantages of the general solution method, an intelligent programming solution method is explored and proposed.

2.3.1.1: The general way in which the task planning problem is solved

Mission planning problems are often faced in the field of planning and need to be solved, and the general solution of mission planning problems[193] is shown in Figure 2-4.

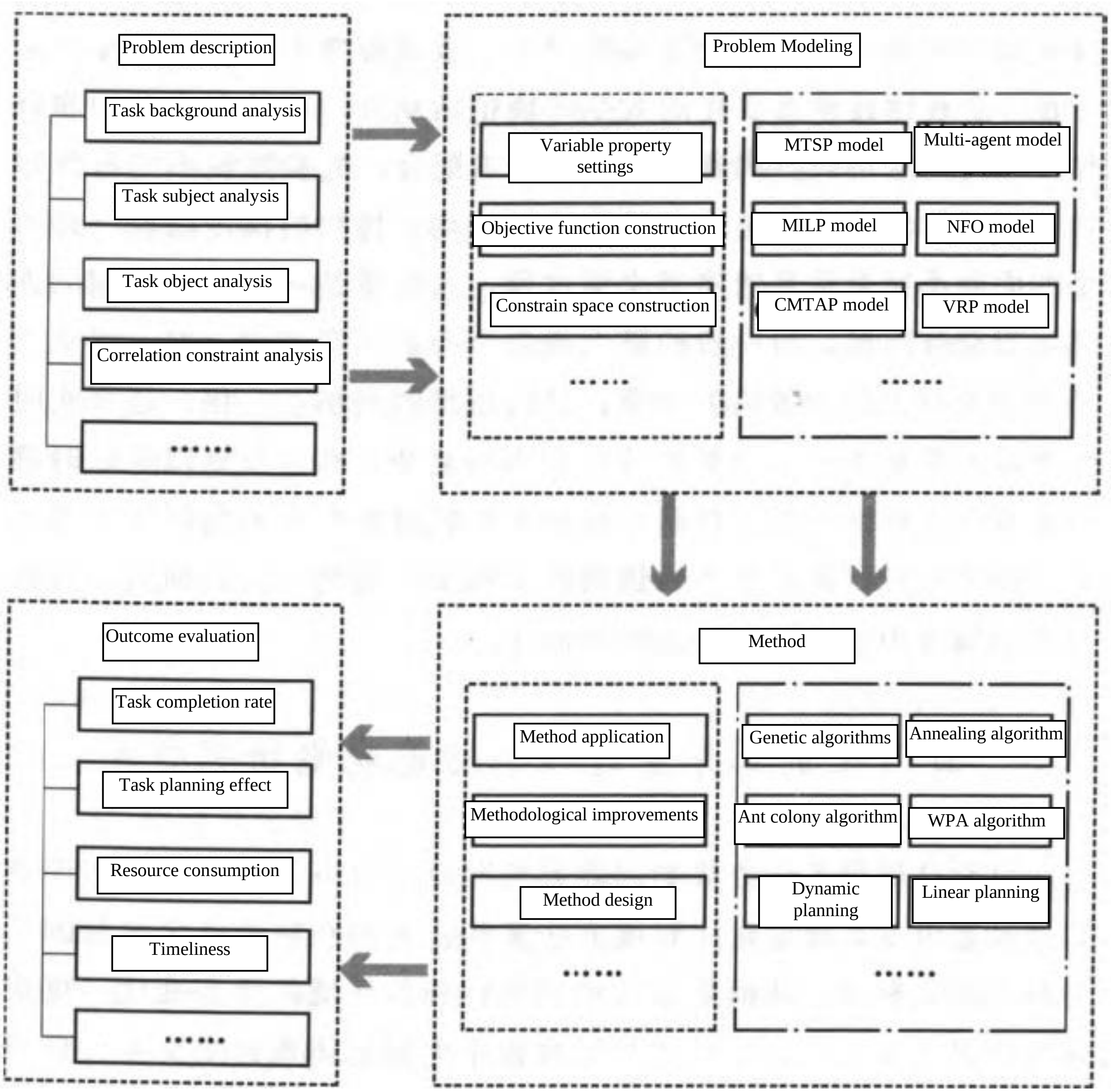


Figure 2-4: A General Approach to Solving a Task Planning Problem

The solution of task planning problems can generally be divided into problem description, problem modeling, method solving and result evaluation. First of all, in the problem description, it is necessary to focus on the analysis of task background, task subject and object analysis, and related constraint analysis. Then, in the process of problem modeling, it is necessary to focus on the setting of variable attributes, the construction of objective functions, and the construction of constraint space. Then, in the method solving, the most appropriate method is mainly used or designed to solve the task planning problem. Finally, in the process of outcome evaluation, it is necessary to effectively evaluate the task planning results from the aspects of task completion rate, task planning effect, resource consumption and timeliness.

The planning of on-orbit servicing missions is likely to encounter outstanding risks such as uneven supply and demand, space debris and non-cooperative targets, and the planning problems are heterogeneous, constrained, complex and dynamic, which adds difficulty to the application of general solution methods.

(1) Obvious isomerism

Heterogeneity reflects the diversity of things, and the diversity of things is one of the necessary factors for complex properties.[193] The heterogeneity of on-orbit servicing mission planning involves multi-objective services, multi-resource allocation, and multi-situation disposal, which adds considerable difficulty to the application of general solution methods.

(2) There are many constraints

For on-orbit servicing mission planning, in addition to various constraints that need to be considered in the process of problem analysis, spacecraft maneuvering capability constraints, payload service capability constraints, on-orbit resource constraints, target state constraints, and space environment constraints should be studied in the modeling process, resulting in many and complex relationships between solving constraints, and some of them have coupling relationships with each other, which greatly increases the difficulty of mathematical modeling.

(3) The space is complex

On-orbit servicing task planning may involve multiple targets at different orbit positions, tasks in different situations, and a variety of different types of service modes with certain timing or other coupling associations between different stages, resulting in a more complex solution space. At the same time, there are many factors that affect the performance of the task planning problem solving method, which may cause a sudden change in the effect and even lead to the interruption of the method solution due to some small changes. All these make the degree of nonlinearization of task planning increase, the diversity of decision optimization is stronger, and the difficulty and complexity of solution method design will become higher.

(4) Dynamically variable

The uncertainty of mission planning and implementation of on-orbit services, such as orbital maneuvering and service implementation, has a certain time series relationship, and the space environment is difficult to fully grasp, which greatly increases the uncertainty of mission planning and implementation. Environmental variables may change at any time, the target state may change at any time, and emergencies may occur at any time, which directly leads to the dynamic changeability of task planning problems, and it is necessary to update tasks in time and carry out temporary planning or real-time planning, which puts forward higher requirements for the adaptability of task planning solutions.

2.3.1.2: Intelligent planning method for task planning problems

The general solution method refers to the one-way solution process from problem description to problem modeling to problem solving. Of course, the results can be obtained in this way, but the progressive characteristics are obvious and the process is relatively fragmented, for example, it is difficult to weigh and apply to dynamic and reciprocating planning problems such as orbital games.

Intelligent planning mainly refers to the use of relevant intelligent technologies to solve several task planning problems, which is the application of artificial intelligence in the field of mission planning, which will be able to give full play to the advantages of artificial intelligence technology in terms of independent learning, self-optimization, easy use, easy convergence and strong adaptability, and more effectively deal with the difficulties of heterogeneous and complex space and dynamic variability in the task planning problems of on-orbit services.

In intelligent planning that interacts with the environment in real time, the machine that learns and implements decisions autonomously is usually called an agent, everything that interacts with it outside the agent is collectively called the environment, and any action or decision that wants to be done is called an act.[194] Based on this, Table 2-1 establishes the correspondence between the elements of the intelligent planning problem and the intelligent planning method of the on-orbit mission.

Table 2-1: Correspondence between the elements of the intelligent planning problem and the intelligent planning method of the in-orbit task

Elements of the On-Orbit Servicing Mission Planning Problem	Elements of the Smart Planning Methodology
Spacecraft in orbit	Agents
State of space	Environment
Tactics	Behavior
Effect returns	Reward

Continued

Elements of the On-Orbit Servicing Mission Planning Problem, continued	Elements of the smart planning methodology
Mission planning process	Autonomous training process
The best strategy for task planning	Consistently reward the greatest behaviors

In intelligent planning, whether it is the core "feedback" mechanism of deep neural networks, or the "execution-reward" process embodied in reinforcement learning, and the idea of "experience playback," they all pay attention to the dynamic interaction with the environment and the response to the target. Figure 2-5 shows the correspondence between the general solution method and the intelligent programming solution. In intelligent programming, the problem description in the general solution method will be replaced by environmental perception and sensing interaction. Problem modeling will be characterized by interaction and learning; the method solution will consist of self-learning, update evolution, and automatic feedback. The evaluation of the results will be reflected by a feedback mechanism and a network update.

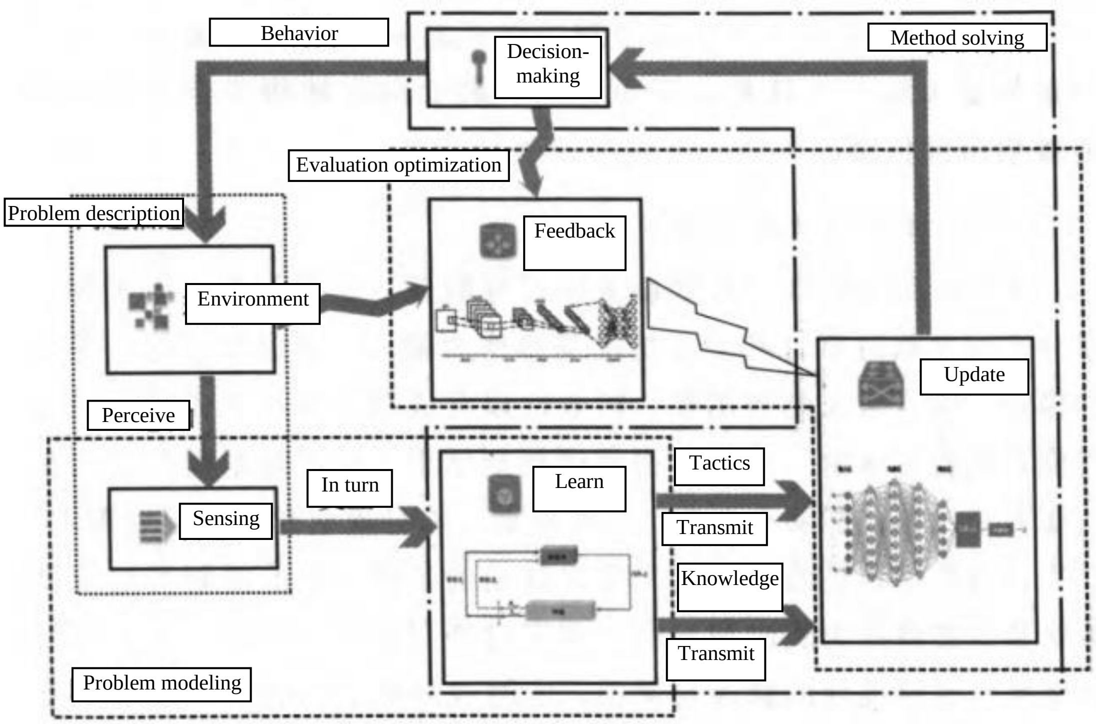


Figure 2-5: Correspondence Between the General Solution Method and the Intelligent Programming Solution

This intelligent planning method embodies the ideas of autonomous interaction, self-learning, automatic feedback, and update and evolution in the task planning process. Among them, autonomous interaction enhances the ability to obtain effective information, and can more effectively deal with various incomplete, incomplete and unstructured spatial situations.

Self-learning is based on the interaction with the current situation and the ability to make independent decisions with intelligent methods. Automatic feedback is to feed back the effect of decision-making behavior in the environment to the agent through the reward mechanism, which helps to improve the ability of overall decision-making and emergency response, and is an important part of forming an efficient, safe and reliable solution. Renewal evolution is based on the accumulation of experience and trial and error mechanism to form an effective intelligent evolution paradigm, which is the foundation and core of realizing the level of brain-like intelligence.

2.3.2: Research Framework for Intelligent Planning

With the continuous expansion of aerospace technology and on-orbit servicing means, the on-orbit servicing mission planning will also change: first, the scope of mission planning will be expanded from pre-planning to on-orbit interim planning and real-time planning; second, the service model has expanded from a single model of "one-to-one" or "one-to-many" to a composite service model combining "one-to-many" and "many-to-one"; third, the planning process has expanded from a single-objective element to a multi-objective all-element; fourth, the planning method has been expanded from manual assistance to independent planning. In order to further adapt to these changes and requirements and give full play to the advantages of intelligent programming solutions, an intelligent planning research framework for the whole process of on-orbit servicing is introduced below.

2.3.2.1: Intelligent planning framework for typical problems

The intelligent planning framework takes the number of targets, the number of orbit roots, the distribution of spacecraft forces, and the information of temporary events and emergencies in the on-orbit servicing as inputs, and through the corresponding mission planning module, comprehensively considers the constraints such as fuel consumption, timeliness and success rate, and solves it independently through intelligent algorithms, so as to obtain the optimal planning results. The intelligent planning framework is oriented to the whole process of on-orbit services, and can effectively meet the needs of pre-planning, ad hoc planning, and real-time planning. Among them, ex-ante planning is the basis of interim planning and real-time planning, and the strategies and schemes obtained from the ex-ante planning framework will be used as input to the interim planning framework and the real-time planning framework, the interim planning framework is the modification and adjustment of the ex-ante planning, and the real-time planning framework is the expansion and improvement of the ex-ante planning and interim planning.

Combined with the main problems of this book, this paper constructs a research framework for the whole process of on-orbit servicing by taking the problem of in-orbit target allocation, the problem of spacecraft orbit temporary avoidance path planning and the real-time planning of spacecraft orbit game strategy in the composite service mode as examples (see Figure 2-6).

(1) On-orbit target allocation under composite service mode

Target allocation is the primary planning problem faced in pre-planning, and it is an important part of on-orbit servicing task planning. In order to improve the success probability of on-orbit servicing and improve the overall service efficiency, spacecraft on-orbit target allocation should not only meet the needs of the composite service model, but also comprehensively consider the implementation benefit and energy efficiency, and take into account factors such as fuel consumption, timeliness and robustness, so as to better select service targets and determine the service sequence for the in-orbit spacecraft.

In this part, the execution benefit and energy efficiency are taken as the optimal allocation objectives, the number of targets, the number of orbits, and the distribution of spacecraft forces are taken as inputs, and the allocation constraints, fuel consumption constraints and timeliness constraints are comprehensively considered, and the balanced development of execution efficiency and energy efficiency is realized through the on-orbit target allocation algorithm under the composite service mode, so as to obtain the optimal allocation strategy under the composite service mode.

(2) Spacecraft orbit temporary avoidance path planning

The problem of temporary orbital avoidance may occur not only during the orbital maneuver of spacecraft, but also during the implementation of service, which is a typical problem to be dealt with in the temporary planning. In the face of space debris attack during the on-orbit servicing operation, although the original plan is difficult to sustain, there is still a certain amount of decision-making time, so it is necessary to temporarily carry out path planning and take orbital maneuver in time.

In this part, orbit dynamics and celestial mechanics are taken as the technical basis, the target allocation strategy and space transfer orbit obtained from prior planning are taken as prerequisites, the detected space debris state information is taken as input, and the factors such as fuel consumption, minimum offset, and braking time are comprehensively considered, and the space avoidance motion is simply expressed with the help of the Frenet coordinate system, and the optimal avoidance path is independently generated through the artificial potential field of spacecraft and space debris to meet different avoidance needs and preferences.

(3) Real-time planning of spacecraft orbit game strategy

Non-cooperative objectives are the core issues that have to be considered in the planning of actual on-orbit servicing missions, and they are also special events that are difficult to fully consider in advance planning. The real-time planning of spacecraft orbit game strategy has a very short reserved decision-making time, and it is necessary to respond quickly to the behavior of non-cooperative targets in a timely manner, so it needs to be dealt with in a more time-sensitive real-time planning way

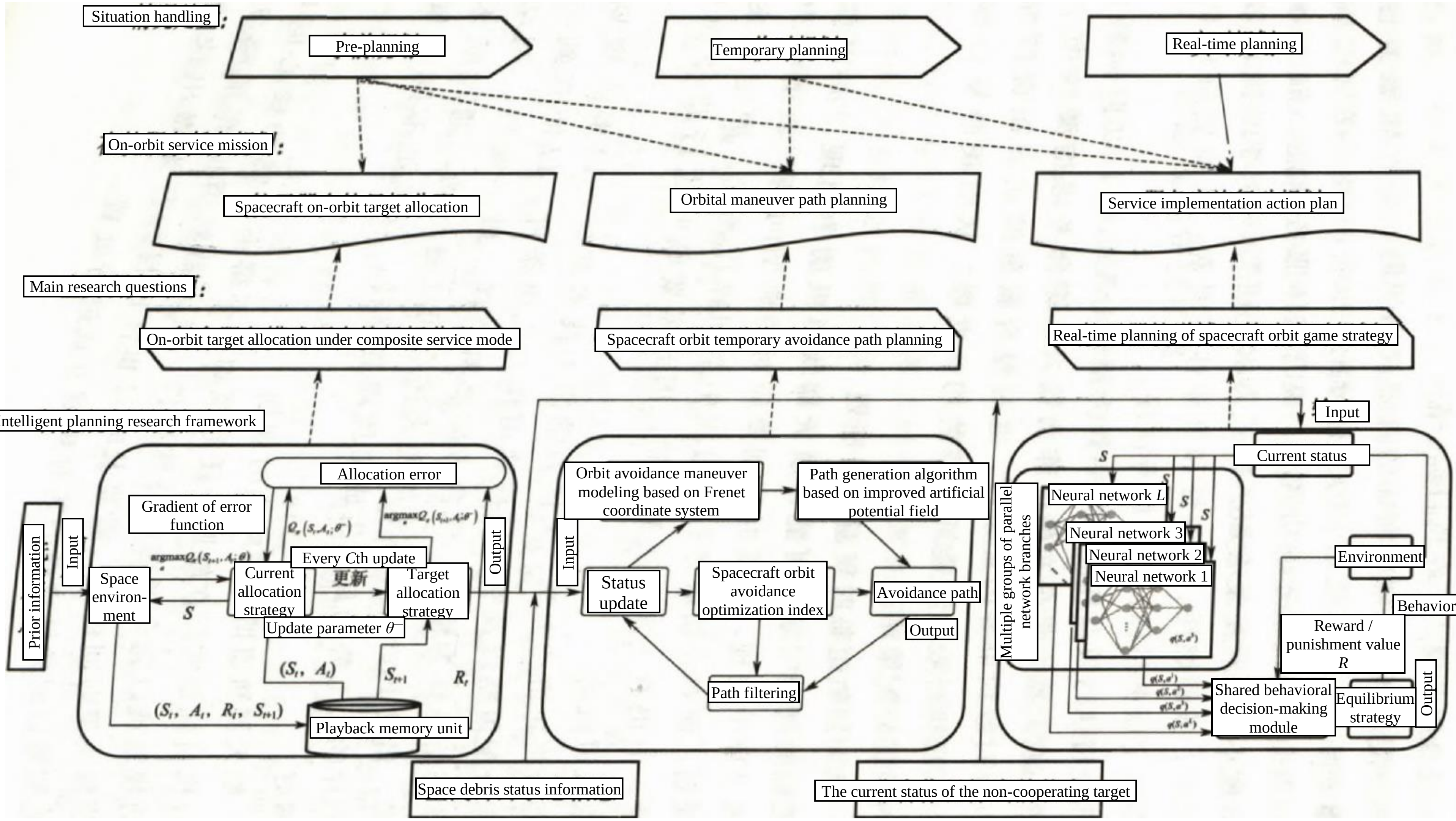


Figure 2-6: A Research Framework for Intelligent Planning for the Whole Process of On-Orbit Service

In this part, the target allocation strategy and space transfer orbit obtained from pre-planning and temporary planning are prerequisites, and only the current finite state and previous behavior of the non-cooperative target are taken as inputs, and the relative motion between the spacecraft and the non-cooperative target is described with the help of differential countermeasures under the condition that the future behavior strategy of the non-cooperative target is unknown, and the game theory and control theory are combined to pass through a branched deep reinforcement learning architecture with multiple sets of parallel neural networks and shared decision-making modules. In this way, the equilibrium strategy of orbit game can be obtained for spacecraft in real time.

2.3.2.2: Analysis of the advantages of intelligent planning

The significance of intelligent planning in on-orbit servicing tasks lies in the fact that it meets the needs of on-orbit servicing task planning, is suitable for the general solution method of mission planning and can overcome the shortcomings of the general solution method. Keeping up with the future development trend of on-orbit services, the application of intelligent planning can bring the following advantages to on-orbit servicing task planning:

(1) Further meet the needs of each index

From the perspective of on-orbit servicing index requirements, the task planning focuses on timeliness, constraints and autonomy, and needs to have more effective on-orbit allocation strategies, more flexible orbit maneuvering capabilities, and more diverse service implementation methods, which undoubtedly improves the planning accuracy of problem solving, increases the difficulty of decision-making, and puts forward higher requirements for the overall planning level. Facing the planning needs of on-orbit servicing tasks, the intelligent planning technology with a higher planning level and faster decision-making mode can achieve the solution effect that cannot be achieved by the general solution method.

(2) Further integration of the model

Although the heterogeneity and complexity of the problem description improve the applicability of mission planning to the actual on-orbit servicing, it also increases the difficulty of the subsequent links of problem solving to a large extent. In the general solution method, the problem modeling, method solving, evaluation and optimization and other links are often carried out separately, and the consideration situation is relatively simple, but now the diverse task requirements, huge constraint space and uncertain developments require the on-orbit servicing to accept a diversified mission pattern and integrated planning mode. Intelligent planning uses suitable problem modeling methods, advanced intelligent methods and independent decision-making methods to closely connect all links and realize the integration of planning models, which can further solve the difficulties existing in general solution methods and improve the overall efficiency and robustness of planning.

(3) Further improve the level of intelligence

Through real-time perception and interaction with the environment, intelligent planning can quickly and fully understand the overall information, self-training, self-learning, and obtain the optimal strategy under the condition of incomplete information, which can effectively deal with the difficulties of on-orbit servicing mission planning, such as complex space and dynamic changeability, and greatly improve the intelligent level of mission planning. At the same time, intelligent planning can also well cope with the uncertainty of target behavior and the dynamic change of tasks, improve the effect of dynamic planning and autonomous decision-making in a targeted manner, and play a vital role in improving the automatic planning ability, autonomous decision-making ability and task execution effect of on-orbit services as a whole.

2.4: CHAPTER SUMMARY

According to the research idea of "problem overview-demand analysis-framework construction," firstly, the relevant concepts of intelligent planning of on-orbit servicing tasks were defined, and the planning process of on-orbit servicing tasks was studied from three perspectives: pre-engagement, temporary and real-time. Second, the demand analysis of on-orbit servicing mission planning is carried out from the aspects of timeliness, controllability and autonomy, and it is pointed out that mission planning will face the difficult problems of "uneven supply and demand," "space debris attack" and "non-cooperative target service." The third is to innovate the task planning method of the whole process of on-orbit servicing, replace the problem description in the task planning with environmental perception and sensing interaction, the problem modeling is characterized by interaction and learning, the method solution is composed of self-learning, update evolution and automatic feedback, and the result evaluation is reflected by the feedback mechanism and network update, and a research framework for intelligent planning for the whole process of on-orbit servicing is proposed, so as to perceive the dynamic changes of tasks and environment more keenly and improve the solution direction more quickly. This paper lays a foundation for the formation of intelligent planning methods for on-orbit servicing tasks.

Under the condition of uneven supply and demand, in the face of multiple service objectives, the use of limited space forces to serve the target is the most prominent in the target allocation process, that is, in the target allocation stage, the allocation constraints, fuel consumption constraints and timeliness constraints are fully considered, and a robust allocation strategy is generated, so as to minimize the possibility of strategy readjustment in the implementation process. The application requirements of the composite service mode should not only treat each service objective differently but also take into account the balanced development of implementation benefit and energy efficiency, which leads to the non-deterministic (NP-hard) characteristics of polynomial complexity of spacecraft on-orbit target allocation and faces a complex nonlinear combination optimization problem.[73]

3.1: TARGET ASSIGNMENT PROBLEM DESCRIPTION AND MODELING

This paper describes the on-orbit target allocation process under the composite service mode, focuses on its unique requirements of the composite service mode and the robustness requirements of the scheme, and points out the nonlinear and complex characteristics of the on-orbit target allocation problem under the composite service mode. The on-orbit target allocation problem in the composite service mode is transformed into a nonlinear combinatorial optimization problem, and the objective function of the problem optimization is established, and the constraint space for problem solving is constructed.

3.1.1: Problem Description

According to statistics, as of January 1, 2021, there were 3,372 satellites in orbit around the world, and more and more satellites in orbit are conducive to the development of aerospace science and technology and may also form more demand for on-orbit services. Taking communication satellites as an example, at present, some countries have built a four-in-one communication satellite system of "broadband, narrowband, protected, and relay," which has high-speed and large-capacity trunk communication, node communication, and high-speed user access and other communication means, and the communication frequency bands cover ultra-high frequency, ultra-high frequency and extremely high frequency, which can provide large-capacity, timely and stable communication services.

However, these communications satellites, with an investment of hundreds of millions or even billions of dollars, are limited by factors such as the fuel they carry and the aging of their components, and their lives are extremely fragile, and it may only take one accident to end their service life. At the same time, although the on-orbit servicing technology is constantly developing, and more and more on-orbit servicing spacecraft are put into use, compared with the huge on-orbit servicing demand, there are still outstanding problems such as "many objects, scattered and limited on-orbit servicing force."

At present, due to the limitation of the number of spacecraft in orbit, the service mode of "one spacecraft serving one space target" or "one spacecraft serving multiple space targets in turn" is usually adopted.[29, 74] However, the service capability of a single spacecraft is limited after all, and it is difficult to meet the practical requirements of a high probability of service success. With the continuous development of on-orbit servicing technology and the increase of in-orbit spacecraft, multiple spacecraft will have the ability to serve a space target at the same time, so that the composite service mode combining "one-to-many" and "many-to-one" will be gradually adopted.

The composite service model is a hybrid model that comprehensively considers the input and benefits of spacecraft for many different types, different orbits and different priorities, and adopts the "one-to-many" and "many-to-one" allocation strategies. Compared with the single service model, this method needs to take into account the amount of spacecraft input and service effect and has higher requirements for the comprehensive decision-making ability of the allocation method and model. In addition, in the uncertain environment, in the face of various unexpected situations in the process of strategy implementation, the optimality and feasibility of the strategy are often contradictory, which has high requirements for the robustness of target allocation. Therefore, the spacecraft on-orbit target allocation process involves the interaction of multiple factors such as multiple objectives, multiple constraints, incomplete information and uncertainty, which leads to the diversified characteristics of the target allocation process, which creates the nonlinearity and complexity of the on-orbit target allocation problem under the composite service mode.

The use of multiple spacecraft to serve multiple space targets is different from the general orbital mission, which is the process of formulating the optimal target allocation strategy through a comprehensive evaluation of the target priority and the cost of spacecraft to each target to complete the on-orbit servicing mission. This contains three meanings: purpose, process and attributes: first, the on-orbit servicing target allocation is the purposeful, limited, and goal-oriented active behavior of the spacecraft; secondly, the allocation of on-orbit servicing targets is a balancing process, which is a process of unifying optimality and feasibility in the face of the adverse effects of various constraints. Finally, the allocation of on-orbit servicing targets is to change the original state of the spacecraft in orbit, that is, to break the existing inertial flight of the spacecraft and even no longer obey Kepler's law, so that the factors and conditions considered are different from usual.

The allocation of spacecraft on-orbit targets needs to consider multiple service objectives, multiple types of constraints, and multiple optimization purposes, which can be regarded as a nonlinear combination optimization problem, and the following points need to be emphasized:

1) Clarify the requirements such as task execution priority, timing relationship, and logical relationship.

2) Grasp the prior information such as service target attributes, orbit position, status, priority, and service success probability.

3) Fully consider factors such as the number and status of spacecraft in orbit and fuel stock.

4) Pay attention to the constraints of designation, robustness, timeliness and fuel consumption.

5) Formulate planning goals such as implementation benefits and energy efficiency.

3.1.2: Target Allocation Modeling

Considering the reality of uneven supply and demand, the target allocation problem is studied for the demand of the composite service model, so as to provide effective auxiliary decision-making for the allocation of spacecraft in orbit. The essence of the research on on-orbit target allocation under the composite service model refers to a scientific paradigm that uses mathematical theory and scientific computing technology to analyze, describe and plan the quantitative aspects of on-orbit target allocation, and grasps the essence and law of target allocation from the relationship between quality and quantity. However, in order to highlight the focus of planning and reveal the essence of the problem, in order to highlight the focus of planning and reveal the essence of the problem, it is necessary to use highly abstract mathematical models to conduct research in the process of mathematical modeling and scientific calculation, and reasonable assumptions need to be made about some related conditions:

1) It is assumed that before the execution of the target assignment task, a certain prior knowledge related to the target is mastered, and the relevant statistical analysis is carried out, so that the prior information such as fuel consumption estimation, target priority, and service success probability can be directly used for the data input of the task allocation model.

2) In the process of spacecraft on-orbit target allocation, it is assumed that the spacecraft has strong space situational awareness, timely command and control, advanced power system and strong on-orbit computing capabilities.

On the basis of the problem description and modeling assumptions, the on-orbit target allocation problem in the composite service mode is transformed into a nonlinear combinatorial optimization problem, the objective function of the problem is established with the goals of execution benefit and energy consumption efficiency, and the constraint space for problem solving is constructed considering the factors such as assignment, fuel consumption, timeliness and robustness.

3.1.2.1: Objective function

Objective function assumptions: n space service targets, denoted as E_1, E_2,...,E_n; spacecraft in orbit should be counted as P_1, P_2,...,P_m; adopt a composite service model, that is, each spacecraft can serve multiple targets in a certain order, and each target will receive services and multiple spacecraft can be assigned to serve; the service range of all payloads of the spacecraft overlaps with each other, i.e., each payload can be assigned to any target.

In pursuit of the balanced development of implementation benefit and energy efficiency as the optimization index, the following objective functions are constructed

$$\max G = \sum_{j=1}^{n} C_j H_j \tag{3-1}$$

where G assigns the total target to the spacecraft's in-orbit target; n is the number of space targets; C_j is the implementation benefit of the target allocation; H_j is the energy efficiency assigned to the target.

Specifically, in the target allocation, it is necessary to ensure that the implementation efficiency is maximized, and the energy efficiency is also maximized.

3.1.2.2: Constraint space

In the process of spacecraft on-orbit target allocation, the constraint space for problem solving is constructed by comprehensively considering the constraints such as assignment, burn-up, timeliness and robustness.

(1) Assign constraints

In order to meet the requirements of the composite service model, each spacecraft can serve multiple targets in turn, and each target can be served by multiple spacecraft, i.e.

$$\begin{cases} PA_j = 1 - \prod_{i=1}^{n} [1 - PA_{ij} X_{ij}] \geqslant PA'_j \\ \sum_{i=1}^{n} X_{ij} \geqslant 1 \\ X_{ij} = 0 \ \boxed{\text{OR}} \ 1 \end{cases} \tag{3-2}$$

where PA_{ij} is the success probability of the ith spacecraft serving the jth space target; PA_j is the cumulative service success probability of spacecraft resources to the jth space target, and its value should meet the service success probability index, and each space target has a corresponding spacecraft and multiple spacecraft can be assigned for service;

PA'_j is the success probability index of the spacecraft service to the j-th space target, and its numerical setting is related to two factors, namely, the priority and number of space targets, and the number and attributes of spacecraft in orbit. X_{ij} is the allocation decision variable, expressed as a Boolean value, i.e.

$$X_{ij} = \begin{cases} 1, \text{ The } i\text{th spacecraft was assigned to the } j\text{th space target} \\ 0, \text{ Other} \end{cases} \quad (3\text{-}3)$$

(2) Fuel consumption constraints

In order for the spacecraft to have freely adjustable fuel resources for each service mission, the remaining available fuel should be greater than or equal to 0 after removing the fuel required for each orbital maneuver and the fuel required for real-time service, i.e.

$$\Delta M_i = M_i^{\text{total}} - M_i^{\text{maneuver}} - M_i^{\text{implement}} \geq 0 \quad (3\text{-}4)$$

where $\triangle M_i$ is the remaining amount of fuel for the ith spacecraft; M_i^{total} is the total amount of fuel for the ith spacecraft; M_i^{manual} is the fuel consumption of the ith spacecraft orbital maneuver; $M_i^{\text{implement}}$ is the fuel consumption for the implementation of the ith spacecraft.

(3) Timeliness constraints

In order to make all on-orbit servicing tasks be completed within the executable time interval and meet certain timeliness requirements, it is necessary to limit the time efficiency, that is

$$\begin{cases} T_{k(l)}^{\text{StartTime}} \geqslant T_{k(l)}^{\text{EarlyTime}} \\ T_{k(l)}^{\text{StartTime}} + T_{k(l)}^{\text{Dur}} \leqslant T_{k(l)}^{\text{LateEndTime}} \end{cases} \quad (3\text{-}5)$$

where $T_{k(l)}^{\text{StartTime}}$ is the start time of phase l in task k; $T_{k(l)}^{\text{EarlyTime}}$ is the earliest start time of task k; $T_{k(l)}^{\text{Dur}}$ is the execution duration of phase l in task k; $T_{k(l)}^{\text{LateEndTime}}$ is the latest end time of task k.

(4) Robustness constraints

Robustness is an important concept in modern control theory and system science, which generally refers to the ability to maintain the normal implementation of the scheme in case of emergencies. The robust constraints of target allocation emphasize the ability to improve the ability of the scheme to "resist" or "absorb" various disturbances and reduce the possibility of readjustment during the implementation of the scheme. For the problem of spacecraft target allocation in orbit, the mass robustness included in robustness is as important as the scheme robustness,[204] therefore, the robustness constraints are divided into two aspects: mass robustness constraints and scheme robustness constraints.

The quality robustness constraint is to constrain the performance index of spacecraft target allocation in orbit, which requires that in case of emergency, one or some performance indicators of the scheme deviate from the performance value of the established scheme not much.[204] Mass robustness is defined as the maximization of the probability of completion of an on-orbit servicing task within a given time period, with the constraint that the probability of completion of the scheme within a given time period is greater than the minimum standard, i.e.

$$P\left(T_{k(l)}^{\mathrm{Dur}} \leq T_{k(l)}^{\mathrm{Dur\_limit}}\right) \geq \bar{P}_{\min} \tag{3-6}$$

where $P(\cdot)$ is the probability function of scheme completion; $\bar{P}_{\min}$ is the minimum standard value of the probability of program completion. $T_{k(l)}^{\mathrm{Dur\_limit}}$ is the time limit for the completion of the Phase l task in Task k.

The robustness constraint of the scheme is to constrain the stability of the scheme itself, and it is required that the deviation from the implementation of the given scheme after an emergency is as small as possible. The scheme robustness is defined as the weighting and minimization of the deviation of the actual execution time of the task from the time of the scheme and the weighting of the task, and its constraint is that the weighted sum of the deviation between the actual execution time and the planned time of the program in the process of program execution needs to be less than the minimum standard, that is

$$\sum_{l=1}^{N} \omega_{k(l)} \cdot E\left| T_{k(l)}^{\mathrm{Dur}} - T_{k(l)}^{\mathrm{Dur\_planning}} \right| \leqslant \Delta T_{\min} \tag{3-7}$$

where E|•|is the desired function; ΔT_{min} is the minimum standard value of the deviation of the program execution time. $\omega_{k(j)}$ is the weighting factor, and $\omega_{k(l)} \in (0,1)$; $T_{k(l)}^{\mathrm{Dur\_planning}}$ is the planned completion time of the lth phase in task k.

3.2: On-Orbit Servicing Target Allocation Index Model

In this section, we will consider and estimate the implementation effect of the target allocation stage in the subsequent orbital maneuvering and service implementation process and focus on the calculation model with "execution benefit" and "energy efficiency" as the indicators, as shown in Figure 3-1.

3.2.1: Implement Efficiency Indicators

The implementation efficiency of spacecraft on-orbit target allocation should be reflected in the subsequent orbital maneuvering stage and service implementation stage. Therefore, the estimation index of the implementation benefit is decomposed into two parts: the implementation benefit in the orbital maneuvering stage and the implementation benefit in the service implementation stage, namely

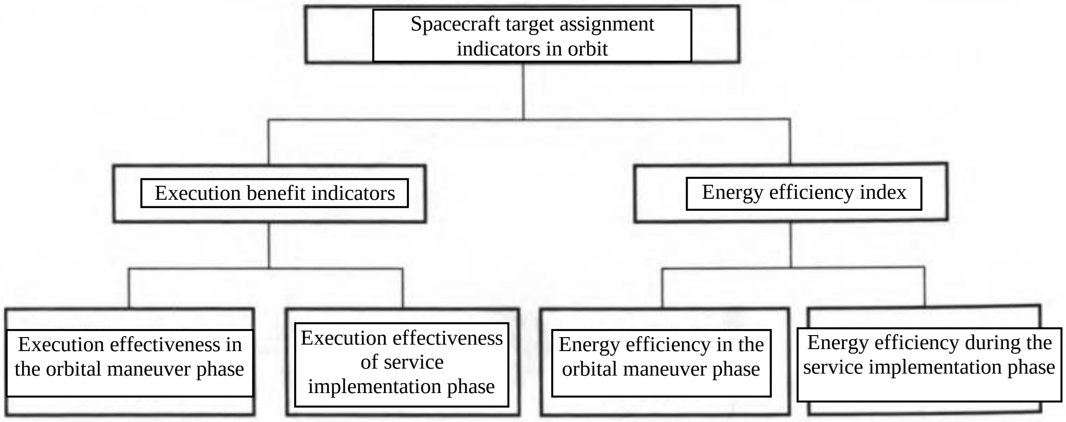


Figure 3-1: On-Orbit Service Target Allocation Index System

$$C_{ij} = C_{ij}^{\text{maneuver}} + C_{ij}^{\text{counter}} \qquad (3\text{-}8)$$

where C_{ij} is the execution benefit index of spacecraft in-orbit target allocation; C_{ij}^{maneuver} is the estimated execution benefit of the allocation plan during the orbital maneuvering phase; C_{ij}^{counter} is the estimated execution benefit of the distribution plan during the service implementation phase.

3.2.1.1: Execution benefits during the orbital maneuvering phase

Orbital maneuvering is the control of the spacecraft to change the original flight orbit and enter another orbit under the action of the control system, which is a necessary process to complete the on-orbit servicing mission. According to the different types and action times of spacecraft thrusters, orbital maneuvering methods can be divided into three categories: pulse thrust, finite thrust and small thrust. In the actual flight mission, orbital maneuvering is a very complex process that takes into account many factors such as fuel consumption, maneuver timeliness, condition constraints, and ground tracking and control, but its basic principle is similar to that of two-impulse orbital maneuvering,[201] Therefore, this section will take impulse thrust orbital maneuvering as an example to analyze the performance benefits in coplanar and non-coplanar orbits.

(1) Coplanar orbital maneuvering

Coplanar orbital maneuvering means that the initial orbit and the target orbit are in the same orbital plane, and the spacecraft does not need to consider the change of orbital inclination between the two orbits during orbital maneuvering.[201] In general, coplanar orbital maneuvers can be further subdivided into common orbital maneuvers and hetero-orbital maneuvers.

Common orbital maneuver (see Figure 3-2) refers to the process in which the spacecraft and the space target are located in the same orbit plane and altitude, and the spacecraft performs orbital maneuver to reach the target orbital position after several circles around the earth.

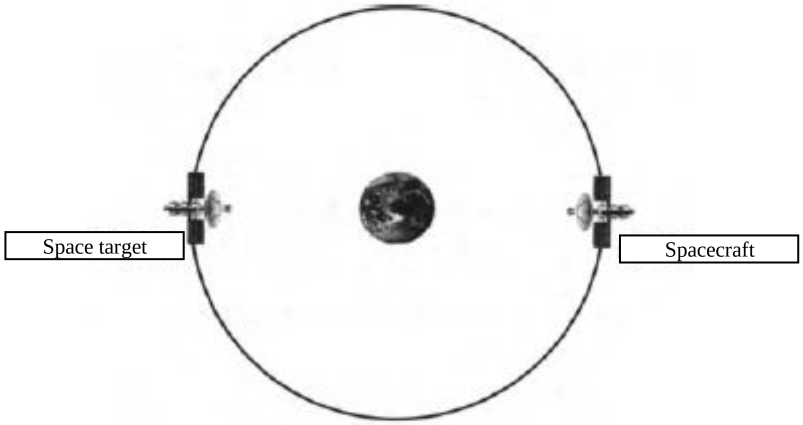


Figure 3-2: Schematic Diagram of Common Orbit Maneuver

A typical double-pulse fixed-time multiturn solution can take a multiturn orbit with zero transfer angle, but the transfer angle of maneuvering flight is a multiple of 360°.[195] In Figure 3-2, the initial lead angle of the space target relative to the spacecraft is β = 180°, and the time for the target to first reach the orbital position shown in Figure 3-2 is half of the known orbital period of the target, and the subsequent time for the target to reach the same position can be given by adding the multiple of the target's orbital period. Therefore, the total flight time of the target to the specified multi-turn is known, and this time is also the time of the spacecraft's multi-turn orbital maneuver at zero transfer angle

$$\Delta t = (360 n_{\mathrm{T}} + \theta - \beta) / \omega_T \quad (3\text{-}9)$$

where Δt is the orbital maneuver time of the spacecraft; the number of full orbits of the n_T space target; θ($0° \leq \theta < 360°$) is the transfer angle of the spacecraft; β is the initial lead angle of the space target relative to the spacecraft; ω_{T} is the average angular velocity of the target orbit.

Off-orbital maneuvering (see Figure 3-3) is the process in which a spacecraft initially lies on another orbit different from the orbit of a space target and reaches the vicinity of the target orbit with two or more pulses within a specified time.[195]

For indefinite time hetero-orbit maneuvers, theoretically speaking, the impulse thrust of the two impulses can realize the transfer from the initial orbit to any target orbit.

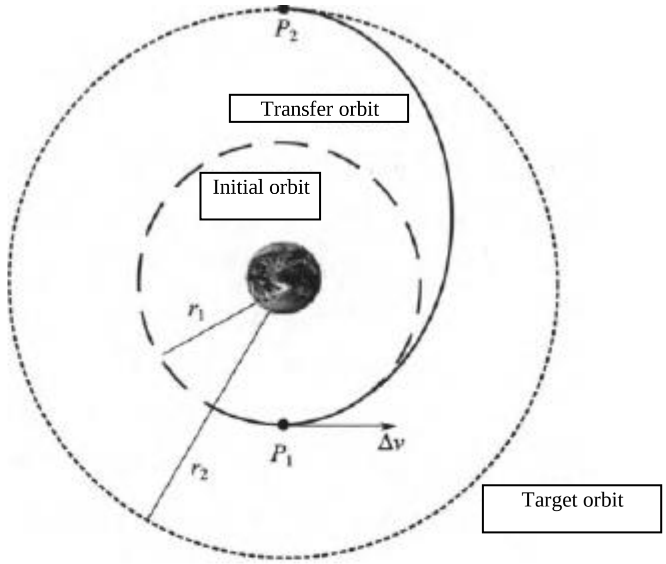


Figure 3-3: Schematic Diagram of Off-Orbit Maneuvering

Minimum impulse, finite time, and double-pulse maneuvering,[196] are the basis for other types of orbital maneuvers. In the double-pulse thrust orbital maneuver shown in Figure 3-3, the first velocity impulse Δv is generated at any point P_1 of the initial orbit with radius r_1, which changes its original orbit and flies to the transfer orbit. In general, the orbital transfer problem is based on the velocity impulse Δv at the periapron point. While changing the original orbit position, it is also necessary to consider the problem of the apex point smoothly entering the target orbit. However, in the orbit service, the orbit entry behavior may not be taken after reaching the vicinity of the target orbital position.

A model of the relationship between the velocity impulse of orbital maneuvering and the orbital radius was constructed.[201]

$$\begin{cases} v_{c1} = \sqrt{\dfrac{\mu}{r_1}} \\ v_{c2} = \sqrt{\dfrac{\mu}{r_2}} \\ \triangle v = v_{c1}\left(\sqrt{\dfrac{2(r_2\,/\,r_1)}{1+(r_2\,/\,r_1)}} - 1\right) \end{cases} \tag{3-10}$$

where v_{c1} and v_{c2} are the velocities of the initial orbit and the target orbit, respectively. $\triangle v$ is the velocity impulse of the periapsis point; r_1 is the initial orbital radius; r_2 is the radius of the target orbit; μ is the gravitational constant of the celestial body.

Based on this, it is possible to estimate the time taken by the spacecraft to maneuver to the target orbit

$$\triangle t = \frac{\pi}{\sqrt{\mu}} \left(\frac{r_1 + r_2}{2}\right)^{\frac{3}{2}} \tag{3-11}$$

(2) Non-coplanar orbital maneuvering

When the initial orbit does not coincide with the orbital plane of the target orbit, the maneuvering process of the spacecraft transferring from the initial orbit to the target orbit is called non-coplanar orbital maneuvering.[201] The focus of this section is to introduce the calculation method of the execution benefit of the orbital maneuver phase, and here we take the typical application of geostationary orbit maneuver as an example.

When a spacecraft maneuvers from an initial orbit with an orbital inclination angle of i_1 and a radius of r_1 to a target orbit with an orbital inclination angle of 0° and a radius of r_2, the maneuvering process can be converted into a Lambert problem for solving. The so-called Lambert problem is a problem of the position vector and the transition time of two points in space relative to the center of gravity, and it is required to determine an orbital maneuver that passes through these two points and the transfer time meets the requirements.[198]

For the elliptical transfer orbit shown in Figure 3-4, according to Lambert's theory, the orbital maneuver time Δt from the initial orbital point P_1 to the target orbital point P_2 depends only on the semi-major axis of the transfer orbit, the sum of the radial radials of the two endpoints from the center of gravity, and the chord length connecting the two points, in the following mathematical form.[198]

$$\sqrt{\mu}\,\Delta t = F\,(a, r_1 + r_2,\ c) \tag{3-12}$$

where μ is the gravitational constant of the earth; Δt is the orbital maneuver time; $F(\bullet)$ is the Lambert expression function; a is the semi-major axis of the transfer orbit; $(r_1 + r_2)$ is the sum of the radials of the two endpoints from the center of gravity; c is the chord length that connects the initial orbital point P_1 the P_2 to the target orbital point.

Based on this, the analytical solution of the Lambert problem for elliptical transfer orbits can be obtained

$$\sqrt{\mu}\,\Delta t = a^{\frac{3}{2}}\,[2N\pi + \alpha - \beta - (\sin\alpha - \sin\beta)] \tag{3-13}$$

where N is the number of rotations; the variables α and β are determined by the equation parameter $(a, r_1 + r_2,\ c)$.

$$\sin\frac{\alpha}{2} = \left(\frac{s}{2a}\right)^{\frac{1}{2}};\ \sin\frac{\beta}{2} = \left(\frac{s-c}{2a}\right)^{\frac{1}{2}} \tag{3-14}$$

where the variable is the half-perimeter of the spatial triangle formed by the initial orbital point P_1 target orbital point P_2 and the gravitational center, and s = (r_1+r_2+c) / 2.

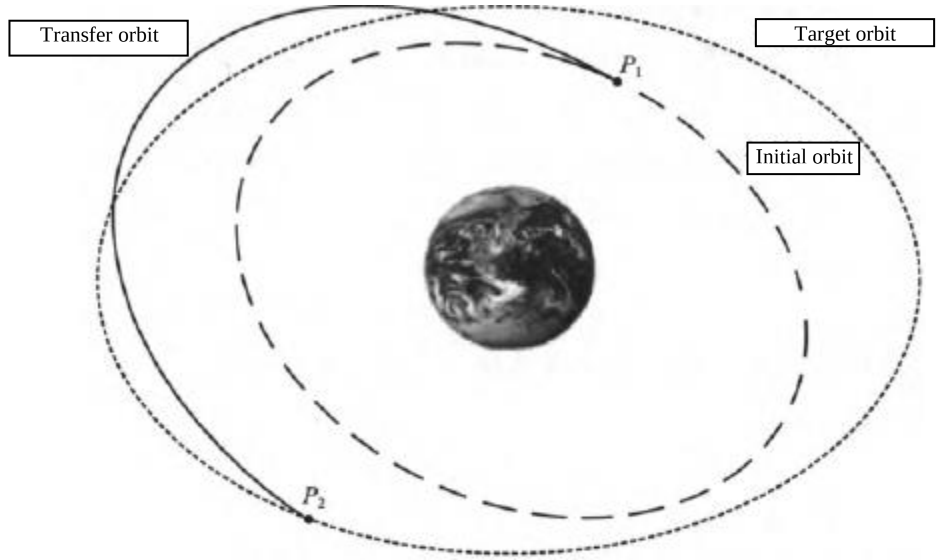


Figure 3-4: Schematic Diagram of Non-Coplanar Orbital Maneuver

Suppose m ={1, 2,...,M} represents the collection of spacecraft capable of on-orbit servicing, and n ={1, 2,...,N} represents the collection of space targets. The W_n indicates the priority of the nth objective. According to different objectives, taking into account the priority and orbital maneuvering time, the benefit index model of the orbital maneuvering stage is constructed

$$C_{ij}^{\text{maneuver}} = \sum_{j=1}^{n} W_j \Delta t_j \left(\sum_{i=1}^{m} X_{ij} PA_{ij} - \prod_{i=1}^{m} (X_{ij} + \Delta) PA_{ij} \right) \quad (3\text{-}15)$$

Where C_{ij}^{maneuver} is the estimated benefit of orbital maneuvering from spacecraft i to target j, W_j is the priority of space target j, Δ is a very small number that avoids $\prod_{i=1}^{m}(\bullet)$ term always being zero, X_{ij} is the allocation decision variable, and Δt_j is the estimated time spent on orbital maneuvering from spacecraft i to target j.

3.2.1.2: Execution benefits in the service implementation phase

The implementation benefit of target allocation in the service implementation stage is mainly reflected in the service benefit. According to the division of spacecraft service mode, the revenue in the service implementation stage is divided into two situations: single-aircraft service income and multi-aircraft service income.

(1) Stand-alone service revenue

If, based on target allocation, only one spacecraft P_i is used to provide service for space target E_j, let CI_i be the spacecraft single-flight service capability index, $LO_i(t)$ be the service capability of the payload carried by the spacecraft at time *t*, and DF_j be the minimum service requirement for space target E_j.[193] The success probability of spacecraft P_i serving space target E_j is

$$PA_{ij}\begin{cases} CI_i & LO_i(t) \geq DF_j \\ 0 & LO_i(t) < DF_j \end{cases} \tag{3-16}$$

If W_j is the service priority of the space target, then the service revenue of the spacecraft P_i to the space target E_j is

$$C_{ij}^{counter} = PA_{ij} \bullet W_j \tag{3-17}$$

(2) Multi-machine service revenue

If m spacecraft are used to carry out services to space targets at the same time according to E_j target allocation, $M = P_1, P_2, \ldots, P_m$ the number of spacecraft participating in on-orbit servicing. Multiple spacecraft can carry the same or different types of payloads to carry out services to space E_j targets, and the service effect will be stronger than that of single aircraft services, so the success probability of *m* spacecraft serving space target E_j is η if the joint service gain parameter is set

$$PA_j^M = \begin{cases} \left[1 - \eta \prod_{M=P_1}^{P_m} \left(1 - \frac{CI_M}{100}\right)\right] \times 100 & \sum_{N=P_1}^{P_N} LO_M(t) \geqslant DF_j \\ 0 & \sum_{M=P_1}^{P_m} LO_M(t) \geqslant DF_j \end{cases} \tag{3-18}$$

If E_j is the service priority of the space target, then the service revenue of the *m* spacecraft to the space target E_j at time *t* is

$$C_{Mj}^{\text{counter}} = PA_j^M \bullet W_j \tag{3-19}$$

3.2.2 Energy Efficiency Indicators

The energy efficiency of spacecraft target allocation in orbit can also be reflected by estimating the energy consumption in the subsequent orbital maneuvering phase and the energy consumption estimation in the service implementation phase.

Therefore, the energy efficiency index is decomposed into two parts: energy consumption in the orbital maneuvering stage and energy consumption in the service implementation stage, namely

$$H_{ij} = H_{ij}^{maneuver} + H_{ij}^{counter} \quad (3\text{-}20)$$

where H_{ij} is the energy efficiency index assigned to the spacecraft's on-orbit target; $H_{ij}^{maneuver}$ is the estimation of energy efficiency during the orbital maneuvering phase; $H_{ij}^{counter}$ is the energy efficiency estimation during the service implementation phase.

(1) Energy efficiency in the orbital maneuver phase

Through the analysis of orbital dynamics and Kepler motion, it can be seen that the velocity impulse Δv increases within a certain range, and the orbital maneuvering path will evolve from an ellipse to a parabola via an elongated ellipse, as shown in Figure 3-5. Under Kepler's law, the method of estimating the orbit maneuver time and each orbit parameter only needs to know the position of the initial orbit and the target orbit and the velocity impulse that can be provided, which belongs to the Gaussian transfer method. Compared with the Hohmann transfer, the Gaussian transfer mode generated by the change of velocity impulse can effectively reduce the maneuvering time between orbits, but at the cost of increasing the velocity impulse, which puts forward higher requirements for the power plant and fuel reserve of the spacecraft.

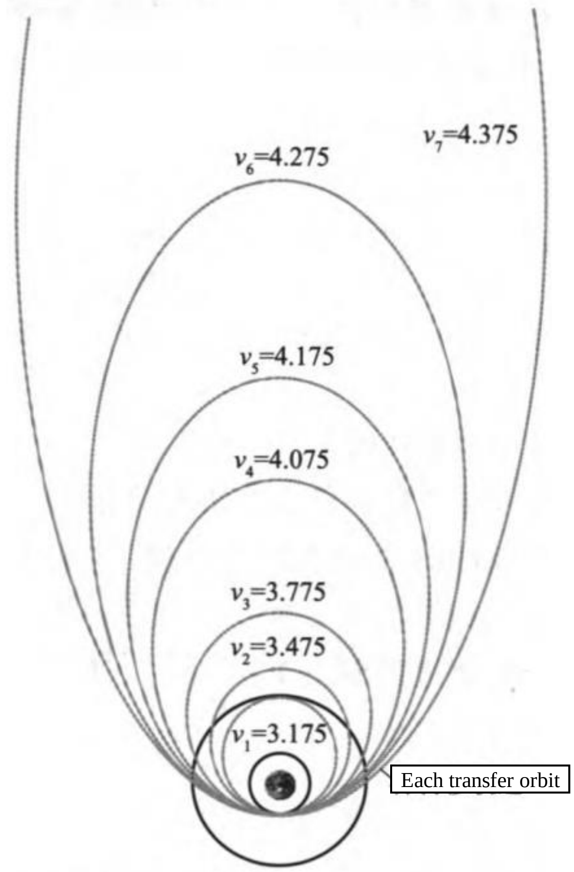


Figure 3-5: Morphological Diagram of Orbital Maneuvering at Different Speeds

The velocity impulse Δv in the Gaussian transfer method is an uncertain quantity, which needs to be determined in advance according to the power unit and fuel reserve. Based on Kepler motion,[201] a model of the relationship between the pericamber point of the transfer orbit and the initial orbital velocity with velocity impulse as a variable can be established.[206]

$$\Delta v = v_{EP} - v_{c1} = v_{EP} - \sqrt{\frac{\mu}{r_1}} \quad (3\text{-}21)$$

where $v_{EP}$is the near camber velocity of the transfer orbit tangent to the initial orbit.

Based on the principle of conservation of momentum, Tsiolkovsky gave a theory of the relationship between velocity impulse and fuel loss,[207, 208] that is, the spacecraft can generate and obtain acceleration at its original operating speed through a counter directional propulsion system that consumes its own fuel mass. Therefore, an impulse mobile fuel consumption estimation model is constructed

$$\Delta m = m_0 \left(1 - e^{-\frac{|\Delta v|}{v_e}}\right) \quad (3\text{-}22)$$

where v_e is the effective exhaust velocity of the spacecraft's impulse engine.

According to the service success probability model equation (3-16) and equation (3-18), the energy efficiency of the orbital maneuver phase can be comprehensively measured by the fuel consumption estimation and the service success probability

$$H_{ij}^{\text{maneuver}} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{1 + \dfrac{W_j^2}{X_{ij} \Delta m_{ij} + \delta_{ij}^2}}{\Delta m_{ij}} \quad (3\text{-}23)$$

where $W_j$is the priority of the spatial target E_j; X_{ij} is the allocation of decision variables; δ_{ij}^2 is random error.

(2) Energy efficiency in the service implementation phase

The dynamic equation of the spacecraft in the service implementation phase can be described by the C-W equation,[209] as shown in Figure 3-6, which forms a coordinate system with the *x*-axis along the geocentric vector, the *y*-axis of the orbital velocity of the space target, and the *z*-axis of the normal direction of the orbital plane. Assuming that the spacecraft is only controlled by the control force, excluding the perturbation, its orbital control force $\boldsymbol{F}_P = (F_x, F_y, F_z)^T$, and the relative equation of motion relative to the space target is[209, 210]

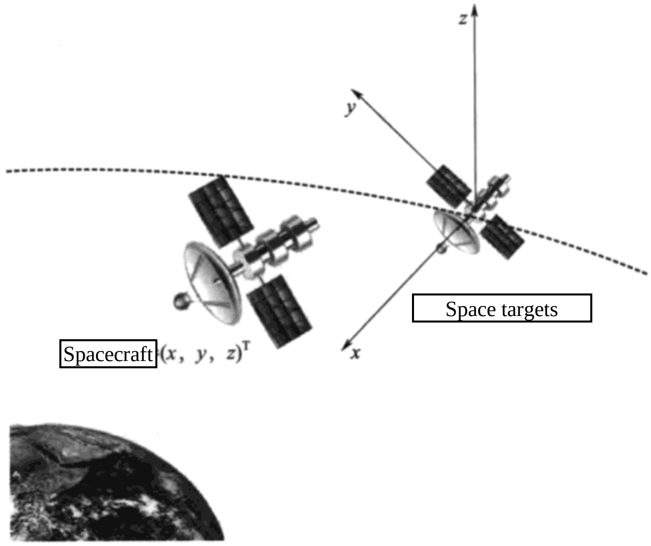


Figure 3-6: Orbital Coordinate System During Service Implementation

$$\begin{cases} \dfrac{F_x}{m_P} = \ddot{x} - 2\omega_E \dot{y} - 3\omega_E^2 x \\ \dfrac{F_y}{m_P} = \ddot{y} + 2\omega_E \dot{x} \\ \dfrac{F_z}{m_P} = \ddot{z} + \omega_E^2 z \end{cases} \tag{3-24}$$

where $\omega_E = \sqrt{\mu/a_E^3}$is the angular velocity of the orbit of the space target; μ is the gravitational constant of the earth, $a_E$is the semi-major axis of the orbit of the space target; $m_p$for spacecraft mass.

During the service implementation phase, the spacecraft needs to increase the velocity impulse, which is generated by consuming the fuel carried by the spacecraft. Describe the equations of motion of the spacecraft as:

$$m_P \frac{\mathrm{d}v_P}{\mathrm{d}t} = \boldsymbol{F}_\mathrm{P} \bullet \mathbf{v}_\mathrm{P} \tag{3-25}$$

where v_P=(v_x, v_y, v_z)$^\mathrm{T}$ is the instantaneous velocity of the spacecraft.

Let $\boldsymbol{P}_P$ indicate the power of the spacecraft, and $\boldsymbol{P}_P$=$\boldsymbol{F}_P \bullet \boldsymbol{v}_P$, then there is

$$\boldsymbol{P}_P = \boldsymbol{v}\left(-\frac{\mathrm{d}m_P}{\mathrm{d}t}\right) \tag{3-26}$$

where $v = (v_x, v_y, v_z)$ is the effective rejection rate of the fuel product.[212]

Thus, the available spacecraft velocity increment is[212]

$$\Delta \boldsymbol{v}_P = \boldsymbol{v} \ln \frac{m_\mathrm{P}^0}{m_\mathrm{P}^\mathrm{f}} \tag{3-27}$$

where m_P^0 and m_P^f are the initial and ending mass of the fuel-carrying spacecraft, respectively.

The estimated model of fuel consumed by the spacecraft during the service implementation phase is:

$$\begin{aligned} \Delta m_{ij} &= m_\mathrm{P}^0 - m_\mathrm{P}^\mathrm{f} \\ &= m_\mathrm{P}^0\left[1 - \exp\left(-\frac{\Delta \boldsymbol{v}_\mathrm{P}}{\boldsymbol{v}}\right)\right] \end{aligned} \tag{3-28}$$

Therefore, considering the priority of the target service and the probability of service success, the energy efficiency model of the service implementation stage is constructed

$$H_{ij}^{\mathrm{counter}} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{1 + \dfrac{W_j^2}{X_{ij}\Delta m_{ij} + \delta_{ij}^2}}{\Delta m_{ij}} \tag{3-29}$$

3.3: Target Allocation Solving Algorithm Based on Improved Deep Q Networks

The in-orbit target allocation under the composite service mode is a process in which the number of targets, the number of orbit roots, and the distribution of spacecraft forces are taken as inputs, and the execution benefit and energy efficiency are taken as the optimal allocation goals, and the two service modes of "one-to-many" and "many-to-one" are considered at the same time. In order to achieve this effect, to further improve the shortcomings of the conventional method that are difficult to directly apply to the composite service mode and the high operation time, this section constructs a two-way training network of Deep Q Networks, improves the convergence and stability of the Deep Q Networks (DQN) algorithm, and introduces a target allocation solving algorithm based on the improved Deep Q Networks.

3.3.1: Target Assignment Deep Q Networks

According to the research framework of intelligent planning for the whole process of on-orbit servicing, for the problem of in-orbit target allocation in the composite service mode, the spacecraft can adopt behaviors according to the initial assumed allocation strategy to realize the interaction with the battlefield environment, and generate a new state under the joint action of behavior and environment while feeding back a certain execution effect, so that the continuous interaction between the spacecraft and the environment can generate a large amount of data. Deep Q Networks is an example of how to use the generated data to select behaviors, and then continue to interact with the environment to generate new data, and use the new data to further improve behavior. Therefore, it is necessary to get rid of the traditional rule-based target allocation method, pursue the balanced development of execution efficiency and energy consumption efficiency in Deep Q Networks, and realize the solution of on-orbit target allocation in the composite service mode through self-learning and self-training.

(1) Target distribution network structure

In the Deep Q Networks network, the optimal allocation strategy D^* is used for the purpose of obtaining the optimal allocation strategy, the success probability of the joint service is taken as the state quantity S_t, the energy efficiency is used as the execution effect (return) R_t, and the assignment factor, fuel consumption, timeliness and robustness are used as the parameters of the allocation strategy θ^-. Firstly, the target allocation scheme D is initialized, its parameter θ is initialized randomly, and the parameter θ^- of the target allocation scheme D^* is initialized by the parameter θ. After that, M iterations of the allocation scheme were performed, that is, M times of Deep Q Networks network autonomous training was performed. During each iteration, the initial state S_1 is initialized X_1 the space environment and encoded φ by a conversion function. Then, Deep Q Networks forward transmission is carried out, and the following specific content is completed in each time step t: According to the current target allocation scheme D, select behavior A_t and execute, and the execution effect (return) is observed R_t and the current space environment X_{t+1}will be S_t, A_t and $X_{t+1}$are combined to get the current latest state $S_{t+1}$and encode with φ, and then φ_t, A_t, R_t, and φ_{t+1} are stored as experience fragments in the playback memory unit M. Then, Deep Q Networks reverse training was performed, that is, the experience fragments were extracted from the playback memory unit in batches to train the parameters of the neural network, the error value of the target assigned to the target was calculated and guided by it, and the backpropagation mechanism was used to transmit it to each layer of the network layer by layer to complete an update of the scheme parameter θ. Finally, after each C update, the target allocation scheme parameter θ^- is updated with the current θ value. This is repeated until the target allocation scheme D^* meets the target effect or reaches the set number of times, and the automatic allocation process is terminated. Figure 3-7 shows the network structure of Deep Q Networks.

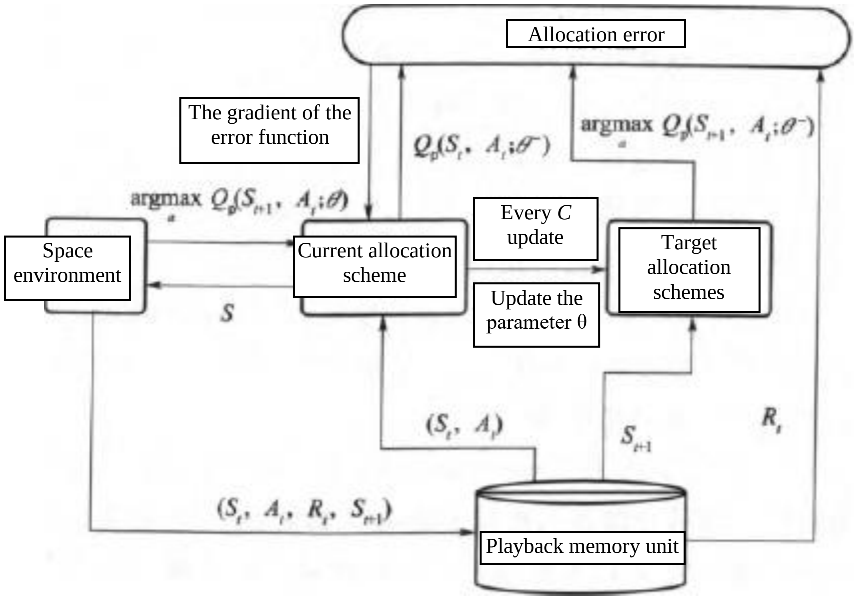


Figure 3-7: Target Assignment Deep Q Networks Network Fabric

(2) Forward transmission of target allocation

In the process of target allocation, while pursuing the high efficiency of target allocation, in order to ensure that each space target has a spacecraft to deal with it and can reach the preset success probability threshold, combined with the overall target formula (3-1) and the constrained space, the forward transmission of the Deep Q Networks network is expressed as

$$\begin{cases} \min\limits_{(X_{ij})} -G \\ \sum\limits_{i=1}^{m} X_{ij} \geqslant 1(\forall i \in m) \\ 1-\prod\limits_{i=1}^{m}[1-PA_{ij}X_{ij}]-PA'_j \geqslant 0(\forall i \in m) \\ \Delta M_i = M_i^{\text{total}} - M_i^{\text{maneuver}} - M_i^{\text{implement}} \geqslant 0 \\ T_{k(l)}^{\text{StarTime}} \geqslant T_{k(l)}^{\text{EarlyTime}} \\ T_{k(l)}^{\text{StarTime}} + T_{k(l)}^{\text{Dur}} \leqslant T_{k(l)}^{\text{LateEndTime}} \\ P(T_{k(l)}^{\text{Dur}} \leqslant T_{k(l)}^{\text{Dur\_limit}}) \geqslant \overline{P}_{\min} \\ \sum\limits_{l=1}^{N} \omega_{k(l)} \cdot E\,|\,T_{k(l)}^{\text{Dur}} - T_{k(l)}^{\text{Dur\_planning}}\,| \leqslant \Delta T_{\min} \end{cases}$$

(3-30)

The penalty function method is used to transform the constrained optimization problem into the following unconstrained optimization problem

$$G=\min\left[\tau\left(\sum_{j=1}^{n}h_j^2+\sum_{j=1}^{n}g_j^2+\sum_{j=1}^{n}k_j^2+\sum_{j=1}^{n}l_j^2+\sum_{j=1}^{n}t_j^2+\sum_{j=1}^{n}r_j^2+\sum_{j=1}^{n}e_j^2\right)-\sum_{j=1}^{n}\frac{W_j}{D_j}\right] \tag{3-31}$$

where W_j is the priority of the jth target; the parameter τ is the penalty coefficient; and the expressions of h_j, g_j, k_j, l_j, t_j, r_j, e_j and D_j are respectively

$$\begin{cases} h_j=\max\left\{1-\sum_{i=1}^{n}X_{ij},0\right\} \\ g_j=\max\left\{PA'_j+\prod_{i=1}^{m}[1-PA_{ij}X_{ij}]-1,0\right\} \\ k_j=\max\{M_i^{\text{maneuver}}+M_i^{\text{implement}}-M_i^{\text{total}},0\} \\ l_j=\max\{T_{k(l)}^{\text{StartTime}}-T_{k(l)}^{\text{LateEndTime}},0\} \\ t_j=\max\{T_{k(l)}^{\text{StartTime}}+T_{k(l)}^{\text{Dur}}-T_{k(l)}^{\text{LateEndTime}},0\} \\ r_j=\max\{\overline{P}_{\min}-P(T_{k(l)}^{\text{Dur}}\leqslant T_{k(l)}^{\text{Dur\_limit}}),0\} \\ e_j=\max\left\{\sum_{l=1}^{N}\omega_{k(l)}\cdot E\,|T_{k(l)}^{\text{Dur}}-T_{k(l)}^{\text{Dur\_planning}}|-\Delta T_{\min},0\right\} \\ D_j=\sum_{j=1}^{m}L_{ij} \end{cases} \tag{3-32}$$

The process of maximizing target G is not only the maximization of current benefits, but also the cumulative benefits of the long term. The goal allocation goal sought is noted as a G_t, defined as a function of the sequence of execution effects (returns), which is the cumulative effect of execution over time

$$G_t \doteq R_{t+1}+R_{t+2}+R_{t+3}+R_{t+4}+\cdots+R_{t+n} \tag{3-33}$$

In the face of a continuous task, $n=\infty$ at the final moment, the maximum comprehensive benefit tends to be infinite, and it will be unnecessary trouble to continue to describe it with equation (3-33). To do this, each item in Eq. (3-33) is "discounted" so that it receives a return weighted by the discount factor

$$G_t \doteq R_{t+1}+\gamma R_{t+2}+\gamma^2 R_{t+3}+\gamma^3 R_{t+4}+\cdots+\gamma^{n-1}R_{t+n}=\sum_{n=1}^{\infty}\gamma^{n-1}R_{t+n} \tag{3-34}$$

where γ is the discount rate, and $\gamma \in [0,1]$.

The introduction of the discount rate results in a return of only γ^{n-1}times the current value at the next *n* moments. When γ = 0, it means that only the next decision is always considered, which is "short-sighted" and only cares about the current benefits. But when γ > 0, as the discount rate increases, more attention will be paid to the future situation, and the long-term vision will be taken into account, which means that the decision will become far-sighted. This process can be illustrated in Figure 3-8.

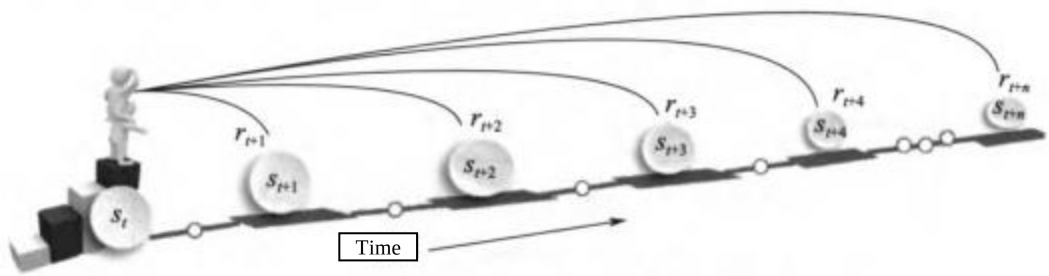


Figure 3-8: Cumulative Benefits of Target Allocation After the Introduction of the Discount Rate

In the opposite equation (3-34), the comprehensive benefits of adjacent moments can be expressed recursively

$$\begin{aligned} G_t &\doteq R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \cdots + \gamma^{n-1} R_{t+n} \\ &= R_{t+1} + \gamma (R_{t+2} + \gamma^1 R_{t+3} + \gamma^2 R_{t+4} + \cdots + \gamma^{n-2} R_{t+n}) \\ &= R_{t+1} + \gamma G_{t+1} \end{aligned} \quad (3\text{-}35)$$

With the cumulative benefit of target allocation, the state value function can be used to evaluate the effect of behavior in a given state. The state-value function is related to a particular way of behaving, which this book calls "allocation strategies.” An allocation strategy is a mapping from a state to the probability of a behavior being executed. When *t* selects the allocation strategy π, then $\pi(a|s)$is the probability of executing the behavior $A_t = a$ in the state $S_t = s$. Therefore, the expected value of the probability of the execution effect obtained in state *s* according to the allocation strategy π selection behavior $a$can be called the state value function $v_\pi(s)$

$$v_\pi(s) \doteq E_\pi[G_t \mid S_t = s] = E_\pi\left[\sum_{n=1}^{\infty} \gamma^{n-1} R_{t+n} \mid S_t = s\right] \quad (3\text{-}36)$$

where $E_\pi[\cdot]$ is the desired function for a random variable given an allocation strategy. Similarly, the value of behavior a selected in state *s* based on the allocation strategy π is called the behavior value function.

The behavioral value function $Q_\pi(S_t, a)$ is the expected return of all possible benefit sequences starting from state s and after performing behavior a, according to the π of the allocation strategy

$$Q_\pi(S_t, a) \doteq E_\pi[G_t \mid S_t = s, A_t = a] = E_\pi\left[\sum_{n=1}^{\infty} \gamma^{n-1} R_{t+n} \mid S_t = s, A_t = a\right] \quad (3\text{-}37)$$

Deep Q Networks' goal allocation problem is to find an allocation strategy through self-training and learning, so that it can get the maximum benefit in the long-term behavior process. When the allocation strategy π always outperforms all other strategies, the allocation strategy π referred to as the optimal strategy π^*. During this period, if an allocation strategy π is considered to be better than any other strategy π', then its execution effect on all states should be better than or equal to the effect of strategy π', that is, for $s \in S$, $v_\pi(s) \geq v_{\pi'}(s)$. Denote the optimal state value function as

$$v^*(s) \doteq \max_\pi v_\pi(s) \quad (3\text{-}38)$$

The optimal strategy also enjoys the optimal behavioral value function, which can be expressed as:

$$q^*(S_t, a) \doteq \max_\pi q_\pi(S_t, a) \quad (3\text{-}39)$$

For the binary set of state and behavior, the optimal state value function $v^*(s)$ can be used to represent the optimal behavior value function $q^*(s, a)$, i.e.,

$$q^*(S_t, a) = E[R_{t+1} + \gamma v^*(S_{t+1}) \mid S_t = s, A_t = a] \quad (3\text{-}40)$$

Therefore, the forward transmission of target allocation is the process of seeking the optimal allocation strategy to maximize the behavioral value function through the cumulative analysis of the benefits of the allocation target.

(3) Target assignment inverse training

Target allocation inverse training is a process of updating parameters and improving the allocation scheme by adjusting the error between the iterative effect of the scheme and the estimation of the value function on the basis of forward transmission.

For some experiences π allocation strategies, as well as the non-terminating states in these experiences, the state value function can be estimated according to equation (3-36) in the reverse training. However, according to the cumulative characteristics of the comprehensive benefits of Eq. (3-34), each estimation needs to wait until the execution effect of the final act is obtained

$$\begin{cases} V(S_t) \leftarrow \frac{1}{n}\sum_{t=1}^{n} G_t \\ V(S_t) \leftarrow \frac{1}{n}(G_n + \sum_{t=1}^{n-1} G_t) \\ V(S_t) \leftarrow \frac{1}{n}\left[G_n + (n-1)\frac{1}{n-1}\sum_{t=1}^{n-1} G_t\right] \\ V(S_t) \leftarrow \frac{1}{n}[G_n + (n-1)V(S_n)] \\ V(S_t) \leftarrow \frac{1}{n}[G_n + nV(S_n) - V(S_n)] \\ V(S_t) \leftarrow V(S_n) + \frac{1}{n}[G_n - V(S_n)] \end{cases} \tag{3-41}$$

where $\frac{1}{n}$ is the constant step parameter.

In order to avoid having to wait until the end of an act to determine $V(S_t)$, the sequential differential pair (3-41) can be improved to obtain an estimate $V(S_t)$ based on the obtained data without waiting for the final result of the interaction

$$V(S_t) \leftarrow V(S_t) + \frac{1}{n}[R_{t+1} - \gamma V(S_{t+1}) - V(S_t)] \tag{3-42}$$

Figure 3-9 provides a complete description of the estimation process for this time series difference. The estimation of time series differences has two distinct comparative advantages:

1) Compared to the Monte Carlo method, this method naturally uses an online, fully incremental approach to prediction, i.e., there is no longer a need to wait until the end of an act.

2) Compared with the dynamic programming method, this method no longer requires an environment model.

In Eq. (3-42), the result of the calculation in parentheses represents an error that determines the difference between the next time $R_{t+1} - \gamma V(S_{t+1})$ the current $V(S_t)$ and is named the Timing Difference (TD) error term

$$\delta_t = R_{t+1} - \gamma V(S_{t+1}) - V(S_t) \tag{3-43}$$

Based on the estimation idea of time series difference, the estimation of the optimal behavior value function can be realized by Q-learning.

Time series difference (TD), which is used to estimate $V(S_t)$
Input: The π of the policy to be estimated Algorithm parameters: step size $\frac{1}{n}\epsilon(0, 1]$ For all $s \in S^+$, arbitrary initialization $V(S_t)$, where V(termination state) = 0 For each loop: Initialize S Each step in the playbook: A← strategy π the decision action made in state S to execute action A, observed R, S' $V(S_t) \leftarrow V(S_t) + \frac{1}{n}[R_{t+1} - \gamma V(S_{t+1}) - V(S_t)]$ $S \leftarrow S'$ Until S is in the terminated state

Figure 3-9: The Estimation Process of Time Series Difference

According to the current state S_t and its behavior A_t, combined with the observed state S_{t+1} and execution effect R_{t+1} at the next moment, the value function of the next optimal behavior can be estimated

$$Q(S_t, a) \leftarrow Q(S_t, A_t) + \frac{1}{n}[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \tag{3-44}$$

Figure 3-10 illustrates the estimation process of Q-learning in its entirety.

In the estimation process of the behavioral value function, the optimal behavioral value function is directly taken as the learning goal, which is no longer related to the allocation strategy used to generate the trajectory of the decision sequence, which will greatly simplify the process of operation and analysis. Therefore, in the reverse training process, this parameter update method based on time series difference (TD) error is expressed as

$$Y_t^Q = R_{t+1} + \gamma \max_a Q(S_{t+1}, a; \theta_t) \tag{3-45}$$

where R_{t+1} is the execution effect of the allocation scheme in the state S_{t+1}; $\gamma \in [0, 1]$ is the discount factor; $Q(S_{t+1}, a; \theta_t)$ is the estimate of the value of behavior after taking behavior a according to the allocation scheme θ_t in state S_{t+1}.

According to Eq. (3-44) and Eq. (3-45), the minimum error function of the reverse training is constructed

Q-learning, which is used to estimate $Q(S_t, a)$
Algorithm parameters: step size $\frac{1}{n}\epsilon(0, 1]$ Arbitrarily initialize Q(s, a) for all $s \in S^+$, $a \in A(s)$ where Q(terminal state, ·)= 0 For each act: Initialize S Each step in the playbook: Using the strategy π obtained from Q, select A outside S to perform the action A observes R, S' $Q(S_t, a) \leftarrow Q(S_t, A_t) + \frac{1}{n}[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$ $S \leftarrow S'$ Until S is in the terminated state

Figure 3-10: The Estimation Process of Q-Learning

$$E = \min [Y_t^{DQ} - Q(S_t, a)] \quad (3\text{-}46)$$

Therefore, the gap between the iterative effect of the scheme and the estimation of the value function is continuously reduced through the target allocation inverse training, and the optimal direction is continuously modified and improved to obtain the optimal allocation scheme.

3.3.2: Deep Q Networks Convergence and Stability Improvements

Typically, Q Teaming in Deep Q Networks is maximized on the basis of the estimate, which can be seen as an implicit estimation of the maximum, and this treatment produces a significant reward bias. In particular, due to the instability of the model during the training process, the bias will cause the judgment bias of the model on the advantages and disadvantages of behavior, which will affect the convergence of the model. In addition, due to the possible prediction error of Deep Q Networks, the behavior value function will be corrected in the direction of the maximum error by using its maximum approximation every time, and the error will be amplified after repeated iterations, resulting in the final convergence of the behavior value function value much higher than the real one, resulting in the overestimation problem affecting the network stability.[214]

To solve this problem, the improved network structure of Deep Q Networks as shown in Figure 3-11 was constructed, and the convergence improvement of Deep Q Networks was realized by splitting the behavior value function into the state value function and the behavior advantage function to alleviate the reward bias problem.

By changing the original single-neural network estimation method to double-Q network estimation, the overestimation problem is alleviated and the stability of Deep Q Networks is improved.

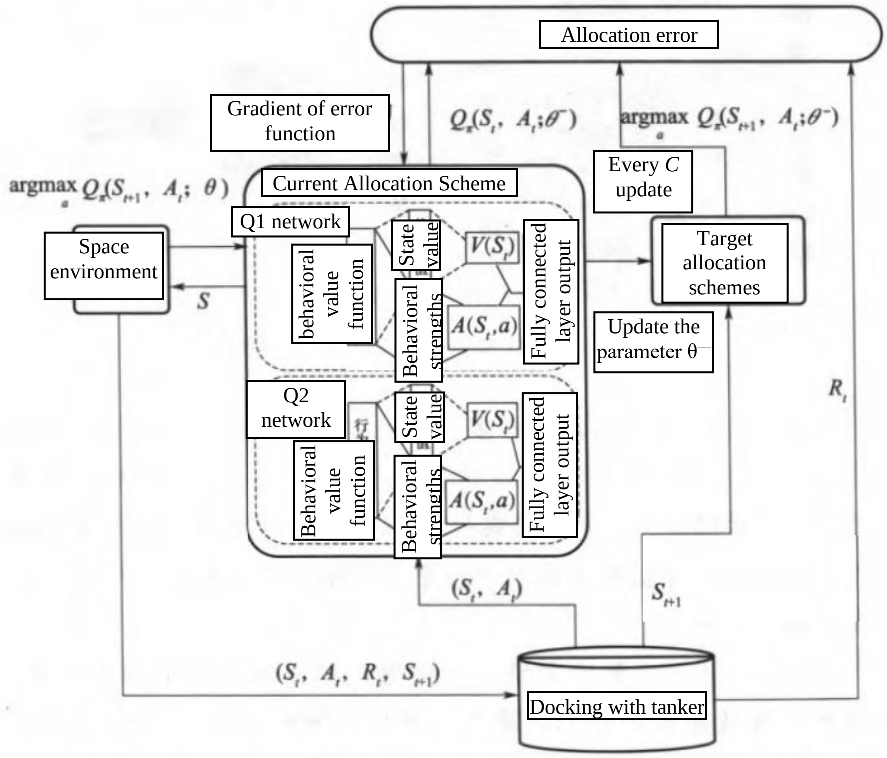


Figure 3-11: Deep Q Network Structure Convergence and Stability Improvements

(1) Deep Q Networks' convergence improvement

Deep Q Networks is an approach that combines reinforcement learning with neural networks, in which the neural network is essentially a multilayer perceptron (MLP).[194] Multilayer perceptrons introduce one or more hidden layers on top of a single-layer neural network, which are located between the input and output layers. Figure 3-12 illustrates the neural network of a multilayer perceptron in Deep Q Networks. In this multilayer perceptron, the number of inputs and outputs is nine sum machines, and in the middle are three layers of hidden layers, each hidden layer is a fully connected structure, that is, the neurons are fully connected with each input in the input layer, and the neurons in the output layer and each neuron in the hidden layer are also fully connected.

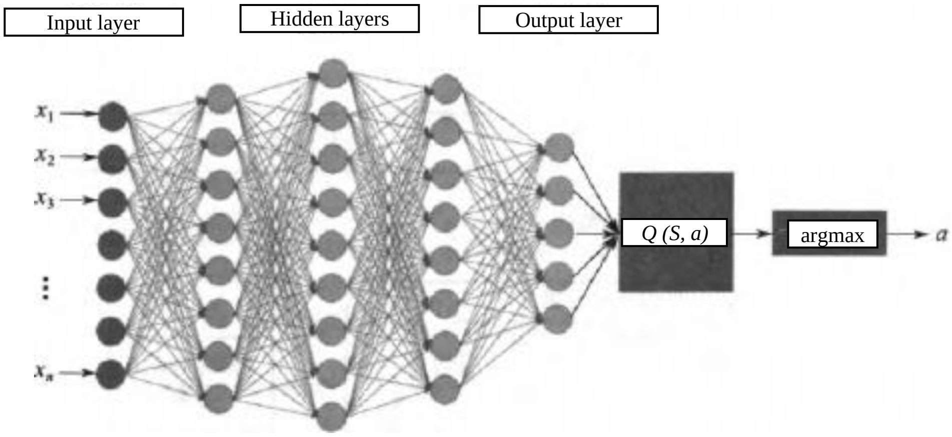


Figure 3-12: Deep Q Networks Multilayer Neural Network Structure

For Deep Q Networks multilayer neural network, the behavior value function $Q(S_t, a)$ can be naturally split into two parts: the state value function $V(S_t)$ and the behavioral advantage function $A(S_t, a)$. Among them, the state value function is independent of behavior; the behavioral dominance function is related to behavior, measures the degree of average return of behavior relative to the state, and can be used to solve the problem of reward bias.

Accordingly, in order to further solve the problem of reward bias, improve the training effect of neural networks, and accelerate the method convergence speed, a competitive neural network is constructed to replace the single-output network model in classical Deep Q Networks. As shown in Figure 3-13, the fully connected layer output of Deep Q Networks' multilayer neural network is divided into a state value $V(S_t)$ and a behavior dominance value $A(S_t, a)$

$$Q(S_t, a) = V(S_t) + A(S_t, a) \qquad (3\text{-}47)$$

According to Eq. (3-47), simply decomposing the value function of the action will not work. Because after the action value function is split, when the behavior is fixed, there are infinite possible combinations of state values and behaviors, and in fact only a small part of the combinations is reasonable and close to the true values. For this reason, it is also necessary to limit the output of the two parts of state value $V(S_t)$ and behavioral advantage value $A(S_t, a)$.

According to the behavioral dominance function $A(S_t, a)$ the expected value is zero.[215, 216]

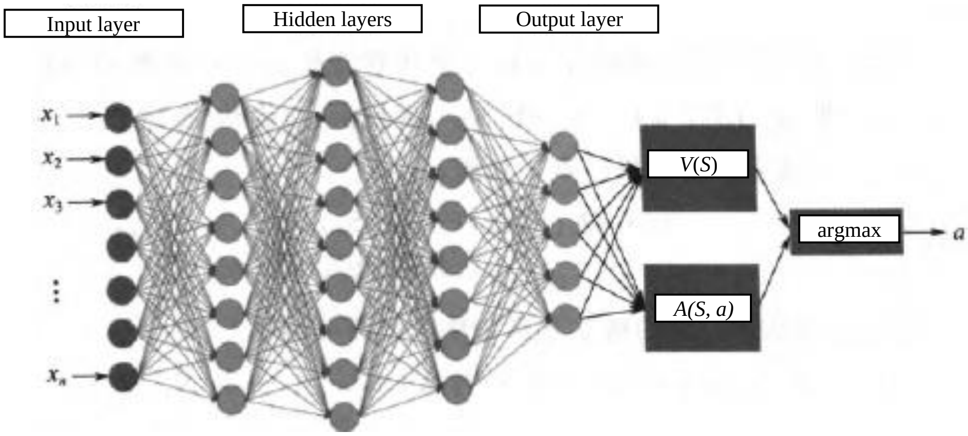


Figure 3-13: Deep Q Networks Competing Neural Network

$$E_a[A(S_t,a)] = E_a[Q(S_t,a) - V(S_t)] = V(S_t) - V(S_t) = 0 \tag{3-48}$$

Limit the behavioral dominance function $A(S_t, a)$by modifying Eq. (3-47) to

$$Q(S_t,a) = V(S_t) + \left(A(S_t,a) - \frac{1}{|A|}\sum_{a'} A(S_t,a')\right) \tag{3-49}$$

At the effect level, if the behavioral advantage value of the current allocation scheme is $A(S_t, a)$ positive, it means that the scheme performs better than all other possible schemes in terms of execution effect; conversely, if the behavioral advantage value $A(S_t, a)$ of the allocation regimen is negative, it means that the potential execution of the regimen is below average.

From the method level, decomposing the value of the behavioral value function into the state value function and the dominant value function, compared with the direct use of the behavior value function value, the processing of removing the mean of all the dominant values keeps the expected value of the behavioral value of the behavior value at 0, so as to ensure the rapid convergence of the model and the efficient output.

(2) Deep Q Networks stability improvements

The overestimation phenomenon in Deep Q Networks is that it contains uniformly distributed random errors within [-ε, ε], which will cause each behavioral value function to be overestimated by γ times.[217]

Considering that in the state where the optimal behavior value estimation function is equal to the optimal state value estimation function, that is, $Q^{*}(S_t, a) = V^{*}(S_t)$, assumes that the behavior value estimation error is an independent uniform random distribution of [–1, 1] and satisfies

$$\begin{aligned}\varepsilon_a &= Q_t(S_t, a) - Q^{*}(S_t, a) \\ &= Q_t(S_t, a) - V^{*}(S_t)\end{aligned} \tag{3-50}$$

It can be obtained according to the independent distribution characteristics of the value estimation error

$$\begin{aligned}P(\max_a \varepsilon_a \leqslant x) &= P(X_1 \leqslant x \wedge X_2 \leqslant x \wedge \cdots \wedge X_m \leqslant x) \\ &= \prod_{a=1}^{m} P(\varepsilon_a \leqslant x)\end{aligned} \tag{3-51}$$

The function $P(\varepsilon_a \leq x)$ is the Cumulative Distribution Function (CDF) of ε_a, which can be divided into

$$P(\varepsilon_a \leqslant x) = \begin{cases} 0 & x \leqslant -1 \\ \dfrac{1+x}{2} & x \in (-1, 1) \\ 1 & x \geqslant 1 \end{cases} \tag{3-52}$$

And then it can be obtained

$$\begin{aligned}P(\max_a \varepsilon_a \leqslant x) &= \prod_{a=1}^{m} P(\varepsilon_a \leqslant x) \\ &= \begin{cases} 0 & x \leqslant -1 \\ \dfrac{1+x}{2} & x \in (-1, 1) \\ 1 & x \geqslant 1 \end{cases}\end{aligned} \tag{3-53}$$

The expectation of the random variable $\max_a x\varepsilon_a$ can be expressed by a cumulative distribution function

$$E[\max_a \varepsilon_a] = \int_{-1}^{1} x f_{\max}(x)\,\mathrm{d}x \tag{3-54}$$

where $f_{\max}(\cdot)$ is the probability density function of the variable x, which is defined as the derivative of the cumulative distribution function

$$f_{\max}(x)=\frac{\mathrm{d}P(\max_a \varepsilon_a \leqslant x)}{\mathrm{d}x} \tag{3-55}$$

So when $x \in [-1, 1]$, there is

$$f_{\max}(x)=\frac{m}{2}\left(\frac{1+x}{2}\right)^{m-1} \tag{3-56}$$

where m is the number of actions that can be taken.

For the integration of Eq. (3-56), we get the expectation of $\max_a x\varepsilon_a$, and we can determine the overestimation multiple γ

$$\begin{aligned}\gamma &= E[\max_a \varepsilon_a] \\ &= \int_{-1}^{1} x f_{\max}(x)\mathrm{d}x \\ &= \left[\left(\frac{x+1}{2}\right)^m \frac{mx-1}{m+1}\right]_{-1}^{1} \\ &= \frac{m-1}{m+1}\end{aligned} \tag{3-57}$$

This phenomenon of overestimation in Deep Q Networks is rooted in the fact that the same sample is used in the process of determining the most valuable behavior and estimating its value. To this end, this book divides these samples into two and independently estimates the value of the behavior, obtaining Q_1 and Q_2. Use one of these estimating Qi to get the maximum behavior

$$A^* = \arg\max_a Q_1(a) \tag{3-58}$$

An estimate that utilizes another estimate Q_2 to determine its value

$$Q_2(A^*) = Q_2[\arg\max_a Q_1(a)] \tag{3-59}$$

Since $E[Q_2(A^*)] = q(A^*)$,[218] therefore, $Q_2(A^*)$ is an unbiased estimate.

Interchanging the process of Eq. (3-58) with Eq. (3-59), i.e., exchanging the roles of Q_1 and Q_2, gives another unbiased estimator

$$Q_1(A^*) = Q_1[\arg\max_a Q_2(a)] \tag{3-60}$$

Eq. (3-59) and Eq. (3-60) constitute a double-Q learning idea, and the whole estimation process is shown in Figure 3-14.

In Deep Q Networks, the same parameter θ_t is used to perform the behavior after selecting the best behavior a^* of state S_{t+1} according to the timing differential target update formula (3-45).

Double-Q learning for estimating $Q_1 \approx Q_2 \approx q^*$

Algorithm parameters: step size $\frac{1}{n}\epsilon(0, 1]$
For all $s \in S^+, a \in A(s)$, arbitrarily initialize Q_1 (s, a) and Q_2 (s, a), where Q(termination state, $\cdot$) = 0
For each loop:
    Initialize S
    Each step in the playbook:
        Based on $Q_1 + Q_2$, select A in S π to use the strategy
        Execution behavior A Observed R, S′
        Execute with a probability of 0.5:
        $Q_1(S_t, a) \leftarrow Q_1(S_t, A_t) + \frac{1}{n}[R_{t+1} + gQ_2(S_{t+1}, \max_a Q_1(S_{t+1}, a)) - Q_1(S_t, A_t)]$
        Or execute:
        $Q_2(S_t, a) \leftarrow Q_2(S_t, A_t) + \frac{1}{n}[R_{t+1} + gQ_1(S_{t+1}, \max_a Q_2(S_{t+1}, a)) - Q_2(S_t, A_t)]$
        $S \leftarrow S'$
    Until S is in the terminated state

Figure 3-14: The Estimation Process of Double-Q Learning

Selection and evaluation.[219] However, this approach would cause the maximum of the estimated value to be regarded as the largest estimate of the true value, thus creating an overestimation phenomenon. Although the goal of Deep Q Networks is to find the optimal strategy, the non-uniform occurrence of the overestimation phenomenon will cause the value function to be overestimated and affect the decision, resulting in the final decision is not optimal, but only suboptimal.[220]

In order to weaken the influence of the maximum error, based on the double-Q learning model, another neural network is introduced to select and update the behavior with different value functions. Therefore, the parameter θ_t is used to select the behavior through Eq. (3-45), and after the optimal behavior a^* is selected, the parameter θ_t^- of another neural network is used to update the behavior[186]

$$Y_t^{DQ} = R_{t+1} + \gamma Q(S_{t+1}, a^*; \theta_t^-) \tag{3-61}$$

According to Eq. (3-61), a new Timing Differential (TD) target update formula can be obtained

$$Y_t^{DQ} = R_{t+1} + \gamma Q(S_{t+1}, \max_a Q(S_{t+1}, a; \theta_t); \theta_t^-) \tag{3-62}$$

Therefore, the double-Q learning model is used to improve the time series difference target update formula and reduce the overestimation of the value function of the behavior, which can help the method to select a better allocation scheme and achieve better results execution, and the stability of the method is improved as a result.

3.3.3: Target Allocation Method Process

After clarifying the network inputs, outputs, key models and training structures, the main process of the spacecraft on-orbit target allocation method is given based on the stability improvement and convergence improvement of Deep Q Networks (see Figure 3-15).

Step 1 The spacecraft in-orbit target allocation scheme D is initialized by stochastic θ.

Step 2 Let $\theta_t = \theta$, according to Eq. (3-45) and Eq. (3-61), the behavior value of the time-series differential target is calculated.

Step 3 Iterate over each event in a loop.

Step 4 The first state S_t of the initialization event is preprocessed to obtain the optimization feature vector corresponding to the state $f(S_t)$.

Step 5 Loop every step of each event.

Step 6 Use the probability ε to execute the behavior A in the current allocation scheme D, and if the small probability exploration event does not occur, the greedy strategy is used to execute the behavior with the largest function value of the current behavior value $a_t = \max_a Q(S_{t+1}, a; \theta_t)$.

Step 7 Execute behavior A_t in the simulator and observe the execution effect R_{t+1}.

Step 8 Set up S_{t+1}=S_t conformity (S_t, A_t, R_{t+1},S_{t+1}) and stored in the playback memory unit.

Step 9 Uniformly and randomly sample a transformed sample data from replay memory unit and store the result as (S_t, A_t, R_{t+1}, S_{t+1}).

Step 10 Determine whether it is the termination state of an event, and if so, the TD target is R_{t+1}, otherwise the TD target Y_t^{DQ} will be calculated according to Eq. (3-62).

Step 11 In the reverse training, a gradient descent ∇E is performed on Eq. (3-46), and the target allocation scheme parameter θ^- is updated by gradient backpropagation of the neural network.

Step 12 If S_{t+1} is in the terminated state, and the current iteration is completed, otherwise go to step 3.

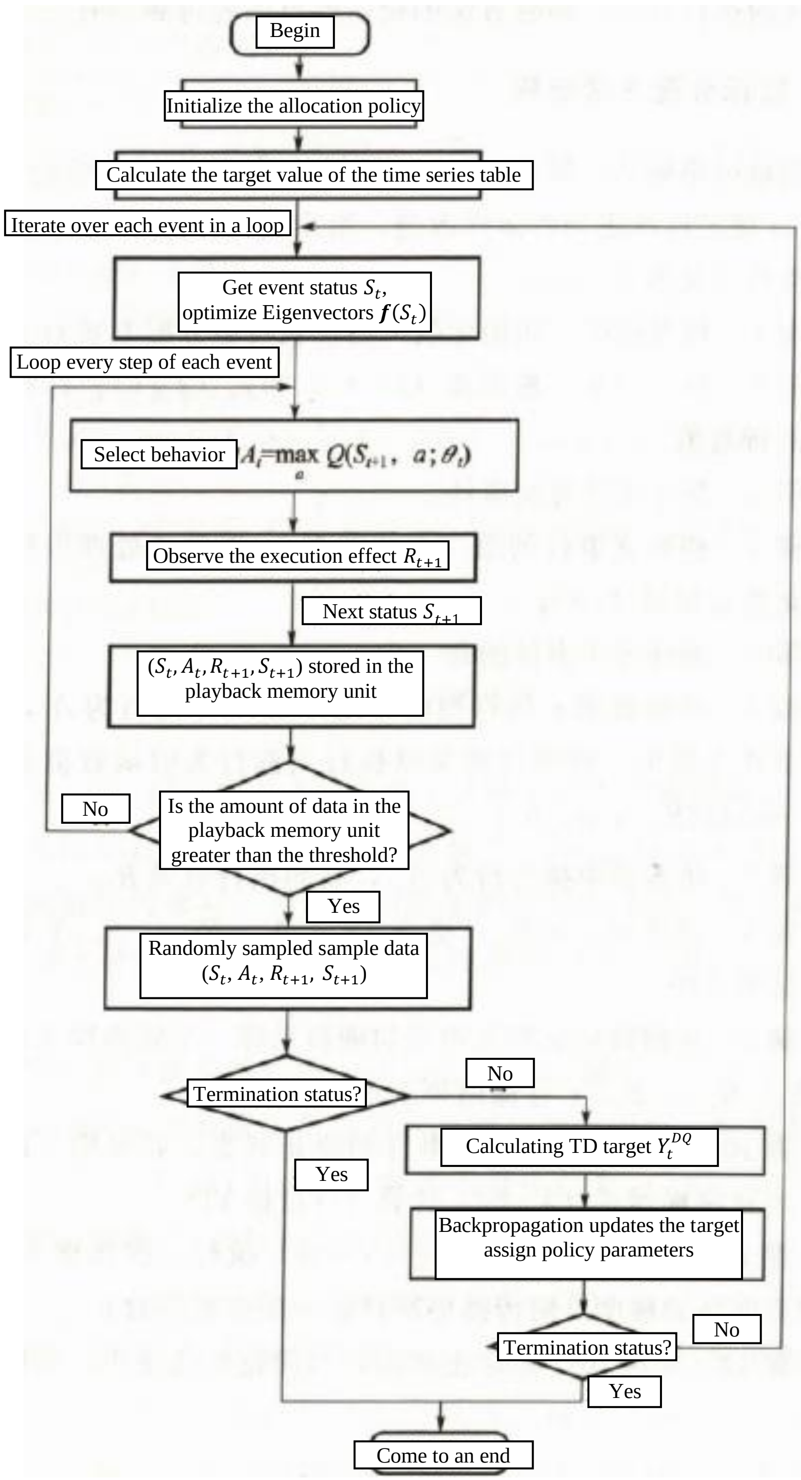


Figure 3-15: Solving Process Based on the Improved Target Allocation Algorithm of Deep Q Networks

3.4: CASE STUDY OF ON-ORBIT TARGET ALLOCATION UNDER COMPOSITE SERVICE MODE

In order to verify the applicability of the on-orbit target allocation model in the composite service mode and the effectiveness of the target allocation solving algorithm based on the improved Deep Q Networks, a case simulation and analysis are carried out in this section.

3.4.1: Case Description

In an on-orbit servicing mission, there are 13 geostationary orbit satellite targets, and their orbital root numbers[225] are shown in Table 3-1. In the table: e is the eccentricity; i is the orbital inclination; Ω is the ecliptic longitude of the ascending node; ω is the pericentricity angle; τ is the flat peripoint angle. Six spacecraft with an orbital radius of 38, 864 km, a mass of 2, 000 kg, an initial true perigee angle of 0°, and a specific impulse of the propulsion system of 300 s can perform on-orbit servicing missions, and their orbital root numbers are shown in Table 3-2. In view of the uneven supply and demand quantity, the large number of scattered targets and the limited service force in orbit, in order to obtain a high service success probability, it is proposed to adopt a composite service model combining "one-to-many" and "many-to-one," and seek the balanced development of implementation benefits and energy consumption efficiency while comprehensively considering the assignment constraints, fuel constraints and timeliness constraints, and improve the service success probability of target 13 and reach 85% on the basis of satisfying the service success rate of 70%. According to the prior information, the service success probability PA of each spacecraft is known, the velocity increment from the spacecraft to each target orbit is determined according to the number of orbital roots, and the fuel consumption $\triangle m$, the success probability threshold $\bar{P}$ and the target priority W are estimated by combining with the Tsiolkovsky formula, respectively.

$$
PA = \begin{bmatrix}
0.68 & 0.69 & 0.67 & 0.70 & 0.66 & 0.68 & 0.70 & 0.76 & 0.73 & 0.74 & 0.70 & 0.75 & 0.66 \\
0.69 & 0.70 & 0.68 & 0.66 & 0.72 & 0.69 & 0.75 & 0.72 & 0.74 & 0.72 & 0.75 & 0.75 & 0.67 \\
0.75 & 0.72 & 0.73 & 0.68 & 0.73 & 0.68 & 0.70 & 0.77 & 0.63 & 0.68 & 0.76 & 0.69 & 0.72 \\
0.69 & 0.70 & 0.69 & 0.74 & 0.67 & 0.76 & 0.74 & 0.69 & 0.76 & 0.73 & 0.65 & 0.77 & 0.69 \\
0.72 & 0.70 & 0.74 & 0.70 & 0.72 & 0.76 & 0.78 & 0.74 & 0.76 & 0.74 & 0.77 & 0.77 & 0.73 \\
0.73 & 0.70 & 0.76 & 0.69 & 0.73 & 0.73 & 0.70 & 0.70 & 0.71 & 0.74 & 0.73 & 0.70 & 0.76
\end{bmatrix}
$$

$$
\Delta m = \begin{bmatrix}
88 & 89 & 87 & 80 & 86 & 88 & 90 & 96 & 93 & 94 & 90 & 95 & 86 \\
89 & 90 & 88 & 86 & 92 & 89 & 95 & 92 & 94 & 92 & 95 & 95 & 87 \\
90 & 92 & 93 & 88 & 86 & 91 & 95 & 97 & 93 & 94 & 96 & 94 & 88 \\
89 & 91 & 92 & 90 & 91 & 92 & 94 & 95 & 96 & 93 & 95 & 97 & 89 \\
92 & 90 & 94 & 88 & 92 & 96 & 98 & 94 & 96 & 94 & 97 & 97 & 90 \\
93 & 90 & 96 & 89 & 93 & 97 & 90 & 95 & 96 & 94 & 93 & 95 & 91
\end{bmatrix}
$$

$$
\overline{P} = [0.69 \quad 0.69 \quad 0.69 \quad 0.69 \quad 0.69 \quad 0.69 \quad 0.75 \quad 0.75 \quad 0.75 \quad 0.75 \quad 0.75 \quad 0.75 \quad 0.85]
$$

$$
W = [0.7 \quad 0.6 \quad 0.7 \quad 0.7 \quad 0.7 \quad 0.7 \quad 0.6 \quad 0.7 \quad 0.6 \quad 0.7 \quad 0.6 \quad 0.6 \quad 0.9]
$$

Table 3-1: Number of orbital roots for space targets[225]

Serial number	The name of the satellite	e	$i/(°)$	$\Omega/(°)$	$h/(°)$	$t/(°)$
1	KAZSAT-2	0.001 3	0.006 8	82.815 0	235.408 0	65.576 3
2	AZERSPACE 1	0.002 1	0.008 6	41.364 2	228.155 3	90.470 3
3	INSAT-3D	0.001 4	0.017 0	251.038 2	8.029 4	100.930 5
4	INTELSAT 907	0.003 1	0.008 1	34.893 2	214.263 7	110.823 7
5	INTELSAT 16	0.003 5	0.008 5	88.034 3	183.858 7	117.095 1
6	RADUGA-1M2	0.003 3	0.008 9	320.714 2	332.982 5	126.556 4
7	INTELSAT	0.000 7	0.003 4	294.454 9	53.883 3	138.072 7
8	NSS-10	0.003 1	0.022 7	10.679 6	246.486 4	190.272 9
9	VINASAT-2	0.002 6	0.008 7	2.251 965	271.291 2	210.311 6
10	SCATHA	0.001 4	0.004 3	309.112 9	243.954 1	208.673 2
11	DSCS II F-13	0.001 2	0.003 3	127.881 6	233.586 7	198.274 6
12	DSCS II F-11	0.001 3	0.003 2	191.458 7	232.678 1	195.084 6
13	ZHONGXING-6B	0.004 9	0.009 5	328.046 2	274.604 7	236.420 9

Table 3-2: Number of orbital roots of spacecraft in orbit

Serial number	Spacecraft in orbit	e	$i/(°)$	$\Omega/(°)$	$h/(°)$	$t/(°)$
1	Spacecraft 1	0.005 6	0.014 0	165.356 8	53.189 9	44.787 2
2	Spacecraft 2	0.002 5	0.004 3	294.454 9	9.029 6	93.272 9
3	Spacecraft 3	0.003 6	0.008 8	76.356 0	214.104 5	63.764 5
4	Spacecraft 4	0.003 5	0.007 1	38.782 1	206.459 9	115.823 9
5	Spacecraft 5	0.004 2	0.008 6	298.035 4	259.858 7	242.093 7
6	Spacecraft 6	0.004 5	0.009 7	348.042 2	278.302 8	231.598 0

3.4.2: Model Operation and Analysis

In this case, in the face of 13 synchronous orbit targets at different orbit positions, only 6 in-orbit spacecraft can participate in on-orbit servicing, and in order to meet the service requirements, it is necessary to reasonably allocate service targets to 6 in-orbit spacecraft.

From the perspective of the least power input and the most fuel-saving economy, the "one-to-many" service model should be adopted as much as possible. However, the on-orbit servicing capability of a single spacecraft is limited, and the success probability of service is relatively low, which cannot meet the requirements of 85% success probability of service objective 13. Therefore, it is necessary to adopt a composite service model combining "one-to-many" and "many-to-one." In the on-orbit target allocation under the composite service mode, while pursuing high execution efficiency and meeting the requirements of the success probability of each service, it is necessary to consider the issues of fuel consumption and energy efficiency, as well as the constraints such as assignment mode, timeliness and robustness. If manual operation is adopted, it is necessary to iterate through various policy combinations, which is cumbersome and time-consuming, which is not conducive to rapid response to demand.

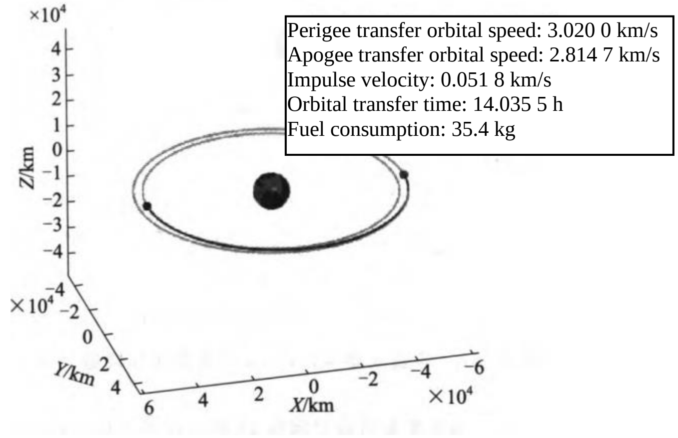
3.4.2.1: Process analysis

In order to obtain a scheme with high timeliness and robustness, and to achieve the unity of optimality and feasibility, it is necessary to fully consider the implementation benefits and energy consumption estimation of each stage of subsequent on-orbit services in the target allocation process.

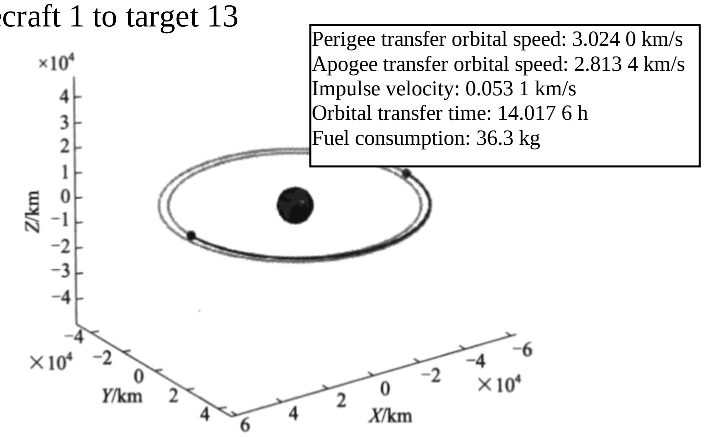
For the estimation of the execution benefit and fuel consumption of the orbital maneuver stage, it is necessary to fully consider the constraints such as designation factors, fuel consumption, timeliness and robustness, and calculate the orbital maneuver duration and fuel consumption of each spacecraft to each space target. Figure 3-16 shows the orbital maneuvering process of spacecraft 1, spacecraft 2, and spacecraft 3 to target 13, and calculates the perigee transfer orbital velocity, apogee transfer orbital velocity, velocity impulse, orbital maneuver duration, and fuel consumption, respectively. For example, spacecraft 1 maneuvered from an initial orbit with an orbital radius of 38, 864 km to a target orbit with an orbital radius of 42, 378 km, with a perigee transfer orbit velocity of 3.02 km/s, an apogee transfer orbit velocity of 2.81 km/s, a velocity impulse of 0.05 km/s, an orbital maneuver duration of 14.03 h, and a fuel consumption of 35.4 kg.

For the estimation of the execution benefit and fuel consumption in the service implementation stage, the execution benefit and fuel consumption estimation of each spacecraft to the space target are calculated by distinguishing between single-aircraft service and multi-aircraft service. Table 3-3 shows the implementation benefits and energy consumption estimates for spacecraft 1, spacecraft 2 and spacecraft 3 to Objective 13. For example, in the service implementation phase, spacecraft 1 has a single execution benefit of 77.4 for target 13, which will increase to 88.3 if it is combined with spacecraft 2, 88.4 if it is combined with spacecraft 3, and 89.8 if it is jointly served by three spacecraft for goal 13.

The energy consumption of spacecraft 1 for target 13 is estimated to be 2.48 kg during the service implementation phase, spacecraft 2 is estimated to be 2.45 kg for target 13 during the service implementation phase, and spacecraft 3 is estimated to consume 2.81 kg for target 13 during the service implementation phase.

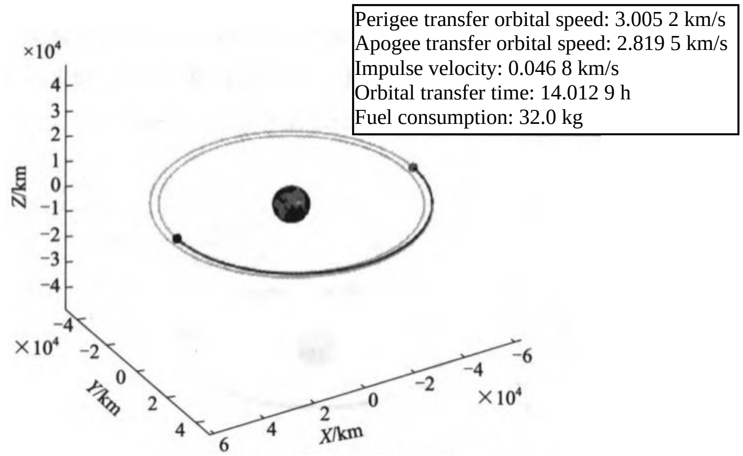


(a) Transfer orbit of spacecraft 1 to target 13



(b) Transfer orbit of spacecraft 2 to target 13

Figure 3-16: The Hohmann Transfer Orbit of Each Spacecraft to Target 13



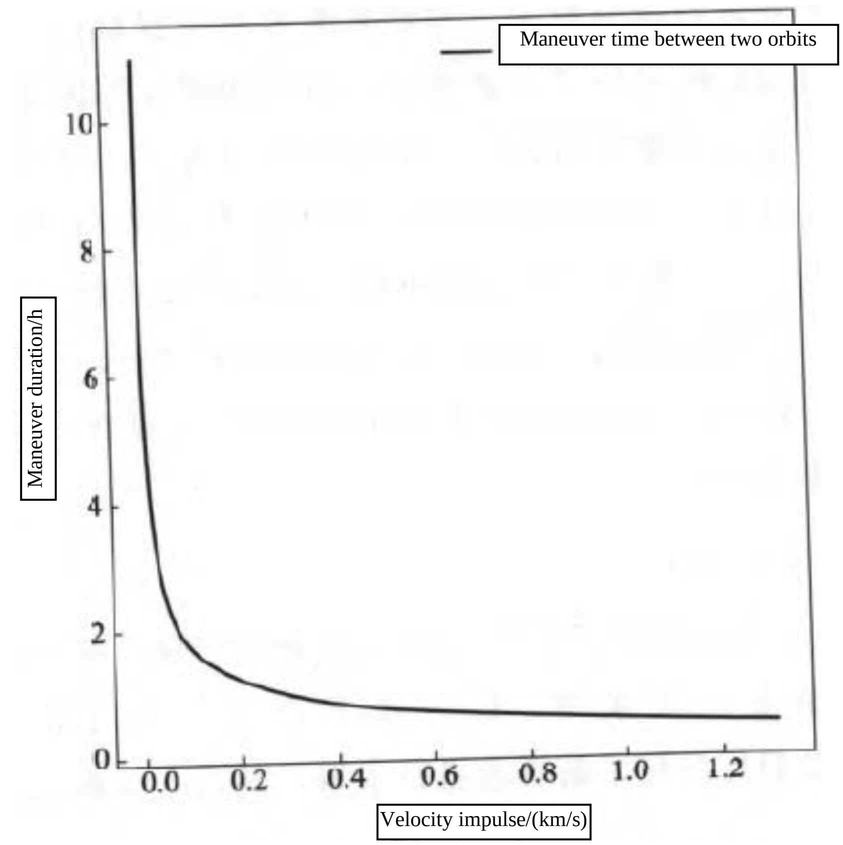
(c) Transfer orbit of spacecraft 3 to target 13

Figure 3-16: The Hohmann Transfer Orbit of Each Spacecraft to Target 13 (continued)

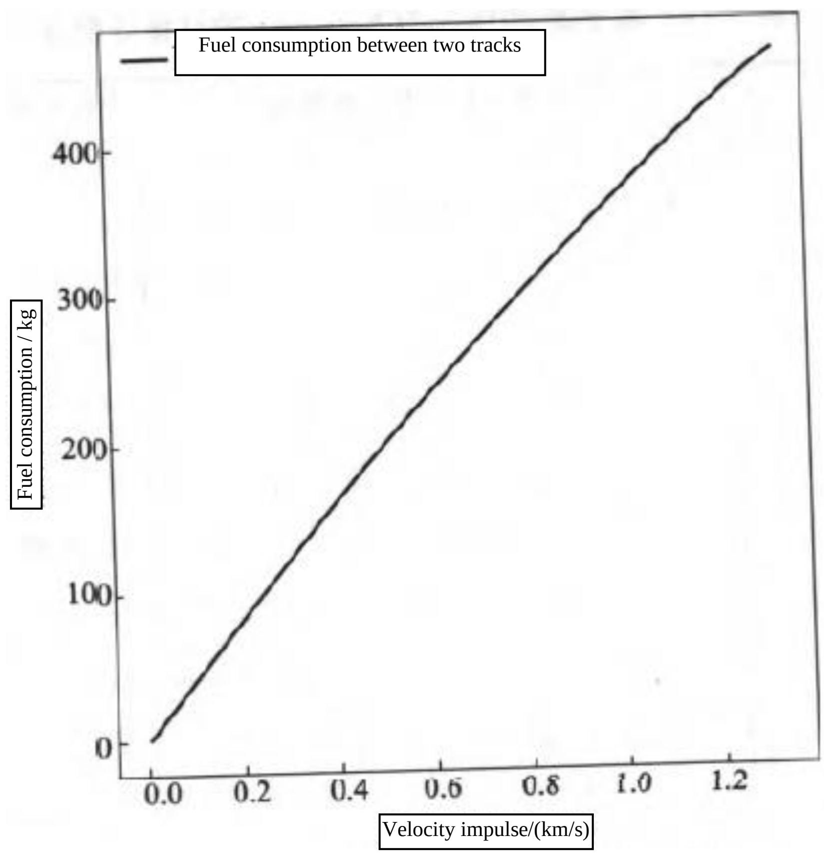
Table 3-3: Estimation of implementation benefits and energy consumption for Goal 13 during the service implementation phase

Spacecraft	Effectiveness of service implementation phase					Estimation of energy consumption during the service implementation phase
	Stand-alone service	Spacecraft 1 and 2 combined	Spacecraft 2, 3 combined	Spacecraft 1, 3 joint	Three-machine combination	Energy consumption estimate to Goal 13 / kg
Spacecraft 1	77.4	88.3	—	88.4	89.8	2.48
Spacecraft 2	78.3		88.5	—		2.45
Spacecraft 3	79.2	—		88.4		2.81

In order to improve the execution efficiency of target allocation, the method of increasing the velocity impulse can be adopted during orbital maneuvering, but this method will also increase the fuel consumption of the spacecraft, resulting in a decrease in energy efficiency. Figure 3-17 shows that the time taken to maneuver to the target decreases nonlinearly with the increase of velocity impulse, and the fuel consumption between the two orbits increases nonlinearly with the increase of velocity impulse. Therefore, it is necessary to estimate and analyze the velocity impulse of each orbital maneuver in the process of target allocation, so as to maximize the execution benefit and improve the energy efficiency as much as possible, promote the balanced development of the execution benefit and energy consumption efficiency, and unify the optimality and feasibility.



(a) The trend of inter-orbit maneuver time with velocity impulse



(b) Trend of fuel consumption with velocity impulse

Figure 3-17: Orbit Transfer Time and Fuel Change

In the allocation of spacecraft on-orbit targets, it is necessary to consider the orbital maneuvering process of each spacecraft for each space target, and measure the orbital maneuvering mode, duration and fuel consumption, so as to select the service target and formulate the orbital maneuver mode for the spacecraft. For this example alone, it is necessary to calculate 78 sets of orbital maneuver data from 6 spacecraft to 13 target satellites. If we consider taking different velocity impulses, the amount of computation will increase to the order of 78^n, which is a large amount of combination and inefficient. Therefore, we give full play to the advantages of fast optimization and self-training of Deep Q Networks and use the target allocation method based on improved Deep Q Networks introduced in this chapter to solve the problem.

3.4.2.2: Policy generation

In order to solve the on-orbit target allocation problem in this composite service mode, according to the problem description and modeling, considering multiple constraints, the target allocation algorithm based on the improved Deep Q Networks introduced in this book is used to solve the problem, and the autonomous allocation process and results are shown in Table 3-4.

Table 3-4: Target allocation process based on improved Deep Q Networks

Serial number	State	Spacecraft target allocation strategy in orbit	Energy efficiency	Execution benefits
①	Initialize	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	0.99	0
②	Process status	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	0.01	0.99
③		$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	0.98	0.58

Continued

Serial number	State	Spacecraft target allocation strategy in orbit	Energy efficiency	Execution benefits
④	Process status	$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1 \\ 1&1&1&1&1&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1&1&1&1&1&0 \end{bmatrix}$	0.92	0.86
⑤	Best strategy	$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 1&1&1&0&1&0&0&1&0&0&1&0&0 \\ 0&0&0&1&0&1&1&0&1&1&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix}$	0.96	0.98

(1) Initialization

Firstly, the assignment strategy of the spacecraft and each space target is initialized by the all-0 matrix, and the orbital maneuvering mode is initialized by the double-pulse maneuvering mode. Then, the target assignment Deep Q Networks network is initialized, the initial state of the network is initialized with the current assigned policy value, and the behavior value Q is initialized with a random number. Finally, the energy efficiency and execution efficiency values of the current allocation strategy ① are obtained through the Deep Q Networks network by taking the current assignment strategy and orbit maneuvering mode as inputs. Among them, since each assignment strategy is 0, that is, there is no substantive assignment task, the execution efficiency is negative.

(2) Autonomous training process

In order to obtain better execution benefits and energy efficiency, the greedy strategy is adopted, and new possibilities are boldly explored while pursuing the balanced development of execution benefits and energy consumption efficiency, and multiple sets of possible results will be obtained in the training process. For example, when the allocation strategy reaches state ②, each spacecraft should correspond to each space target, which greatly reduces the overall energy efficiency. State ③ does not meet the requirement of 85% success probability of service objective 13, which makes the implementation efficiency low; although state ④ takes into account the energy efficiency and execution benefits, it is not the optimal solution of the model.

(3) Optimal distribution scheme

After multiple rounds of iteration and optimization of Deep Q Networks' network autonomous training, it finally converges to the optimal allocation strategy ⑤. The optimal allocation strategy in the composite service mode is as follows:

Spacecraft 3 serves: Target 1, Target 2, Target 3, Target 5, Target 8, and Target 11.

Spacecraft 4 serves: Target 4, Target 6, Target 7, Target 9, Target 10, Target 12.

Spacecraft 5 combined with spacecraft 6 service: Target 13.

In summary, based on the improved target allocation algorithm of Deep Q Networks, the number of targets, the number of orbital roots, and the distribution of spacecraft forces in orbit are taken as inputs, combined with the prior information such as fuel consumption estimation, target priority, and service success probability, and comprehensively considering the assignment constraints, fuel constraints, and timeliness constraints, the two-way training network of Deep Q Networks is used to achieve the balanced development of execution benefit and energy efficiency, and the optimal allocation strategy is obtained ⑤. The optimal allocation strategy not only satisfies the minimum space force input requirements, but also meets the corresponding service success probability requirements, and adopts the "many-to-one" service mode for target 13, and the joint service success probability of spacecraft 5 and spacecraft 6 is 93.5%, which meets the index requirement of 85% service success probability.

3.4.3: Algorithm Comparison

When solving the on-orbit target allocation problem in the composite service mode with multiple optimization objectives and multiple constraints, the scientific and reasonable solution algorithm will be conducive to improving the operation efficiency of the model and the accuracy of the results. Therefore, the comparative analysis between different algorithms is a part of the research on target allocation methods.

(1) Comparison of algorithm applicability

The target allocation problem is a type of nonlinear combinatorial optimization problem, and the optimal solution can only be obtained by using the enumeration method. Algorithms such as implicit enumeration, cut plane and branch delimitation are also often used to solve target allocation problems, but they have the shortcomings that the solution rate decreases with the increase of constraints, and it is difficult to apply to more complex target allocation problems.[222] Genetic algorithms, particle swarm algorithms, ant colony algorithms, and artificial intelligence algorithms have shown good application advantages in target allocation problems due to their extremely low requirements for the continuity, convexity, and analytical expression of the objective function, and their high adaptability to uncertainty.

According to the current research status at home and abroad, although the research literature on target allocation is relatively rich and the research on methods is relatively in-depth, the research in the field of aerospace is still relatively lacking.

For the solution of the on-orbit target allocation problem under the composite service model, it is necessary to use the "one-to-many" and "many-to-one" service modes at the same time, and the influence of different modes on the execution efficiency and energy efficiency needs to be considered, and it is subject to the constraints of assignment constraints, fuel constraints and timeliness constraints. This requires the solution algorithm not only to consider the two service modes at the same time, but also to make trade-offs through benefit comparison in the allocation process. Table 3-5 selects two solutions that have been used in the aerospace field for comparative analysis of their applicability.

Table 3-5: Applicability of Different Target Allocation Algorithms

The name of the algorithm	Characteristics	Advantage	Applicability analysis
Genetic algorithms	1) Work in a coded way, effectively simulating genetic ideas 2) The probabilistic transfer rule is used, which has good global optimization characteristics	1) The optimal solution of the target allocation problem can be obtained 2) The allocation results are independent of the initial environment and have strong robustness	The solution mode is fixed, there are many control variables, and the convergence speed is slow, so the "many-to-one" mode needs to be converted to multiple "one-to-one" modes for solving
Ant colony algorithm	1) It has a positive feedback mechanism to effectively simulate the foraging behavior of ant colonies 2) The initial value is solved randomly, and the optimal is reached through iteration	1) Population-based evolutionary solution method, easy to implement 2) It is easy to perform and easy to combine with other methods	Parameter initialization has a great influence on the solution direction of the algorithm, and eventually the ants will tend to choose the same path, and it is difficult to use the "one-to-many" and "many-to-one" modes at the same time
Target allocation solving algorithm based on improved Deep Q Networks	1) It has a two-way training network mechanism of forward transmission and backward feedback 2) Greedy strategy: boldly explore unknown strategies	1) Not limited by rules, through self-training to achieve independent optimization 2) It is not limited by dimensions, and is suitable for solving complex decision optimization problems	It has the computing advantages of forward transmission and reverse training of neural network, and can consider the two modes of "one-to-many" and "many-to-one" at the same time. The decision-making mechanism with trial and error reward can make the implementation benefit and energy efficiency indicators develop in a balanced manner

The genetic algorithm can give full play to the evolutionary principle of "natural selection, survival of the fittest," and has a good application effect for solving the on-orbit target allocation problem in the "one-to-many" mode. However, due to the limitations of heredity, crossover and mutation rules, the solution mode of genetic algorithm is relatively fixed, and in the face of the on-orbit target allocation problem in the "many-to-one" mode, it needs to be converted into multiple "one-to-one" target allocation problems, which increases the complexity of the solution model.

For the composite service model combining "one-to-many" and "many-to-one," the genetic algorithm itself will be difficult to solve directly, and the composite service model needs to be transformed or decomposed into two independent problems in order to be effectively solved.

The ant colony algorithm is an optimization algorithm that effectively simulates the swarm intelligence behavior of ant colony foraging, and has been effectively used in spacecraft scheduling problems.[223, 224] However, the optimization process of ant colony algorithm strategy directly depends on the accumulation of pheromones on each search path, which makes it difficult to solve the problem of on-orbit target allocation in the composite service mode at the same time.

The improved Deep Q Networks target allocation solving algorithm is a method that uses observation to make decisions and further improves the strategy through continuous interaction with the environment. The algorithm gives full play to the computing advantages of forward transmission and reverse training of neural network, and can consider the two allocation modes of "one-to-many" and "many-to-one" at the same time. With a decision-making mechanism of trial and error reward, the implementation benefit and energy efficiency indicators are developed in a balanced manner, which is conducive to the independent trade-off of the two modes of "one-to-many" and "many-to-one."

(2) Comparison of the amount of algorithm computation

In order to solve the problem of on-orbit target allocation under the composite service model, an attempt was made to draw on the research results of on-orbit servicing target allocation[225, 208, 227-229] to solve the problem, but it was found that these methods need to solve the target allocation problem of the "one-to-one" or "one-to-many" mode separately, which is not suitable for the composite service mode involved in the case. In order to compare and analyze the computational amount of different algorithms, we have to consider the "one-to-many" target allocation problem in the case separately and use three algorithms to solve it.

The simulation operation relies on 1.6 GHz and 1.8 GHz dual-core CPU, 8 GB RAM computing hardware, is performed using the python language PyCharm compilation environment, and the amount of computation of each method is shown in Figure 3-18. Among them, the ant colony algorithm uses the global search method to calculate the overhead, and the different starting directions lead to large fluctuations in the operation time, with an average time of 0.32 s. The genetic algorithm failed to use the feedback information, the training time was relatively long, and the operation time fluctuated greatly due to random cross-variation, with an average time of 0.19 s. Based on the improved target allocation solving algorithm of Deep Q Networks, the neural network is used to achieve the shortest autonomous training time, and the operation time fluctuates in a small range due to the use of exploration and utilization strategies, with an average time of 0.06 s.

Therefore, the target allocation solving algorithm based on the improved Deep Q Networks can give full play to the computing advantages of forward transmission and reverse training of neural networks and use the decision-making mechanism of reinforcement learning trial and error reward, which is more efficient and more suitable for the on-orbit target allocation problem in the composite service mode.

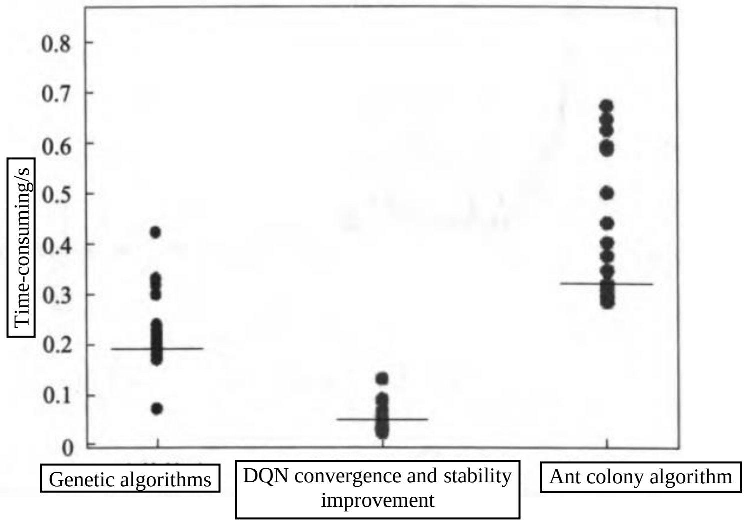


Figure 3-18: Comparison of the Computational Costs of the Three Algorithms

At the same time, the classical Deep Q Networks method is used to solve the problem, and the same results are obtained, which confirms the accuracy of the results. As shown in Figure 3-19, the difference between the two methods makes the algorithm differentiate and process the fully connected layer, which makes it possible to achieve a training effect of 0.01 error after only 70 learning times, and the error function value of the whole training process also decreases at a rate nearly doubled, and the improvement effect in convergence is obvious.

As shown in Figure 3-20, the algorithm introduces another neural network into the behavior estimation to ensure that the reward value rises rapidly and fluctuates less, and the optimal reward value can remain around 0.197 8 after only 33 times of autonomous training, which fully reflects the improved advantage of stability.

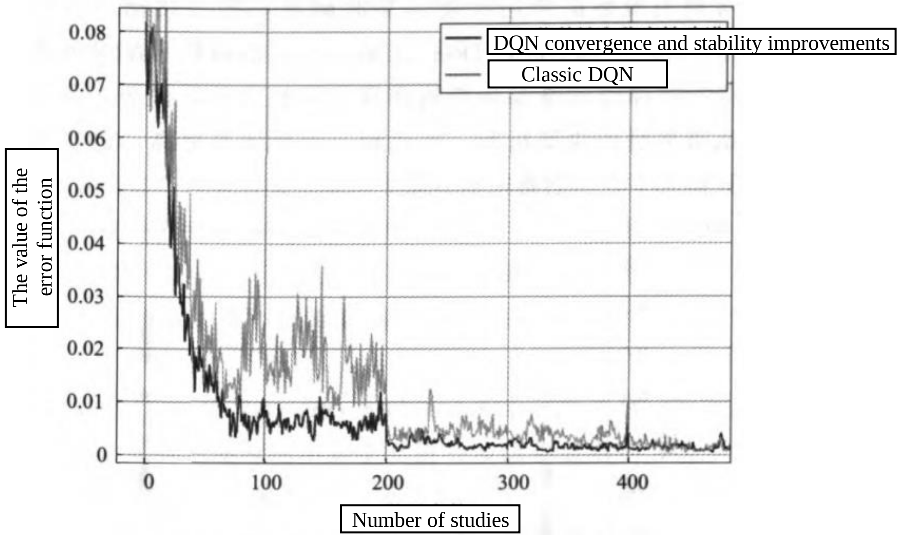


Figure 3-19: Comparison of Error Function Values Between the Two Methods (see color insert)

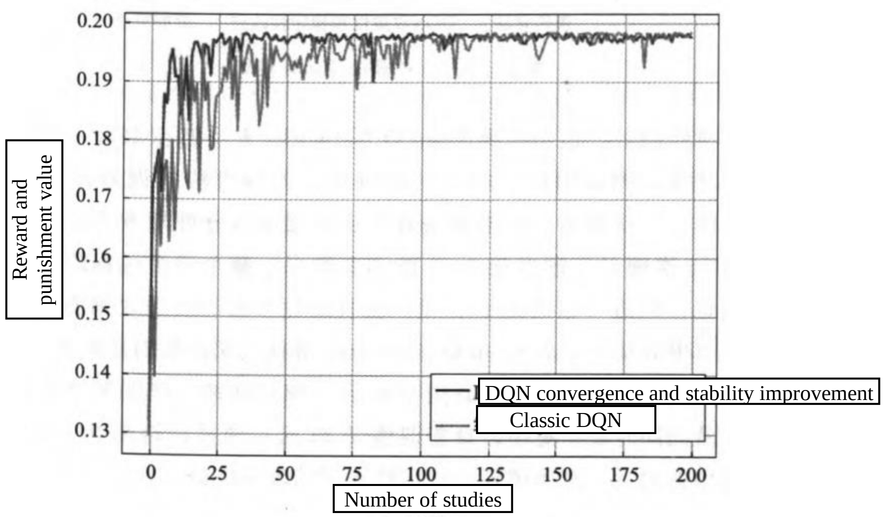


Figure 3-20: Comparison of the Reward Values of the Two Methods (see color insert)

3.5: CHAPTER SUMMARY

In view of the fact that there are many scattered targets and limited on-orbit servicing force, the problem of on-orbit target allocation under the composite service mode is studied, and the target allocation solution algorithm based on the improved Deep Q Networks is introduced. Firstly, in the face of the "nonlinear combinatorial optimization" feature of spacecraft on-orbit target allocation, the on-orbit target allocation problem under the composite service mode is described. Secondly, in order to meet the characteristics of the on-orbit servicing target allocation problem and make up for the reward bias and overestimation problems of the classical method, the convergence and stability of the Deep Q Networks method are improved. Afterwards, the balanced development of the execution benefit and energy efficiency was hindered, and the target allocation dual training network was built. The case analysis shows that the method can effectively deal with the adverse effects of uneven supply and demand, give full play to the advantages of the composite service model, achieve the balanced development of execution benefit and energy efficiency, and achieve the comprehensive goal of less force investment and higher expected success rate, which can provide effective auxiliary decision-making for on-orbit services. The comparison of the algorithms shows that compared with other commonly used algorithms, the training time of the algorithm in this book is shortened by 80% on average, and compared with the classical Deep Q Networks algorithm, the training error reduction rate is twice as fast, and the reward value fluctuates less while rising rapidly, which further improves the shortcomings that the conventional method is difficult to directly apply to the composite service mode and the computing time is high.

According to the optimal allocation strategy obtained by the on-orbit target allocation method in the composite service mode, the in-orbit spacecraft can perform the on-orbit servicing task by planning and designing the space transfer orbit in advance and formulating the service implementation plan. In the course of the operation, when the spacecraft flying along the set transfer orbit suddenly encounters space debris, in order to avoid space debris, it is necessary to temporarily carry out path planning and take orbital avoidance maneuvers in time. Temporary orbital avoidance should not only successfully avoid space debris, but also consider the timeliness of orbital avoidance and the optimality of fuel consumption, and at the same time, it is also limited by factors such as orbit offset, avoidance mode and braking timeliness, resulting in a complex multi-restriction shortest path problem.[230]

4.1: ORBITAL AVOIDANCE PROBLEM DESCRIPTION AND MODELING

This section analyzes the shortcomings of the pre-planning method of orbital maneuver path, describes the space debris that spacecraft may encounter and the avoidance methods adopted, and clarifies the focus of orbital avoidance path planning. Combined with the relative motion of spacecraft and space debris, the orbit avoidance control model is constructed, and the orbit avoidance constraints are established by considering the factors of control, consumption, deviation and agility.

4.1.1: Problem Description

Spacecraft orbital maneuvering is a flight process that changes the original free-flight trajectory and enters another orbit under the action of the control system, which is a necessary process to complete the long-distance on-orbit servicing mission. Orbital maneuver path planning is usually based on certain task constraints, orbital dynamics and spatio-temporal systems, and uses control theory, mathematical programming theory and optimization methods to design a transfer orbit that meets the mission requirements and has the lowest cost. However, in the process of actual orbital maneuvering, the pre-planned transfer orbit (set transfer orbit) is often difficult to meet the actual needs, and the main reasons are as follows:

1) Uncertainties are difficult to grasp. In the actual orbital maneuvering path planning, it is often necessary to comprehensively consider various factors such as spacecraft maneuverability, space environment, debris threat and maneuvering mission, but it is very difficult to establish a planning model that can include all these uncertainties in advance.

2) It is difficult to know about emergencies. In the process of orbital maneuvering and service implementation of spacecraft, unknown space debris attacks may occur at any time, and it is difficult to know and predict these unexpected situations in advance planning, so it is difficult to fully apply the transfer orbit planned in advance.

3) The objective function is difficult to determine. The optimal transfer trajectory may change depending on the environmental situation, mission requirements, and unexpected circumstances, so the objective function determined in advance often does not reflect the mission requirements of a particular situation.

4) It is difficult to choose the orbit. For the pre-set objective function, there may be multiple transition orbit with the same cost in some cases, and the pre-set trade-off strategy is difficult to adapt to various temporary situations.

The temporary events and uncertainties that spacecraft are most likely to face during orbital maneuvering and service implementation mainly come from space debris. In order to avoid the temporary occurrence of space debris, spacecraft flying along a given transfer orbit need to adopt orbital avoidance maneuvers so that at a certain moment they deviate from the space debris radially or tangentially. Avoidance maneuver is an emergency maneuver taken by a spacecraft to avoid space debris during the process of traveling along a given transfer orbit, and whether it is smooth or not is directly related to the realization of the entire on-orbit servicing intention and affects the success or failure of follow-up actions. In the meantime, it is necessary to take a holistic view of the overall situation, judge the hour and size up the situation, give resolute commands, handle matters calmly, and always maintain the initiative in action.

The spacecraft temporarily adopts orbital avoidance maneuver, and it will no longer be able to use the traditional orbital planning method carried out in advance on the ground but will need to carry out the temporary planning of orbital avoidance path. As an in-orbit active defense technology, orbital temporary avoidance path planning can effectively improve the safety, reliability and survivability of spacecraft in orbit. In the process of planning the temporary avoidance path of the orbit, it is necessary to pay attention to the following points:

1) Grasp the timing of avoidance. In the process of orbital temporary avoidance path planning, it is necessary to pay attention to grasp the timing of initiating avoidance and choose the time when there is still a distance to approach space debris, and the spacecraft has considered the cost of avoidance and is fully prepared.

2) Selection of evasion strategies. We need to carefully choose the avoidance strategy and method, and try to make the situation of all parties favorable to us, but we should not be indecisive so as not to lose the most favorable opportunity to avoid avoidance. 3) Different preference settings. Different avoidance preferences can be set before the orbital temporary avoidance path planning, that is, different factors can be considered to achieve space debris avoidance, such as fuel consumption as the preference, and the minimum fuel consumption to achieve space debris avoidance. The setting of different avoidance preferences will be able to meet more task requirements and more task scenarios, but this requires more flexibility of the path planning method.

In summary, it can be seen that the optimal avoidance path should not only guide the spacecraft to avoid space debris smoothly, but also consider the fuel consumption of the avoidance maneuver, and take into account the avoidance safety, minimum offset and braking time, which belongs to a class of multi-restriction shortest path problems.

4.1.2: Modeling Spacecraft Orbit Avoidance Problems

In order to highlight the focus of research and reveal the essence of the problem, it is necessary to use a highly abstract mathematical model to study the orbit avoidance problem in the process of mathematical modeling and scientific calculation.

1) Weaken the influence of the Earth's rotation on the orbital maneuver of spacecraft and ignore the problem of multi-body perturbation for the time being.

2) It is assumed that the chasing spacecraft and space debris are in low Earth orbit throughout the spacecraft orbit avoidance process.

3) In the process of orbit avoidance, it is assumed that the spacecraft has strong space situational awareness, timely command and control, advanced power system and strong on-orbit computing capabilities.

4.1.2.1: Orbital avoidance control model

If a spacecraft flying along a given transfer orbit suddenly encounters space debris, and the geocentric vector diameter of the spacecraft is $\boldsymbol{r}_{\mathrm{P}}$ and the geocentric vector diameter of the space debris is $\boldsymbol{r}_{\mathrm{O}}$, then the position vector of the spacecraft relative to the space debris is ψ

$$\boldsymbol{\Psi} = \boldsymbol{r}_{\mathrm{P}} - \boldsymbol{r}_{\mathrm{O}} \qquad (4\text{-}1)$$

Assuming that both the spacecraft and the space debris are subjected to a central gravitational field inversely proportional to the square of the distance, there is

$$\begin{cases} \dfrac{d^2 \boldsymbol{r}_P}{dt^2} = \boldsymbol{a}_P - \mu \dfrac{\boldsymbol{r}_P}{r_P^3} \\ \dfrac{d^2 \boldsymbol{r}_O}{dt^2} = -\mu \dfrac{\boldsymbol{r}_O}{r_O^3} \end{cases} \tag{4-2}$$

where a_P is the acceleration vector of spacecraft orbit avoidance, $\boldsymbol{a}_P = [\boldsymbol{a}_{Px}, \boldsymbol{a}_{Py}, \boldsymbol{a}_{Pz}]$; and μ is the gravitational constant.

Subtract the two formulas of Eq. (4-1) to get it

$$\frac{d^2 \boldsymbol{\psi}}{dt^2} = \frac{d^2 \boldsymbol{r}_P}{dt^2} - \frac{d^2 \boldsymbol{r}_O}{dt^2} = \boldsymbol{a}_P - \mu \frac{\boldsymbol{r}_P}{r_P^3} + \mu \frac{\boldsymbol{r}_O}{r_O^3} \tag{4-3}$$

where $d^2\boldsymbol{\psi}/dt^2$ is the difference between the absolute acceleration of the spacecraft and the space debris.

According to the relationship between relative motion and absolute motion,[201] it can be seen that

$$\frac{d^2 \boldsymbol{\psi}}{dt^2} = \frac{\delta^2 \boldsymbol{\psi}}{\delta t^2} + 2\boldsymbol{\omega} \frac{\delta \boldsymbol{\psi}}{\delta t} + \boldsymbol{\omega}^2 + \boldsymbol{\omega}\boldsymbol{\psi} + \boldsymbol{\varepsilon}\boldsymbol{\psi} \tag{4-4}$$

where $\delta\boldsymbol{\psi}/\delta t$ and storage $\delta^2\boldsymbol{\psi}/\delta t^2$ are the relative velocity and acceleration of the spacecraft and space debris, respectively; $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$ are the rotational angular velocity and angular acceleration of space debris, respectively.

By combining Eq. (4-3) and Eq. (4-4), the relative motion model of the spacecraft and space debris vector form can be obtained

$$\frac{\delta^2 \boldsymbol{\psi}}{\delta t^2} + 2\boldsymbol{\omega} \frac{\delta \boldsymbol{\psi}}{\delta t} + \boldsymbol{\omega}^2 + \boldsymbol{\omega}\boldsymbol{\psi} + \boldsymbol{\varepsilon}\boldsymbol{\psi} = \boldsymbol{a}_P + \mu \frac{\boldsymbol{r}_O - (\boldsymbol{r}_O/\boldsymbol{r}_P)^3 \boldsymbol{r}_P}{r_O^3} \tag{4-5}$$

Let x^0, y^0 and z^0 be the unit vectors of the three axes of the space Cartesian coordinate system, respectively, and we can obtain

$$\begin{cases} \boldsymbol{\psi} = x\boldsymbol{x}^0 + y\boldsymbol{y}^0 + z\boldsymbol{z}^0 \\ \boldsymbol{r}_P = (x + r_O)\boldsymbol{x}^0 + y\boldsymbol{y}^0 + z\boldsymbol{z}^0 \\ \boldsymbol{r}_O = r_O \boldsymbol{x}^0 \\ \dfrac{\delta \boldsymbol{\psi}}{\delta t} = \dot{x}\boldsymbol{x}^0 + \dot{y}\boldsymbol{y}^0 + \dot{z}\boldsymbol{z}^0 \\ \dfrac{\delta^2 \boldsymbol{\psi}}{\delta t^2} = \ddot{x}\boldsymbol{x}^0 + \ddot{y}\boldsymbol{y}^0 + \ddot{z}\boldsymbol{z}^0 \\ \boldsymbol{\omega} = \omega \boldsymbol{z}^0, \boldsymbol{\varepsilon} = \varepsilon \boldsymbol{z}^0 \end{cases} \tag{4-6}$$

When the relative distance between the spacecraft and the space debris $\boldsymbol{\psi}$ much smaller than the orbital radius of the spacecraft

r_P, the gravitational acceleration difference between the spacecraft and the space debris can be processed linearly, which can be obtained from the second equation of Eq. (4-6).

$$\begin{aligned} r_P &= [(r_O + x)^2 + y^2 + z^2]^{\frac{1}{2}} \\ &= (r_O^2 + 2r_O x + \boldsymbol{\psi}^2)^{\frac{1}{2}} \end{aligned} \tag{4-7}$$

Thus, there is

$$\left(\frac{r_O}{r_P}\right)^3 = \left[1 + \frac{2x}{r_O} + \left(\frac{\boldsymbol{\psi}}{r_O}\right)^2\right]^{-\frac{3}{2}} \tag{4-8}$$

Therefore, by substituting Eq. (4-6) and Eq. (4-8) into Eq. (4-5), the control model of spacecraft to avoid space debris can be obtained

$$\begin{cases} \ddot{x} - 2\omega \dot{y} - \omega^2 x - \varepsilon y - \dfrac{2\mu x}{r_O^3} = a_{Px} \\ \ddot{y} + 2\omega \dot{x} - \omega^2 y + \varepsilon x + \dfrac{\mu y}{r_O^3} = a_{Py} \\ \ddot{z} + \dfrac{\mu z}{r_O^3} = a_{Pz} \end{cases} \tag{4-9}$$

4.1.2.2: Orbital circumvention constraints

The purpose of spacecraft orbit temporary avoidance path planning is to plan the optimal avoidance path for spacecraft under the constraints of space environment, spacecraft flight control, orbit mechanics, etc. In addition, the avoidance orbit path planning also needs to focus on the issue of orbit keeping, that is, to ensure the successful avoidance of space debris while returning to the set transfer orbit as soon as possible, and to minimize the time and fuel consumption of the evasion maneuver. Therefore, according to the establishment of the above orbit avoidance control model and the requirements of avoidance path planning, the following constraints are mainly considered for spacecraft orbit avoidance problem:

(1) Circumvention of control constraints

Considering the finite nature of the fuel carried by the spacecraft and the fact that the acceleration of the avoidance maneuver cannot exceed the rated load capacity of the propulsion engine, the acceleration change rate of the spacecraft avoidance maneuver is conditionally constrained

$$\sum_{n=1}^{N} \Delta J_n \leqslant AC \tag{4-10}$$

where ΔJ is the rate of change of acceleration of the spacecraft's orbital avoidance maneuver; N is the number of times the propulsion engine is controlled; AC is the rated rate of change of acceleration of the spacecraft thruster.

(2) Orbital deviation constraints

In the process of space debris avoidance, spacecraft should consider orbit keeping, that is, it should return to the set transfer orbit as soon as possible after successful avoidance, so as to ensure that the mission can be completed as planned and reduce the impact of uncertainties

$$\int_{t_1}^{t_2} \Delta D_t \, dt \leqslant Dis \tag{4-11}$$

where ΔD_t is the offset of the spacecraft from the set transfer orbit per unit time; Dis is the maximum cumulative deviation within the maneuvering control range of the spacecraft.

(3) Braking agility constraints

Both spacecraft and space debris fly at high speed, and the braking sensitivity of the spacecraft should be paid attention to in the orbit avoidance process to avoid instability

$$\sum_{n=1}^{N} \Delta T_n \leqslant Terr \tag{4-12}$$

where ΔT_n is the braking agility of the spacecraft to avoid maneuvering; $Terr$ is the rated cumulative braking error.

(4) Fuel consumption constraints

Fuel constraint means that the energy consumption used for orbital avoidance must not exceed the total amount of fuel left over from the spacecraft for orbital maneuvering and service execution

$$\int_{t_1}^{t_2} \dot{m} \, dt \leqslant m - \Delta m_{\text{maneuver}} - \Delta m_{\text{counter}} \tag{4-13}$$

where $\dot{m}$ is the fuel consumption per unit time of the spacecraft's orbital avoidance maneuver; m is the total amount of fuel carried by the spacecraft; $\Delta m_{maneuver}$ is fuel consumption estimates for the orbital maneuvering phase; $\Delta m_{counter}$ is fuel consumption estimates for the service implementation phase.

4.2: Orbit Avoidance Indicator Model Based on Frenet Coordinate System

In the establishment of the spacecraft orbit avoidance index model, in addition to considering the avoidance of space debris, it is also necessary to continue to travel along the set transfer orbit as much as possible, and there will be both the relative motion of the spacecraft and the target and the absolute motion of the set transfer orbit, and the coordinate expression is relatively complex, which increases the difficulty of modeling.

Therefore, this section firstly constructs a Frenet coordinate system for the orbital maneuvering characteristics of spacecraft to avoid space debris, so as to solve the problem that the relative position of the spacecraft and the set transfer orbit in path planning is not easy to express. Secondly, the conversion relationship between the Frenet coordinate system and the Cartesian coordinate system is given, which is convenient for the application of the module. Finally, considering the factors of avoidance safety, orbit keeping, braking aging and fuel consumption, the orbit avoidance path planning index model is constructed to meet different tasks and preferences.

4.2.1: Frenet-Based Spatial Motion Coordinate System

When spacecraft avoids space debris, it usually adopts the avoidance strategy of lateral offset, longitudinal offset, or offset along the vertical orbital plane in the orbital plane. The velocity increment required for the lateral offset strategy is smaller and consumes less fuel than the other two evasion strategies.[107] Therefore, this section takes the lateral offset strategy as an example to study the autonomous planning of spacecraft orbital avoidance paths, and other strategies can also be used as a reference.

The path planning process of orbit avoidance involves the approximation or even rendezvous with space debris, and it is difficult to clearly distinguish the relative motion relationship between the two if the absolute orbital parameters (e.g., the number of six orbits) are used to describe the relative motion process.[233] Therefore, this kind of problem is mostly described in conjunction with the relative equation of motion. The relative motion of space can be represented by a variety of coordinate systems, among which the Earth Centered Inertial (ECI) coordinate system[1] is more conducive to the description of the orbit around the geocentric, the Line of Sight (LOS) coordinate system[2] is conducive to the description of the relative motion of the spacecraft and the space target in inertial space, and the spacecraft Body Fixed (BF) coordinate system[3] is conducive to the description of the spacecraft's own rotational motion.

---

[1] Geocentric inertial coordinate system is a spatial Cartesian coordinate system established with the Earth's center of mass as the origin, or a geodetic coordinate system established with the Earth's ellipsoid plane coincident with the Earth's center of mass as the datum. In the coordinate system, the intersection of the meridian plane and the equatorial plane is the X-axis, the Z-axis coincides with the earth's axis of rotation, and the Y-axis is perpendicular to the XZ plane to form the right-hand system.

[2] Line-of-sight coordinate system, also known as Range Horizontal Vertical (RHV) coordinate system, its X-axis is along the direction of the target line of sight, the Y-axis is in the horizontal plane, and the Z-axis is determined by the right-hand spiral theorem.

[3] The coordinate system of the spacecraft body, also known as the aircraft coordinate system, the missile body coordinate system or the rocket body coordinate system, etc., has its coordinate origin O at the center of mass of the spacecraft, the X axis points to the head along the longitudinal axis of the spacecraft, the Y axis is perpendicular to the X axis in the longitudinal symmetry plane of the spacecraft, and the Z axis and the X axis and the Y axis constitute a right-hand Cartesian coordinate system.

However, spacecraft avoidance maneuvers should not only consider the absolute motion along the given transfer orbit, but also take into account the relative motion of space debris avoidance, and the use of the above coordinate system representation is cumbersome and not conducive to calculation.

In order to solve the problem that the relative position of the spacecraft and the set transfer orbit is not easy to express in path planning, the Frenet coordinate system shown in Figure 4-1[234, 235] was constructed to describe the evasive maneuver process of the spacecraft. The Frenet coordinate system refers to the right-hand Cartesian coordinate system that uses the spacecraft as the reference point and the set transfer orbit as the reference line, and the tangential vector $\vec{t}_r$ and normal vector $\vec{n}_r$ of the spacecraft's flight trajectory are formed along the reference line. In the coordinate system, the longitudinal offset *s* (i.e., the displacement along the direction of the reference line) and the transverse offset *d* (i.e., the displacement along the normal direction deviating from the reference line) are used to describe the position coordinates *(s, d)* of any point in the process of spacecraft evasion maneuver.

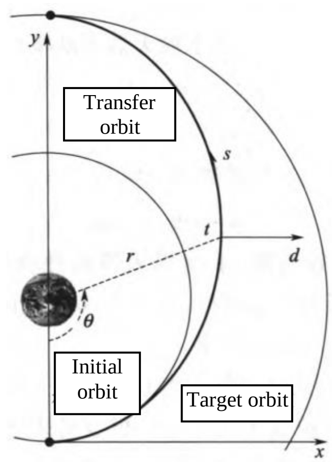


Figure 4-1: Space Motion of a Spacecraft Based on the Frenet Coordinate System

In the Frenet coordinate system, the relationship between the tangential vector self and the normal vector sink is as follows[236]

$$\begin{bmatrix} \dfrac{d\vec{t}_r}{ds} \\ \dfrac{d\vec{n}_r}{ds} \end{bmatrix} = \begin{bmatrix} 0 & k \\ -k & 0 \end{bmatrix} \begin{bmatrix} \vec{t}_r \\ \vec{n}_r \end{bmatrix} \tag{4-14}$$

where k is the curvature of the current point.

Curvature is a measure of how fast a tangent changes its rate along a reference line[237]

$$k = \lim_{\Delta s} \left| \frac{\Delta\theta}{\Delta s} \right| = \frac{d\vec{t}_r}{ds} \tag{4-15}$$

where θ is the tangential angle.

In the Frenet coordinate system, the evasive maneuver of the spacecraft can be expressed as a vector-valued function with respect to the arc length[100]

$$\vec{r}(s) = x(s)\vec{i} + y(s)\vec{j} \tag{4-16}$$

Therefore, with the set transfer orbit as the dynamic reference line, the coordinates of the spacecraft at time t are

$$\vec{x}(s(t), d(t)) = \vec{r}(s(t)) + d(t)\vec{n}_r(s(t)) \tag{4-17}$$

Therefore, the spacecraft motion model in the Frenet coordinate system can be expressed as

$$\begin{cases} \dot{s} = \dfrac{v\cos\phi}{1 - dk_r} \\ \dot{d} = v\sin\phi \\ \dot{\phi} = vk_v - \dot{s}k_r \end{cases} \tag{4-18}$$

where v is the flight speed of the spacecraft; ϕ is the heading deviation of the corresponding reference line of the spacecraft, $\phi = \vartheta - \vartheta_r$, ϑ is the heading angle of the spacecraft, ϑ_r is the heading angle of the reference line; k_v and k_r, respectively, are the steering curvature of the spacecraft and the curvature of the reference line.

The description of spacecraft avoidance maneuver based on the Frenet coordinate system is only related to the selection of reference lines, and has nothing to do with the absolute position of the spacecraft, so that it is easier to express the deviation of the spacecraft along the set transfer orbit. The representation of the spacecraft avoidance maneuver path is decomposed into two directions related to the set transfer orbit (reference line), which not only conforms to the reality of spacecraft avoidance maneuver, simplifies the path planning model, reduces the complexity of the space motion model, but also is simple in the way of expression, convenient for solving the state differential equation, and conducive to improving the computational efficiency.

4.2.2: Cartesian Transformation of the Frenet Coordinate System

The temporary avoidance path planning of spacecraft orbit based on the Frenet coordinate system will output separate transverse and longitudinal paths, and the output of the final avoidance path needs to be directly applied by the execution module, so it is necessary to output the path planning results obtained in the Frenet coordinate system to the Cartesian coordinate system.[1][239, 240]

[1] Cartesian Coordinates is a collective term for Cartesian coordinate system and oblique coordinate system. The two axes of number that intersect at the origin form a planar radial coordinate system. If the units of measurement on the two number axes are equal, the radial coordinate system is called the Cartesian coordinate system. A Cartesian coordinate system in which two number axes are perpendicular to each other is called a Cartesian Cartesian coordinate system, otherwise it is called a Cartesian oblique coordinate system.

In order for the avoidance path to be feasible and operable, the resulting first-order differentiation of the curvature of the path with respect to time and the first-order differentiation of the acceleration of the spacecraft evasion maneuver with respect to time must be continuous. For this reason, when the initial state and its first- and second-order differential functions are defined, $d(t)$ and $s(t)$ are characterized by the fifth-order polynomial[1].

$$d(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4 + \alpha_5 t^5 \qquad (4\text{-}19)$$

$$\dot{d}(t) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2 + 4\alpha_4 t^3 + 5\alpha_5 t^4 \qquad (4\text{-}20)$$

$$\ddot{d}(t) = 2\alpha_2 + 6\alpha_3 t + 12\alpha_4 t^2 + 20\alpha_5 t^3 \qquad (4\text{-}21)$$

$$s(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4 + \alpha_5 t^5 \qquad (4\text{-}22)$$

$$\dot{s}(t) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2 + 4\alpha_4 t^3 + 5\alpha_5 t^4 \qquad (4\text{-}23)$$

$$\ddot{s}(t) = 2\alpha_2 + 6\alpha_3 t + 12\alpha_4 t^2 + 20\alpha_5 t^3 \qquad (4\text{-}24)$$

where s is the longitudinal offset of the spacecraft; $\dot{s}$ is the longitudinal offset rate of the spacecraft; $\ddot{s}$ is the longitudinal offset acceleration of the spacecraft; d is the lateral offset of the spacecraft; $\dot{d}$ is the lateral offset rate of the spacecraft; $\ddot{d}$ is the lateral offset acceleration of the spacecraft.

Eq. (4-19) to Eq. (4-24) are expressed in matrix form

[1] There are many higher-order polynomials used to describe paths, such as cubic polynomials, fifth polynomials, and seventh polynomials. Among them, although the cubic polynomial can ensure a certain smoothness of the path and the continuity of position and velocity, it cannot specify the acceleration boundary condition, and the smoothness degree is easily affected by the motion characteristics and inertial load. As the order increases, the planning time will also increase, which affects the timeliness of the planning model. Therefore, the spacecraft avoidance motion process described by the fifth polynomial is the optimal choice for the first-order differentiation of each parameter with respect to time.[240]

$$\begin{bmatrix}1 & t & t^2 & t^3 & t^4 & t^5 \\ 0 & 1 & 2t & 3t^2 & 4t^3 & 5t^4 \\ 0 & 0 & 2 & 6t & 12t^2 & 20t^3 \\ 0 & t & t^2 & t^3 & t^4 & t^5 \\ 0 & 1 & 2t & 3t^2 & 4t^3 & 5t^4 \\ 0 & 0 & 2 & 6t & 12t^2 & 20t^3\end{bmatrix}\begin{bmatrix}\alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5\end{bmatrix}=\begin{bmatrix}d(t) \\ \dot{d}(t) \\ \ddot{d}(t) \\ s(t) \\ \dot{s}(t) \\ \ddot{s}(t)\end{bmatrix} \quad (4\text{-}25)$$

causes

$$\boldsymbol{W}=\begin{bmatrix}1 & t & t^2 & t^3 & t^4 & t^5 \\ 0 & 1 & 2t & 3t^2 & 4t^3 & 5t^4 \\ 0 & 0 & 2 & 6t & 12t^2 & 20t^3 \\ 0 & t & t^2 & t^3 & t^4 & t^5 \\ 0 & 1 & 2t & 3t^2 & 4t^3 & 5t^4 \\ 0 & 0 & 2 & 6t & 12t^2 & 20t^3\end{bmatrix}, \boldsymbol{P}=\begin{bmatrix}d(t) \\ \dot{d}(t) \\ \ddot{d}(t) \\ s(t) \\ \dot{s}(t) \\ \ddot{s}(t)\end{bmatrix} \quad (4\text{-}26)$$

So there is

$$\boldsymbol{C} = \boldsymbol{W}^{-1}\boldsymbol{P} \quad (4\text{-}27)$$

According to the fifth-order polynomial representation of $[d, \dot{d}, \ddot{d}, s, \dot{s}, \ddot{s}]$, the Frenet coordinates are converted to Cartesian coordinates, and the state of the spacecraft at any time t can be expressed as $[\vec{x}, \theta_x, k_x, v_x, a_P]$. Among them, $\vec{x}$ is the current position of the spacecraft, which can be represented by longitudinal offset and lateral offset $\vec{x}$=(s, d); θ_x is the current azimuth angle of the spacecraft; k_x is the curvature of the maneuvering path; v_x for the speed of maneuvering of the spacecraft; a_P is avoidance maneuver acceleration for spacecraft.

As shown in Figure 4-2, when the spacecraft takes an evasive maneuver, its flight trajectory no longer coincides with the set transfer orbit. In Cartesian coordinates, the position vector projected from the current position O of the spacecraft to the P point in a given transfer orbit is denoted as $\vec{r}$(s)=(x_r, y_r), and the spatial distance between the current position of the spacecraft and the projection point is

$$d = \pm\sqrt{(x-x_r)^2+(y-y_r)^2} \quad (4\text{-}28)$$

where $(y - y_r)\cos\theta_r - (x - x_r)\sin\theta_r > 0$, and distance d are positive; otherwise, the distance d is negative.

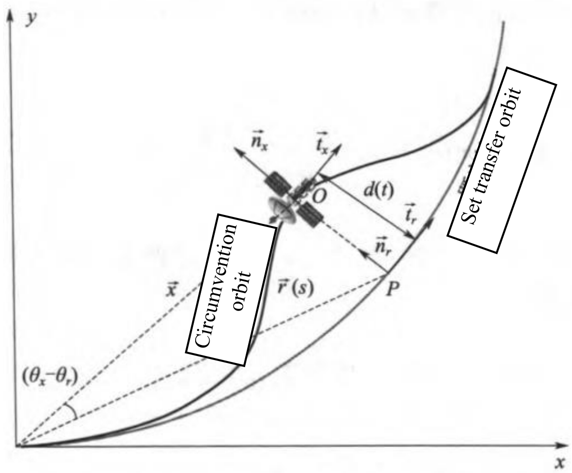


Figure 4-2: Cartesian Conversion of the Frenet Coordinate System

$\vec{t}_x$ and $\vec{n}_x$ is the unit orthogonal vector of the current position of the spacecraft, $\vec{t}_r$ and $\vec{n}_r$ is the unit orthogonal vector of the projection point on a given transfer orbit. From this, the lateral offset of the spacecraft can be expressed as

$$d = [\vec{x} - \vec{r}(s)]^{\mathrm{T}} \cdot \vec{n}_r \quad (4\text{-}29)$$

The spacecraft lateral velocity is expressed as

$$\begin{aligned} \dot{d} &= [\dot{\vec{x}} - \dot{\vec{r}}(s)]^{\mathrm{T}} \cdot \vec{n}_r + [\vec{x} - \vec{r}(s)]^{\mathrm{T}} \cdot \dot{\vec{n}}_r \\ &= v_x \vec{t}_x \vec{n}_r \\ &= v_x \sin(\theta_x - \theta_r) \end{aligned} \quad (4\text{-}30)$$

where, $\dot{\vec{n}}_r = -k_r \vec{t}_r$.

It can be obtained from Eq. (4-30).

$$\theta_x = \theta_r + \arcsin\left(\frac{\dot{d}}{\sqrt{[1-k_r d]^2 \dot{s}^2 + \dot{d}^2}}\right) \quad (4\text{-}31)$$

The first derivative of the spacecraft's lateral offset versus the longitudinal offset is

$$\dot{d} = \frac{\mathrm{d}}{\mathrm{d}s}(d) = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}s}(d) = \frac{\dot{d}}{\dot{s}} = \frac{1}{s} v_x \sin(\theta_x - \theta_r) \quad (4\text{-}32)$$

The second derivative of the spacecraft's lateral offset versus the longitudinal offset is

$$\begin{aligned}\ddot{d} &= \frac{\mathrm{d}}{\mathrm{d}s}(\dot{d}) \\ &= -[\dot{k}_r d + k_r \dot{d}]\tan(\theta_x - \theta_r) + \frac{1 - k_r d}{\cos^2(\theta_x - \theta_r)}\left[k_x \frac{1 - k_r d}{\cos^2(\theta_x - \theta_r)} - k_r\right]\end{aligned} \tag{4-33}$$

Furthermore, from Eq. (4-32) and Eq. (4-33), the conversion relationship between Frenet coordinates and Cartesian coordinates can be obtained

$$v_x = \dot{s}\,\frac{1 - k_r d}{\cos(\theta_x - \theta_r)} \tag{4-34}$$

$$\begin{aligned}a_P &= \dot{v}_x \\ &= \ddot{s}\,\frac{1 - k_r d}{\cos(\theta_x - \theta_r)} + \frac{\dot{s}^2}{\cos(\theta_x - \theta_r)} \\ &\quad [(1 - k_r d)\tan(\theta_x - \theta_r) \cdot (\theta_x - \theta_r)' - (k'_r d + k_r d')]\end{aligned} \tag{4-35}$$

4.2.3: Spacecraft Orbit Avoidance Optimization Index

In order to obtain the optimal avoidance path, a number of indicators should be considered in spacecraft orbit avoidance. First, it is necessary to ensure that the maneuver path smoothly avoids space debris; secondly, it is necessary to make the avoidance path take into account the orbit maintenance; then, the spacecraft braking time should be considered; finally, save fuel as much as possible. Based on this, the spacecraft orbit avoidance index system (see Figure 4-3) and model are constructed to plan and obtain the optimal avoidance path.

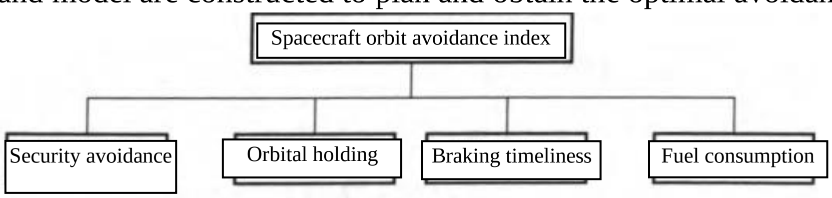


Figure 4-3: Spacecraft Orbit Avoidance Index System

$$\min Q=\gamma_J J_t(d(\Delta t))+\gamma_d \Delta d+\gamma_t \Delta t+\gamma_r |\ddot{d}(\Delta t)| \quad (4\text{-}36)$$

In the formula, the $J_t(\bullet)$ term penalizes the path direction with a large rate of change of maneuver acceleration to ensure the safety of space debris avoidance. Δd term to reduce the lateral offset and promote the spacecraft to travel along the set transfer orbit; Δt term to reduce braking time and improve spacecraft braking agility; $|\ddot{d}(\Delta t)|$ term to control lateral acceleration and reduce fuel consumption; γ_J, γ_d, γ_t, γ_r is the weight value of the global optimization function, and $\gamma_J+\gamma_d+\gamma_t+\gamma_r = 1$ is satisfied.

(1) Circumvent security

In the process of avoidance maneuver, if the acceleration change rate of the spacecraft is large, the state of the spacecraft will be unstable, and the safety risk of the spacecraft in the process of avoiding space debris will be increased. The orbital avoidance path can be decomposed into a one-dimensional problem of longitudinal $\vec{s}$ and normal $\vec{d}$ by the Frenet coordinate system, so that a one-dimensional integral model can be constructed according to the Jerk function, which is in the form of

$$\dot{\vec{u}}(t)=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}\vec{u}(t)+\begin{bmatrix}0\\0\\1\end{bmatrix}\dddot{f}(t) \quad (4\text{-}37)$$

where $\vec{u}(t)=[f(t),\dot{f}(t),\ddot{f}(t)]^T$ is the lateral offset; $\dddot{f}(t)$ is the Jerk function, then the horizontal Jerk component is $\dddot{d}(t)$.

For Eq. (4-37), the initial state of $S_0=[f(t_0),\dot{f}(t_0),\ddot{f}(t_0)]$ at the time t_0 and the state of space debris at $S_1=[f(t_1),\dot{f}(t_1),\ddot{f}(t_1)]$ at the time t_1 are known to have optimal trajectories $J_{f(t)}$

$$J_{f(t)}=\int_{t_0}^{t_1} g(\dddot{f}(t))\,\mathrm{d}t+h(\vec{u}(t),t)_{t_1} \quad (4\text{-}38)$$

where $g(\dddot{f}(t))$ can be used to evaluate the stationarity of the avoided orbital path; $h(\vec{u}(t),t)$ a function configured for the target, which can be used to evaluate the path of the avoided orbit.

Figure 4-4 shows a schematic diagram of the path of the avoided orbit at different accelerations. In the figure, the different colored areas show the rate of change of acceleration, and the spacecraft will no longer have obstacles in the direction of flight after avoiding space debris, and the spacecraft will choose the fastest maneuvering path to return to the set transfer orbit, so that the spacecraft can return to the set transfer orbit.

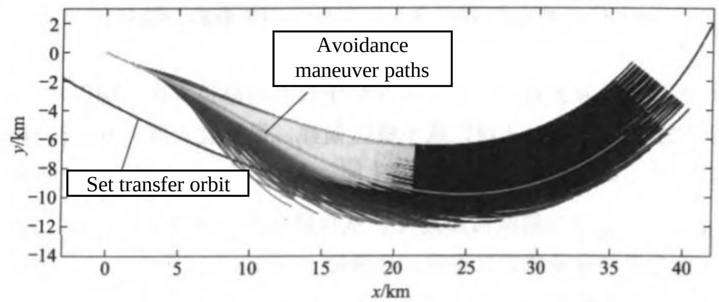


Figure 4-4: Acceleration Rate of Change Curve of Orbital Avoidance Path (see color insert)

(2) Orbit holding

The lateral offset of the spacecraft is mainly responsible for the avoidance of space debris and the timely recovery of the path after avoidance. Equation (4-25) is expressed in the form of a matrix equation as:

$$\begin{bmatrix} d(t) \\ \dot{d}(t) \\ \ddot{d}(t) \end{bmatrix} = \begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} t^3 & t^4 & t^5 \\ 3t^2 & 4t^3 & 5t^4 \\ 6t & 12t^2 & 20t^3 \end{bmatrix} \cdot \begin{bmatrix} \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix}$$

$$= \boldsymbol{M}_{d1}(t) \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \boldsymbol{M}_{d2}(t) \cdot \begin{bmatrix} \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} \quad (4\text{-}39)$$

To simplify the calculation, let $t_0 = 0, t_1 = \tau$. When t = 0, $\alpha_0 = d(0)$, $\alpha_1 = \dot{d}(0)$, $\alpha_2 = \frac{\ddot{d}(0)}{2}$. When $t = \tau > 0$, the coefficient α_3, α_4, α_5 can be solved according to the following equation

$$\begin{bmatrix} \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} = \boldsymbol{M}_{d2}^{-1}(\tau)(\tau) \cdot \left(\begin{bmatrix} d(\tau) \\ \dot{d}(\tau) \\ \ddot{d}(\tau) \end{bmatrix} - \boldsymbol{M}_{d1}(\tau) \cdot \begin{bmatrix} d(0) \\ \dot{d}(0) \\ \frac{\ddot{d}(0)}{2} \end{bmatrix} \right) \quad (4\text{-}40)$$

For spacecraft lateral planning, when the initial state of the t_0 time $S_0 = \left[d(t_0), \dot{d}(t_0), \ddot{d}(t_0)\right]$ the state of space debris at the t_1 time $S_1 = \left[d(t_1), \dot{d}(t_1), \ddot{d}(t_1)\right]$ is known, the lateral offset path of the spacecraft can be solved by Eq. (4-39) to Eq. (4-40).

Figure 4-5 shows a schematic diagram of the avoided orbital path caused by different lateral offsets. In the figure, the curve where the blue point is located is the established transition orbit, and the curve where the green point is located is the optimal lateral avoidance path. In addition, the surrounding black curves are valid paths that meet the basic requirements, and the gray curves are invalid paths.

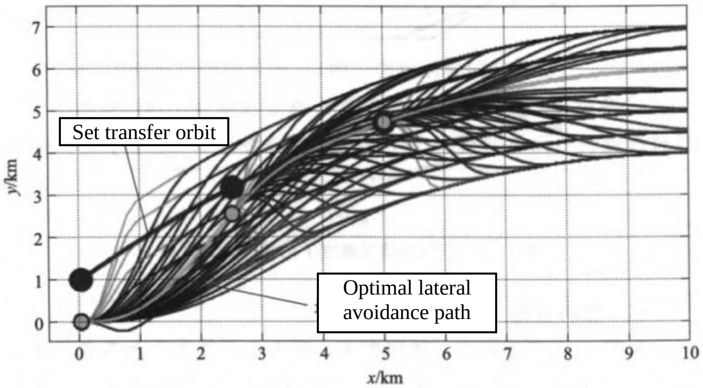


Figure 4-5: Lateral Offset Curve of the Orbital Avoidance Path (see color insert)

(3) Brake timeliness

If the brake of the spacecraft is not timely, it may lead to overshoot or instability of the spacecraft evasive maneuver. If the spacecraft braking is not sensitive, the orbital avoidance path will no longer meet Bellman's optimization principle[1], so that the current path will no longer be the optimal path at the next moment. In particular, overshoot or instability becomes noticeable when the path of the next moment is too far away from the path predicted at the previous moment.

[1] Bellman's optimization principle refers to the optimal path from (i_0, j_0) to (i_f, j_f) through (i, j), which is a path from the origin (i_0, j_0) to the node (i, j). Concatenation to the end point (i_f, j_f) optimal path.

Figure 4-6 illustrates two different braking sensitivities under the same planning strategy, and the orbital avoidance path will depend on the sensitivity sampling time. Among them, $n_i (i = 1, 2, \ldots, N)$ is the starting point of subsequent path planning, and Figure 4-6(a) is the avoidance orbit path with a high braking sensitivity ΔT_a. Figure 4-6(b) shows an avoidance orbit path with a lower braking sensitivity ΔT_b, which fluctuates up and down the reference line and may lead to instability.

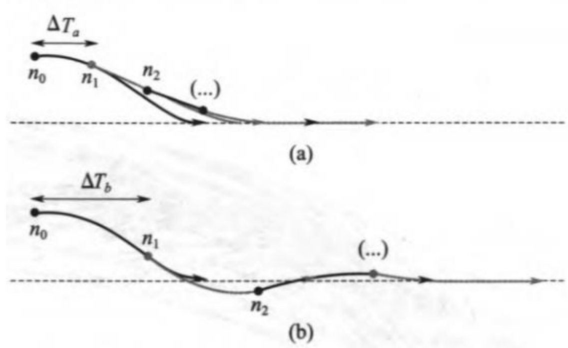


Figure 4-6: Orbit Path Avoidance at Different Braking Sensitivities

(4) Fuel consumption

In the two-body model shown in Figure 4-1, the spacecraft flies along an elliptical transfer orbit with a gravitational constant of μ, taking the central celestial body as the reference point, the normal direction of the near apex as the starting point of the polar angle, and the direction of maneuver of the spacecraft as the positive polar angle. In the meantime, in the process of avoiding space debris, the lateral offset strategy is adopted, and the thrust direction is always in the plane of the transfer orbit and along the normal direction of the transfer orbit $\vec{d}$, and the dynamic equation is[242]

$$\begin{cases} \ddot{r} - r\dot{\theta}^2 + \dfrac{\mu}{r^2} = 0 \\ 2\dot{r}\dot{\theta} + r\ddot{\theta} = u\dfrac{\ddot{d}(t)}{\tan\varphi} \end{cases} \tag{4-41}$$

where r is the orbital radius; θ is the polar angle of the spacecraft; $u = [0, 1]$ is the thrust switch function; φ is the steering angle of the spacecraft.

Based on the characteristics of electric propulsion engine[1] with high specific impulse and low fuel consumption,[243, 244] the continuous thrust maneuvering mode is adopted, and the relationship between thrust and fuel consumption is [245]

$$m\,|\ddot{d}(t)| = -\dot{m}V_e \tag{4-42}$$

where m is the mass of the spacecraft; $\dot{m}$ for mass flow; V_e is the effective exhaust velocity of the continuous thrust engine.

Considering the effect of fuel consumption on the mass of the spacecraft, the fuel consumption expression can be obtained by integrating equation (4-42)[244, 245]

$$\begin{cases} m_f = m e^{-\int_{t_1}^{t_2} |\ddot{d}(t)|\,dt/V_e} \\ \Delta m = m - m_f \end{cases} \tag{4-43}$$

where m_f is the remaining mass of the spacecraft; Δm is an estimate of fuel consumption.

When considering that fuel consumption is small relative to the mass of the spacecraft, it can be assumed that the mass of the spacecraft remains constant throughout the evasive maneuver. After integrating the equation (4-42), the estimation of fuel consumption is obtained

$$\Delta m = \int_{t_1}^{t_2} \dot{m}\,\mathrm{d}t = \frac{m}{V_e}\int_{t_1}^{t_2} |\ddot{d}(t)|\,\mathrm{d}t \tag{4-44}$$

After analysis, when $\int_{t_1}^{t_2}|\ddot{d}(t)|dt \ll V_e$ Eq. (4-43) and Eq. (4-44) are approximately equal.

4.3: Path Generation Algorithm Based on Improved Artificial Potential Field

Spacecraft orbital temporary avoidance should not only avoid space debris, but also consider factors such as fuel consumption, minimum offset, and braking time, which is a typical multi-restriction and shortest path problem. While the artificial potential field method has the advantages of clear mathematical description, fast operation, small amount of calculation and low hardware requirements, it also has disadvantages such as unattainable target and local minimal traps.

[1] An electric propulsion engine, also known as an electric rocket engine, is a device that can generate thrust without relying on chemical combustion. Electric propulsion engines no longer need to use solid or liquid fuels, eliminating the need for complex storage tanks, pipelines, engine combustion chambers, nozzles, corresponding cooling mechanisms, etc., which can greatly reduce the fuel carrying capacity of spacecraft.

In order to further improve the shortcomings of conventional methods, such as the difficulty of satisfying different avoidance preferences at the same time and the weak orbit consideration, this section proposes a path generation algorithm based on improved artificial potential field by constructing an artificial potential field model with reference line traction, long-distance point repulsion ignorance, and obstacle point gravity weakening, so as to realize the autonomous avoidance of space debris by spacecraft.[246]

4.3.1: Artificial Potential Fields for Space Debris Avoidance

Always take the set transfer orbit as the reference line, adjust the action area of each potential field, construct a continuous differentiable potential field function, avoid premature trajectory deviation and local oscillation, and realize the autonomous avoidance of space debris driven by the comprehensive potential field.

(1) Space artificial potential field method

A virtual force field method proposed by Khatib and known as the artificial potential field method[247] is to use the space potential field force to study the relative motion of the object, and control the avoidance motion of the object through the changing position potential field, as shown in Figure 4-7. The basic idea of the artificial potential field method is to search for an optimal path without collision through the joint action of the gravitational potential field at the target position and the repulsive potential field of the obstacle. Compared with other evasion methods, the artificial potential field method has the comparative advantages of simple mathematical description, small amount of calculation, high practicability and smooth path.

The process of spacecraft evasive maneuver is represented by an artificial potential field. In the artificial potential field, the target orbit will produce a gravitational potential field on the spacecraft, and the space debris will produce a repulsive potential field on the spacecraft, as shown in Figure 4-8. The resultant field of the gravitational potential field and the repulsion potential field will determine the direction and rate of maneuver of the spacecraft.

Thus, the comprehensive potential field of the spacecraft at any space position can be expressed as[247]

$$\boldsymbol{U}(\boldsymbol{x}) = \boldsymbol{U}_{att}(\boldsymbol{x}) + \boldsymbol{U}_{rep}(\boldsymbol{x}) \quad (4\text{-}45)$$

where x is the current position state vector of the spacecraft; $\boldsymbol{U}(\boldsymbol{x})$ is the comprehensive potential field of the spacecraft in space; $\boldsymbol{U}_{att}(\boldsymbol{x})$ gravitational potential field of the target orbit position on the spacecraft; $\boldsymbol{U}_{rep}(\boldsymbol{x})$ is the repulsive potential field of space debris on the spacecraft.

As shown in Figure 4-9, the gravitational potential field generated by the target orbital position of the spacecraft is related to the distance between the spacecraft, and the greater the distance between the two, the greater the potential energy, and vice versa.

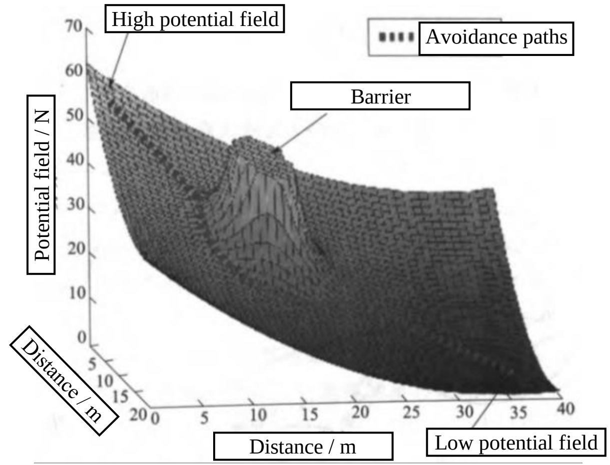


Figure 4-7: Schematic Diagram of Artificial Potential Field Avoidance Path (see color insert)

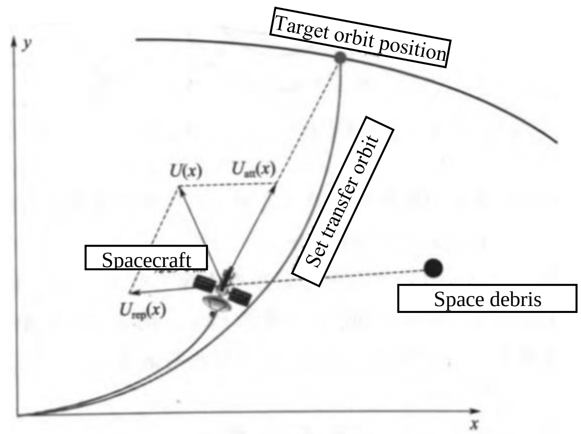


Figure 4-8: Schematic Diagram of the Force of a Spacecraft in an Artificial Potential Field

Therefore, the gravitational potential field is proportional to the distance between the two, and the gravitational potential field can be expressed as

$$\boldsymbol{U}_{att}(\boldsymbol{x}) = \frac{1}{2}k_{att}d^2(\boldsymbol{x}, \boldsymbol{x}_g) \quad (4\text{-}46)$$

where k_{att} is the gravitational potential field gain coefficient: $\boldsymbol{x}$ is the current space position vector of the spacecraft; $\boldsymbol{x}_g$ spatial position vector for the target orbit position; $d^2(\bullet)$ is the Euclidean distance calculation function.

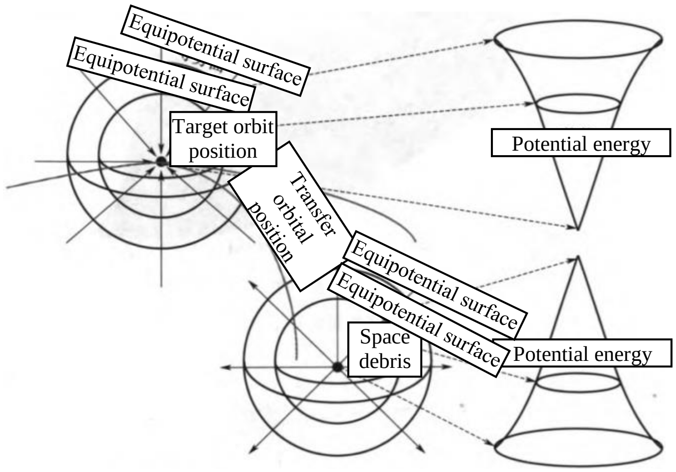


Figure 4-9: Artificial Gravitational Potential Field and Repulsive Potential Field in Spacecraft Orbital Avoidance

The gravitational force generated by this gravitational potential field on the spacecraft is the negative gradient of gravitational potential energy

$$F_{att}(\boldsymbol{x}) = -\nabla \boldsymbol{U}_{att}(\boldsymbol{x}) = -k_{att}d^2(\boldsymbol{x}, \boldsymbol{x}_g) \quad (4\text{-}47)$$

At the same time, space debris will create a repulsive potential field for the spacecraft. The magnitude of the repulsive potential field is determined by the spatial distance between the spacecraft and the space debris, and the smaller the distance between the two, the larger the repulsive potential field, and vice versa. From this, the repulsive potential field can be expressed as:

$$\boldsymbol{U}_{rep}(\boldsymbol{x}) = \begin{cases} \frac{1}{2}k_{rep}\left(\frac{1}{d(\boldsymbol{x}, \boldsymbol{x}_o)} - \frac{1}{d_n}\right) & d(\boldsymbol{x}, \boldsymbol{x}_o) \leqslant d_n \\ 0 & d(\boldsymbol{x}, \boldsymbol{x}_o) > d_n \end{cases} \quad (4\text{-}48)$$

where k_{rep} is the repulsion potential field gain coefficient; $\boldsymbol{x}_0$ is the current position vector of space debris;

d_n is the range of the non-repulsive region.

The repulsion force generated by the repulsive potential field is the negative gradient of the repulsive potential energy:

$$
\begin{aligned}
F_{rep}(\boldsymbol{x}) &= -\nabla \boldsymbol{U}_{rep}(\boldsymbol{x}) \\
&= \begin{cases} k_{rep}\left(\dfrac{1}{d(\boldsymbol{x},\boldsymbol{x}_o)} - \dfrac{1}{d_n}\right)\dfrac{1}{d^2(\boldsymbol{x},\boldsymbol{x}_o)}\dfrac{\partial d(\boldsymbol{x},\boldsymbol{x}_o)}{\partial x} & d(\boldsymbol{x},\boldsymbol{x}_o) \leqslant d_n \\ 0 & d(\boldsymbol{x},\boldsymbol{x}_o) > d_n \end{cases}
\end{aligned} \quad (4\text{-}49)
$$

The artificial potential field method can well describe the spatial avoidance behavior with mathematical expressions, and provides a concrete and effective solution for the avoidance path planning method, which has significant advantages:

1) Real-time evasion. Compared with other path planning methods, the artificial potential field method has good real-time avoidance performance, which can focus the center of gravity of path planning modeling on the flight direction of spacecraft and space debris, and can achieve a better real-time avoidance effect with less calculation.

2) Path smoothing. In the process of spacecraft avoidance of space debris, the artificial potential field method can collect the state information of spacecraft and space debris with the surrounding space environment, and reflect it in the repulsive potential field and gravitational potential field between the two sides. In path planning, the artificial potential field method only needs to combine the current position with the comprehensive potential field to obtain a smooth and safe path, and there is no need to perform operations such as path smoothing and obstacle avoidance detection like other algorithms, and the application advantages are obvious.

However, the traditional artificial potential field method has also exposed its own structural defects in a large number of practical applications, that is, the existence of local optimal solutions. The so-called local optimal solution is the equivalent inverse effect between the guidance force potential field and the repulsion potential field, resulting in the spacecraft staying in situ or oscillating back and forth, and unable to continue to travel towards the target orbit. There are two main manifestations of the local optimal solution:

1) The goal is unattainable. In the process of space debris avoidance, if the space debris is located near the target orbit at this moment, when the spacecraft is far away from the target orbit and the space debris, the gravitational potential field is larger, and the repulsion potential field is smaller, and the repulsion potential field is less than the gravitational potential field, then the spacecraft will travel towards the target orbit under the action of the resultant potential field; however, as the spacecraft gets closer to the target orbit and space debris, the gravitational potential field becomes smaller and smaller, and the repulsion potential field becomes larger and larger, and when the gravitational potential field and the repulsion potential field are reversed, this may cause the spacecraft to oscillate or stay at the current position and fail to reach the target orbital position, as shown in Figure 4-10.

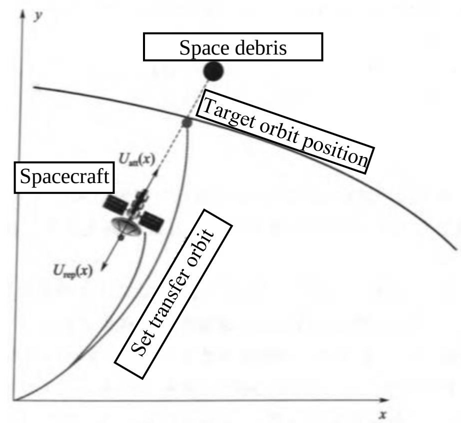


Figure 4-10: Schematic Diagram of Unattainable Targets

2) Local minima traps. When the spacecraft is close to the space debris, if the target orbit is at the other end of the space debris, and the angle between the gravitational potential field and the repulsion potential field of the spacecraft is 180° and the magnitude is equal, the spacecraft will stay at the point where the net force is zero, that is, the local minimality trap, and finally cannot successfully reach the target orbital position, as shown in Figure 4-11.

This book divides local minimal traps into two categories:

a) Local minima problems. The first type of local minima trap is the problem of the potential field encountering the local minimum[1], which is the minimum value in any direction when it is in the trap, and its comprehensive potential field is concave. The local minimum can be determined according to theorem 4.1.

Theorem 4.1[248] gives $\boldsymbol{U}(\boldsymbol{x})$ a second-order continuous differentiable function of $\boldsymbol{R}^n \rightarrow \boldsymbol{R}$, and $g \in \boldsymbol{R}^n$ is the station point of $\boldsymbol{U}(\boldsymbol{x})$.

[1] Local minima: If there is a $\varepsilon > 0$, such that there is $|x - x^*| < \varepsilon$ for any x that satisfies $f(x^*) \leq f(x)$, then put the function value of the point x^* corresponding to $f(x^*)$ is called a local minimum of the function f.

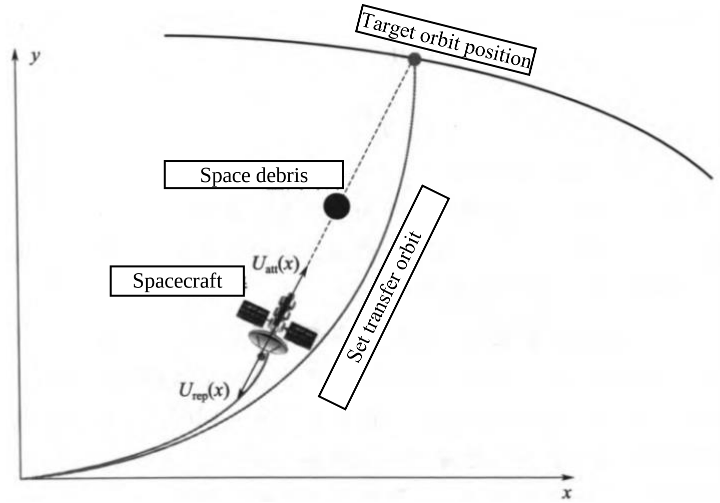


Figure 4-11: Schematic Diagram of a Local Minimal Trap

If the Hessian matrix at g is positively determined, then g is the local minima; conversely, g is not necessarily an extremum.

b) Saddle point issues. The second type of local minima trap is the saddle point[1] problem, i.e., the comprehensive potential field in one or more directions is a local minimum. As the spacecraft maneuvers in the direction of these local minimals, it may linger near the saddle point due to insufficient potential field and cannot continue traveling.

Let the direction vector with angle $\varphi \in [0, 2\pi)$ as the parameter

$$\boldsymbol{L} = \boldsymbol{l}_x \cos\varphi + \boldsymbol{l}_y \sin\varphi \qquad (4\text{-}50)$$

where $\boldsymbol{l}_x$ and $\boldsymbol{l}_y$ are the unit vectors of the direction of motion.

The position and direction of the saddle point in the second type of trap can be obtained by solving the first and second partial derivatives of vector $\boldsymbol{l}_x$ and vector $\boldsymbol{l}_y$ and satisfying all the spatial positions and corresponding angle parameters of Eq. (4-51) and satisfying the φ of all the spatial positions and corresponding angle parameters

[1] Saddle point: In differential equations, the singularity that is stable in one direction and unstable in the other direction is called the saddle point.

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{L}}\boldsymbol{U} = 0 \\ \frac{\mathrm{d}^2}{\mathrm{d}^2\boldsymbol{L}}\boldsymbol{U} > 0 \end{cases} \tag{4-51}$$

(2) Improved model of comprehensive potential field

The artificial potential field method usually takes the end point as the gravitational source and the obstacle as the repulsion source, and the gravitational potential field and the repulsive potential field are synthesized into a comprehensive potential field in space, which drives the moving body to reach the end point while avoiding the obstacle along the weakening direction of the potential field.[249]

In the application of spacecraft evasive maneuver, it is necessary to avoid the planned path going straight to the end point, and to continue to travel in the direction of the set transfer orbit as much as possible. The repulsive potential field should be weakened at a distance from space debris to avoid premature trajectory deviation. At the same time, the gravitational potential field should be weakened when approaching space debris to avoid local oscillation. As shown in Figure 4-12, this book constructs a comprehensive potential field model with reference line traction, ignorance of repulsion at distant points, and attenuation of gravitational force at obstacle points

$$\boldsymbol{U}(\boldsymbol{x},\Delta s) = k_{refer}\boldsymbol{U}_{refer}(\Delta s) + k_{att}\boldsymbol{U}_{att}(\boldsymbol{x}) + k_{rep}\boldsymbol{U}_{rep}(\boldsymbol{x}) \tag{4-52}$$

where $\boldsymbol{U}(\boldsymbol{x},\Delta s)$ is the combined potential field of the current position x spacecraft; $\boldsymbol{U}_{refer}(\Delta s)$ is the reference line potential field; k_{refer} is the reference line potential field coefficient; $\boldsymbol{U}_{att}(\boldsymbol{x})$ is the gravitational potential field: k_{att} is the gravitational potential field coefficient; $\boldsymbol{U}_{rep}(\boldsymbol{x})$ is the repulsive potential field; k_{rep} is the repulsion potential field coefficient.

(3) Reference line potential field model

In order to meet the needs of the spacecraft to travel in the direction of the reference line, this book proposes a method to replace the target point potential field with the reference line potential field, so that the spacecraft can closely follow the reference line on the way to avoid space debris. The reference line potential field will constrain the circumvention of the orbital direction and ensure that the target orbit is flown towards the target orbital position, i.e., the value of the forward potential field is lower than that of the rear potential field, which is described using a Gaussian-like function[250] surface as follows

$$\boldsymbol{U}_{refer}(\Delta s) = (s_{total} - \Delta s)exp\left(1 + \frac{\Delta d^2}{2\delta^2}\right)\boldsymbol{s} \tag{4-53}$$

where s_{total} is the entire transfer orbit range; Δs is the range from the starting point x_{st} to the current position x; Δd is the distance along the normal deviation from the reference line; is the normal convergence coefficient. δ[250]

(4) Gravitational potential field model

In order to smoothly avoid space debris and avoid local oscillations, the gravitational potential field function is improved, and a ring region that weakens the gravitational potential field is set near space debris.[248]

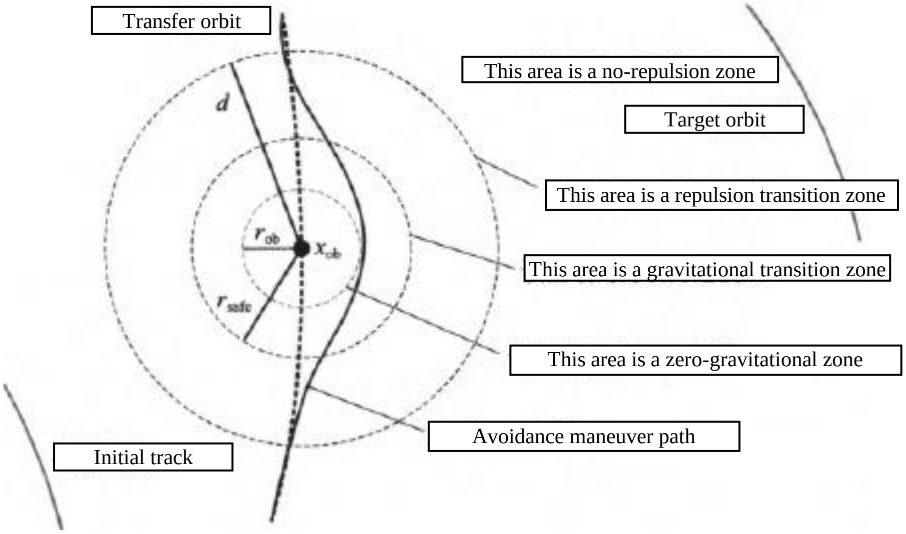


Figure 4-12: Illustration of a Spacecraft's Avoidance of Maneuvering Artificial Potential Field (see color insert)

$$U_{att}(\boldsymbol{x}) = -\eta\left(1 - \frac{\|\boldsymbol{x} - \boldsymbol{x}_{goal}\|}{\|\boldsymbol{x}_{ob} - \boldsymbol{x}_{goal}\|}\right)^n,$$

$$\eta = \begin{cases} 0, & \|\boldsymbol{x} - \boldsymbol{x}_{ob}\| < r_{ob} \\ \frac{1}{2}\left[\sin\left(\frac{\|\boldsymbol{x} - \boldsymbol{x}_{ob} - r_{ob}\|}{r_{safe} - r_{ob}} \cdot \pi - \frac{\pi}{2}\right) + 1\right], & r_{ob} \leqslant \|\boldsymbol{x} - \boldsymbol{x}_{ob}\| < r_{safe} \\ 1, & \text{Other} \end{cases} \tag{4-54}$$

where η is the coefficient of attraction field related to the distance of space debris; $\boldsymbol{x}_{ob}$ is the vector for the current position of the space debris; $\boldsymbol{x}_{goal}$ is the position vector of the intersection point between the transfer orbit and the target orbit; n is a positive integer; r_{ob} is the range of threats from space debris; r_{safe} is the extent of the gravitational attenuation zone; $\|\cdot\|$ is 2-norm.

(5) Repulsive potential field model

In order to keep up with the transfer trajectory and avoid premature trajectory deviation, the potential field transition strategy is used to improve the repulsion potential field function

$$U_{rep}(\boldsymbol{x}) = -\lambda\left(\frac{r_{ob}}{\|\boldsymbol{x} - \boldsymbol{x}_{ob}\|}\right)^q,$$

$$
\lambda=\begin{cases}1, \|\boldsymbol{x}-\boldsymbol{x}_{ob}\|<r_{safe} \\ \frac{1}{2}\left[\cos\left(\frac{\|\boldsymbol{x}-\boldsymbol{x}_{ob}\|-r_{safe}}{d-r_{safe}}\cdot\pi\right)+1\right], r_{safe}\leqslant\|\boldsymbol{x}-\boldsymbol{x}_{ob}\|<d \\ 0, \|\boldsymbol{x}-\boldsymbol{x}_{ob}\|\geqslant d\end{cases} \tag{4-55}
$$

where λ is the repulsion potential field coefficient; q is a positive integer; d is the range of the repulsion transition zone.

(6) Continuous differentiable proofs

In general, the potential field force can be obtained by solving the negative gradient of the artificial potential field model, but the comprehensive potential field model is required to be continuous and differentiable. However, the coefficients of the potential field model constructed in this chapter are determined by the relative position of the spacecraft, especially the coefficients of the gravitational potential field model and the repulsion potential field model are piecewise functions related to the relative position. In order to test that the continuous potential field force can be obtained from the constructed potential field model, the continuous differentiability of each piecewise function is proved in this section.

For the gravitational potential field function $\boldsymbol{U}_{att}(\boldsymbol{x})$, its continuous differentiability is mainly seen in the piecewise function η in Eq. (4-54).

$$
\eta=\begin{cases}0, \|\boldsymbol{x}-\boldsymbol{x}_{ob}\|<r_{ob} \\ \frac{1}{2}\left[\sin\left(\frac{\|\boldsymbol{x}-\boldsymbol{x}_{ob}\|-r_{ob}}{r_{safe}-r_{ob}}\cdot\pi-\frac{\pi}{2}\right)+1\right], r_{ob}\leqslant\|\boldsymbol{x}-\boldsymbol{x}_{ob}\|<r_{safe} \\ 1, \text{Other}\end{cases} \tag{4-56}
$$

Let $x = \|x - x_{ob}\| - r_{ob}$, then

$$
\begin{aligned}
&\lim_{\Delta x \to 0} \frac{\eta(x+\Delta x)-\eta(\boldsymbol{x})}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{1}{2}\left[\sin\left(\frac{x+\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)+1\right]-\frac{1}{2}\left[\sin\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)+1\right]}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{1}{2}\sin\left(\frac{x+\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)-\frac{1}{2}\sin\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{1}{2}\left[\sin\left(\frac{\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi\right)\cos\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)+\cos\left(\frac{\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi\right)\sin\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)\right]-\frac{1}{2}\sin\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{1}{2}\cos\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)\cos\left(\frac{\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi\right)\frac{\pi}{r_{\text{safe}}-r_{\text{ob}}}-\frac{1}{2}\sin\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right)\sin\left(\frac{\Delta x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi\right)\frac{\pi}{r_{\text{safe}}-r_{\text{ob}}} \\
&= \frac{1}{2}\frac{\pi}{r_{\text{safe}}-r_{\text{ob}}}\cos\left(\frac{x}{r_{\text{safe}}-r_{\text{ob}}}\cdot\pi-\frac{\pi}{2}\right) \\
&= \eta'(x)
\end{aligned}
\tag{4-57}
$$

Again let $y = \|\boldsymbol{x} - \boldsymbol{x}_{ob}\| - r_{safe}$, there is

$$
\lim_{\Delta y \to 0} \frac{\eta(y + r_{\text{safe}} + \Delta y) - \eta(y + r_{\text{safe}})}{\Delta y}
$$

$$
= \lim_{\Delta y \to 0} \frac{\frac{1}{2}\left[\sin\left(\frac{y + r_{\text{safe}} + \Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right) + 1\right] - \frac{1}{2}\left[\sin\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right) + 1\right]}{\Delta y}
$$

$$
= \lim_{\Delta y \to 0} \frac{\frac{1}{2}\sin\left(\frac{y + r_{\text{safe}} + \Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right) - \frac{1}{2}\sin\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)}{\Delta y}
$$

$$
= \lim_{\Delta y \to 0} \frac{\frac{1}{2}\left[\sin\left(\frac{\Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi\right)\cos\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right) + \cos\left(\frac{\Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi\right)\sin\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)\right]}{\Delta y} -
$$

$$
\lim_{\Delta y \to 0} \frac{1}{2\Delta y}\sin\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)
$$

$$
= \lim_{\Delta y \to 0} \frac{1}{2}\cos\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)\cos\left(\frac{\Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi\right)\frac{\pi}{r_{\text{safe}} - r_{\text{ob}}} -
$$

$$
\frac{1}{2}\sin\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)\sin\left(\frac{\Delta y}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi\right)\frac{\pi}{r_{\text{safe}} - r_{\text{ob}}}
$$

$$
= \frac{1}{2}\,\frac{\pi}{r_{\text{safe}} - r_{\text{ob}}}\cos\left(\frac{y + r_{\text{safe}}}{r_{\text{safe}} - r_{\text{ob}}} \cdot \pi - \frac{\pi}{2}\right)
$$

$$
= \eta'(y + r_{\text{safe}})
$$

(4-58)

Therefore, it is proved that the η of the piecewise function is continuously differentiable, so that the gravitational potential field model $\boldsymbol{U}_{att}(\boldsymbol{x})$ is also continuously differentiable. In the same way, the continuous differentiability of the repulsion potential field model $\boldsymbol{U}_{rep}(\boldsymbol{x})$ can be proved.

4.3.2: Path Planning to Avoid Space Debris

Using the normal projection of the comprehensive potential field, drawing on Jerk's description, the avoidance path to avoid space debris with the minimum lateral offset is obtained.

In order to maintain the established orbital transfer velocity of the spacecraft along the reference line, only the projection of the comprehensive potential field in the normal direction d is considered to generate a lateral offset and realize the autonomous avoidance of space debris

$$U_d \boldsymbol{d} = \boldsymbol{U} - U_r \boldsymbol{r} \qquad (4\text{-}59)$$

where $\boldsymbol{r}$ is the tangential vector of the reference line; U_r is the tangential component value of the comprehensive potential field; $\boldsymbol{d}$ is the reference line normal vector; U_d is the value of the normal component of the comprehensive potential field.

In order to ensure that the avoidance path driven by the normal component of the comprehensive potential field is smooth and smooth, this book uses the Jerk function to describe the rate of change of acceleration with reference to Ref. 239. The relationship model between the transverse acceleration rate of change $\dddot{d}(\tau)$ and the normal component of the comprehensive potential field and the description of Jerk is constructed

$$U_d(t) = \omega \int_{t_0}^{t_1} \dddot{d}(\tau)\,\mathrm{d}\tau \qquad (4\text{-}60)$$

$$J_t(d(t)) = \int_{t_0}^{t_1} \dddot{d}(\tau)^2\,\mathrm{d}\tau \qquad (4\text{-}61)$$

where ω is the efficiency coefficient of the normal component of the comprehensive potential field; $J_t(d(t))$ is a Jerk description of the lateral displacement; $d(t)$ is the lateral offset.

According to Ref. 99 and 252, the task of path planning is to find the lateral offset that can make $J_t(d(t))$ the smallest, and the solution of any Jerk optimization problem can be expressed by a 5th degree polynomial of the form (4-25). From this, polynomial representations of lateral offset $\Delta d = d(t)$, lateral velocity $\dot{d}(t)$ and lateral acceleration $\ddot{d}(t)$ can be obtained

$$\begin{cases} d(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4 + \alpha_5 t^5 \\ \dot{d}(t) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2 + 4\alpha_4 t^3 + 5\alpha_5 t^4 \\ \ddot{d}(t) = 2\alpha_2 + 6\alpha_3 t + 12\alpha_4 t^2 + 20\alpha_5 t^3 \end{cases} \qquad (4\text{-}62)$$

where α_0, α_1, α_2, α_3, α_4, α_5 is the polynomial coefficient, let $t_0 = 0$ and $\Delta t = t_1 - t_0$ respectively, and substitute the equation (4-62) to obtain their values.

In addition, the lateral acceleration is constrained by the spacecraft's steering maneuvering ability, i.e., the heading angular rate, while promoting the lateral offset of the spacecraft

$$\ddot{\mathrm{d}}(\mathrm{t}) = \begin{cases} \ddot{\mathrm{d}}(\mathrm{t}), \ddot{\mathrm{d}}(\mathrm{t})/\mathrm{v}(\mathrm{t}) \leq \dot{\varphi}_{\max} \\ \dot{\varphi}_{\max}\mathrm{v}(\mathrm{t}), \ddot{\mathrm{d}}(\mathrm{t})/\mathrm{v}(\mathrm{t}) > \dot{\varphi}_{\max} \end{cases} \quad (4\text{-}63)$$

where $v(t)$ is the spacecraft velocity; $\dot{\varphi}_{max}$ is the maximum angular rate of the spacecraft's steering maneuver.

Let $\Delta t = t_1 - t_0$ be the braking time of the spacecraft, and the lateral offset under different braking times can be calculated by equation (4-62) under the drive of $U_d(\Delta t)$. By accumulating the lateral displacement along the reference line over time, a path that can avoid space debris can be obtained.

4.3.3: Dynamic Optimization of Avoidance Paths

In the process of space debris avoidance maneuver, the spacecraft considers the avoidance safety, orbit keeping, braking time and fuel consumption according to the index model established above, and the path planning method based on the improved artificial potential field can obtain the set of candidate avoidance orbit paths, and then the optimal avoidance path can be screened out by using the global objective function formula (4-36). In the global objective function Q, each index item and the weighting coefficient jointly determine the choice of the optimal path for spacecraft orbit avoidance, and the comprehensive index will determine which aspect of the global objective function Q focuses on optimization. For example, when the weighting coefficient of the evasion safety term accounts for a large proportion of the global optimization function, the path optimization will be more inclined to choose the avoidance path with a stable rate of change of maneuvering acceleration. When the weighting coefficient of the orbit holding term accounts for a large proportion of the global optimization function, the avoidance path that can recover to the set transfer orbit as soon as possible will be chosen, but this may increase the risk of collision with space debris during the spacecraft evasion maneuver.

For the spacecraft orbit temporary avoidance path planning problem considering multiple factors, the dynamic optimization steps are as follows:

Step 1: Circumvent demand

According to the current environmental factors and the state of space debris, the focus terms of the spacecraft's orbital avoidance path are obtained, that is, the weight coefficients in the global objective function Q are clarified.

Step 2: Path solving

In order to reduce the computational cost of the method, the lateral offset range of the spacecraft orbit avoidance path is limited, and the spacecraft avoidance maneuver does not exceed this range, that is, the lateral offset d is limited to the $[d_{min}, d_{max}]$ region. The transverse offset discrete sequence $\{d_{10}$, …, d_{1i}, ……, $d_{1m}|1 \leq i \leq m\}$ is generated from the transverse offset sampling, and then the avoidance maneuver parameters satisfying the description of the fifth polynomial (4-62) are solved. Then, according to the improved artificial potential field method from Eq. (4-52) to Eq. (4-59), the path set satisfying the avoidance condition is calculated within the transverse offset range.

Step 3: Path generation

In the process of path generation, if the prediction time is too short, the ability to respond to emergencies will be insufficient. If the prediction time is too long, the reliability of the avoidance path may be reduced and computing resources will be wasted. To do this, the sampling time ΔT is set to discretize the avoidance path of the $t_0 \sim t_1$ time period. Equations (4-60) to (4-63) are used to generate avoidance paths according to the lateral offset, and the position and state information of each candidate avoidance orbit path are determined. Finally, the candidate avoidance paths in the Frenet coordinate system are converted into the Cartesian coordinate system.

Step 4: Filter paths

According to the optimal avoidance path evaluation index, the avoidance paths that do not meet the requirements of evasion safety, orbit keeping, braking restrictions and fuel consumption are filtered.

Step 5: Status Update

The spacecraft flies to the next state according to the optimal avoidance path and updates the current state information. The global objective function determines whether to update the weight coefficient according to the orbit avoidance index, otherwise, the path is planned according to the original target.

4.4: Case Study of Spacecraft Orbit Temporary Avoidance Path Planning

In order to test the effectiveness of the method, a path generation algorithm based on improved artificial potential field is used to test the effectiveness of the spacecraft orbit temporary avoidance path planning case. In order to illustrate the application advantages of the improved artificial potential field method, the classical artificial potential field method and the commonly used Dijkstra and Rapidly Extended Random Tree (RRT) path planning algorithms, the experimental results are compared.

4.4.1: Case Description

The on-orbit spacecraft 6 with a mass of 2, 500 kg and a maximum thrust acceleration of 1.3 m/s2 maneuvered along a given transfer trajectory to target 13 according to the optimal target allocation strategy in orbit in the composite service mode described above. During the orbital maneuver, a space debris temporarily appears in the plane of the transfer orbit. At this time, the space debris position $[33\,854 \;\; 5\,146]^T$ km and velocity $[-0.763 \; 0.763]^T$ km/s. In order to ensure that the on-orbit servicing operation can be continued, the spacecraft carries out orbital temporary avoidance path planning according to different avoidance preferences. In the process of solving the path generation algorithm based on Frenet and the improved artificial potential field, the model parameters are set to $k_{refer}= 0.2, k_{att}= 0.4, k_{rep}= 0.4, \delta= 0.96, \eta= 0.4, n = 1, q = 1, \mu= 3.986\times10^5 km^3/s^2, r_{safe}= 300km, d = 600km, \lambda= 0.4$.

4.4.2: Model Calculation and Analysis

4.4.2.1: Simulation effect analysis

In the face of temporary space debris, in order to ensure the safety of the spacecraft and continue to complete the mission, the spacecraft orbit temporary avoidance path planning method was used to carry out the optimal path planning. According to the method flow, the Frenet coordinate system is used to express the orbital avoidance motion of the spacecraft, and the improved artificial potential field method is used to construct the artificial potential field of the spacecraft to avoid space debris. When the space debris is far away from the spacecraft, the gravitational potential field is dominant, and the spacecraft continues to travel along the set transfer orbit. As shown in Figure 4-13, when space debris enters the repulsion transition zone, the spacecraft is subjected to an increasingly strong repulsive potential field, which will drive the spacecraft away from the set transfer orbit (yellow curve in the figure).

As shown in Figure 4-14, when the space debris (blue dot in the figure) intersects with the spacecraft's set transfer orbit, the spacecraft (red triangle) can avoid the risk of collision with each other with a certain offset.

As shown in Figure 4-15, the spacecraft is guided by the spacecraft comprehensive potential field model from leaving the set transfer orbit to successfully avoiding space debris and then continuing along the set transfer orbit, and the optimal avoidance path is automatically calculated and generated by the spacecraft orbit temporary avoidance path planning method, which can meet the needs of the spacecraft's active defense and autonomous avoidance.

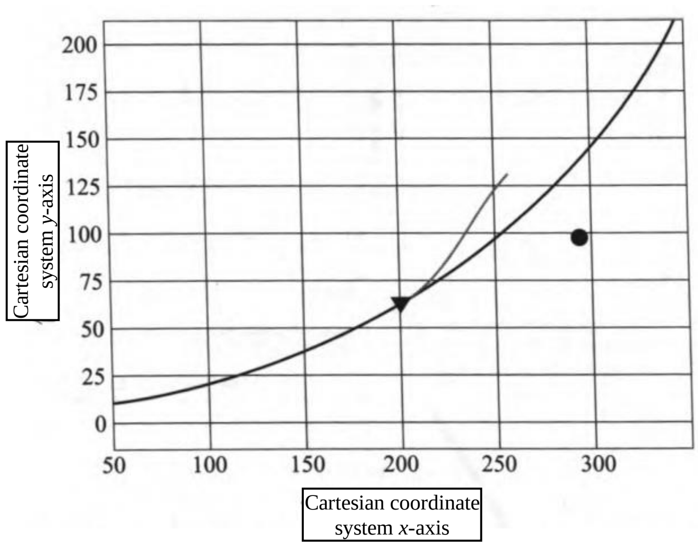


Figure 4-13; Simulation Diagram of a Spacecraft Leaving a Given Orbit (see color insert)

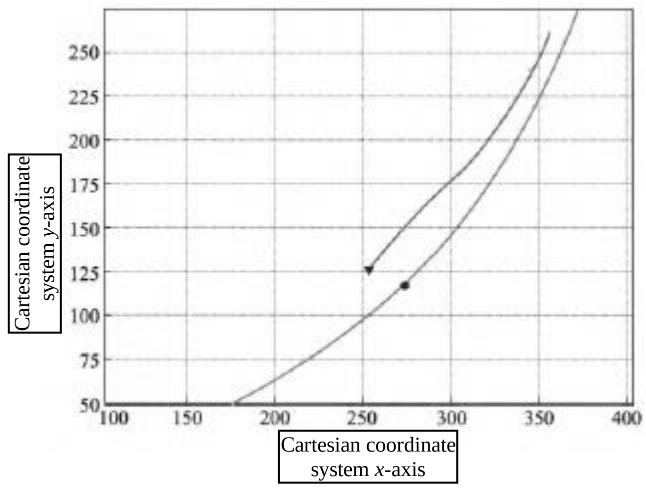


Figure 4-14: Simulation Diagram of Spacecraft Successfully Avoiding Space Debris (see color insert)

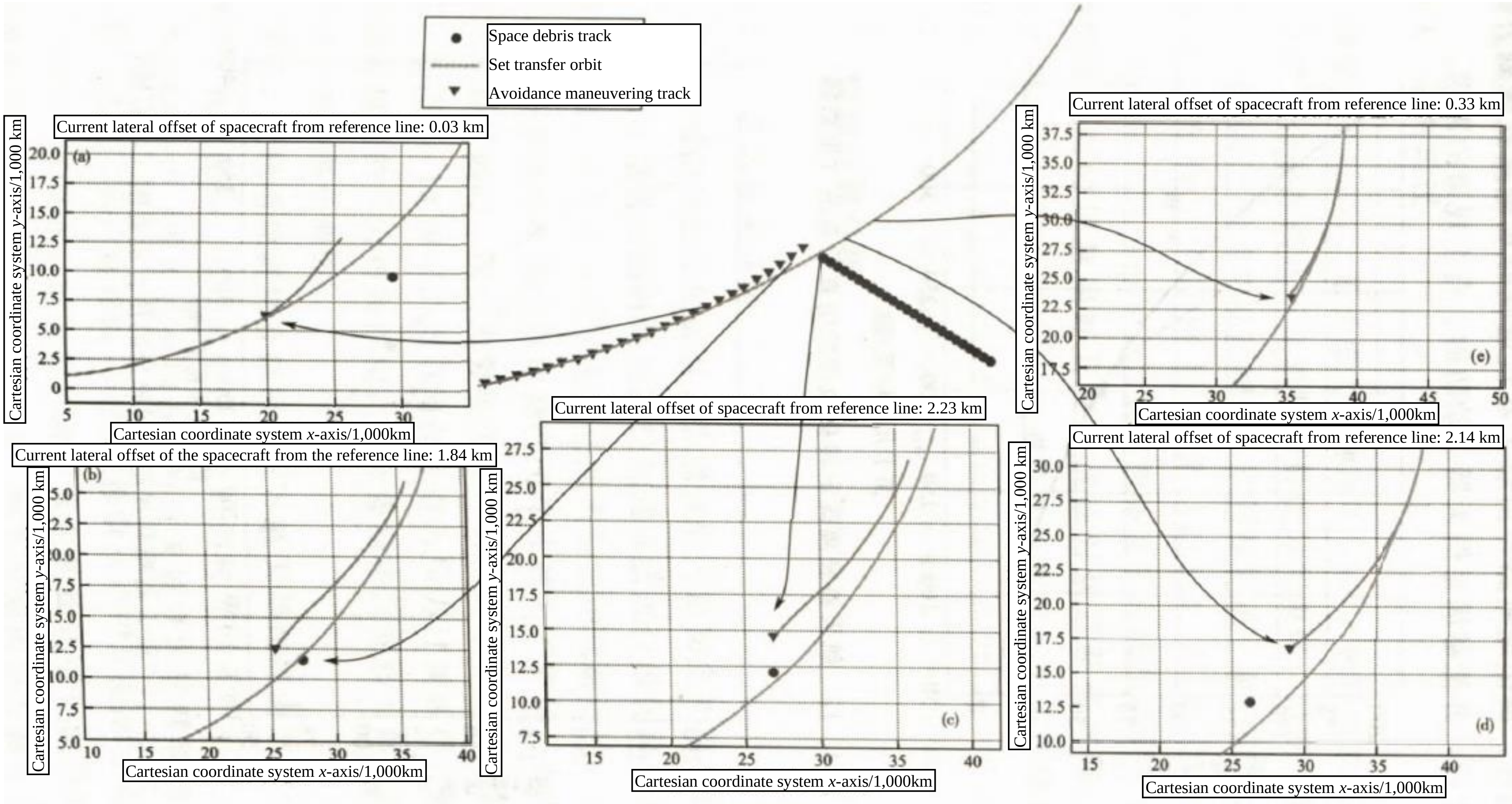


Figure 4-15: Overall Simulation Rendering of Spacecraft Space Debris Avoidance (see color insert)

4.4.2.2: Parameter control analysis

In order to better demonstrate the influence of different indicators of avoidance safety, fuel consumption, braking efficiency and minimum offset on path planning, the application advantages of the test method can meet the needs and preferences of different tasks, and the single-objective optimization parameter control and multi-objective optimization parameter control are simulated and analyzed, respectively.

(1) Single-objective optimization method

Considering single-objective optimization in spacecraft orbit avoidance, that is, only considering a single index in equation (4-36) (e.g., only considering constraints such as fuel consumption or safety avoidance), the spacecraft orbit temporary avoidance path planning method is used to simulate and calculate, and the lateral acceleration changes under the constraints of each index are obtained, as shown in Figure 4-16.

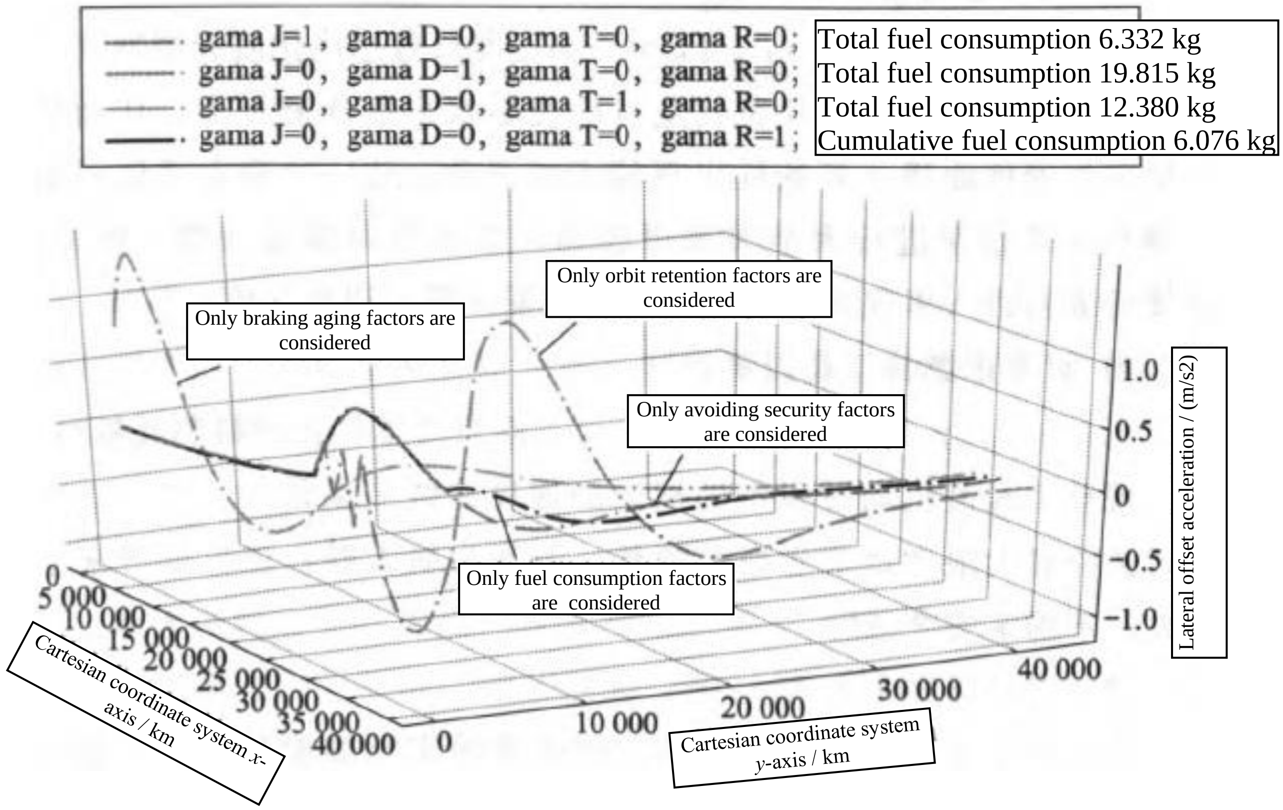


Figure 4-16: Lateral Offset Acceleration in Single-Objective Optimization Mode (see color insert)

1) Only consider avoiding security factors.

When $\gamma_J = 1, \gamma_D = 0, \gamma_T = 0, \gamma_R = 0$, only the Jerk function is optimized, there is a large fluctuation in the lateral offset acceleration of the spacecraft when it is close to the space debris during the evasion maneuver, which causes the spacecraft to obtain a large lateral offset to avoid the collision threat of space debris.

Although this treatment can ensure that the spacecraft can successfully avoid space debris, it requires the spacecraft to have a large lateral thrust if it suddenly adopts an evasive behavior when the target is approaching.

2) Only orbital retention factors are considered.

When $\gamma_J = 0, \gamma_D = 1, \gamma_T = 0, \gamma_R = 0$, that is, only the minimum lateral offset is considered, the spacecraft needs to adjust the direction and magnitude of the lateral acceleration many times in order to continue to travel along the set transfer orbit as much as possible during the evasive maneuver, resulting in many large fluctuations in the lateral acceleration curve. Although this treatment can make the evasion path better consistent with the set transfer orbit, it requires multiple lateral acceleration adjustments, which consumes more fuel and increases the risk of collision between the spacecraft and space debris.

3) Only the brake aging factor is considered.

When $\gamma_J = 0, \gamma_D = 0, \gamma_T = 1, \gamma_R = 0$, that is, only the optimization of the spacecraft's braking time is considered, the spacecraft will take evasive actions as soon as possible after learning of the debris attack crisis, so that the lateral acceleration curve will fluctuate greatly at the beginning. This treatment method makes the spacecraft perform a high-thrust evasion maneuver at the beginning, deviate from the orbit prematurely and be difficult to recover, resulting in the spacecraft entering the airspace unknown in advance, and the orbit safety is difficult to be guaranteed.

4) Only fuel consumption factors are considered.

When $\gamma_J = 0, \gamma_D = 0, \gamma_T = 0, \gamma_R = 1$, that is, only the optimization of fuel consumption is considered, the lateral acceleration curve fluctuation in the whole evasive maneuver can be minimized. Although this treatment method can save fuel consumption to the greatest extent, it brings a large uncertain safety risk to the spacecraft orbital maneuver.

(2) Multi-objective optimization method

In order to better meet the actual needs and preferences of spacecraft to avoid space debris, it is necessary to comprehensively consider the factors of avoidance safety, orbit keeping, braking timeliness and fuel consumption, and adopt the multi-objective optimization method according to Eq. (4-36) to obtain the corresponding optimal avoidance maneuver path under a certain combination of optimization weights. The spacecraft orbit temporary avoidance path planning method is used to simulate and calculate, and the lateral acceleration change under multi-factor constraints is obtained, as shown in the curve in Figure 4-17.

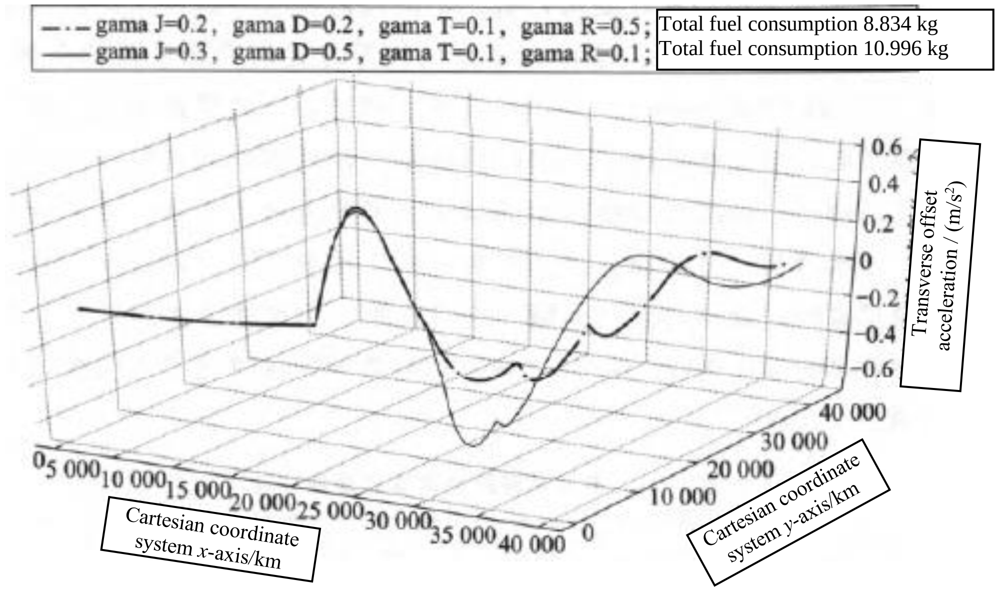


Figure 4-17: Lateral Offset Acceleration in Multi-Objective Optimization Mode (see color insert)

Considering the factors of each index, for equation (4-36), for example, according to $\gamma_J = 0.2, \gamma_D = 0.2, \gamma_T = 0.1, \gamma_R = 0.5$ or $\gamma_J = 0.3, \gamma_D = 0.5, \gamma_T = 0.1, \gamma_R = 0.1$ to assign the weight of each index, the spacecraft lateral offset acceleration curve will fluctuate less and the peak value will be smaller, so that the entire avoidance path can take into account factors such as fuel consumption, minimum offset and braking time while successfully avoiding space debris. It can better meet different practical needs.

4.4.3: Algorithm Comparison

The spacecraft orbit temporary avoidance path planning should not only successfully avoid space debris, take into account the absolute motion of flying along the transfer orbit and the relative motion of the space target, but also meet different avoidance needs and preferences, so the appropriate path generation algorithm is particularly important. The applicability and computational cost of the commonly used path planning algorithms are compared and analyzed, which is helpful to clarify the advantages and scope of application of the algorithms.

(1) Comparison of algorithm applicability

Spacecraft orbit temporary avoidance path planning can be regarded as the development of an optimal or suboptimal avoidance path from the starting orbit to the target orbit in the maneuverable area with reference to a certain index (such as the least cost, the shortest distance and the lowest computation time, etc.), and its essence is to obtain the optimal or feasible solution under a certain avoidance preference.

The orbital temporary avoidance path planning problem under different avoidance needs and preferences can often be converted into a single-objective problem by weighting and other methods and then solved by mathematical programming. However, the objective function and constraint space of the orbital temporary avoidance path planning problem may be nonlinear or non-differentiable under different avoidance needs and preferences, that is, with the increase of the problem model, the solution will become more complex, and the amount of computation will increase exponentially, which will affect the solving efficiency of the algorithm. For the orbital temporary avoidance path planning problem under different avoidance needs and preferences, heuristic algorithm and intelligent algorithm provide a more efficient way to solve the problem. Table 4-1 compares the applicability of two commonly used algorithms.

Table 4-1: Comparison of the applicability of different avoidance path algorithms

The name of the algorithm	Peculiarity	Advantage	Applicability analysis
Dijkstra algorithm	1) A single-source path algorithm from one vertex to the rest of the vertices 2) It expands from the center to the outer layer, and the divergence is good	1) The algorithm is very concise and can effectively obtain the optimal solution 2) There are few parameters, and the convergence speed is fast in the early stage	The path search efficiency decreases with the increase of data nodes, and the search mode is fixed, so it is not easy to add multi-preference constraints
RRT algorithm	1) Sampling-based search algorithm, suitable for solving high-dimensional space problems 2) The search process is similar to the process of growing branches and spreading around	1) Few parameters, simple structure and strong search ability 2) It is easy to add incomplete constraints, and has flexible search capabilities in complex environments	The node utilization rate is low, the path is unstable, and it is difficult to take into account the minimum offset from the set transfer orbit in the path generation process
Path generation algorithm based on improved artificial potential field	1) Spatial planning methods that do not rely on graphical representations 2) A virtual force field method, in which the resultant force generated by gravity and repulsion force realizes obstacle avoidance	1) The planning speed is fast, the real-time avoidance is strong, and the path is smooth 2) The structure is simple, the amount of calculation is small, and it is convenient for the real-time control of the bottom layer	The gravitational potential field can refer to the set transfer orbit, and the comprehensive potential field can consider different avoidance needs and preferences, and it is easy to obtain the avoidance path under different needs and preferences

Dijkstra's algorithm is a typical single-source shortest path algorithm, which is mainly used to calculate the shortest path from one node to all other nodes. Due to the limitation of this extended search method from the center to the outside, the path search efficiency decreases with the increase of data nodes, and it is not easy to add multi-preference constraints to the relatively fixed search mode, which is difficult to directly solve the problem of orbit temporary avoidance path planning under different avoidance needs and preferences.

The Rapid Expansion Random Tree (RRT) algorithm is a sampling-based search algorithm that starts from the root node of the tree and spreads around and returns the path as a feasible path whenever a leaf node touches the target. The RRT algorithm avoids the modeling of the space by detecting the collision of the sampling points in the state space and can effectively solve the path planning problems of high-dimensional space and complex constraints. However, the RRT algorithm has a low utilization rate of nodes and adopts the method of random sampling for a large space such as spacecraft orbit avoidance, and the planned path is more random, and the optimal path acquisition is unstable. The RRT algorithm is difficult to generate an avoidance path from the perspective that the original scheme should be continued as much as possible in the temporary planning, and the generated avoidance path is difficult to take into account the absolute motion of the flight along the transfer orbit and the relative motion of the avoidance space debris.

Based on the path generation algorithm of improved artificial potential field, the comprehensive potential field is constructed with the ideas of reference line traction, long-distance point repulsion ignorance, and obstacle point gravitational attenuation, and the optimal avoidance path is obtained by searching for the descending direction of the potential function. In this process, the target point potential field is replaced by a reference line potential field, so that the spacecraft can closely follow the set transfer orbit on the way to avoid space debris, and the absolute motion of the flight along the transfer orbit is well considered. The comprehensive potential field can be combined with different indicators such as avoidance safety, fuel consumption, minimum offset and braking time, and it is easier to obtain the optimal avoidance path under different needs and preferences.

(2) Comparison of the amount of algorithm computation

In order to illustrate the comparative advantages of the path generation algorithm based on the improved artificial potential field, it is simulated and compared with the classical artificial potential field method and the commonly used shortest path algorithm. Because the commonly used avoidance algorithms such as Dijkstra and RRT cannot take into account the indicators of avoidance safety, fuel consumption, minimum offset and braking time at the same time, it is difficult to directly solve the problem of temporary avoidance path planning of spacecraft orbit. In order to effectively compare the performance of each algorithm, Figure 4-18 designs a separate barrier avoidance scenario, that is, it is necessary to independently plan the shortest path from the dot to the fork point. On the 1.6 GHz, 1.8 GHz dual-core CPU 8 GB RAM computing hardware, the same PyCharm simulation compilation environment is used to solve the problem by using four algorithms.

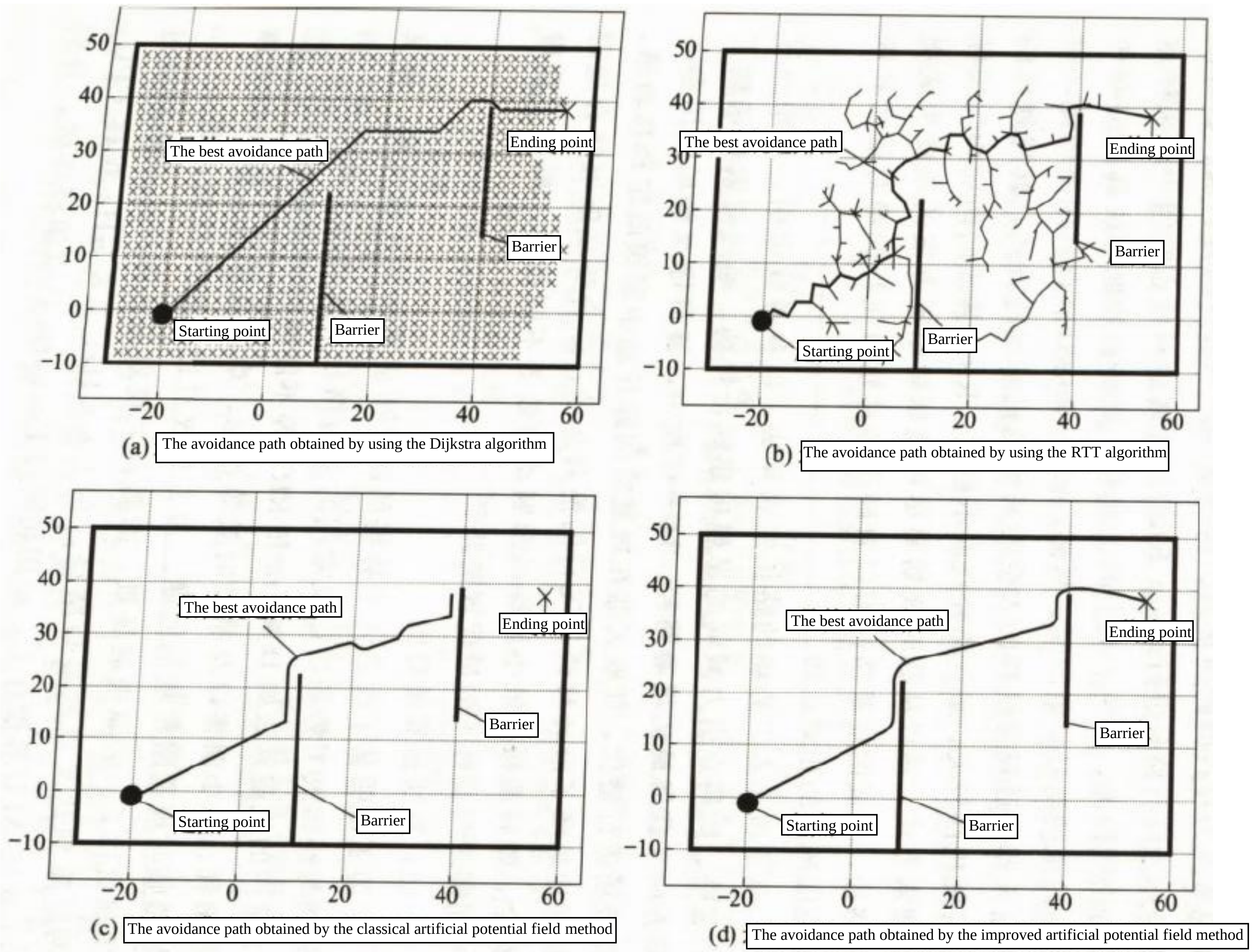


Figure 4-18: Solving Effects of Different Algorithms for Barrier Avoidance Problem (see color insert)

The typical breadth-first search method (Dijkstra algorithm[253, 254]) can be used to obtain the avoidance path shown in Figure 4-18(a). In the process of solving the Dijkstra algorithm, it is necessary to preset the search area (the black box in the figure) and expand the starting point to the outer layer (the small fork in the figure represents the searched node), so that the search time is 0.48 s, and the resulting path is angular, and the path length is 93.3. The evasion path shown in Figure 4-18(b) can be obtained by using an incremental, probabilistic-complete, and suboptimal path planning algorithm (RRT algorithm[255]). The RRT algorithm adopts the search tree method with the initial state as the root node and the target node as the leaf node (the branch in the figure is the searched area), which takes an average time of 0.62 s, and the resulting path is tortuous and has a length of 111.4. The avoidance path shown in Figure 4-18(c) can be obtained by using the classical artificial potential field method. The simulation results show that the path will fluctuate during the approach to the target, and when the gravitational potential field and the repulsive potential field are reversed and blocked by obstacles, local oscillations will occur (the bold part of the curve in the figure), and the target is unreachable. Using the improved artificial potential field method, it takes an average time of 0.29 s to reach the target position, and the obtained path is shown in Figure 4-18(d), with a path length of 91.1 and a better smoothing effect. Compared with the two commonly used path planning algorithms, the average time is shortened by 47%, the distance is saved by 11%, and there is no need for path smoothing, obstacle avoidance detection, etc., which has certain comparative advantages.

4.5: Chapter Summary

To avoid space debris, spacecraft need to take into account the absolute motion of flying along the transfer orbit and the relative motion of space debris avoidance, and it is difficult to plan the path independently, and there are few published research results at home and abroad. In order to solve the above problems, this chapter introduces a path generation algorithm that combines the Frenet coordinate system with the improved artificial potential field. Firstly, the "multi-restriction shortest path" feature presented in the face of spacecraft orbit temporary avoidance is described, and the orbit temporary avoidance path planning problem in the face of space debris is described. Secondly, a space motion coordinate system based on Frenet is constructed, which can take into account the absolute motion of flight along the transfer orbit and the relative motion of space debris avoidance.

Finally, the artificial potential field function is improved, the action area of each potential field is adjusted, and a comprehensive potential field model with reference line traction, long-distance point repulsion neglect and obstacle point gravitational weakening is constructed, which avoids the premature trajectory deviation and local oscillation phenomenon of the traditional artificial potential field method. The case analysis shows that the proposed method further solves the problem that the relative position of the spacecraft and the set transfer orbit is not easy to express in path planning and realizes the simple representation of space avoidance motion. It avoids the premature trajectory deviation and local oscillation of the traditional artificial potential field method and realizes the autonomous avoidance of space debris. It can consider avoidance safety, fuel consumption, minimum offset and braking aging factors at the same time to meet different avoidance needs and preferences. The comparison of the algorithms shows that compared with other commonly used algorithms, the average time taken by the algorithm in this book is shortened by 47% and the distance is reduced by 11%, which further improves the shortcomings of the conventional methods that it is difficult to meet different avoidance preferences at the same time and the orbit balance is weak.

The on-orbit servicing of spacecraft to non-cooperative targets is a deep integration of optimal control and dynamic game, which can be described as a dynamic game problem,[165] which is an orbital game process in which the optimal behavior is adopted under the condition that only the self-state and the opponent's current finite state are known, and the opponent's future behavior strategy is not known. Compared with the traditional space action planning, the spacecraft orbit game needs to consider the control of both sides of the game, and its problem dimension is increased by 1 times, and the behavior of both parties is also considered, and the continuous dynamic interaction characteristics are obvious, resulting in a continuous dynamic interaction problem of bilateral control in strategic planning.[166]

5.1: ORBITAL GAME PROBLEM DESCRIPTION AND MODELING

From the perspective of game theory, the basic elements of the game, such as participants, behaviors, information, strategies and equilibrium, are used to describe the orbital game problem between spacecraft and non-cooperative targets. The dynamic model of orbital game was established, the Nash equilibrium and strategy combination of orbital game were defined, and the objective function of strategy planning was constructed considering distance and time factors.

5.1.1: Problem Description

If the space service target is a non-cooperative target, the spacecraft may face the non-cooperative behavior of the target when approaching the target, which will develop into an orbital game problem between spacecraft in orbit.[257] The orbital game problem between spacecraft reflects the interaction behavior between spacecraft to some extent, and it is appropriate to use zero-sum game theory considering the strategic choice of both parties to describe the orbital game process between such spacecraft.[127, 173] The zero-sum game theory can analyze the influence of different strategy choices on the final result, and can provide an effective pursuit strategy for spacecraft, and its application to the orbit game problem between spacecraft can more appropriately solve the strategy choice problem of spacecraft to deal with non-cooperative goals.

To use zero-sum game theory to deal with orbital game problems between spacecraft, it is necessary to redefine the elements of participants, behavior, information, rationality, payment and equilibrium in game theory from the perspective of spacecraft.

1) Participants in the orbital game: Participants are also known as insiders, and they are the two sides of the game in the situation of non-cooperative countermeasures, that is, spacecraft and space targets that can make independent decisions and maneuver quickly in the orbital game. Participants are the decision-making subjects in the orbital game, and each pursues the maximization of utility or return. In the spacecraft orbit game, the game composed of spacecraft and space target is called a two-person game.

2) Behavior in orbital games: Behavior is the decision-making variable of participants at a certain point in time in orbital games. For each participant, there are multiple possible behaviors to choose from at different moments, and different behavioral choices will lead to different outcomes of the game.

3) Information in the orbital game: Information refers to all the knowledge about the game that the participants in the orbital game have, such as information about the target type, orbit position, state, orbit parameters, etc. According to the sufficiency of the available information, the information can be divided into complete information and incomplete information, which will have an important impact on the outcome of the game.

4) Rationality in orbit game: In the spacecraft orbit game, it is assumed that both participants are rational[1] and will choose behaviors to maximize their interests under certain constraints, that is, in the face of two behavior choices that cannot coexist, each participant will choose the behavior that can make the objective function optimal (maximum or minimum).

5) Strategy in orbital games: Strategy can be understood as a contingent action plan of the participants,[258] which is the rule of behavior, which stipulates the timing of the behavior and stipulates how the participants should act under what circumstances. The purpose of spacecraft orbit game planning is to obtain the best strategy to deal with the problem of on-orbit services for non-cooperative targets.

6) Payouts in orbital games: Payouts are the rewards that each participant receives under the chosen strategy. In the spacecraft orbit game, the payment depends not only on the strategy of spacecraft selection, but also on the behavior strategy of space target, and the purpose of planning is to maximize the combination payment through the selection of strategy combinations.

[1] In game theory, the definition of "rationality" is different from the definition of philosophy and other social sciences, and this book mainly uses the definitions of rationality in game theory by John Von Neumann and Oskay Morgenstern as references.

7) Equilibrium in orbital game: Equilibrium is a combination of strategies composed of the optimal strategies of spacecraft and space target, which is a stable state of the game between the two sides, in which both sides are no longer willing to unilaterally change their strategies. This combination of strategies is unique because when both the spacecraft and the space target adopt their own strategies, neither of them can get more out of changing their strategies on their own.

The orbital game problem from the perspective of spacecraft[260] is a dynamic game process in which the optimal behavior is adopted and the space approximation is finally completed under the condition that only the self-state and the current finite state of the space target are known, and the future behavior strategy of the space target is unknown. However, in addition to operating in a continuous and dynamically changing space environment, the space target in the orbit game problem also has typical non-cooperation, that is, it has the characteristics of non-communication at the information level, non-cooperation in maneuvering behavior, and incomplete prior knowledge.

In the face of non-cooperative goals, spacecraft only know their own situation and the current limited state of the target, and the future behavior strategy of the space target is unknown, and how to use incomplete situation information and limited knowledge and experience to deal with the orbit game between spacecraft and obtain the most reasonable behavior strategy in the current state is the key problem to be solved by the mission planning method.

5.1.2: Problem Modeling

In order to facilitate the analysis of the problem and highlight the key points, the following assumptions are made before modeling the spacecraft orbit game problem:

1) Weaken the influence of the Earth's rotation on the orbital maneuver of spacecraft, and ignore the problem of multi-body perturbation for the time being.

2) In the whole orbit game between spacecraft, both spacecraft and space targets are located in low earth orbit.

3) In the process of spacecraft orbit game, both spacecraft and space target have certain timely command and control and space situational awareness capabilities, and both sides can obtain the position and velocity information at the current moment and in the past time.

4) In the process of spacecraft orbit game, both spacecraft and space target have certain in-orbit computing capabilities and power systems;

in the chase and escape game, both sides use continuous thrust, and the maximum unit mass acceleration is known.

5) The orbit game between spacecraft is a game countermeasure with incomplete and imperfect information, and neither side of the game can know the form of the other party's payment function and payment matrix exactly, that is, the spacecraft only knows its own state and the current finite state of the space target, and does not know the future behavior and strategy of the space target.

The orbital game between spacecraft is a deep integration of orbit optimal control and dynamic game,[261, 165, 263, 169] which is a typical sequential decision oriented to incomplete information.[1] The essence of the process is a continuous dynamic interaction problem of bilateral control. The two sides have conflicting behavioral objectives, with the spacecraft aiming to approach the space target, and the space target changing its original orbit to behave non-cooperatively. Compared with the traditional space rendezvous and docking problems, the orbit game with non-cooperative targets needs to consider bilateral orbit control and sequential decision game, which increases the problem dimension, and at the same time, it also needs to consider the information conditions of both parties in real time, which makes the game countermeasure model more complex. In addition, compared with the general differential countermeasures, the spatial dynamics model is more complex and the orbital maneuvering conditions are more restrictive, so it is a special kind of game problem. This kind of continuous dynamic conflict between two parties can be mathematically described by differential countermeasures[167, 168] and solved with the help of Nash equilibrium.

(1) Orbital game dynamics model

In the two-body model, the central celestial body is used as the reference point, *P* represents the spacecraft in orbit, and *E* represents the space target, and the spatial position relationship between the two is shown in Figure 5-1. In the figure, a reference star in the same orbital plane is used as the coordinate origin *O*, the direction of the line between the reference star and the central celestial body is the *x*-axis, the direction of the orbital velocity in the orbital plane is the *y*-axis, and the *z*-perpendicular to the transfer orbital plane forms a right-hand system with the *x*-axis and *y*-axis. The relative distance between the spacecraft and the space target is much smaller than the orbital radius of the space target, and its dynamic model can be described as follows.[259]

[1] Sequential decision-making refers to the decision-making method used for the optimization of random or uncertain dynamic systems by arranging them in chronological order to obtain various decisions (strategies) in sequence.

$$\begin{cases} \ddot{x}_i(t) = 2\dfrac{\mu}{r^3(t)}x_i(t) + 2\omega(t)\dot{y}_i(t) + \dot{\omega}(t)y_i(t) + \omega^2(t)x_i(t) + \dfrac{T_i}{m_i}u_i^x(t) \\ \ddot{y}_i(t) = -\dfrac{\mu}{r^3(t)}y_i(t) + 2\omega(t)\dot{x}_i(t) + \dot{\omega}(t)x_i(t) + \omega^2(t)y_i(t) + \dfrac{T_i}{m_i}u_i^y(t), \\ \ddot{z}_i(t) = -\omega^2(t)z_i(t) + \dfrac{T_i}{m_i}u_i^z(t) \end{cases} \tag{5-1}$$

where $x_i(t)$, $y_i(t)$ and $z_i(t)(i = \mathrm{P, E})$ represent the components of the spacecraft and the space target in the x-axis, y-axis and z-axis directions, respectively; $\dot{x}_i(t)$, $\dot{y}_i(t)$ and $\dot{z}_i(t)$ represent the first derivative of the coordinate component against time t, respectively; $\ddot{x}_i(t)$, $\ddot{y}_i(t)$ and $\ddot{z}_i(t)$ represent the second derivative of the coordinate component to time t, respectively; $r(t)$ is the orbital radius of the reference star; ω is the reference angular velocity; μ is the gravitational constant of the earth; T_i is continuous thrust; m_i is quality; u_i^x, u_i^y and u_i^z if the behavior control quantities in the three directions, the subscripts i = P and i = E represent the spacecraft and space target, respectively.

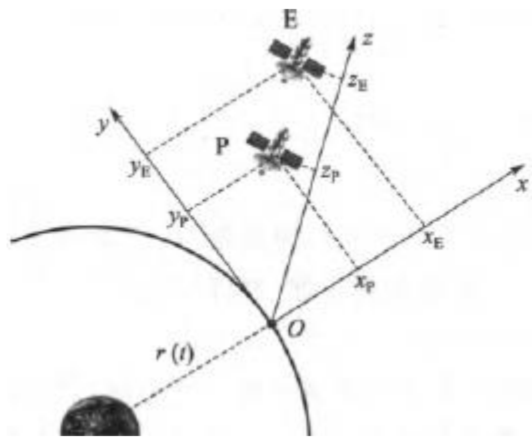


Figure 5-1: Schematic Representation of the Coordinates of a Spacecraft and Space Target

In the survival differential countermeasures,[173] both the spacecraft and the space target adopt the maximum thrust, and the actual behavior control quantity of both sides is the thrust direction angle, i.e., $u_P = [\theta_P, \delta_P]$ and $u_E = [\theta_E, \delta_E]$. Figure 5-2 shows the pitch angle 3 between the thrust force and the orbital plane and the thrust angle θ in the orbital plane.

At this point, the orbital game dynamics model can be expressed as:

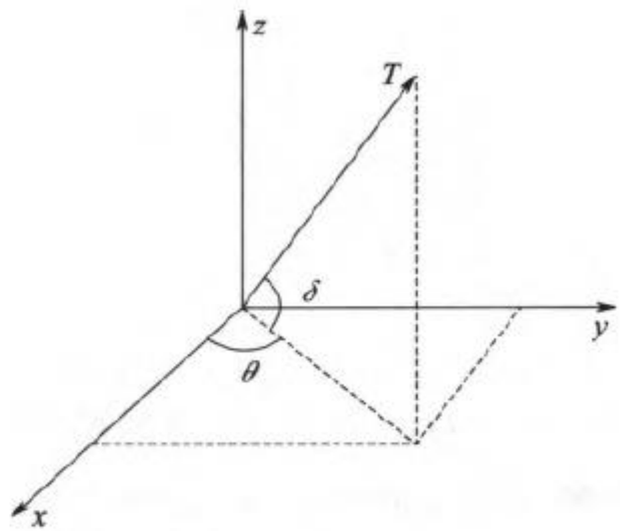


Figure 5-2: Schematic Diagram of the Direction Angle of the Behavior Control Quantity

$$\begin{cases} \ddot{x}_i(t) = 2\dfrac{\mu}{r^3(t)}x_i(t) + 2\omega(t)\dot{y}_i(t) + \dot{\omega}(t)y_i(t) + \omega^2(t)x_i(t) + \dfrac{T_i}{m_i}\cos\delta_i\cos\theta_i \\ \ddot{y}(t) = -\dfrac{\mu}{r^3(t)}y_i(t) + 2\omega(t)\dot{x}_i(t) + \dot{\omega}(t)x_i(t) + \omega^2(t)y_i(t) + \dfrac{T_i}{m_i}\cos\delta_i\sin\theta_i \\ \ddot{z}_i(t) = -\omega^2(t)z_i(t) + \dfrac{T_i}{m_i}\sin\delta_i \end{cases} \tag{5-2}$$

where δ_i is the pitch angle between the thrust and the orbital plane; θ_i is the thrust angle in the orbital plane; the subscripts i = P and i = E indicate spacecraft and space target, respectively.

(2) Orbital game strategy combination

The spacecraft orbit game is dynamic in nature, with the spacecraft interacting with the space target and selecting behaviors at the same time. In this process, both parties are faced with sequential decision-making of behavior, and the general criteria for the decision-making process are defined as:

Definition 5.1 There are 2 participants in a dynamic random game that can be played by $\Gamma = \langle S, u^1, u^2, r^1, r^2, p \rangle$ where S is the state space, u^i is a participant $i(i = 1, 2)$ behavior, $r^i: S \times u^1 \times u^2 \rightarrow R$ be a participant, $i(i = 1, 2)$'s reward value, $p: S \times u^1 \times u^2 \rightarrow \Delta(S)$ is the transition probability, and $\Delta(S)$ is the set of probability distributions over the state space S.

For the bilaterally controlled orbital game problem, the observation process of discrete time points $t = 0, 1, 2, \ldots, n$ is considered.

At each moment, the state is denoted by s_t, and it is assumed that s_t is the value in the set S, that is, the orbital game decision-making process of $s_t \in S$ spacecraft is controlled by the spacecraft and the space target, respectively, and is called participant 1 and participant 2. In state s, each participant independently chooses the behavior $u^1 \in U^1$, $u^2 \in U^2$ and receives the reward $r^1(s, u^1, u^2,)$, $r^2(s, u^1, u^2,)$. For all s, u^1, u^2, the dynamic random game when $r^1(s, u^1, u^2) + r^2(s, u^1, u^2) = 0$ is called a "zero-sum game;"[1] When $r^1(s, u^1, u^2) + r^2(s, u^1, u^2) \neq 0$, the dynamic random game is called the "general sum number game."

In a given state s, participants independently choose behaviors u^1, u^2, and receive a reward $r^i(s, u^1, u^2)$. Then, according to the transition probability, the state is transitioned to the next state s', and the constraints are satisfied

$$\sum_{s' \in S} p(s' \mid s, u^1, u^2) = 1 \tag{5-3}$$

In the spacecraft orbit game, for a given initial state, both the spacecraft and the space target try to obtain the most favorable behavioral strategy for themselves

$$v^1(s, \pi^1, \pi^2) = \sum_{t=0}^{\infty} \beta^t E(r_t^1 \mid \pi^1, \pi^2, s_0 = s) \tag{5-4}$$

$$v^2(s, \pi^1, \pi^2) = \sum_{t=0}^{\infty} \beta^t E(r_t^2 \mid \pi^1, \pi^2, s_0 = s) \tag{5-5}$$

where π^i is the strategy adopted by participant $i(i = 1, 2)$ is the strategy adopted by the participant r_t; $\beta \in [0, 1)$ is the discount factor; s_0 is the initial state; r_t is the reward at time t.

Nash equilibrium is a joint strategy, which means that each participant's equilibrium strategy is the best response to the other participants' strategies. For the spacecraft orbit game, the strategy of each participant is defined over the entire time frame of the game.

Definition 5.2 For dynamic random game r, the Nash equilibrium is for all s ∈ S are all set up in the strategy portfolio (π_*^1, π_*^2)

$$v^1(s, \pi_*^1, \pi_*^2) \geq v^1(s, \pi_*^1, \pi_*^2) \tag{5-6}$$

[1] "Zero-sum game" is a concept in game theory, which is a non-cooperative game, which refers to the two sides participating in the game, under strict competition, the gain of one party must mean the loss of the other party, and the sum of the gains and losses of all parties in the game is always "zero."

$$v^2(s,\pi^1_*,\pi^2_*) \geqslant v^2(s,\pi^1_*,\pi^2_*) \tag{5-7}$$

(3) Orbital game objective function

Spacecraft and space targets have different goals and expectations in the orbital game, and have different preferences for outcomes, so they can optimize their expectations through strategies or behaviors. Since the desired optimization is the goal of each participant, it is also called the objective function of the participants.

For the objective function of the spacecraft orbit game, the Euclidean distances of both sides of the game are first considered

$$j(u_P,u_E) = \| [x_P - x_E, y_P - y_E, z_P - z_E]^T_{|t=t_f} \|_2^2 \tag{5-8}$$

where, $\|\bullet\|_2$ is the Euclidean norm; t_f is the moment at which the thrust action of t_0 ends for the start moment.

For continuous thrust, the fuel consumption is directly proportional to the thrust action time, and the longer the thrust action time, the more fuel consumption is consumed. Therefore, the thrust action time interval is also taken as a part of the objective function of the orbit game, and the objective function of the time-distance comprehensive optimal control is constructed on the basis of Eq. (5-8).

$$J(u_P,u_E) = k\int_{t_0}^{t_f} dt + (1-k)\int_{t_0}^{t_f} j(u_P,u_E)\, dt \tag{5-9}$$

where $J(\bullet)$ is the objective function of both sides of the game; k is the proportional weight, and $k \in [0, 1]$.

In the orbital game, the spacecraft and the space target take behavior by independently optimizing the objective function according to the current state. In the meantime, the spacecraft will seek to obtain behavioral strategies that minimize the objective function, while the space target will expect to obtain behavioral strategies that maximize the objective function. According to the Nash Equilibrium (NE) theory in game theory,[1][231, 205] the behavior strategy of both parties achieves a Nash equilibrium if and only if the following inequalities are satisfied

$$J(u_P^*,u_E) \leqslant J(u_P^*,u_E^*) \leqslant J(u_P,u_E^*) \tag{5-10}$$

where u_P is the behavior strategy of the spacecraft; u_E is the behavioral strategy of the spatial target; u_P^* is the Nash equilibrium strategy for spacecraft; u_E^* is a Nash equilibrium strategy for spatial targets.

The Nash equilibrium strategy in the two-sided game is a combination composed of the optimal strategies of both sides, that is, a special strategy combination.

[1] Nash equilibrium, also known as non-cooperative game equilibrium, is a game theory named after John Nash.

For the Nash equilibrium strategy, when both parties adopt their own strategies in this combination, neither side can achieve better results by taking the other strategies independently. That is to say, the spacecraft chooses the Nash equilibrium behavior strategy u_P^*, and the space target adopts irrational behavior, that is, any behavior other than the Nash equilibrium u_E, which will lead to the suboptimal objective function of the space target.

Consider two participants in the spacecraft orbit game problem, which have the same objective function but are in different directions of optimization. This means that although the objective function is the same, one actor wants to minimize the objective function, while the other actor strives to maximize the objective function. For the problem of bilateral control, both sides of the game have their own behavioral strategies. Let the table u_P, u_E show the behavior strategies of both sides of the game, so the spacecraft orbit game strategy can be expressed as

$$[u_P, u_E]^T = \arg \min_{u_P} \max_{u_E} J \tag{5-11}$$

Therefore, the purpose of solving the spacecraft orbit game problem is to find a set of behavioral strategies to satisfy the Nash equilibrium, even if the following equation is true

$$J^* = \min_{u_P^*} \max_{u_E^*} J = \max_{u_E^*} \min_{u_P^*} J \tag{5-12}$$

5.2: Game Strategy Solving Algorithm Based on Branched Deep Reinforcement Learning

The solution of differential countermeasures for spacecraft orbit game problems has always been a difficult and thorny problem because of the complexity of differential equations, nonlinear constraints, and many state variables.[141, 183] In order to further improve the shortcomings of conventional methods that are difficult to deal with the problem of bilateral control and the small convergence domain of equilibrium strategy, on the basis of testing the existence and consistency of orbital game strategy, the game strategy solving process is regarded as a kind of Markov decision-making process of bilateral control, and a fuzzy reasoning model for continuous space solution is constructed, and the application limitation of deep reinforcement learning in continuous space is improved, and a game strategy solving algorithm based on branched deep reinforcement learning is proposed.[262]

5.2.1: Existentiality and Consistency Testing of Game Strategy

Through the description and modeling of the spacecraft orbit game problem, the zero-sum differential countermeasures are used to describe the spacecraft orbit game process, which can not only effectively express the spacecraft game strategy, but also reflect the countermeasures of space targets, and fully reflect the bilateral control and sequential game process between spacecraft. However, the equilibrium value of the general differential countermeasure problem is not uniquely determined and may not exist or may be multiple. In order to illustrate the rationality of using differential countermeasures to describe the spacecraft orbit game problem, this section analyzes the existence and consistency of the Nash equilibrium solution in the spacecraft orbit game problem and clarifies the scope of application of the method.

(1) The existence test of orbital game strategy

In the orbital game problem, both sides have opposite return goals and strive to adopt the most beneficial behavior strategy, and the Nash equilibrium strategy is not the only one. In this regard, it is necessary to test the existence of the Nash equilibrium strategy of the orbital game between spacecraft and space targets, that is, to clarify under what conditions and circumstances the Nash equilibrium strategy must exist.

Hypothesis 5.1 Behavioral policy set u_P and u_E is a compact set, an objective function in the metric space J: $u_p \times u_E \rightarrow R$ at $u_p \times u_E$ on continuous.

Definition 5.3[127] For the orbital game problem, if the strategy is fixed separately $u_E \in U_E$ and $u_P \in U_P$, which defines the optimal behavioral policy set of spacecraft and space target as

$$S^P(u_P^*) = \{u_P \in U_P : J(u_P, u_E^*) = \min_{u_P} J(u_P, u_E)\} \quad (5\text{-}13)$$

$$S^E(u_P^*) = \{u_E \in U_E : J(u_P^*, u_E) = \max_{u_E} J(u_P, u_E)\} \quad (5\text{-}14)$$

Definition 5.4 The upper countermeasure value $V+$ and the lower countermeasure value $V-$ of the orbital game problem are defined as

$$V^+ = \min_{u_P} \max_{u_E} J(u_P, u_E) = \min_{u_P} J(u_P, u_E^*(u_P)) \quad (5\text{-}15)$$

$$V^- = \max_{u_E} \min_{u_P} J(u_P, u_E) = \max_{u_E} J(u_P^*(u_E), u_E) \quad (5\text{-}16)$$

For any n>0 there will be a corresponding strategy $u_E^n \in U_E$ that makes the following equation true[175]

$$V^- - n \leqslant J(u_P(u_E^n), u_E^n) \quad (5\text{-}17)$$

Thus, it can be concluded that the value of the upper countermeasure is greater than or equal to the value of the lower countermeasure

$$V^+ = \min_{u_P} \max_{u_E} J(u_P, u_E) \geqslant \min_{u_P} J(u_P, u_E^n) = J(u_P(u_E^n), u_E^n) \geqslant V^- \quad (5\text{-}18)$$

In the same way, for any n>0 there is a corresponding strategy $u_P^n \in U_P$ such that the value of the lower countermeasure is greater than or equal to the value of the upper countermeasure

$$V^- = \max_{u_E} \min_{u_P} J(u_P, u_E) \geqslant \max_{u_E}(u_P^n, u_E) = J(u_P^n, u_E(u_P^n)) \geqslant V^+ \tag{5-19}$$

From this, it can be seen that there are strategies $u_P^n \in U_P$ and $u_E^n \in U_E$

$$V^+ = J(u_P^n, u_E(u_P^n)) = J(u_P(u_E^n), u_E^n) = V^- \tag{5-20}$$

When u_P^n ($n \in [1, N]$) is a series of behavioral strategies of a spacecraft, let $u_E(u_P^n)$ be the behavioral strategy of the space target corresponding to u_P^n. For any $n \geq 1$, there is a behavioral strategy that is $u_E^* \in U_E$ satisfied

$$J(u_P^n, u_E^*) \leqslant J(u_P^n, u_E(u_P^n)) = V^+ \tag{5-21}$$

According to the continuity of the objective function J of the orbital game, it can be seen that u_E^* is the Nash equilibrium strategy corresponding to u_P^n

$$J(u_P^n, u_E^*) = \max_{u_E} J(u_P^n, u_E) = J(u_P^n, u_E(u_P^n)) \tag{5-22}$$

In the same way, for the space target, u_E^* ($n \in [1,N]$) the behavior strategy $u_P(u_E^n)$ is the corresponding behavior strategy of the spacecraft, and for any $n \geq 1$, there is a behavior strategy $u_P^* \in U_P$ satisfy

$$J(u_P^*, u_E^n) \geqslant J(u_P(u_E^n), u_E^n) = V^- \tag{5-23}$$

According to the continuity of the objective function J of the orbital game, it can be seen that u_P^* is the Nash equilibrium strategy corresponding to u_E^n

$$J(u_P^*, u_E^n) = \min_{u_P} J(u_P, u_E^n) = J(u_P(u_E^n), u_E^n) \tag{5-24}$$

Therefore, it is concluded that under the condition of hypothesis 5.1, when $V^+ = V^-$, the Nash equilibrium solution of the orbital game problem exists.

(2) Consistency test of orbital game strategy

In the planning process of the optimal strategy of spacecraft orbit game, the Nash equilibrium strategy is not the only one determined, and there may be multiple equilibrium problems. Therefore, it is necessary to further analyze the consistency of the policy values on the basis of the existence of Nash equilibrium strategies and clarify the relationship between the corresponding countermeasure values of different Nash equilibriums.

In order to better illustrate the characteristics of the orbital game, the following assumptions and theorems are given.

Hypothesis 5.2 $f(x)$ is a bounded function satisfying Lipschitz's continuous[1], i.e., there is L>0, there is

$$|f(t,x_1,u_P,u_E)-f(t,x_2,u_P,u_E)|\leqslant L|x_1-x_2| \tag{5-25}$$

And there is a positive real number $K \in R^+$ satisfies it

$$|f(t,x,u_P,u_E)|\leqslant K(1+|x|) \tag{5-26}$$

Theorem 5.1[127] If $V^+(t, x)$ is the value of the upper countermeasure at state x at time t, then for any $h \in [0, t_f - t_0]$ there is

$$V^+(t_0,x_0) = \min_{u_P^\pi \in U_P} \max_{u_E^\pi \in U_E} \left\{\int_{t_0}^{t_0+h} \ell(t,x_t,u_P,u_E)\,dt + V^+(t_0+h,x_t(t_0+h))\right\} \tag{5-27}$$

where $x_t(t_0, x_0, u_P, u_E)$ is the x-axis direction component in Eq. (5-2) with (t_0, x_0) as the initial condition; $\ell(t, x_t, u_P, u_E)$ is the process component of the objective function equation (5-9), and together with the terminal component $\phi(x_t(t_0))$ forms the objective function $J(u_P, u_E)$

$$J(u_P,u_E)=\phi(x_t(t_0))+\int_{t_0}^{t_0+h}\ell(t,x_t,u_P,u_E)\,dt \tag{5-28}$$

Let equation (5-28) be equal to 0, i.e.

$$\min_{u_P^\pi \in U_P} \max_{u_E^\pi \in U_E} \left\{\frac{1}{h}\int_{t_0}^{t_0+h}\ell(t_0,x(t_0),u_P,u_E)\,dt + \frac{V^+(t_0+h,x_t(t_0+h))-V^+(t_0,x_t(t_0))}{h}\right\}=0 \tag{5-29}$$

If V^+ is a first-order differentiable function, there is when $h \to 0$

$$\min_{u_P^\pi \in U_P} \max_{u_E^\pi \in U_E} \{\ell(t_0,x(t_0),u_P,u_E) + \partial_t V^+ + DV^+ \cdot f(t_0,x(t_0),u_P,u_E)\}=0 \tag{5-30}$$

where $\partial_t(\cdot)$ is the first-order partial derivative of the function for time t; D is the first-order partial derivative of the function to the state quantity x.

[1] Lipschitz definition of continuity: There is a function $f(x)$ if there is a constant L, such that $f(x)$ satisfies the following conditions for any two values (real or complex) on the defined domain: $|f(x_1) - f(x2)| \le L\,|x1 - x2|$. Then the function $f(x)$ satisfies the Lipschitz continuity, and L is called the Lipschitz constant of $f(x)$.

Eq. (5-30) is referred to as the Hamilton-Bellman-Isaacs equation in differential countermeasures.[159, 226]

Define H^+ as:

$$H^+(t, x, \zeta) = \min_{u_P^\pi \in U_P} \max_{u_E^\pi \in U_E} \{\ell(t_0, x_0, u_P, u_E) + \zeta f(t_0, x_0, u_P, u_E)\} \tag{5-31}$$

Then the Hamilton-Bellman-Isaacs equation can be expressed as

$$\partial_t V^+ + H^+(t, x, DV^+) = 0 \tag{5-32}$$

According to hypothesis 5.1, $V^+(t, x)$ is a continuous function, so $V^+(t, x) - \varphi(t, x)$ has a local maximum at the point (t_0, x_0).

$$V^+(t_0, x_0) - \varphi(t_0, x_0) \geqslant V^+(t, x) - \varphi(t, x) \tag{5-33}$$

According to Theorem 5.1, there is h_0 for any $h \in (0, h_0)$.

$$\min_{u_P^\pi \in U_P} \max_{u_E^\pi \in U_E} \left\{\int_{t_0}^{t_0+h} \ell(t, x_t, u_P, u_E)\,dt + \varphi(t_0 + h, x_t(t_0 + h)) - \varphi(t_0, x_0)\right\} \geqslant 0 \tag{5-34}$$

where, for any given behavior u_P has

$$\max_{u_E^\pi \in U_E} \left\{\int_{t_0}^{t_0+h} \ell(t, x_t, u_P, u_E)\,dt + \varphi(t_0 + h, x_t(t_0 + h)) - \varphi(t_0, x_0)\right\} \geqslant 0 \tag{5-35}$$

Thus, for arbitrary ε and h, there is a u_E^* satisfaction

$$\int_{t_0}^{t_0+h} \ell(t, x_t, u_P, u_E^*)\,dt + \varphi(t_0 + h, x_t(t_0 + h)) - \varphi(t_0, x_0) \geqslant \varepsilon h \tag{5-36}$$

From the boundedness of the hypothetical 5.2 function, ℓ is a consistent continuous function, which can be obtained

$$\left| \int_{t_0}^{t_0+h} \ell(t, x_t, u_P, u_E^*)\,dt - \int_{t_0}^{t_0+h} \ell(t_0, x_0, u_P, u_E^*)\,dt \right|$$

$$\leqslant \int_{t_0}^{t_0+h} |\ell(t, x_t, u_P, u_E^*) - \ell(t_0, x_0, u_P, u_E^*)|\,dt$$

$$\leqslant \int_{t_0}^{t_0+h} \mathrm{lip}(\ell)(|t - t_0| + |x - x_0|)\,dt$$

$$\leqslant \int_{t_0}^{t_0+h} \mathrm{lip}(\ell)(|t - t_0| + \|f\|_\infty |x - x_0|)\,dt = o(h) \tag{5-37}$$

Since φ is a function on Eq. (5-27), there is

$$\varphi(t_0 + h, x(t_0 + h)) - \varphi(t_0, x(t_0))$$

$$= \int_{t_0}^{t_0+h} \{\partial_t \varphi(t, x_t) + D\varphi(t, x_t) \cdot f(t_t, x_t, u_P, u_E^*)\}\,dt \tag{5-38}$$

It can be known according to the Lipschitz continuity of $f(t_t, x_t, u_P, u_E^*)$ and the continuity of $\partial_t\varphi(t, x_t)$ and $D\varphi(t, x_t)$

$$\left| \int_{t_0}^{t_0+h} \partial_t \varphi(t, x_t)\,dt - h\,\partial_t \varphi(t, x_t) \right| \leqslant o(h) \tag{5-39}$$

$$\left| \int_{t_0}^{t_0+h} D\varphi(t, x_t) \cdot f(t_t, x_t, u_P, u_E^*)\,dt - \int_{t_0}^{t_0+h} D\varphi(t_0, x_0) \cdot f(t_0, x_0, u_P, u_E^*)\,dt \right| \leqslant o(h) \tag{5-40}$$

Substituting Eq. (5-37), Eq. (5-39) and Eq. (5-40) into Eq. (5-36) obtains

$$\int_{t_0}^{t_0+h} \{f(t_0, x_0, u_P, u_E^*) + D\varphi(t_0, x_0) \cdot f(t_0, x_0, u_P, u_E^*)\}\,dt +$$
$$h\,\partial_t \varphi(t_0, x_0) \geqslant o(h) - \varepsilon h \tag{5-41}$$

Generalize to general behavioral strategies

$$\int_{t_0}^{t_0+h} \max_{u_E^* \in U_E} \{f(t_0, x_0, u_P, u_E) + D\varphi(t_0, x_0) \cdot f(t_0, x_0, u_P, u_E)\}\,dt +$$
$$h\,\partial_t \varphi(t_0, x_0) \geqslant o(h) - \varepsilon h \tag{5-42}$$

Divide both sides of Eq. (5-42) by $h(h>0)$.

$$\frac{1}{h}\int_{t_0}^{t_0+h}\max_{u_E^n\in U_E}\{f(t_0,x_0,u_P,u_E)+D\varphi(t_0,x_0)\cdot f(t_0,x_0,u_P,u_E)\}\,dt+\partial_t\varphi(t_0,x_0)\geqslant\frac{o(h)}{h}-\varepsilon$$

(5-43)

When $\varepsilon\to 0$ and $h\to 0$, available

$$\max_{u_E^n\in U_E}\{\ell(t_0,x_0,u_P,u_E)+\partial_t\varphi(t_0,x_0)+D\varphi(t_0,x_0)\cdot f(t_0,x_0,u_P,u_E)\}\geqslant 0$$

(5-44)

Because of the arbitrariness of u_P, there is

$$\min_{u_P^n\in U_P}\max_{u_E^n\in U_E}\{\ell(t_0,x_0,u_P,u_E)+\partial_t\varphi(t_0,x_0)+D\varphi(t_0,x_0)\cdot f(t_0,x_0,u_P,u_E)\}\geqslant 0$$

(5-45)

According to the weak solvability of the Hamilton-Bellman-Isaacs equation,[264, 266] there is a local maximum for any test function $\varphi(t, x)\in([0, T]\times R^n)$ at point (t_0, x_0) for $V^+(t_0, x_0)-\varphi(t, x)$, and satisfies Eq. (5 – 45), then V^+ is a valid solution of Eq. (5 – 32).

Similarly, for the next countermeasure value $V^-(t, x)$ at time t state x, the Hamilton-Bellman-Isaacs equation can be expressed as

$$\partial_t V^- + H^-(t,x,DV^-)=0 \qquad (5\text{-}46)$$

thereinto

$$H^-(t,x,\zeta)=\max_{u_E^n\in U_E}\min_{u_P^n\in U_P}\{\ell(t_0,x_0,u_P,u_E)+\zeta f(t_0,x_0,u_P,u_E)\}$$

(5-47)

For any test function $\varphi(t, x)\in([0, T]\times R^n)$ at point (t_1, x_1) with a $V(t, x)-\varphi(t, x)$ local minima and satisfies Eq. (5 – 48), then V^- is a valid solution of Eq. (5 – 47).

$$\max_{u_E^n\in U_E}\min_{u_P^n\in U_P}\{\ell(t_1,x_1,u_P,u_E)+\partial_t\varphi(t_1,x_1)+D\varphi(t_1,x_1)\cdot f(t_1,x_1,u_P,u_E)\}\leqslant 0$$

(5-48)

Under the conditions of hypothesis 5.1 and hypothesis 5.2 in this book, it can be seen that Eq. (5-49) holds according to the theoretical derivation and proof of Ref. 167

$$H^{+}(t,x,\zeta)=H^{-}(t,x,\zeta),\ \forall(t,x,\zeta)\in[t_0,t_f]\times R^{2n}\times R^{2n} \quad (5\text{-}49)$$

According to Eq. (5-18) and Eq. (5-19), we can see that $V+\geq V^{-}$ and $V+\leq V^{-}$ always hold. Therefore, for any valid countermeasure value, $V = V+ = V^{-}$.

Therefore, it is concluded that under the conditions of hypothesis 5.1 and hypothesis 5.2, if there are multiple Nash equilibrium strategies for the orbital game problem, then different Nash equilibrium strategies will correspond to the same countermeasure value.

5.2.2: Markov Decision for Game Strategy Solving

In order to facilitate the deep reinforcement learning algorithm to solve the game strategy, the game strategy solving can be regarded as a bilaterally controlled Markov decision-making process.[1] The general definition of the Markov decision-making process is:

Definition 5.5[194] The Markov decision-making process can be made by: $\langle S, U, R, P\rangle$ where S is the state space and U is the behavior space, R: $S\times U \rightarrow R$ is a reward function for participants, P: $S\times U \rightarrow \Delta$ is the transition probability function, and $\Delta(S)$ is the set of probability distributions over the state space S.

In the Markov decision-making process, the goal of the spacecraft (participant) is to find a strategy π maximize the sum of the desired rewards.

$$v(s,\pi)=\sum_{t=0}^{\infty}\beta^{t}E(R_t\mid\pi,s_0=s) \quad (5\text{-}50)$$

where $\beta\in[0, 1)$ is the discount factor; s_0 is the initial state; R_t is the reward at time t.

Break down Eq. (5-50) into:

[1] Markov decision process, named after the Russian mathematician Andrey Markov, is a mathematical model of sequential decision based on Markov chain, which mainly simulates stochastic strategies and returns with time in a dynamic environment with Markov properties.

$$v(s,\pi)=R(s,u_{\pi})+\beta\sum_{s'}p(s'|s,u_{\pi})v(s',\pi) \tag{5-51}$$

where u_{π} is the behavior that is selected according to the policy π.

There is an optimal strategy π^* for arbitrary states $s \in S$ that makes the following Bellman equation true[194]

$$v(s,\pi^*)=\max_{u}\left\{R(s,u)+\beta\sum_{s'}p(s'|s,u)v(s',\pi^*)\right\} \tag{5-52}$$

where $v(s, \pi^*)$ is the optimal strategy for state $s \in S$.

If the spacecraft knows the reward function and the state transition function, it can obtain the strategy π^* through some iterative search methods. When the spacecraft does not know the reward function or the probability of state transition, the dynamic random game problem arises. This requires the spacecraft to interact with the environment in real time to get the best strategy. The spacecraft can solve its optimal strategy by learning the reward function and the state transition function and then using Eq. (5-52). This approach is known as "model-based reinforcement learning." However, spacecraft can also learn their optimal strategy directly without knowing the reward function or state transition function and call this approach "model-free reinforcement learning." [182]

According to Eq. (5-52), basic model-free reinforcement learning can be defined as:

$$Q^*(s,u)=R(s,u)+\beta\sum_{s'}p(s'|s,u)v(s',\pi^*) \tag{5-53}$$

According to the definition of Eq. (5-53), $Q^*(s, u)$ will take action in state s and follow the optimal strategy to get the corresponding reward. It can be obtained according to Eq. (5-52).

$$v(s,\pi^*)=\max_{s}Q^*(s,u) \tag{5-54}$$

It can be seen that if $Q^*(s, u)$ is known, the optimal strategy π^* can be found; at the same time, according to this strategy, the optimal behavior can always be obtained by maximizing $Q^*(s, u)$ in the state s.

In model-free reinforcement learning, the spacecraft arbitrarily initializes $Q(s, u)$ in state $s \in S$ and behavior $u \in U$. At each moment, the spacecraft will select a behavior and observe its reward, and then the q value will be updated according to the formula

$$Q_{t+1}(s,u)=(1-\alpha_t)Q_t(s,u)+\alpha_t\left[R_t+\beta\max_{b}Q_t(s',b)\right] \tag{5-55}$$

where $\alpha_t \in [0, 1)$ is the learning rate, which usually needs to decay over time in order for the method to converge.

5.2.3: Fuzzy Inference Models for Continuous Space Solving

However, the traditional deep reinforcement learning method may lead to the Curse of Dimensionality due to its intractability, large continuous state space and behavioral space.1[180] In order to avoid this problem, according to the conclusion that "fuzzy inference is a universal approximator that can approximate arbitrary nonlinear functions with any accuracy,"[221] a fuzzy inference model of spatial behavior is constructed in this section to realize the mapping transformation of continuous state through fuzzy inference to continuous behavior output, which is conducive to giving full play to the advantages of discrete behavior methods of deep reinforcement learning.

Takagi-Sugeno-Kang (TSK),2[160] as the most commonly used fuzzy inference model, after characterizing the continuous state space or behavior space through the membership function (MF),[162] The IF-THEN fuzzy rule3 can be used to obtain the mapping relationship between the fuzzy set and the output linear function[232]

R_l: IF x_1 is A_1^l AND x_2 is A_2^l AND⋯AND x_i is A_1^l THEN $u_l = c_l$ (5-56)

where R_l is the lth rule in the fuzzy inference model (l = 1, ⋯, L); x_i is the input variable passed to the fuzzy model (i = 1, ⋯, n); A_i^l is the fuzzy set corresponding to the input variable x_i; u_l rule R_l output function; c_l is a constant that describes the center of a fuzzy set.[104]

Figure 5-3 illustrates the fuzzy inference model of spatial behavior when the input quantity is n = 2 and the membership function y = 3. The model has a five-layer network structure, in which small circles represent variable nodes and small boxes represent operation nodes.

1 The Curse of Dimensionality usually refers to a phenomenon in vector operations where the computational workload increases exponentially with the increase in dimensions.

2 Takagi-Sugeno-Kang is a well-known fuzzy inference model, which is usually inferred with the help of the "IF-THEN" rule, and is widely used in the fields of system recognition, data mining, pattern recognition, and image processing because of its good approximation performance.

3 The IF-THEN fuzzy rule is in the form if x is A then y is B, "x is A" is called the premise, and "y is B" is called the conclusion, where A and B are the linguistic values defined by the fuzzy set on the domains x and y.

In general, if there are "continuous spatial variables $x_i(i = 1, \cdots, n)$ as input, after using y membership functions for each variable x_i, the exact input u can be obtained through the fuzziness and disambiguation process, and the functions of each layer are as follows:

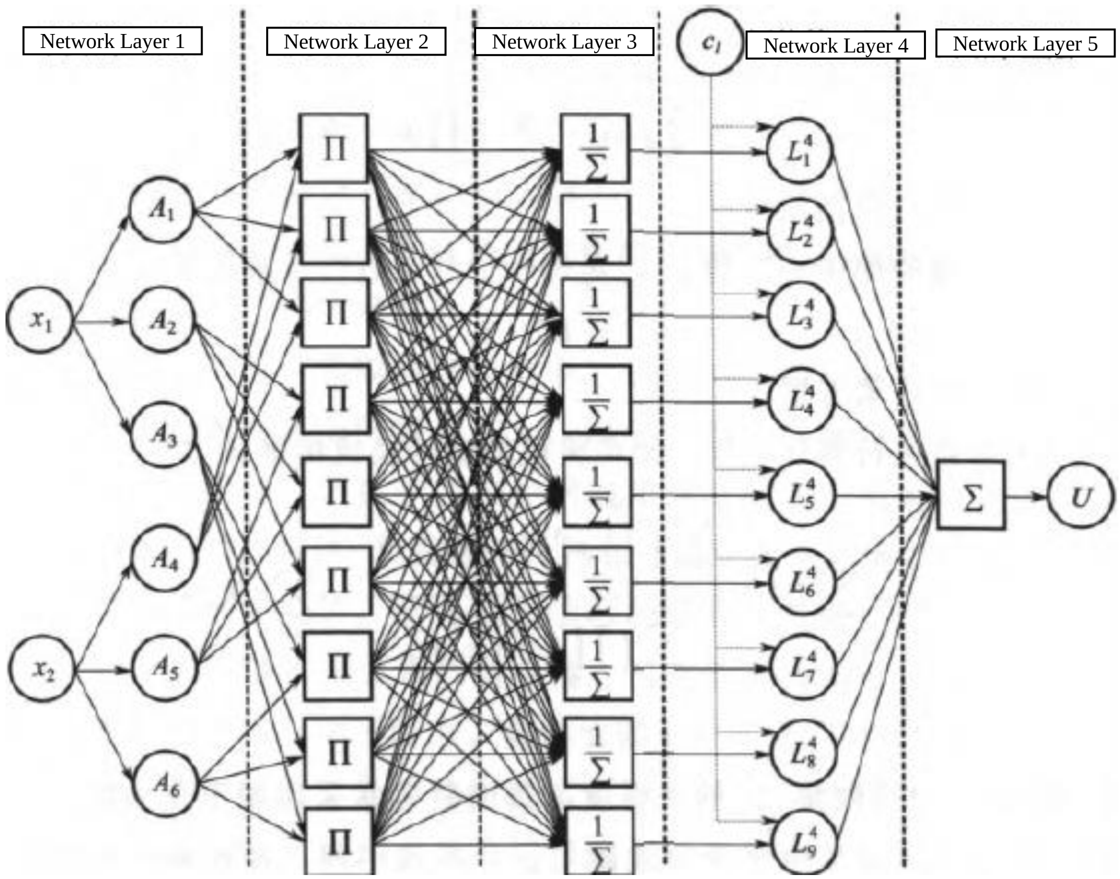


Figure 5-3: Fuzzy Interference Model of Spatial Behavior

(1) Network Layer 1

After the input variables are processed by fuzzy functions, there are $(n \cdot y)$ adaptive output nodes. According to Eq. (5-56), the output of each node is the membership degree $\mu^{A_i^l}$ of its input variable x_i.

(2) Network Layer 2

Direct product inference is adopted for fuzzy sets,[162] that is, cross-multiplication of each membership is performed on $L(L = y^n)$ operation nodes

$$L_l^2 = \prod_{i=1}^{n} \mu^{A_i^l}(x_i) \tag{5-57}$$

(3) Layer 3 of the network

In order to achieve the weighted average disambiguation, the membership degree was normalized

$$L_l^3 = \Psi^l = \frac{L_l^2}{\sum_{l=1}^{L} L_l^2} = \frac{\prod_{i=1}^{n} \mu^{A_i^l}(x_i)}{\sum_{l=1}^{L}\left(\prod_{i=1}^{n} \mu^{A_i^l}(x_i)\right)} \tag{5-58}$$

(4) Layer 4 of the network

The fuzzy set center constant c_l is introduced, and the point multiplication operation is performed on each node

$$L_l^4 = L_l^3 \cdot c_l \tag{5-59}$$

(5) Layer 5 of the network

By accumulating nodes, the amount of ambiguity can be converted to an exact amount.[238]

$$L^5 = u = \sum_{l=1}^{L} L_l^4 = \frac{\sum_{l=1}^{L}\left(\prod_{i=1}^{N} \mu^{A_i^l}(x_i) \cdot c_l\right)}{\sum_{l=1}^{L}\left(\prod_{i=1}^{N} \mu^{A_i^l}(x_i)\right)} = \sum_{l=1}^{L}(\Psi^l \cdot c_l) \tag{5-60}$$

where $\mu^{A_i^l}$ is the membership degree of the fuzzy set A_i^l, and its function is usually described graphically. Among them, Gaussian membership function is widely used in fuzzy inference models due to its simple formula and high computational efficiency.

The Gaussian membership function can be expressed as

$$\mu^{A_i^l}(x_i) = \exp\left(-\left(\frac{x_i - m_i^l}{\sigma_i^l}\right)^2\right) \tag{5-61}$$

Where m_i^l is the mean of the Gaussian membership function; σ_i^l is the variance of the Gaussian membership function.

5.2.4: Branched Deep Reinforcement Learning of Orbital Games

Although Deep Reinforcement Learning is an effective combination of neural network and reinforcement learning, when directly applied to the spatial behavior fuzzy inference model, it will face the problem of the combination of the number of behaviors and the mapping rules, which greatly weakens the behavior control decision-making ability after discretization processing.

In addition, the naïve distribution of value functions and the representation of strategies across multiple independent function approximators likewise encounter many difficulties, leading to convergence problems.[170]

To solve this problem, a new branched deep reinforcement learning architecture is proposed. The representation of the state behavior value function is distributed on multiple network branches, and the independent training and fast processing of discrete behaviors are realized through multiple groups of parallel neural networks. While sharing a behavior decision-making module, the state behavior value function is decomposed into a state function and a dominance function to achieve an implicit centralized coordination. The game interaction process between spacecraft and space target is given, and the stability of the method and the convergence of good strategies can be realized after appropriate training.

(1) Multiple groups of parallel network branches

According to the L rule in the spatial behavior fuzzy inference model, the representation of the state behavior value function is distributed on multiple network branches, and the L-group parallel neural network is constructed. A multi-group parallel neural network is a multi-group parallel neural network added to a single neural network. Similar to a single-group neural network,[251] a parallel neural network is autonomously trained and makes independent decisions in constant interaction with the environment. Combined with the game and feedback mechanism of reinforcement learning, multiple groups of parallel neural networks will have stronger autonomy, flexibility and coordination, which will greatly improve the independent learning ability of discrete behaviors and enhance the overall exploration ability of the environment.

Figure 5-4 shows the multiple groups of parallel neural networks in the branched deep reinforcement learning architecture. When the state information is input into the L group parallel neural network, the forward transmission is carried out independently through the excitation function and the gradient descent reverse training is carried out, and the discrete behavior function (referred to as q function) can be obtained by output.

(2) Shared behavioral decision-making module

For a fuzzy inference model with n input quantities and y membership functions, directly using traditional reinforcement learning methods requires considering y^n possible q functions at the same time. This makes reinforcement learning methods tricky and even difficult to explore effectively in multi-discrete behavior applications.[213]

In the shared behavior decision-making module constructed, this book improves the traditional reinforcement learning method. Figure 5-5 shows a schematic diagram of shared behavior decision-making based on improved reinforcement learning.

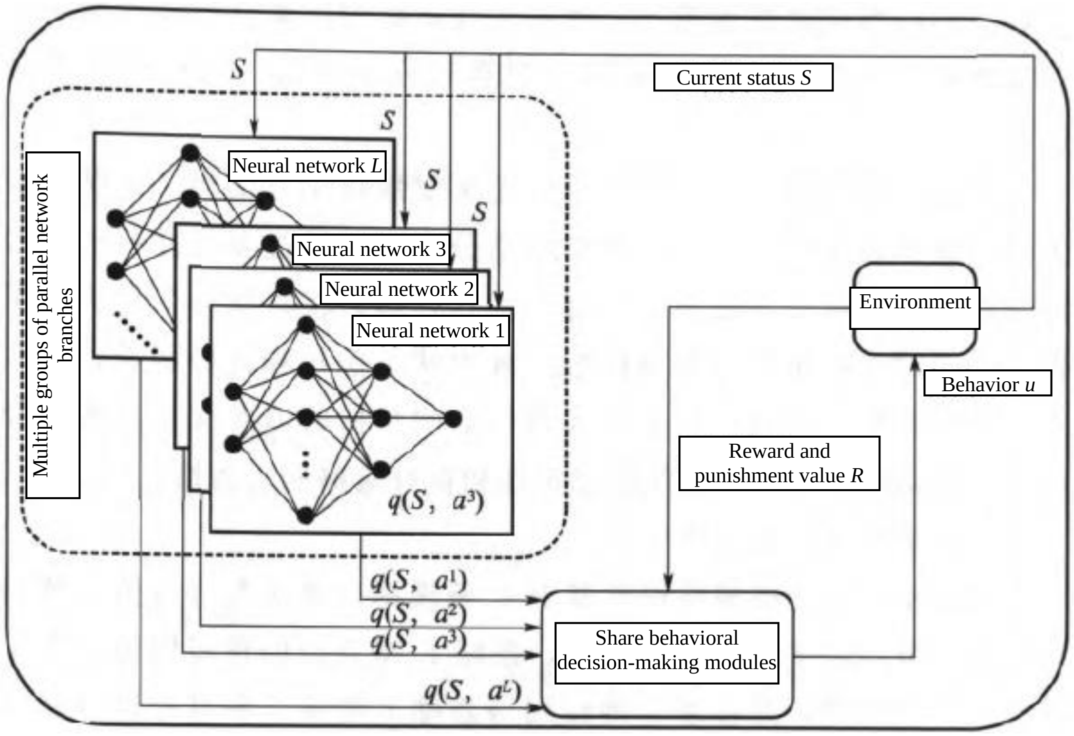


Figure 5-4: Schematic Diagram of Branched Deep Reinforcement Learning Architecture

The main idea is to decompose the q function of multiple sets of parallel neural network calculations into state functions and dominance functions to evaluate the state values and the behavioral advantages of each independent branch respectively, and finally through a special aggregation layer, the state function and the decomposed dominance function are combined to obtain a continuous spatial behavior strategy.

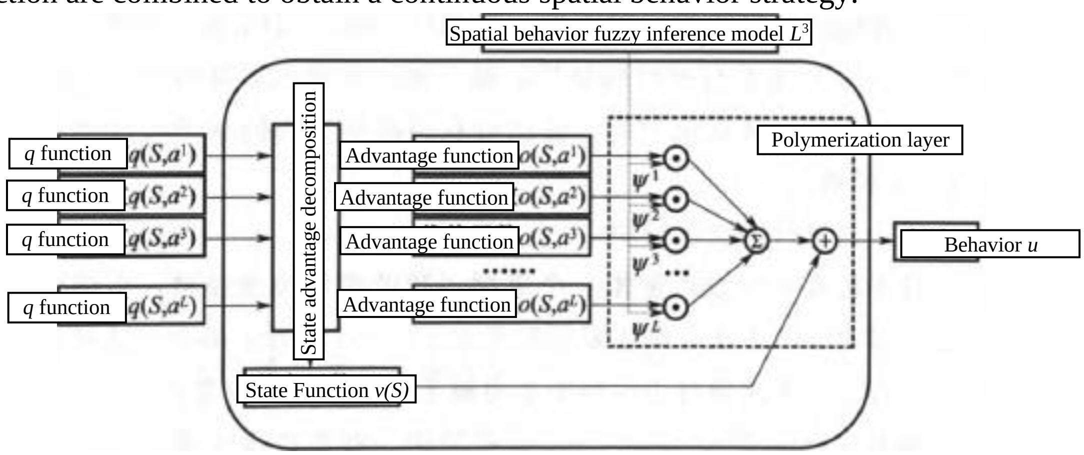


Figure 5-5: Schematic Diagram of Shared Behavior Decision-Making Based on Improved Reinforcement Learning

Here's how:

At the state input end, the fuzzy rules need to be adjusted slightly, and when the spatial behavior fuzzy inference model is mapped to $L(L = y^n)$ IF-THEN fuzzy rules, the c_l in the formula (5-56) is replaced by a^l

R_l: IF x_1 is A_1^l AND x_2 is A_2^l AND⋯AND x_i is A_1^l THEN $u_l = a^l$ (5-62)

where a^l is the behavior corresponding to rule l in the discrete set of behaviors $a = \{a^1, a^2,\ldots, a^L\}$.

In the behavior selection stage, in order to effectively solve the problem of exploration and utilization in reinforcement learning, that is, to continue to use the current optimal strategy to maintain high returns, and dare to try some new behaviors to seek greater rewards, the ε-greedy strategy is adopted for behavioral a^l. The strategy defines a random selection from a discrete set of behaviors with a probability of ε and a selection of the best behavior with a probability of $1–\varepsilon$.

$$a^l = \begin{cases} \boxed{\text{Random selection is performed in } a} & \text{Prob}(\varepsilon) \\ \arg\max\limits_{a^l \in a} q(S, a^l) & \text{Prob}(1-\varepsilon) \end{cases} \quad (5\text{-}63)$$

where S is the position state of the current spacecraft; $q(S, a^l)$ is a function of the corresponding rules/spacecraft behavior $a^l \in a$.

The q function is defined as the expected value G_t after the execution of behavior a from state S under the ε-greedy strategy, and the expectation of the q function under the ε-greedy strategy is called the state function[194]

$$q_t(S, a^l) = E\,[G_t \mid S_t = S, a_t = a, \varepsilon-\text{greedy}] \quad (5\text{-}64)$$

$$v_t(S) = E_{a\sim\varepsilon\text{-greedy}}\,[q_t(S, a^l)] \quad (5\text{-}65)$$

A state function measures the state of a behavior in a particular state, while a q function measures the value of choosing a particular behavior in that state. Based on this, the difference between the q function and the state function is defined as the dominance function

$$o_t(S, a^l) = q_t(S, a^l) – v_t(S) \quad (5\text{-}66)$$

Theoretically, the dominance function is the remainder after subtracting the state value from the q function to obtain a relative measure of the importance of each action, and satisfies $E_{a\sim\varepsilon\text{-greedy}}\,[o_t(S,a^l)] = 0$. However, since the q function is only a value estimate of state-behavior, it is not possible to specify the state value and the dominance value.

Therefore, the expectation value of the dominance function is 0, that is, when the optimal behavior $a^*$is obtained, $q_t(S, a^*) = v_t(S)$, $v_t(S)$ will realize the estimation of the state function, and at the same time, $o_t(S, a^l)$ will also realize the estimation of the dominance function. The q function can then be decomposed into a state function $v_t(S)$ and a dominance function $o_t(S, a^l)$.

$$q_t(S, a^l) = v_t(S) + \left(o_t(S, a^l) - \max_{a^l \in a} o_t(S, a^l)\right) \quad (5\text{-}67)$$

At the behavior output end, the state functions that are irrelevant to the behavior selection can be separated, and only the optimal operation of each dominant function is required and then combined with the formula (5-59) through the fully connected layer output. This processing not only alleviates the computation of the q function, but also effectively avoids the problem of the combination of the number of behaviors and the mapping rules.

$$u^*(\overline{x}_t) = v_t(S) + \sum_{l=1}^{L}\left(\Psi_t^l \cdot o_t(S, a^l) - \Psi_t^l \cdot \max_{a^l \in a} o_t(S, a^l)\right) \quad (5\text{-}68)$$

where $u^*(\bar{x}_t)$ is the global behavior with the optimal q value in the L rule.

In the self-learning stage, the time difference (TD) error function is defined to realize feedback self-learning under the traction of reward and punishment values

$$p_t = R_{t+1} + \gamma u^*(\overline{x}_{t+1}) - u(\overline{x}_t) \quad (5\text{-}69)$$

Where $\gamma \in [0, 1]$ is the discount factor; R_{t+1} is the reward and punishment value that can be obtained at t+1 time, and defines $R_{t+1} = 2e^{-u^2} - 1$.

In the q function update phase, it is updated through independent iterative training

$$q_{t+1}(l, a^l) = q_t(l, a^l) + \eta p_t \Psi_t^l \quad (5\text{-}70)$$

where η is the reinforcement learning rate.

5.2.5: Game Interaction Between Spacecraft and Space Targets

For the orbit game problem between spacecraft and space target described by differential strategy, the continuous behavior of space is represented by a fuzzy inference model, and then the orbit game method based on branched deep reinforcement learning is used to obtain the continuous behavior output. Taking the perspective of spacecraft as an example, the dynamic game interaction process between the two sides is illustrated, as shown in Figure 5-6.

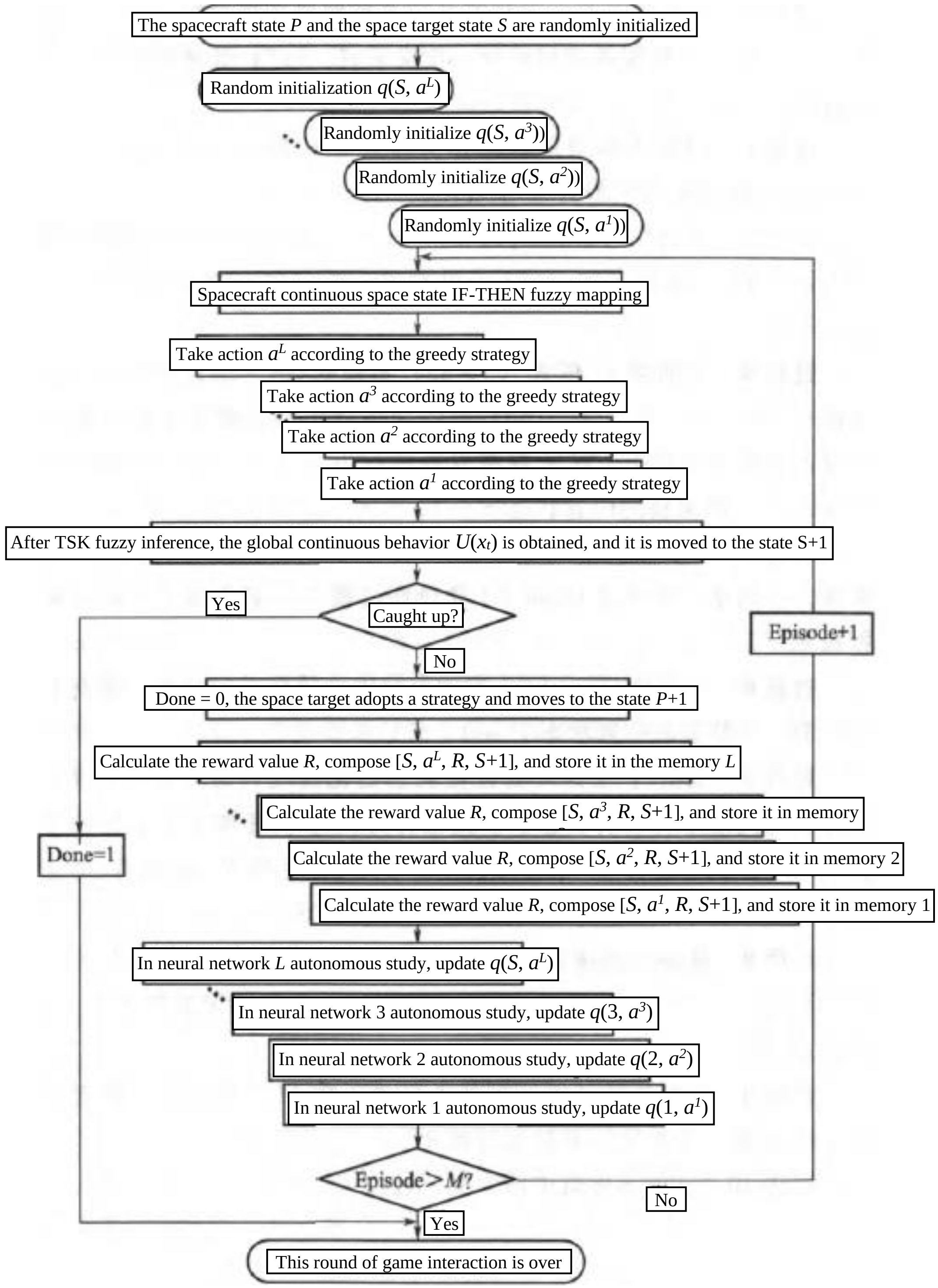


Figure 5-6: The Game Interaction Process Between a Spacecraft and a Space Target

Process 1 According to the current state of the spacecraft, the input quantity n of the fuzzy inference model is defined, and the membership function y is set. According to the number of fuzzy rules, $L(L = y^n)$ is defined neural networks, and the q function of each network is randomly initialized.

Process 2 Take the current state of the spacecraft as input $\bar{x} = (x1, x2, \ldots, x_n)$, pass through IF-THEN fuzzy mapping is an L rule.

Process 3 Respectively in the same as the first $l = \{1, 2, \ldots, L\}$ rules correspond to the neural network, and the q function is calculated $q(S, a^l)$,According to Eq. (5-63), the discrete behavior a^l ($l = 1, 2, \ldots, L$) is selected.

Process 4 The c_l term in Eq. (5-56) is replaced by a^l Eq. (5-62) respectively, and the behavior $u(\bar{x}_t)$ that the spacecraft will take in the current state is obtained by the fuzzy inference model from Eq. (5-57) to Eq. (5-60) and the extraction of the dominance function. The spacecraft takes the behavior $u(\bar{x}_t)$ to move to the new position state S+1.

Process 5 Calculate the Euclidean distance between the spacecraft and the space target to determine whether the interception conditions are met. If satisfied, let the variable Done = 1 and go to process 10; if not, go to process 6.

Process 6 Let the variable Done = 0, and the spatial target will take the most beneficial behavior according to the strategy and move to the new position state P+1.

Process 7 According to the behavior, as well as the change of position state, the reward and punishment value R is calculated. In each branch network, the current state S, the discrete behavior a^l, the reward and punishment value R, and the next state $S+1$ are combined $[S, a^l, R, S+1]$ matrix form and stored in memory.[194, 44]

Process 8 Autonomous reinforcement learning is carried out in the shared behavior decision module, according to Eq. (5-66) to Eq. (5-70), the error function is used as the traction p_t to adopt a certain learning rate η and update the q function.

Process 9 Determine whether the number of steps has reached the maximum number of action steps M. If reached, go to process 10; otherwise, the number of steps adds 1 and goes to process 2.

Process 10 End the interactive process of this round of orbital game.

5.3: A Case Study of Real-Time Planning of Spacecraft Orbit Game Strategy

In order to verify the applicability of the spacecraft differential countermeasure model constructed in this chapter, as well as the effectiveness and superiority of the strategic planning method based on branching deep reinforcement learning, a case simulation and analysis were carried out.

5.3.1: Problem Description

The scenario of the case is constructed with reference to relevant literature,[136, 177] and the spacecraft and the space target (non-cooperative target) are both in the vicinity of low earth orbit, the countermeasure time is short, and the instantaneous state information is fully known. The mass of the spacecraft is 2,500 kg, the maximum acceleration is 0.05 g, g = 9.8×10^{-3} km/s^2. The space target mass is 2,500 kg, the maximum acceleration is 0.01g. In the orbital game, the spacecraft's chasing behavior $u_P = [\theta_P, \delta_P]$ and the space target behavior $u_E = [\theta_E, \delta_E]$ are both compact sets in the measurement space, and their respective objective functions $J(u_P, u_E)$ are continuous on $u_P \times u_E$. Both sides adopt the most beneficial behavior strategy based on the Nash equilibrium, that is, only when both parties adopt their own strategies in this strategy portfolio, neither side can obtain greater benefits by taking other strategies independently. Considering that the fuel consumption is small relative to the mass of the spacecraft and the space target, it is assumed that the mass of the spacecraft and the space target remains unchanged during the whole maneuvering process. Let the gravitational constant of the earth be $\mu = 3.986 \times 10^5$ km^3/s^2 and the maximum number of action steps in the game process M = 3,600, and the relative initial state parameters of their orbital game are shown in Table 5-1.

Table 5-1: Initial values of spacecraft and space targets in relative coordinates

Initial value	x/km	y/km	z/km	$\dot{x}$/(km/s)	$\dot{y}$/(km/s)	$\dot{z}$/(km/s)
Spacecraft	0	0	0	—0.049 6	0.045 8	0
Target	-50	30	0	—0.021 8	0.021 8	0

The difference in space angle φ between spacecraft P and space target E consists of the pitch angle difference Δ and the thrust angle difference Δδ in the orbital plane, i.e., $\varphi = [\Delta\delta, \Delta\theta]$

$$\begin{cases}\Delta\delta = \delta_E - \delta_P \\ \Delta\theta = \arctan\left(\frac{y_E - y_P}{x_E - x_P}\right) - \theta_P\end{cases} \tag{5-71}$$

The rate of change of the angular difference $\dot{\varphi}$ is as follows

$$\dot{\varphi} = \frac{\varphi - \varphi'}{T} \tag{5-72}$$

where φ' is the angular difference of the previous state; T is the sampling time.

When a spacecraft approaches a space target, the space target behaves non-cooperatively. In order to better reflect the whole game interaction process, the angle difference is φ and its rate of change $\dot{\varphi}$ as the state quantities $S = (\varphi,\dot{\varphi})$ and $P(\varphi,\dot{\varphi})$ of spacecraft P and space target E. In order to avoid the dimensionality disaster, the input $n = 2$ and the membership function $y = 3$ are set to construct the fuzzy inference model. The ambiguity set in the angular difference and its rate of change φ is represented by {negative (N), zero (Z), positive (P)}.

5.3.2: Model Operation Analysis

In order to better illustrate the advantages of the method in self-training and self-optimization, the orbital game process after 0, 500, and 1,000 times of self-learning is illustrated as an example.

(1) Spacecraft orbit game after 0 learning

Figure 5-7 shows the orbit game between the spacecraft and the space target after 0 times of self-learning, that is, the trajectory change of the method is directly applied to the orbit game problem without learning. Among them, although the spacecraft is driven by the objective function, because its q function is randomly generated and there is no prior knowledge, the behavior is uncertain and floats back and forth, and the space target is not affected and continues to travel in the original orbital direction. Eventually, the spacecraft is getting farther and farther away from the space target and cannot complete the mission.

(2) Spacecraft orbit game after 500 learning sessions

Figure 5-8 shows the orbital game between the spacecraft and the space target after 500 times of self-learning. After 500 times of self-learning, the spacecraft can approach the direction of the space target, and the space target adopts non-cooperative behavior to change the established orbit on the way, and the two sides continue to play games, and the spacecraft catches up with the space target after 358 s.

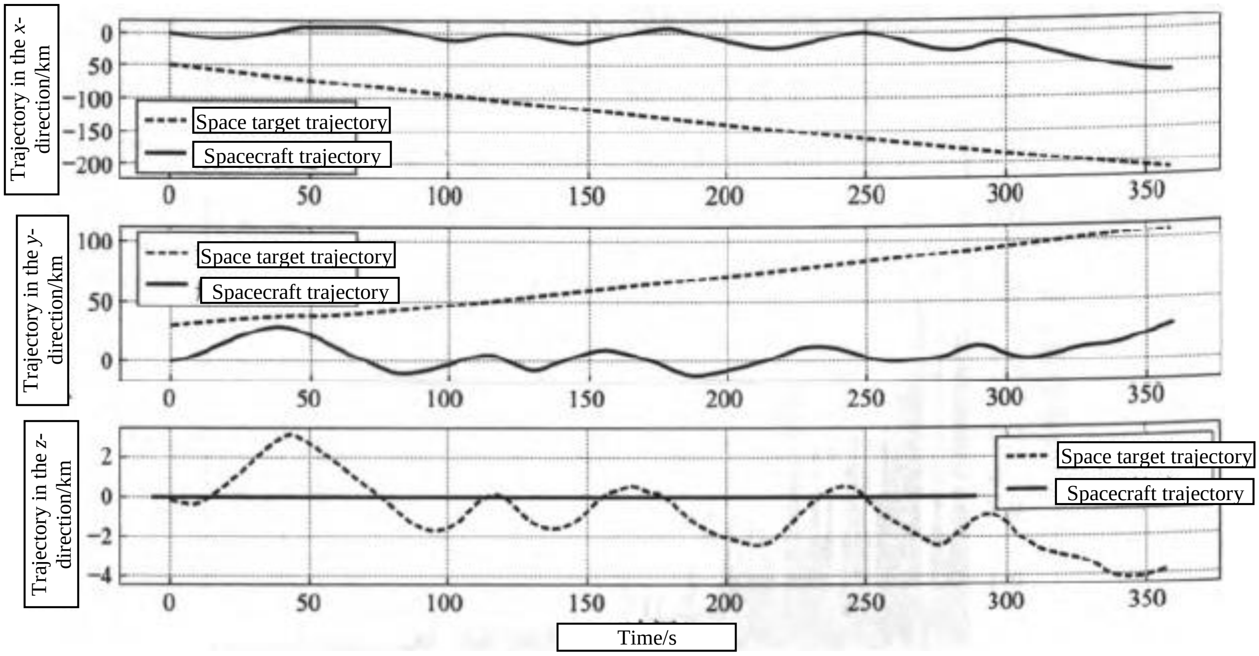


Figure 5-7: Orbital Game Trajectory After Learning 0 Times

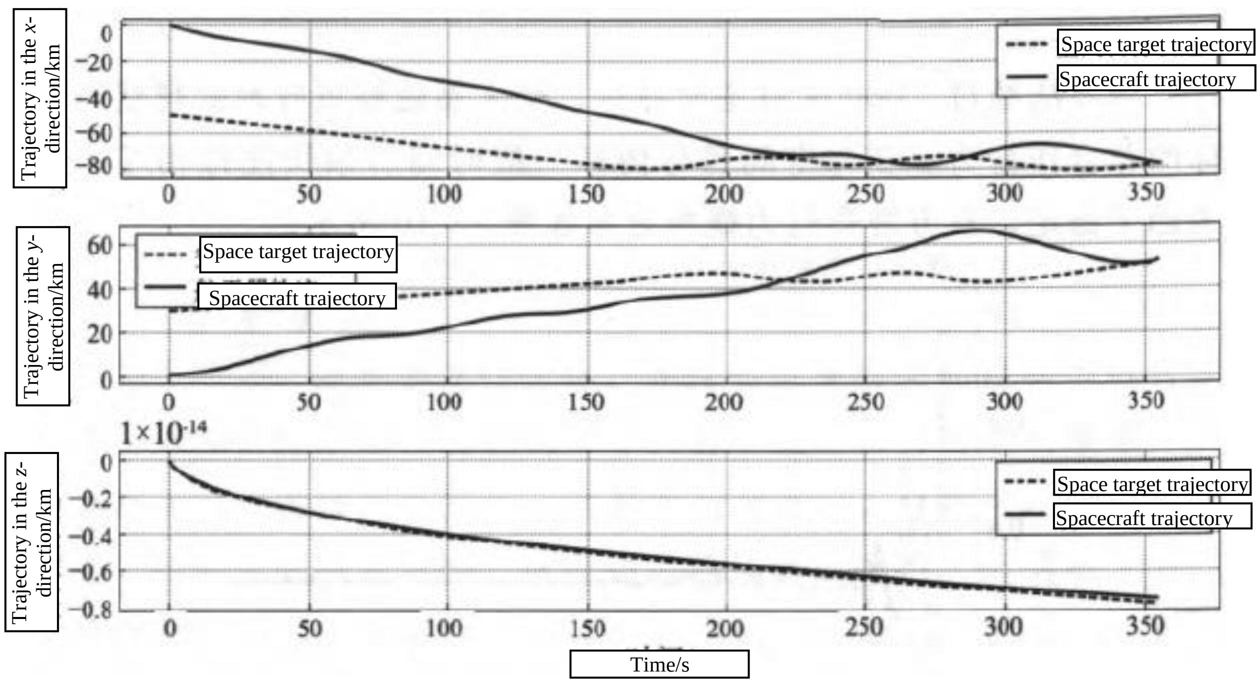


Figure 5-8: Orbital Game Trajectory after 500 Studies

(3) Figure 5-9 of the spacecraft orbit game after 1,000 learning times shows the variation of the q function error with the number of training times in self-learning, and with the increasing number of training times, the q function error becomes lower and lower, and converges to the optimal behavior strategy quickly, so as to achieve the Nash equilibrium of the orbital game.

However, due to the adoption of greedy strategy, there is still a slight fluctuation in the later error.

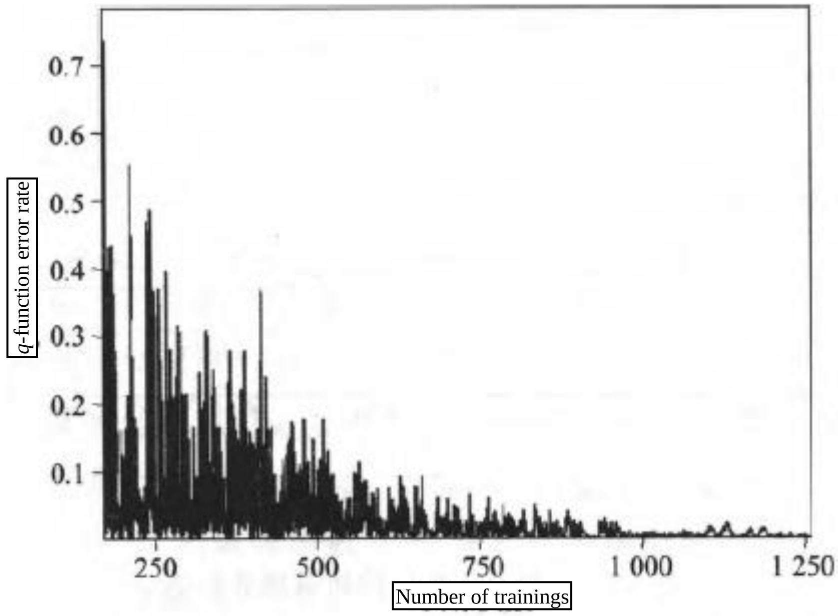


Figure 5-9: q-Function Error as a Function of the Number of Trainings

After 1,000 self-learning sessions, the spacecraft can better deal with the non-cooperative behavior of the space target, and the mutual behavior tends to stabilize quickly after a period of time of game with the space target, as shown in Figure 5-10.

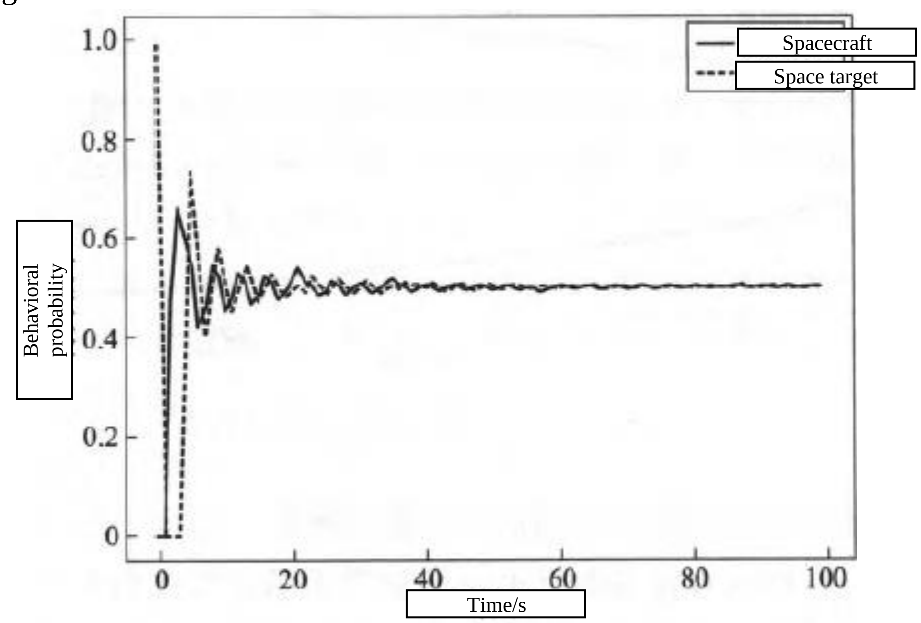


Figure 5-10: Probability Distribution of Game Behavior

Therefore, driven by the equilibrium strategy, the spacecraft can select the best trajectory and catch up with the space target in the shortest time of 336 s, as shown in Figure 5-11, and the orbital game trajectory is shown in Figure 5-12. It can be seen from the diagram that the trajectories of the two sides in the z direction do not change significantly, which is consistent with the conclusion that "the best pursuit strategy of spacecraft and space target in the game process should occur in coplanar orbit" obtained in relevant literature.[177, 13]

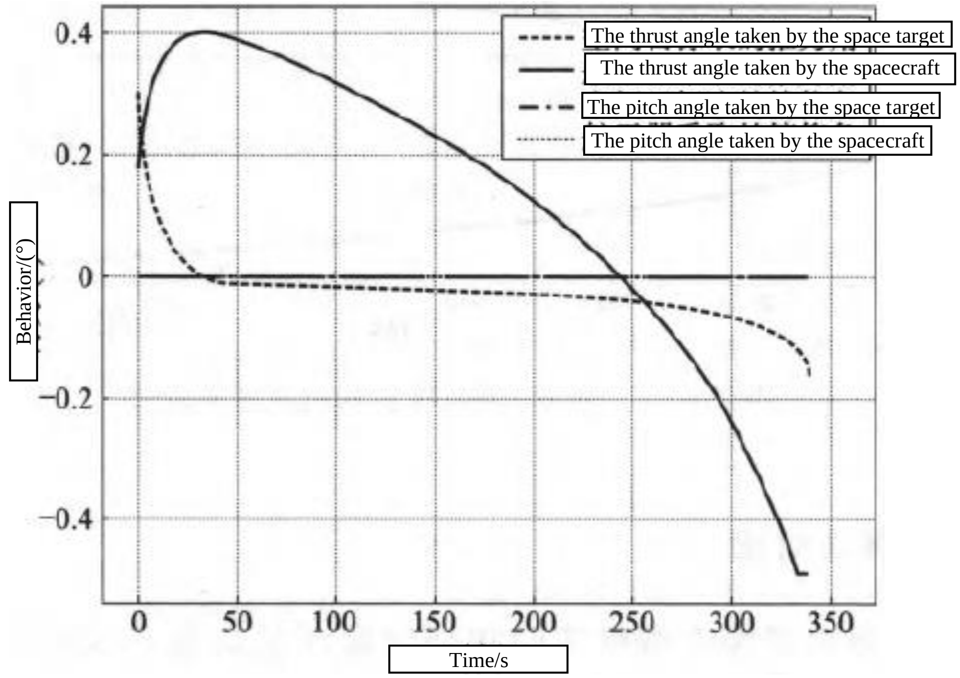


Figure 5-11: The Amount of Behavior Control after 1,000 Sessions

To sum up, in the face of non-cooperative targets, under the condition of only knowing its own state and the current limited state of the target, and not knowing the future behavior of the target, if an effective strategy is not obtained in time, the spacecraft will not be able to correctly predict the direction of the target's behavior at the next moment, and can only passively chase behind, and will never be able to reach the target, thus missing the best rendezvous time. Using the game strategy solving algorithm based on branch deep reinforcement learning, the behaviors that the spacecraft can take at each moment are expressed on multiple network branches, and the different strategies are independently trained and processed by multiple groups of parallel neural networks, and then different strategies are centrally coordinated through the shared decision-making module. In the case of unknown direction of the next behavior of the space target, the spacecraft can obtain the optimal strategy under Nash equilibrium, so that it can rendezvous with the target within 336 s.

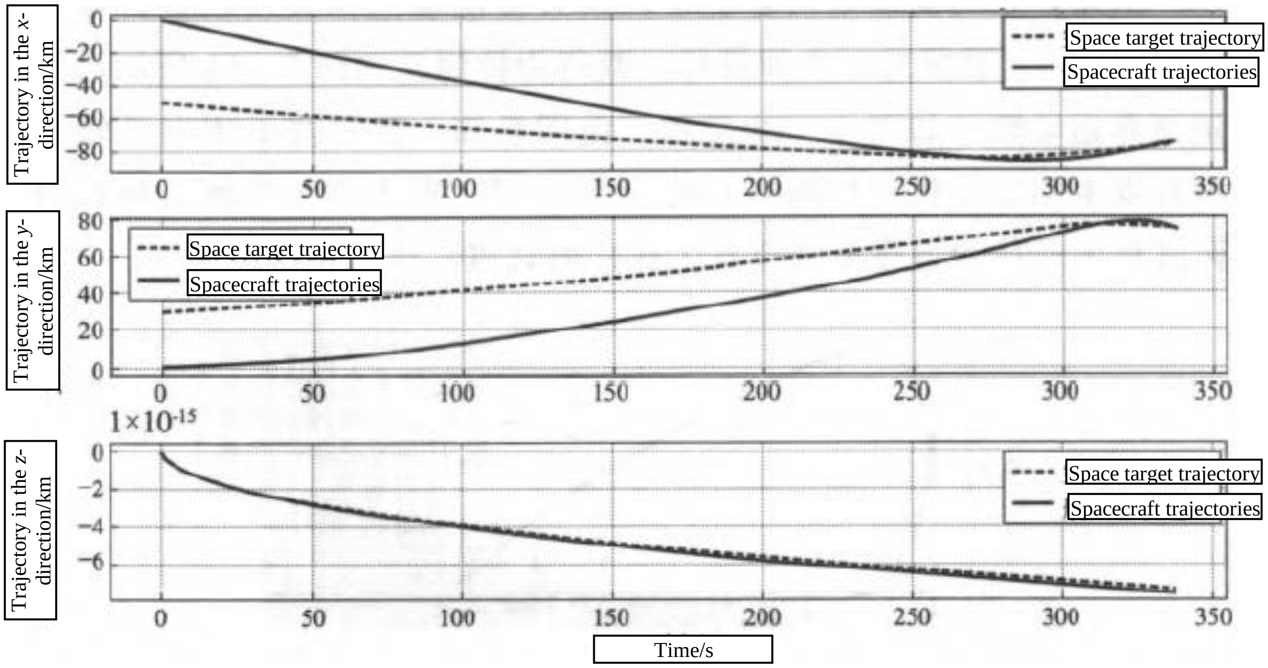


Figure 5-12: Orbital Game Trajectory After 1,000 Sessions

5.3.3: Algorithm Comparison

The real-time planning problem of spacecraft orbit game strategy involves high-dimensional nonlinear differential models, which can be seen from the differential countermeasure theory, and it is difficult to directly solve such high-dimensional nonlinear differential countermeasure models due to the constraints of the differential model order and nonlinear terms, so it is necessary to find a strategy solution algorithm for the purpose of interaction, which is helpful to provide an effective theoretical and decision-making basis for spacecraft orbit game.

(1) Comparison of algorithm applicability

In the orbital game between spacecraft, both sides will choose the behavior strategy that is most beneficial to them. For the general chase optimal control problem, the proportional guidance method can usually be used. The proportional guidance method aims at chasing and can give the optimal control strategy of the chasing party, which has the advantages of simple form and easy implementation. However, in the face of such a continuous space dynamic interaction problem with changeable maneuver laws, non-communication at the information level, and incomplete prior knowledge, it is necessary to consider the strategic choices of both parties in the orbital game at the same time, and the proportional guidance method is difficult to apply because it cannot cope with the behaviors of both parties at the same time.[127] The target method, the matching point method and the intelligent optimization algorithm are more suitable for the numerical solution of the dynamic interaction problem in continuous space and have better application effect in the solution of spacecraft orbit game strategy.

Table 5-2 compares and analyzes the applicability of these three types of solving algorithms.

Table 5-2: Comparison of the applicability of different strategy solving algorithms

The name of the algorithm	Characteristics	Advantage	Applicability analysis
Targeting	1) The boundary value solving algorithm with the improved Newton iterative method as the core 2) Taking the necessity of the existence of the saddle point as the starting point, the boundary value condition is satisfied through the continuous correction of the initial condition	1) The calculation speed is fast and the calculation results are accurate 2) The saddle point of the differential countermeasure can be directly obtained while obtaining the high-precision numerical solution	It has high requirements for the initial value and a small convergence domain for the high-dimensional boundary value problem, which makes it difficult to deal with the boundary value problem of nonlinear or control quantity coupling
Matching method	1) Transform the differential countermeasure problem into an optimal control problem 2) By constantly correcting the initial values, the differential equation satisfies the boundary value condition	1) Good convergence 2) It is conducive to mathematical analysis and convenient for mathematical calculation	The solution time and computational accuracy depend on the initial value, and the core idea has theoretical defects, and it is difficult to explain that the convergence strategy is the saddle point solution of differential countermeasures
Game strategy solving algorithm based on branched deep reinforcement learning	1) Multiple groups of parallel neural networks, which are conducive to independent training and fast processing 2) Shared decision-making module, which helps to centralize the coordination of different strategies	1) Self-training and independent decision-making 2) Able to interact with the environment continuously, suitable for dynamic decision-making problems	It is not limited by the initial value, and is suitable for the boundary value problem of control quantity coupling, and can converge quickly on the premise of ensuring the existence and uniqueness of the equilibrium strategy

The target method is a numerical solution method that takes the necessity of the saddle point of the differential countermeasure problem as the starting point and continuously corrects the initial value to make the differential equation meet the boundary value condition. The target method can not only obtain a high-precision numerical solution, but also directly obtain the optimal trajectory and the corresponding saddle point. However, the target method is based on the Newton iterative method, which depends on the selection of initial values, and the convergence domain of the equilibrium strategy is small, and it may be difficult to obtain the convergent chase strategy.[127]

The coordination method is a method that transforms the differential countermeasure problem into an optimal control problem and solves it numerically by optimizing methods such as quadratic programming or gradient recovery. Although the coordination method is better than the target method in terms of convergence, its solution time and calculation accuracy are heavily dependent on the initial value, and it also faces the problems of dimensional explosion and small convergence domain for the spacecraft orbit game problem. In addition, the core idea of the coordination method to transform the differential countermeasure problem into the optimal control problem only gives a feasibility analysis from a numerical point perspective, which has theoretical defects, and it is difficult to explain that the convergence strategy is the saddle point solution of the differential countermeasures of both sides of the orbital game.[127, 177]

The game strategy solving algorithm based on branching deep reinforcement learning is an intelligent algorithm with multiple sets of parallel neural networks and shared decision-making modules. The algorithm avoids the initial value guessing problem of the indirect method, distributes the representation of orbital game behavior on multiple network branches, realizes the independent training and rapid processing of the optimal strategy through multiple groups of parallel neural networks, and then realizes the centralized coordination of different strategies through the shared decision-making module, and quickly converges under the premise of ensuring the existence and uniqueness of the equilibrium strategy, so as to effectively deal with the continuous dynamic interaction problem of such bilateral control.

(2) Comparison of algorithm computation

Both the target method and the matching method are indirect methods, which are usually combined with heuristic intelligence algorithms such as genetic algorithms in order to avoid the huge and complex amount of computation in the actual application process. However, the game strategy solving algorithm based on branched deep reinforcement learning belongs to the category of intelligent algorithms, and its model structure and operation process are quite different from those of the target method and the matching method, and the direct comparison of the computation amount of the three cannot explain the advantages and disadvantages of each method. In order to effectively analyze the advantages and disadvantages of the game strategy solving algorithm based on branching deep reinforcement learning in terms of computational efficiency, the simulation was compared with the Q-learning algorithm,[232] the reinforcement learning algorithm based on qualification traces,[104] and the reward-based genetic algorithm.[41]

The simulation of the case is carried out on 1.6 GHz, 1.8 GHz dual-core CPU, 8 GB RAM computing hardware, using the PyCharm simulation compilation environment. In the branched deep reinforcement learning architecture, the number of neural network layers is 3, the number of neurons in the hidden layer is 10, the activation function is sigmoid, and the parameter settings are shown in Table 5-3.

Table 5-3: Parameter settings

Parameter	Assignment
Memory bank E	6,000
Training step M	3,600
The exploration rate is ε	0.3

Continued

Parameter	Assignment
The discount factor γ	0.99
The learning rate η	0.3
Sample time T	1 s

The game strategy solving algorithm based on branching deep reinforcement learning was used to simulate, and the simulation results were compared with the Q-learning algorithm, the reinforcement learning algorithm based on qualification traces and the genetic algorithm based on reward. Each algorithm undergoes 100 times of self-learning, and the capture time and training time are shown in Table 5-4. Among them, the Q-learning algorithm takes a long time to learn due to the need to store multiple feature vectors in the chain and iteratively update multiple Q tables at the same time. The reinforcement learning algorithm based on qualification traces unifies the temporal difference and Monte Carlo, and only needs to orbit one trace vector, and no longer needs to store multiple feature vectors, which greatly reduces the self-learning time, but its short-term memory characteristics prolong the actual pursuit time. Although the reward-based genetic algorithm has high practical application performance, it comes at the cost of longer autonomous training time. While giving full play to the advantages of reinforcement learning algorithms, this method uses multiple sets of neural networks for parallel training, which greatly reduces the time spent on self-learning and ensures that the orbiting task can be completed in a short time.

Table 5-4: Time and learning time for different methods to complete the tracking task

The name of the algorithm	Initial relative position of the space target/km				Study duration/s
	(-6, 7)	(-7,-7)	(2, 4)	(5, 5)	
Book algorithms	9.8	10.1	4.3	8.1	21.6
Q-learning algorithms	10.9	12.9	4.7	8.8	258.6
Reinforcement learning algorithm based on qualification traces	12.6	15.6	8.5	11.3	32.0
Reward-based genetic algorithms	9.7	10.5	4.5	8.4	460.8

Through simulation comparison, the method in this book has a comparative advantage in the application of continuous space behavior decision-making. The ε-greedy strategy was also adopted, and the book method and the traditional deep reinforcement learning method were used to learn 1,000 times independently.

The TensorBoard[1] module of TensorFlow was used to detect the learning process, and the reward and punishment values were sampled every 3 times. Figure 5-13 shows the learning curve generated by TensorBoard, that is, the cumulative change of reward and punishment values with the number of learning sessions. From the curve distribution, it can be seen that compared with the traditional deep reinforcement learning method, the reward and punishment value of the method in this book increases more obviously and more steadily.

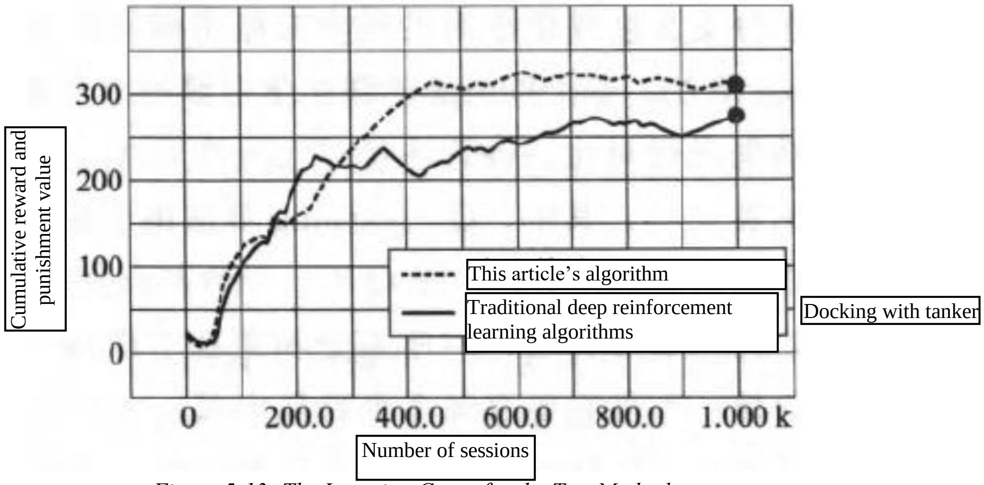


Figure 5-13: The Learning Curve for the Two Methods

5.4: CHAPTER SUMMARY

In order to further solve the real-time planning problem of spacecraft orbit game strategy and alleviate the application limitation of deep reinforcement learning in continuous space, this chapter introduces a game strategy solving algorithm based on branched deep reinforcement learning. Firstly, in the face of the "continuous dynamic interaction" characteristic of orbit game between spacecraft, a sequential decision game model is constructed, and the real-time planning problem of spacecraft orbit game strategy is described. Secondly, the motion model between low-earth orbit spacecraft is constructed, and the Nash equilibrium strategy of orbital game is given, which translates the orbital game problem between orbital spacecraft into a differential countermeasure problem. Thirdly, a fuzzy inference model for continuous space solution is constructed to realize the mapping transformation of continuous state through fuzzy inference to continuous behavior output, which effectively avoids the problem of dimensionality disaster in continuous space in which traditional deep reinforcement learning copes with continuous space.

[1] TensorBoard is a powerful visualization tool that comes with TensorFlow and a web application suite.

Finally, a new branch deep reinforcement learning architecture is introduced, which realizes the branch training and sharing decision-making of behavior strategies and avoids the problem of the combination growth of the number of behaviors and mapping rules. The case analysis shows that the method realizes the combination of optimal control and game theory, improves the learning ability of deep reinforcement learning for discrete behaviors, and further solves the problem that the differential countermeasure model is highly nonlinear and difficult to solve by classical optimal control theory, and can obtain the Nash equilibrium strategy of spacecraft orbit game in real time. The comparison of the algorithms shows that compared with other commonly used algorithms, the algorithm in this book reduces the training time by 90% on average, and the growth rate of reward and punishment value is 17% faster and more stable, which further improves the shortcomings of conventional methods that are difficult to deal with the bilateral control problem and the convergence domain of the equilibrium strategy is small.

This book is oriented to the practical application of intelligent planning of on-orbit servicing tasks, analyzes the foundation, requirements and key points of system construction, and clarifies the application requirements of on-orbit services. Carry out the design of the task planning system to realize the systematization, modularization, easy operation and easy expansion characteristics of the system; the case application of the task planning system was carried out to effectively test the feasibility and effectiveness of the intelligent planning method in this book.

6.1: ON-ORBIT SERVICING MISSION PLANNING SYSTEM REQUIREMENTS

The on-orbit servicing mission planning system can be applied to both ground simulation training and on-orbit servicing applications. The on-orbit servicing mission planning system can not only test the mission planning method, but more importantly, it is of great practical significance for popularizing technological achievements, improving the on-orbit servicing mission planning system, and demonstrating and guiding the development and construction of the independent planning system for spacecraft on-orbit servicing missions.

6.1.1: The Foundation of the Mission Planning System

The on-orbit servicing mission planning system serves the aerospace field and serves the diversified missions and tasks in space. Only by understanding the characteristics of mission planning and the background of system construction can it be possible to develop practical software that meets the characteristics of future on-orbit services and meets the needs of aerospace applications. Only on the basis of fully grasping the essential characteristics and development needs and aiming at the key issues of on-orbit servicing mission planning, can we find the magic weapon to overcome the difficult problems.

(1) Planning features

Due to the special environment, implementation method and implementation mode, the on-orbit servicing task planning has unique characteristics and rules: first, it is to win by pre-winning, that is, to win by strategy, focusing on pre-design and pre-planning; the second is to fight slow with fast, that is, to win at speed, and strive to implement and plan automatically; the third is to control the dark with light, that is, to win with information, consider multiple constraints, and optimize with multiple constraints;

the fourth is to refine and coarse, that is, to win with precision, and seek precise implementation.

(2) Construction background

Based on the characteristics of mission planning, combined with the development of on-orbit services, the construction background of the on-orbit servicing mission planning system is: first, the needs of spacecraft on-orbit target allocation, such as target allocation and force allocation; the second is the need for orbital maneuver path planning, such as orbit design, trajectory optimization and orbit avoidance; the third is to serve the needs of action planning, such as orbital games, control implementation and strategic planning.

(3) Essential characteristics

The on-orbit servicing mission planning is the result of the "dual drive" of aerospace application and intelligent technology development, and its essential characteristics can be summarized as "four characteristics and four modernizations.” "Four characteristics": first, forward-looking, that is, planning in advance, reflecting the precautions; the second is integration, that is, multi-objective planning, which reflects thoughtfulness: the third is autonomy, that is, intelligent planning, which embodies all-intelligence and all-round; fourth, precision, that is, precise planning, embodied without losing the slightest. "Four modernizations": first, standardization, mainly manifested in data information and model formulas; the second is process, which is mainly manifested in the proceduralization of command, task and execution; the third is automation, which is mainly manifested in the automatic implementation of task planning; fourth, intelligence, which is mainly reflected in the autonomous decision-making of interactive games.

(4) Development needs

In view of the characteristics of aerospace technology and the continuous expansion of on-orbit servicing styles, the development needs of on-orbit servicing mission planning are mainly reflected as follows: first, the target allocation is expanded from a "one-to-many" or "many-to-one" single mode to a composite service model; second, the task planning has expanded from pre-planning to temporary planning and real-time planning; third, the evaluation and optimization has expanded from single-objective elements to multi-objective all-elements; fourth, the strategy generation has expanded from manual assistance to intelligent planning.

6.1.2: The Need for the Construction of the Task Planning System

The on-orbit servicing task is an inexhaustible driving force to promote the construction of the mission planning system. It is difficult to maximize the effectiveness of the mission planning system if it is isolated and rigid or built behind closed doors without the actual needs of on-orbit servicing tasks. Only by bundling the mission planning technology with the actual needs, keeping up with the development, technology, and application needs, aiming at the timeliness and autonomy of the future on-orbit servicing, closely integrating system thinking, technical thinking, and innovative thinking, and developing the system according to the essential needs of the on-orbit servicing mission, can the planning technology be adapted to the on-orbit servicing mission.

(1) The strategic needs of overall planning

On-orbit servicing is a national-level behavior with a strong global dimension. The on-orbit servicing mission planning needs to look at the international space form and national development needs, grasp the mission background, service objectives and index requirements, and have high requirements for mission planning.

(2) The precise needs of local planning

The on-orbit servicing mission requires an accurate grasp of the space situation, a fine analysis of the space environment, a precise planning of valuable space resources, and an accurate implementation of orbit transfer.

(3) Timeliness requirements for on-orbit dynamic programming

The timeliness of on-orbit servicing tasks is strong, the transfer speed is fast, the maneuverability is strong, and the timeliness from information acquisition to service implementation is limited, and the service time window is high, which has obvious requirements for the timeliness of on-orbit dynamic planning.

(4) Autonomous needs for online real-time planning

On-orbit servicing mission planning is a complex process from ground to on-orbit, from scheduled to determined, from offline to online, and from pre-engagement to real-time. The whole task planning process includes accepting the task formulation plan, improving and refining the task allocation, adjusting and revising the strategic plan, and the demand for automatic improvement, iterative refinement, and independent generation of the task plan is becoming increasingly obvious.

6.1.3: The Key to the Construction of the Task Planning System

The construction of the on-orbit servicing mission planning system, as the application of mission planning methods and technologies, needs to further change the ideological concept, further improve the mechanism and system, and further break through the key technologies. At present, it is necessary to grasp the focus of system construction, continuously promote technological innovation under the conditions of informatization and intelligence, and promote the solid progress of aerospace technology application.

(1) Further change of ideology

Ideology is the forerunner for the construction and application of the task planning system. The first is from manual planning to automatic planning, reflecting the transformation from subjective to objective; compared with manual planning, computer automatic planning emphasizes more on the basis of constraints, and pays more attention to the use of scientific methods and means to obtain the optimal realization method, which is more objective. The second is from single-objective planning to multi-objective planning, reflecting the transformation from local to global;

compared with single-objective planning, multi-objective planning considers more on-orbit resources, more target situations, and more constraints, and pays more attention to task planning covering all elements and processes, which is more realistic. the third is from static planning to dynamic planning, reflecting the transformation from static to dynamic; compared with static planning, dynamic planning emphasizes the decisive role of time in decision-making, pays more attention to the sequence of spacecraft, and is more dynamic.

(2) Further improve the system and mechanism

A sound system and mechanism is the basic guarantee for the effective application of the mission planning system. The first is the control system of accusation separation; in order to adapt to the characteristics of the application of aerospace technology under the conditions of informatization, establish a management system for separating accusations, optimize the grouping of joint accusations, and reasonably distinguish the functions and authority of accusations, so as to lay the foundation for scientific planning. the second is the administrative system of the main use and the main construction; according to the particularity of the on-orbit servicing and the uniqueness of the aerospace field, it is necessary to establish and improve the system for the construction, management and use of the mission planning system that is driven by the needs of the charging department, led by the practical operation unit, and developed and managed by the industrial department, so as to form a closed-loop loop from demand to use and then to development, so as to ensure that the developed mission planning system can be used, easy to use, and effective. The third is the technical system of the upper and lower cohesion; top-down unification and connection of the system environment, software platform, interface standards, data transmission, and model methods of the task planning system at all levels to ensure that the systems at all levels are easy to connect, communicate, use, and collaborate.

(3) Further breakthroughs in key technologies

Advanced key technologies are an important support for the long-term development of the mission planning system. The first is the target allocation technology for the on-orbit servicing composite service model; this paper focuses on the target allocation technology that can effectively solve such nonlinear combinatorial optimization problems, so as to effectively distinguish and treat each service objective, take into account the execution benefit and energy consumption efficiency, and quickly find the optimal solution of the problem, so that the target allocation can help to win with ingenuity and improve the efficiency of spacecraft decision-making in orbit. the second is the orbit avoidance path planning technology to deal with the sudden approach of space debris; this paper focuses on the path planning technology that can effectively solve the problem of multi-limited shortest path, so as to effectively take into account target avoidance and orbit keeping, meet the needs and preferences of different tasks, and make orbit avoidance boost agility and win, and improve the active defense capability of spacecraft in orbit. the third is the orbital game strategy planning technology in the face of non-cooperative goals; this paper focuses on the strategic planning techniques that can effectively solve such dynamic game problems, so as to effectively solve the problem that the differential countermeasure model is highly nonlinear and difficult to solve by classical optimal control theory, so that the dynamic game can help intelligent control win and improve the autonomous service level of spacecraft in orbit.

6.2: Design of On-Orbit Servicing Mission Planning System

Based on the extensive application of mission planning in the field of aerospace system, an on-orbit servicing mission planning system is designed. At the level of system architecture, design application integration and network integration; at the level of system function architecture, build the support layer, functional layer and application layer; in terms of system control page, design system login, system function, data entry and effect display pages.

6.2.1: System Framework

On-orbit servicing mission planning system integration includes application integration and network integration. Application integration includes platform integration and environment integration, and network integration includes logical and physical networks. Figure 6-1 shows the framework of system integration.

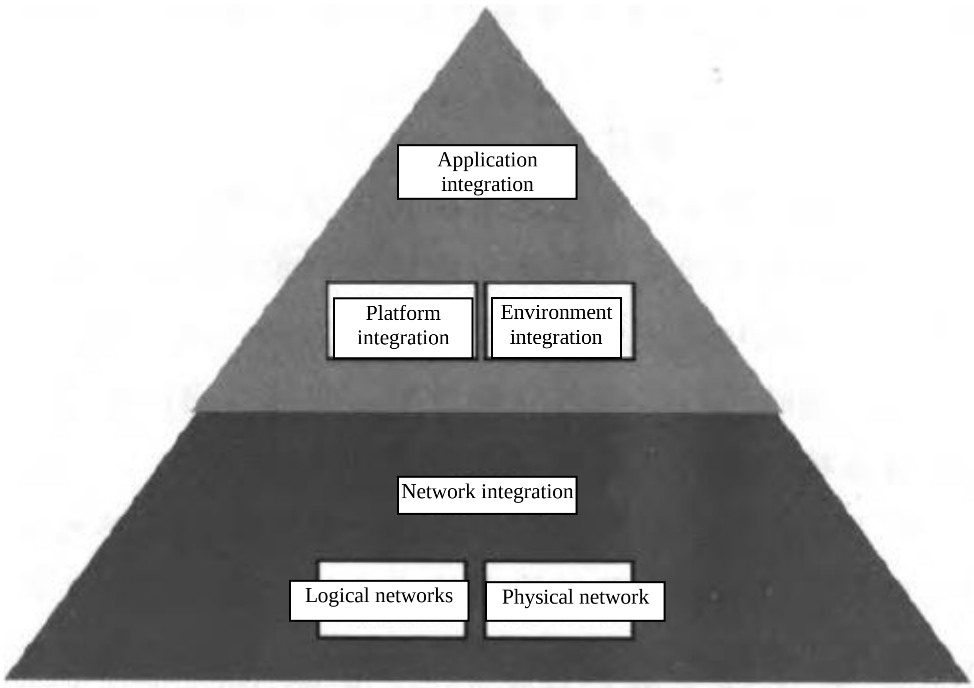


Figure 6-1: Framework of the On-Orbit Servicing Mission Planning System

(1) Application integration

Application integration is to organically integrate computing and display devices based on various platforms and different schemes into a seamless, parallel, independent and easily accessible space environment according to the purpose of system application and can be divided into functions according to application mode.

In accordance with the principle of "scientific, reasonable, streamlined, easy to use and easy to manage," the application integration environment of the on-orbit servicing mission planning system follows the construction idea of "experiment, deduction and drill." As shown in Figure 6-2, there are space target simulation areas, spacecraft control simulation areas, and deduction and evaluation areas, and each area is equipped with seats for data information processing, mission planning, ground control simulation, orbit control simulation, and result processing and analysis.

(2) Network integration

Considering that the on-orbit servicing mission planning system has multi-seat nodes and scalable server nodes, the system generally adopts a layer 2 network structure. Among them, the access layer switch has a low bandwidth, which is connected to each seat and storage separately, and is connected to the core switch. The core switches have high bandwidth and are networked with access layer switches, servers, and external systems. Logically, each seat is divided into three network segments, one for the blade server and one for the storage/application server, according to the space target simulation area, the spacecraft control simulation area and the deduction and evaluation area.

Figure 6-3 shows the logical network structure of the on-orbit servicing mission planning system, and its core layer is the backbone of high-speed network switching, which is crucial to the performance of the entire network, and its particularity mainly includes high-speed switching, high reliability, and low latency. The access layer is the seat access point that provides further operations on the seats through filtering and access control lists.

All network transmissions in the on-orbit servicing mission planning system are wired. An access layer switch and several seat clients are deployed in the space target simulation area, an access layer switch and several seat clients are deployed in the spacecraft control simulation area, an access layer switch and several seat clients are deployed in the deduction and evaluation area, and a core switch, a set of blade servers, an access layer switch, two servers (storage and application) and a set of solid-state storage are deployed in the equipment room.

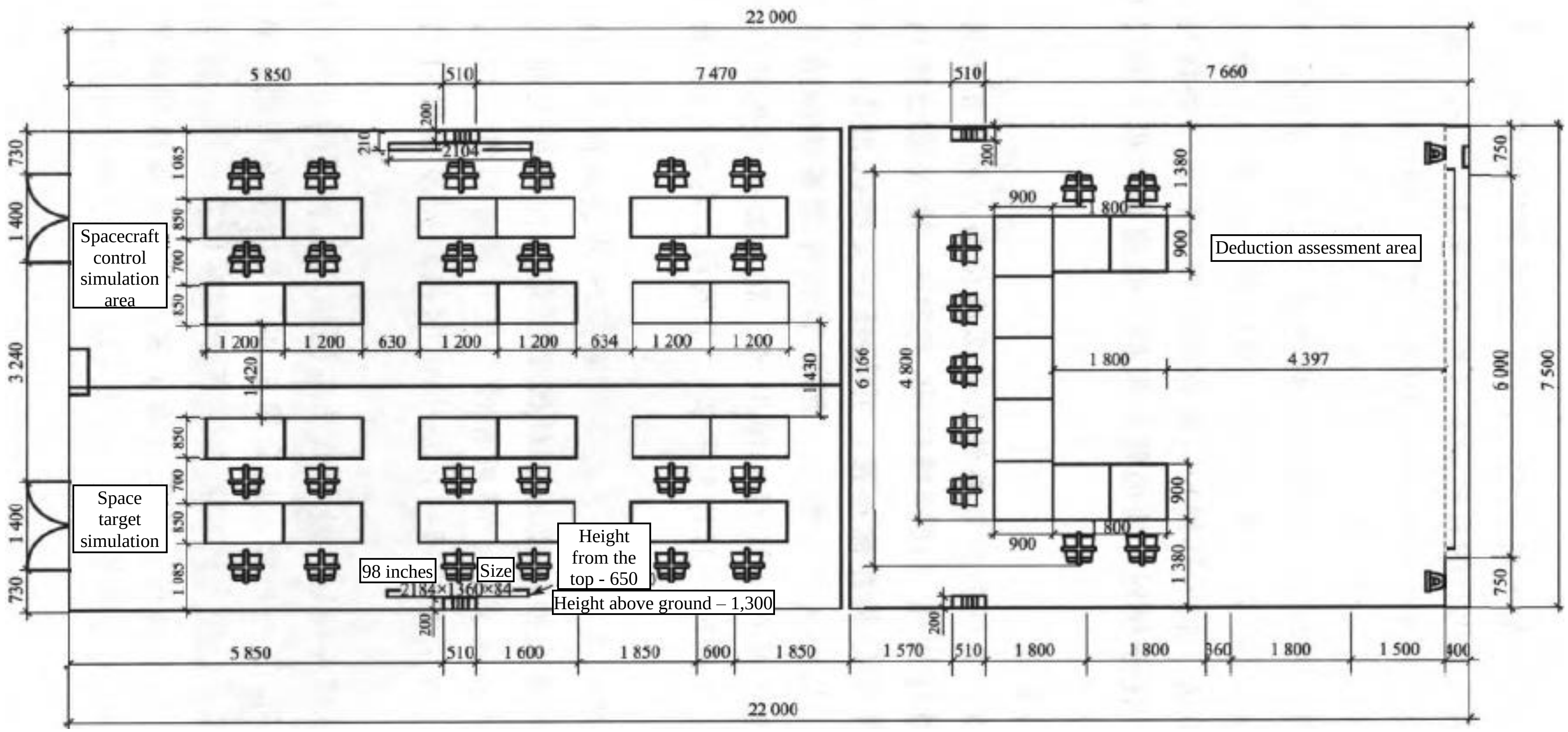


Figure 6-2: Schematic Diagram of the Seat Layout of the Application Integration Environment

The access layer switch is connected to the client and the customer, the access layer switch and the storage/application server, the access layer switch and storage, the blade server and the core switch are connected by twisted pairs, and the core switch is connected to the access layer switch by optical fiber.

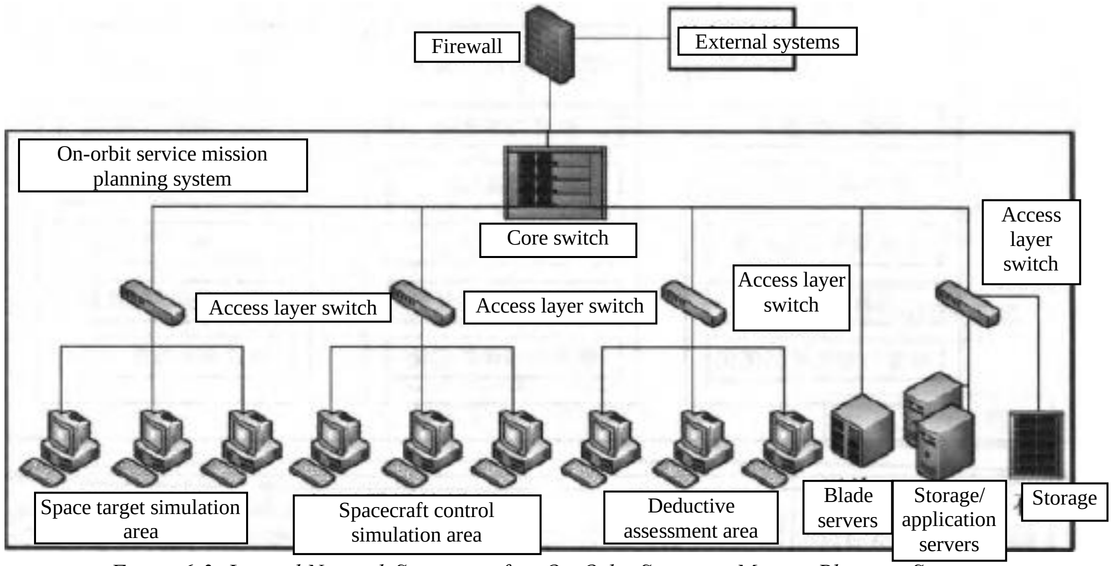


Figure 6-3: Logical Network Structure of an On-Orbit Servicing Mission Planning System

6.2.2: System Functional Architecture

(1) System architecture

Based on the application requirements and core methods of on-orbit servicing task planning, the system architecture composition is determined. The overall architecture of the orbital service mission planning system is divided into support layer, functional layer and application layer. The support layer provides relevant databases for on-orbit servicing mission planning, including on-orbit servicing knowledge base, space target database, space environment database, orbital dynamics model library, and basic model database. The functional layer provides necessary and key planning methods and processing means for on-orbit servicing tasks, including spacecraft on-orbit target allocation, orbital maneuver path planning, and service implementation action planning. The application layer is the on-orbit servicing task planning scenario of the system application and the on-orbit task styles that the system can support. Figure 6-4 shows the architecture of the on-orbit servicing task planning system.

(2) Functional flow

According to the requirements of on-orbit servicing tasks, according to the architecture design of the mission planning system, the functional process of on-orbit servicing task planning is divided into three stages:

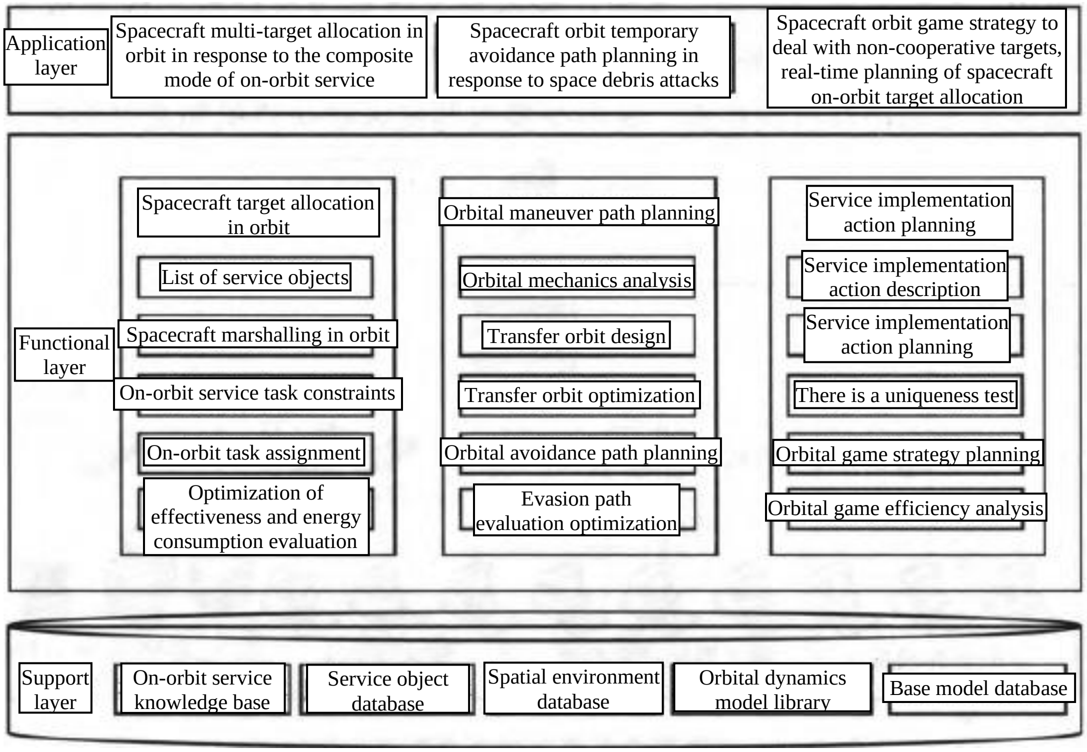


Figure 6-4: On-Orbit Servicing Mission Planning System Architecture

Planning Preparation Phase, Mission Planning Phase, and Assessment Optimization Phase. The planning preparation stage is to form the space environment, object data and task list according to the received mission before the on-orbit servicing mission planning and provide data and support for mission planning. In the mission planning stage, on the basis of determining the service objectives and understanding and mastering the space environment, the target allocation and orbit design are completed, and the pre-planning of strategies such as avoiding orbit path and orbit game is completed for the unexpected situations that may be encountered in the mission execution. In the evaluation and optimization stage, through the evaluation and feedback mechanism based on the simulation of spacecraft and space target services, a closed loop of "planning-evaluation-optimization-repeated planning" is formed, the planning strategy and scheme are improved, and a mission planning system of automatic planning, automatic evaluation and automatic optimization is formed. Figure 6-5 shows the on-orbit servicing task planning process.

(3) Control process

Figure 6-6 shows the main control process of the on-orbit servicing task planning system. 1) Monitor the system operation status data to ensure that the current state of the system is normal and stable; 2) Create a running instance to clarify the planning content of the on-orbit servicing task; 3) Select the hypothetical document, clarify the background of the task, and construct the hypothetical case;

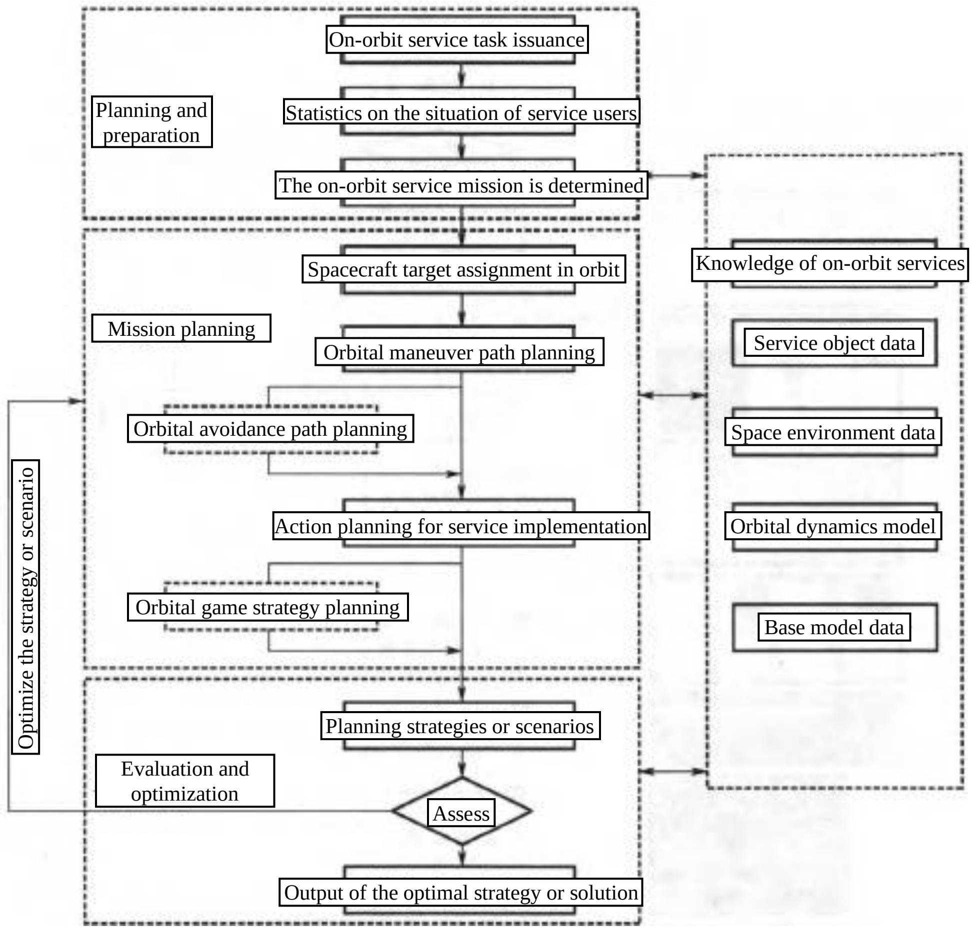


Figure 6-5: On-Orbit Servicing Task Planning Flow

4) Correlation with related systems, according to the type of on-orbit servicing experiment or ground mission planning simulation training, associated with ground operation control system, on-board system or digital satellite system and other related systems; 5) Start the running process, execute the relevant planning methods according to the task, and start the model process in a distributed manner; 6) When starting/pausing/accelerating/decelerating/jumping, select or turn on the ultra-real-time simulation mode according to the task requirements and simulation process; 7) Send scenario intervention events, and implement spacecraft orbit temporary avoidance path planning or spacecraft orbit game strategy real-time planning according to temporary emergencies such as space debris attack.

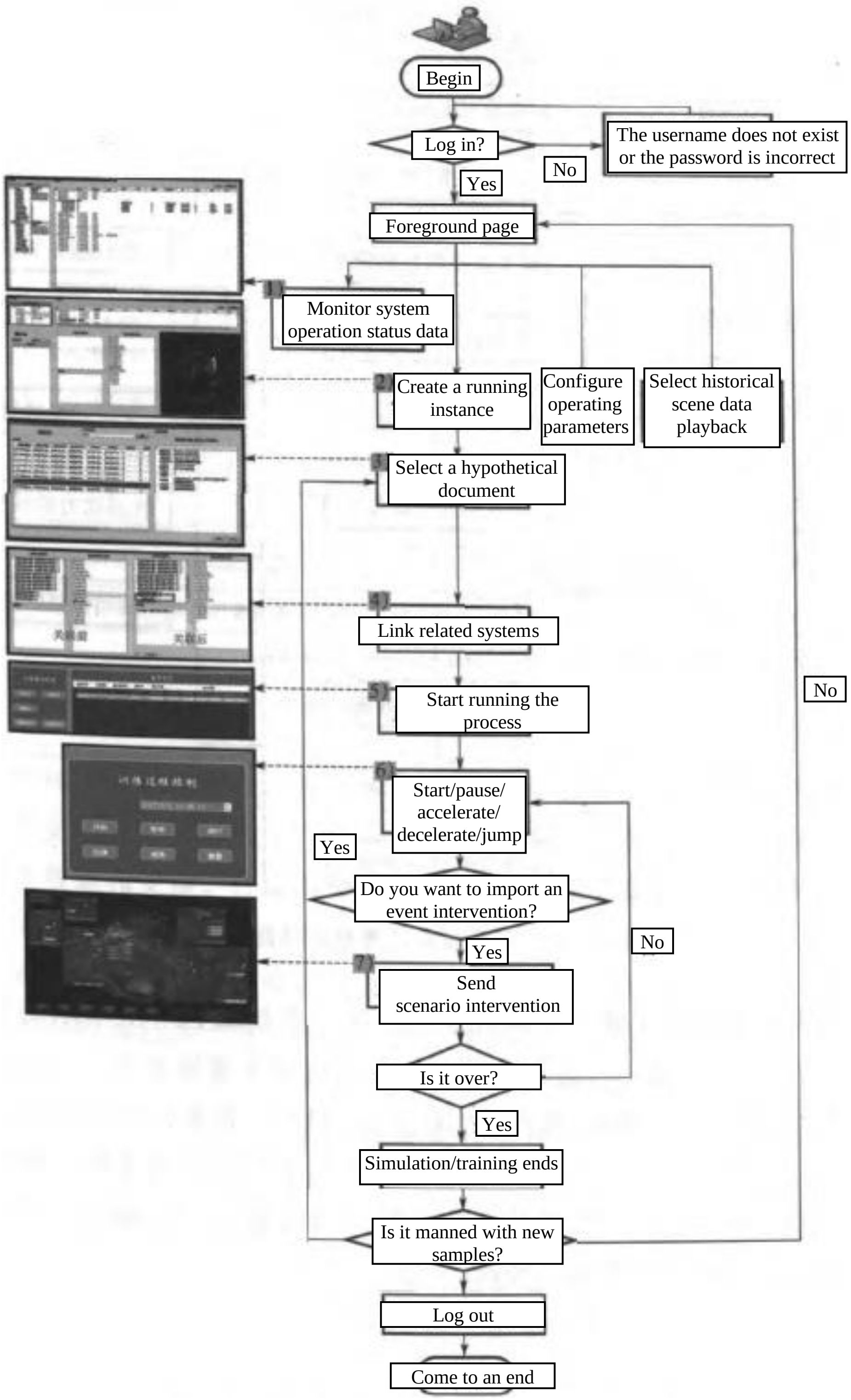


Figure 6-6: The Operation Process of the On-Orbit Servicing Task Planning System

6.3: Application of On-Orbit Servicing Task Planning System

The service models of future on-orbit servicing missions are complex and diverse, and unexpected situations can catch people off guard. To a certain extent, the advantages and disadvantages of the on-orbit servicing task system depend on the practical application of service task planning, so as to ensure that all kinds of mission planning are carried out in an orderly manner without interfering with each other, and maximize the advantages of automaticity and autonomy. The purpose of the system application is to verify the research results of on-orbit servicing mission planning experimentally, so as to test its feasibility, applicability and effectiveness in target allocation, orbit avoidance and orbit game.

6.3.1: Background Assumptions

In 20XX, with the continuous improvement of space research, development and application capabilities, countries have successively developed and launched a large number of spacecraft for various mission requirements, and the structure and composition of spacecraft have become increasingly complex, and their performance and technical level have been continuously improved. In this case, in order to ensure that the in-orbit spacecraft can operate in orbit in a more durable, stable and high-quality space environment, the corresponding on-orbit servicing tasks have been carried out.

6.3.1.1: Service Objectives

According to the demand for on-orbit services, there are 10 important geostationary orbit satellites that urgently need to carry out on-orbit services, and the number of orbits of each satellite is shown in Table 6-1. In the table, e is the eccentricity; i is the orbital inclination; Ω is the ecliptic longitude of the ascending node; ω is the pericentricity angle; τ is the true perisis angle.

Table 6-1: Number of Orbital Roots for Service Targets

Serial number	Target codename	e	i/(°)	Ω/(°)	ω/(°)	τ/(°)
1	Goal 1	0.001 3	0	0	235.408 0	65.576 3
2	Goal 2	0.001 1	0	0	228.155 3	90.470 3
3	Goal 3	0.001 4	0	0	8.029 4	100.930 5
4	Goal 4	0.000 9	0	0	214.263 7	110.823 7
5	Goal 5	0.000 5	0	0	183.858 7	117.095 1
6	Goal 6	0.000 3	0	0	332.982 5	126.556 4

Continued

Serial number	Target codename	*e*	*i*/(°)	*Ω*/(°)	*ω*/(°)	*τ*/(°)
7	Goal 7	0.000 7	0	0	53.883 3	138.072 7
8	Goal 8	0.001 2	0	0	246.486 4	190.272 9
9	Goal 9	0.001 3	0	0	271.291 2	210.311 6
10	Goal 10	0.001 9	0	0	274.604 7	236.420 9

After the preliminary analysis of the service objectives, the service priority of each service objective is estimated, as shown in Table 6-2.

Table 6-2: Service Priority Estimation

Serial number	Target codename	Priority
1	Goal 1	0.5
2	Goal 2	0.5
3	Goal 3	0.5
4	Goal 4	0.6
5	Goal 5	0.6
6	Goal 6	0.6
7	Goal 7	0.5
8	Goal 8	0.5
9	Goal 9	0.6
10	Goal 10	0.9

6.3.1.2: Spacecraft in orbit

In accordance with the on-orbit servicing requirements, spacecraft 1, spacecraft 2 and spacecraft 3 were organized to participate in this mission, and the number of roots in each orbit is shown in Table 6-3. The mass of the spacecraft in orbit is 2,500 kg, the initial true perigee angle is 0°, and the specific impulse of the propulsion system is 300 s. Table 6-4 shows the service success probability, service capability index, and success probability threshold of each spacecraft.

Table 6-3: Number of orbital roots of a spacecraft in orbit

Serial number	Spacecraft serial number	e	i/(°)	Ω/(°)	ω/(°)	τ/(°)
1	Spacecraft 1	0.001	0	0	296	25.39
2	Spacecraft 2	0.001	0	0	274	23.13
3	Spacecraft 3	0.001	0	0	280	24.96

Table 6-4: In-orbit spacecraft service capability data

Serial number	Satellite code	The probability of service success			Service Capability Index			Success probability threshold		
		Space craft 1	Space craft 2	Space craft 3	Space craft 1	Space craft 2	Space craft 3	Space craft 1	Space craft 2	Space craft 3
1	Goal 1	0.70	0.68	0.73	90	89	93	0.69	0.69	0.69
2	Goal 2	0.72	0.71	0.70	92	91	90	0.69	0.69	0.69
3	Goal 3	0.73	0.72	0.76	93	91	94	0.69	0.69	0.69
4	Goal 4	0.67	0.66	0.6	88	86	87	0.69	0.69	0.69
5	Goal 5	0.66	0.66	0.73	86	87	87	0.69	0.69	0.69
6	Goal 6	0.71	0.70	0.77	91	89	90	0.69	0.69	0.69
7	Goal 7	0.75	0.73	0.70	95	93	93	0.75	0.75	0.75
8	Goal 8	0.77	0.75	0.80	97	95	96	0.75	0.75	0.75
9	Goal 9	0.73	0.67	0.81	93	93	91	0.75	0.75	0.75
10	Goal 10	0.85	0.86	0.85	94	95	96	0.85	0.85	0.85

6.3.1.3 Situation settings

From the perspective of testing the effectiveness of the method in this book with multiple situations and multiple conditions, three special cases are set for the whole on-orbit servicing mission planning application, namely orbital coplanarity, space debris attack and non-cooperative target.

(1) Orbital coplanarity

In the process of spacecraft target allocation in orbit, many factors such as velocity impulse, fuel consumption and execution efficiency need to be comprehensively considered, among which whether the orbit of the spacecraft is coplanar with the target orbit will affect the generation of target allocation strategy. For this reason, two cases of coplanarity and non-coplanarity between the spacecraft and the target are considered in the application simulation, respectively.

(2) Space debris infestation

Spacecraft is at risk of space debris attack at any time during orbital maneuvering and service implementation, so space debris attack is set up in the application simulation. Suppose that a space debris appears in the plane of the spacecraft transfer orbit temporarily, and the flight speed is $[-0.763\ 0.763]^T$ km/s, approaching the spacecraft transfer orbit. In addition, avoidance preference is an important factor affecting the generation of spacecraft orbit avoidance paths, so the avoidance path generation under different avoidance preference requirements is considered in the case simulation.

(3) Non-cooperative goals

During the implementation of the service, the spacecraft is very likely to face non-cooperative targets, so the situation of non-cooperative targets is set in the application simulation. In addition, the different initial states of orbit positions of spacecraft and space targets are important factors restricting the generation of strategies, so the real-time planning ability of spacecraft orbit game strategy is tested from a variety of orbit states and different initial states in the application simulation.

6.3.2: Simulation and Validation

This section will focus on three types of special situations in the process of on-orbit servicing, use the planning methods proposed in this book, and verify the feasibility and effectiveness of the methods through the application of simulation.

6.3.2.1: Spacecraft target assignment in orbit

Although the allocation of spacecraft targets in orbit is determined according to the specific on-orbit servicing tasks, the subsequent orbit transfer process and cost need to be fully considered in the target allocation stage, so as to generate a robust allocation scheme. In order to effectively verify the method in this book, the following task scenarios are set up from the aspects of coplanarity and non-coplanarity between spacecraft and target to test the effectiveness of the method to solve the target allocation problem.

(1) The spacecraft is coplanar with the target

When spacecraft 1, spacecraft 2, spacecraft 3 and 10 space targets are in the same orbital plane, only the coplanar orbital transfer between spacecraft and space targets needs to be considered in the allocation process of orbital targets. According to the theory and method of orbital mechanics, the transfer orbits between coplanar orbits can be designed and simulated in advance. In this study, taking the pulsed thrust orbit maneuvering mode as an example, Figure 6-7 designs and simulates the optimal transfer orbit of spacecraft 1 from the initial orbit to the geostationary orbit where the target is located.

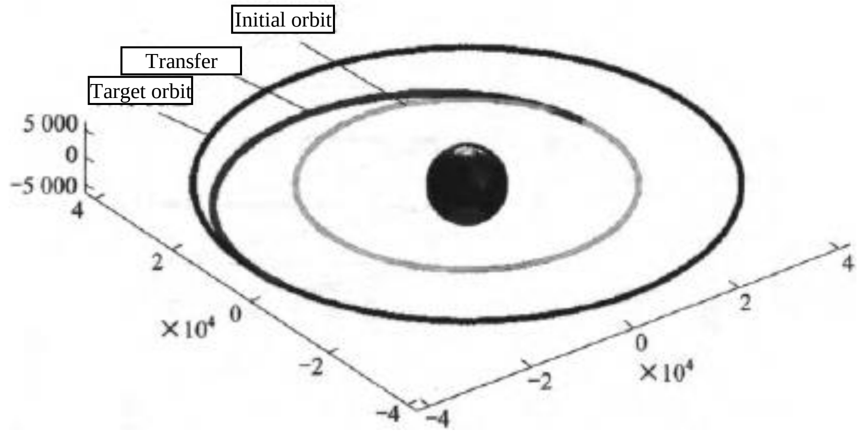


Figure 6-7: Coplanar Transfer Orbit of Spacecraft 1 to Target (see color insert)

Spacecraft 1 maneuvers to the orbit where the target is located according to the transfer orbit, and the change of its orbital root number is shown in Figure 6-8. Among them, because spacecraft 1 is coplanar with the target orbit, the right ascension longitude and orbital inclination of the transfer orbit parameters remain 0, and the other parameters change with the change of the transfer orbit trajectory.

According to the optimal transfer orbit from spacecraft 1 to the target orbit, the transfer parameters of each orbit are calculated, as shown in Table 6-5.

Table 6-5: Velocity and burnup parameters of spacecraft 1 to target coplanar orbit

Initial orbital velocity / (km/s)	Target orbital velocity / (km/s)	Pericamber velocity impulse / (km/s)	Apoapsis velocity impulse / (km/s)	First pulse fuel consumption / kg	Second pulse fuel consumption / kg	Total fuel consumption / kg
3.875 1	3.067 1	0.422	0.375 1	333.951	259.222	593.173

Since the orbits of spacecraft 2 and spacecraft 3 are also coplanar with the target orbits, their transfer orbit morphology and orbital parameter changes are roughly similar to those of spacecraft 1, so they will not be repeated here. Similarly, according to their optimal transfer orbits, the orbital transition parameters of spacecraft 2 to the target orbit are shown in Table 6-6, and the orbital transition parameters of spacecraft 3 to the target orbit are shown in Table 6-7.

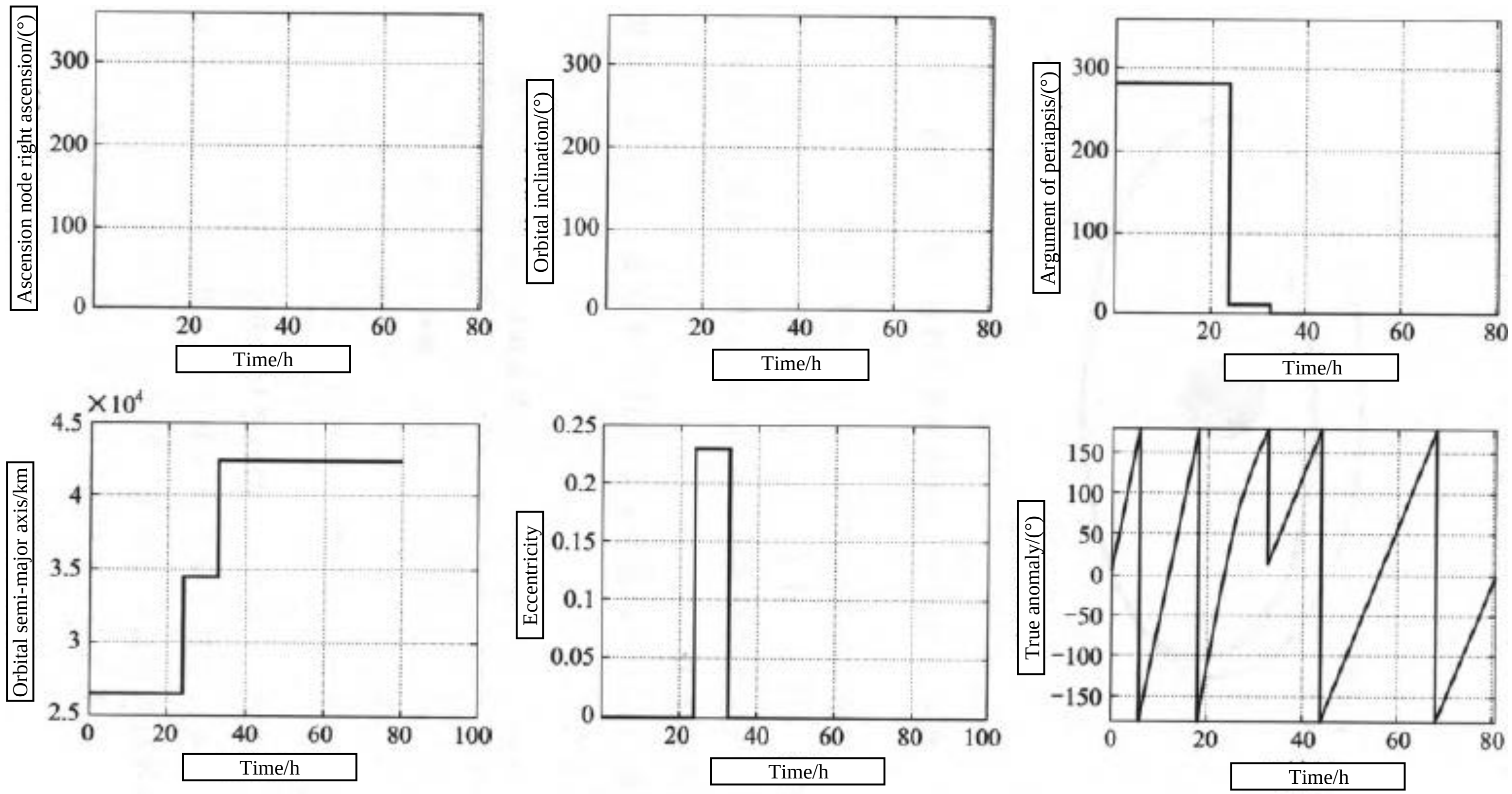


Figure 6-8: Changes in the Number of Orbital Roots of Coplanar Transfer Orbitals (see color insert)

Table 6-6: Velocity and burn-up parameters of spacecraft 2 to target coplanar orbit

Initial orbital velocity / (km/s)	Target orbital velocity / (km/s)	Pericamber velocity impulse / (km/s)	Apocalyptus velocity impulse / (km/s)	First pulse fuel consumption / kg	Second pulse fuel consumption / kg	Total fuel consumption/kg
3.794	3.0671	0.380	0.342	303.080	240.768	543.848

Table 6-7 Parameters of velocity and burnup consumption of spacecraft 3 to target coplanar orbit

Initial orbital velocity / (km/s)	Target orbital velocity / (km/s)	Pericamber velocity impulse / (km/s)	Apocalyptus velocity impulse / (km/s)	First pulse fuel consumption / kg	Second pulse fuel consumption / kg	Total fuel consumption/kg
3.784	3.0671	0.373	0.336	297.629	237.405	535.034

The above-mentioned orbital parameters are incorporated into the spacecraft on-orbit target allocation process, and the execution benefit and energy efficiency are taken as the optimization goals as described in Chapter 3, and the spacecraft on-orbit target allocation method is autonomously allocated. As shown in Table 6-8, the target assignment state ① is initialized with an all-0 matrix, and the target assignment model is substituted, but no resources are invested and the mission requirements are not met, and then the spacecraft is autonomously trained through the on-orbit target assignment process. In the training process, state ② high resource input makes the energy efficiency low, and the comprehensive benefit reaches the lowest value; state ③ meets the constraints, but the comprehensive benefit is not the maximum. After multiple rounds of self-learning and iteration, the method converges to state ④, and the strategy provided not only satisfies the mission elements, but also maximizes the comprehensive benefits, which is the optimal allocation strategy in this state, that is, objectives 1 to 6 are served by spacecraft 1, objectives 7 to 9 are served by spacecraft 2, and goal 10 is served by spacecraft 3.

Table 6-8: Spacecraft target allocation process in orbit

State		Target Allocation Strategy	Energy efficiency	Comprehensive benefits
1	Initialize	$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix}$	0.99	0

Continued

State		Target Allocation Strategy	Energy efficiency	Comprehensive benefits
2	Process status	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	0.01	0.99
3		$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$	0.90	0.91
4	Best strategy	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	0.97	0.98

Conclusion: When the spacecraft and the space target are located in the same orbital plane, the spacecraft orbit target allocation only needs to focus on the influence of spacecraft orbit transfer velocity, impulse and fuel consumption, and the optimal allocation of the target can be achieved according to the spacecraft on-orbit target allocation method proposed in this book.

(2) The spacecraft is non-coplanar with the target

When a spacecraft is located in a different orbital plane from the space target, the non-coplanar orbital transfer problem needs to be considered in the spacecraft on-orbit target allocation process. Here, the non-coplanar orbit of spacecraft 3 and the geostationary orbit where the target is located is taken as an example. Set spacecraft 3 in non-coplanar orbit and update its orbital root number as shown in Table 6-9. According to the theory and method of orbital mechanics, Figure 6-9 designs and simulates the non-coplanar transfer orbit of spacecraft 3 from the initial orbit to the geostationary orbit where the target is located.

Table 6-9: Spacecraft 3 orbital root number update

Serial number	Spacecraft serial number	e	i/(°)	Ω/(°)	ω/(°)	τ/(°)
1	Spacecraft 3	0.001	20	10	270	25.35

Spacecraft 3 maneuvers from a non-coplanar orbit to a geostationary orbit where the target is located, and its orbital root number changes are shown in Figure 6-10.

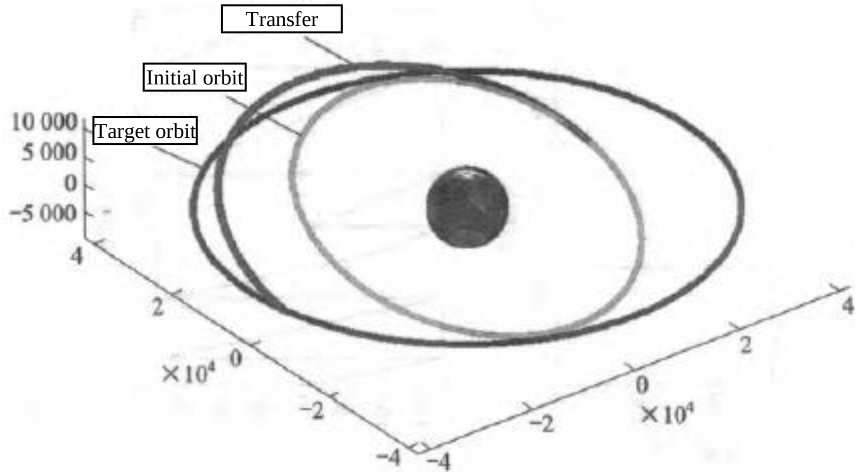


Figure 6-9: Non-Coplanar Transfer Orbit of Spacecraft 3 to Target

Among them, since spacecraft 3 is non-coplanar with the target orbit, the ascension longitude and orbital inclination of the transfer orbit parameters have a process of transformation from non-co-coplanar, and other parameters will change with the change of the transfer orbit trajectory.

According to the optimal transfer orbit from spacecraft 3 to the target orbit, the transfer parameters of each orbit are calculated, as shown in Table 6-10.

Table 6-10: Velocity and burnup parameters of spacecraft 3 to target non-coplanar orbit

Initial orbital velocity / (km/s)	Target orbital velocity / (km/s)	Pericamber velocity impulse / (km/s)	Apoapsis velocity impulse / (km/s)	First pulse fuel consumption / kg	Second pulse fuel consumption / kg	Total fuel consumption/kg
3.784	3.067 1	1.059	0.786	783.225	514.41	1 297.635

Through the calculation of the spacecraft on-orbit target allocation method, it is found that the implementation benefit and fuel efficiency value obtained by the previous allocation strategy are reduced. This is caused by the fact that spacecraft 3 is non-coplanar with the target orbit, and its orbital transfer requires higher velocity impulse and consumes more fuel. Through method recalculation, the optimal allocation strategy in the current state (see Table 6-11) is obtained, that is, target 10 is served by spacecraft 2, and the remaining targets are served by spacecraft 1.

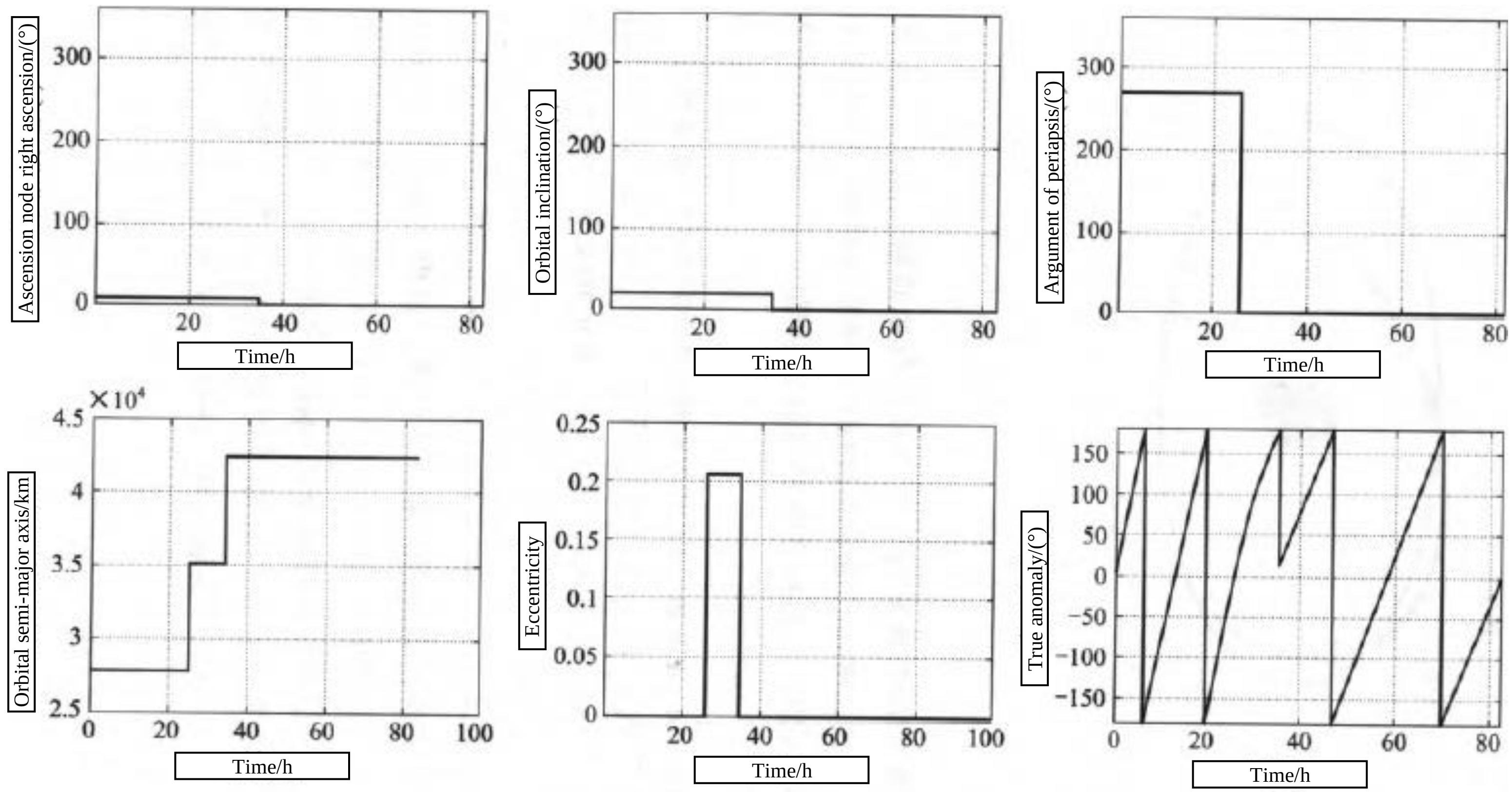


Figure 6-10: Changes in the Number of Orbital Roots for Non-Coplanar Transfer Orbitals (see color inserts)

Table 6-11: Spacecraft target allocation process in orbit

State		Target Allocation Strategy	Energy efficiency	Comprehensive benefits
1	Initialize	$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix}$	0.99	0
2	Process status	$\begin{bmatrix} 1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1 \end{bmatrix}$	0.01	0.99
3		$\begin{bmatrix} 1&1&1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&1&1&1&0 \\ 0&0&0&0&0&0&0&0&0&1 \end{bmatrix}$	0.92	0.89
4	Best strategy	$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&1 \\ 1&1&1&1&1&1&1&1&1&0 \\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix}$	0.98	0.97

Conclusion: Non-coplanar orbit transfer has a great impact on the generation of spacecraft on-orbit target allocation strategy due to the need for greater velocity impulse and more fuel consumption. The methods proposed in this book can cope with different scenarios and multiple orbit states and formulate the optimal target allocation strategy for in-orbit spacecraft.

Through the on-orbit servicing mission planning system, the on-orbit servicing assignment relationship of the spacecraft to each target can be constructed, and the dynamic effect of on-orbit servicing target allocation can be displayed. As shown in Figure 6-11, the red curve is the orbit of the spacecraft, the yellow curve is the orbit of the target satellite, and the remaining curves are the service response relationship between the spacecraft and each target.

6.3.2.2: Spacecraft orbit temporary avoidance path planning

According to the target allocation strategy, spacecraft 3 received the command from the ground center, left the original orbit and entered the space transfer orbit and approached target 10. During the orbital maneuver and service implementation, spacecraft 3 will suddenly encounter the attack of unknown space debris according to the situation. Therefore, it is necessary to use the spacecraft orbit temporary avoidance path planning method to plan the optimal emergency avoidance path for the spacecraft in time. In order to effectively verify the method of this book, this subsection sets different task preferences from the aspects of single factor and compound factor to test the effectiveness of the method in solving the orbital avoidance problem.

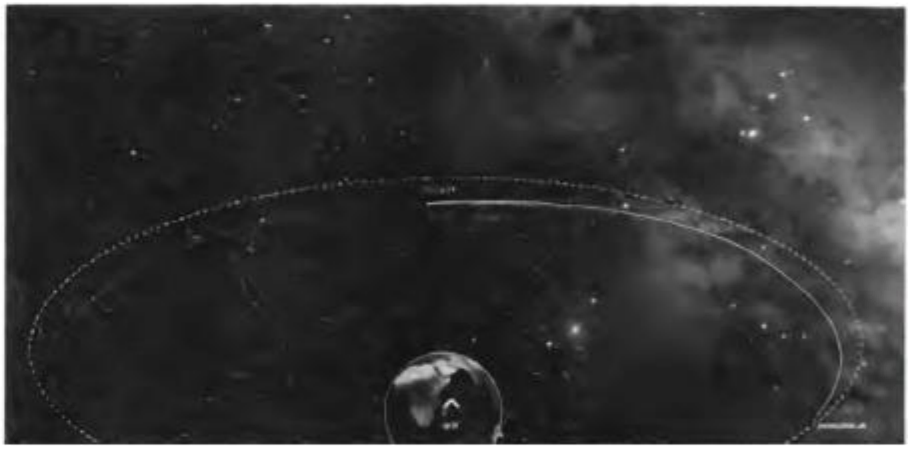


Figure 6-11; Example of Target Allocation Effect in the On-Orbit Servicing Mission Planning System (see color insert)

(1) Path planning to satisfy one-factor preference

For the single-factor preference, the spacecraft orbit temporary avoidance path planning will focus on the factors such as avoidance safety, orbit keeping, braking time, and fuel consumption, so as to obtain the optimal avoidance path to meet different preferences.

1) Focus on avoiding security factors.

In the process of space debris avoidance, the spacecraft focuses on the safety factor of avoidance and controls the change rate of the acceleration of the avoidance maneuver to improve the flight stability of the spacecraft. Therefore, in the spacecraft orbit temporary avoidance path planning, the index weights will be $\gamma_J = 1$, $\gamma_D = 0$, $\gamma_T = 0$, $\gamma_R = 0$ to obtain the optimal avoidance path focusing on the safety avoidance factors, as shown in Figure 6-12. The avoidance path can well consider the safety factors of avoidance, successfully avoid space debris and travel in the direction of the reference line and is expected to consume 6.332 kg of fuel.

2) Focus on orbit retention factors.

In the process of spacecraft avoidance of space debris, the key consideration is the orbit holding factor, that is, the minimization of lateral offset to reduce the deviation of the spacecraft from the set transfer orbit. Therefore, in the spacecraft orbit temporary avoidance path planning, the index weights will be $\gamma_J = 0$, $\gamma_D = 1$, $\gamma_T = 0$, $\gamma_R = 0$ to obtain the optimal avoidance path focusing on orbit holding factors, as shown in Figure 6-13.

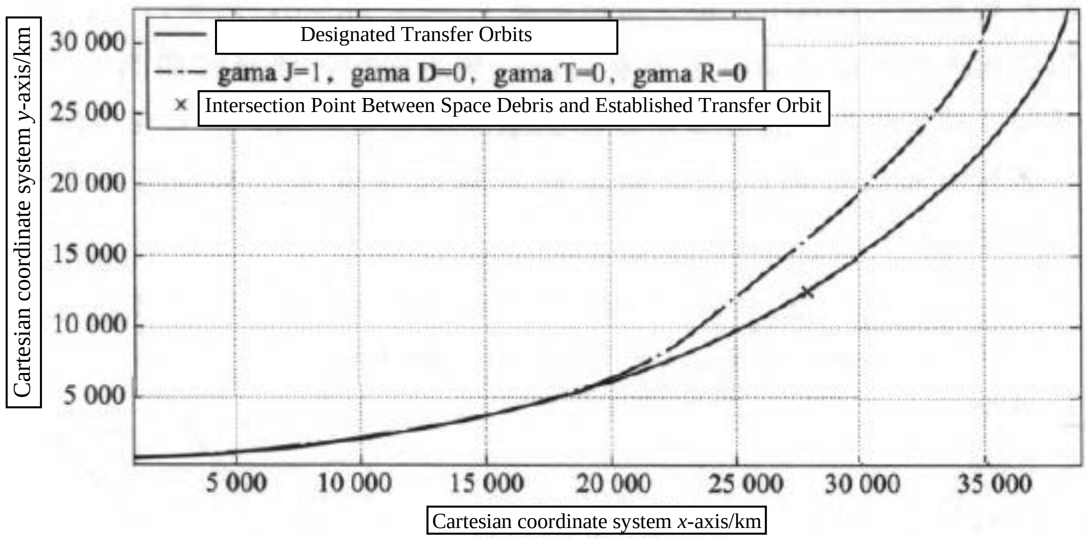


Figure 6-12: The Optimal Avoidance Path Focuses on Security Avoidance

The avoidance path is able to control the lateral offset well, and the space debris can be successfully avoided while continuing along the set transfer trajectory as much as possible, and the estimated fuel consumption is 19.815 kg.

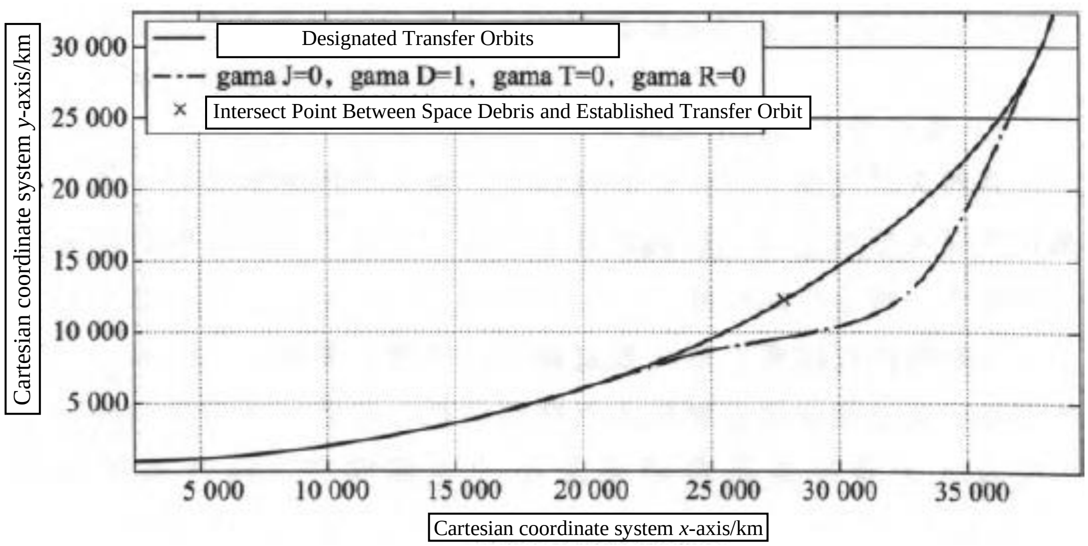


Figure 6-13: The Optimal Avoidance Path Focuses on the Orbit Holding Factor

3) Focus on braking aging factors.

In the process of spacecraft avoidance of space debris, the braking time factor should be considered, that is, the spacecraft braking time should be optimized to take evasive actions as soon as possible.

Therefore, in the spacecraft orbit temporary avoidance path planning, the index weights will be $\gamma_J = 0$, $\gamma_D = 0$, $\gamma_T = 1$, $\gamma_R = 0$ to obtain the optimal avoidance path that focuses on braking aging factors, as shown in Figure 6-14. The avoidance path can well control the spacecraft braking time, and take evasive actions as soon as possible after the discovery of space debris to ensure the safety of the spacecraft itself, and the fuel consumption is expected to be 12.380kg

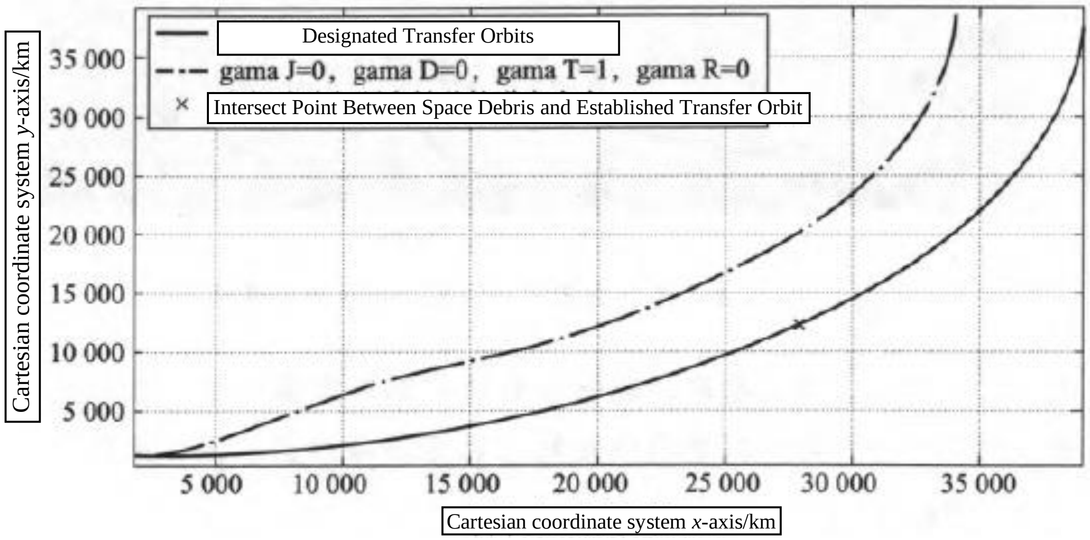


Figure 6-14: The Optimal Avoidance Path Focuses on the Braking Aging Factor

4) Focus on fuel consumption factors.

In the process of spacecraft avoidance of space debris, the fuel consumption factor is mainly considered, that is, the fuel consumption of the spacecraft thruster is optimized. Therefore, in the spacecraft orbit temporary avoidance path planning, the index weights will be $\gamma_J = 0$, $\gamma_D = 0$, $\gamma_T = 0$, $\gamma_R = 1$ to obtain the optimal avoidance path that focuses on fuel consumption, as shown in Figure 6-15. The avoidance path can control the fuel consumption of the spacecraft well, and the avoidance path with the least fuel consumption is selected for the spacecraft while successfully avoiding space debris, and the fuel consumption is expected to be 6.076 kg.

As shown in Figure 6-16, the avoidance paths that satisfy the preferences of each single factor are placed in the same scenario. Comparing the avoidance paths, it can be seen that the avoidance paths obtained by focusing on the safety avoidance factors can better control the acceleration change rate of spacecraft maneuvering and make the avoidance maneuver more stable. The avoidance path obtained by focusing on the orbit holding factor can better control the distance of the spacecraft from the set transfer orbit and reduce the uncertainty risk coefficient of the spacecraft in the unknown airspace.

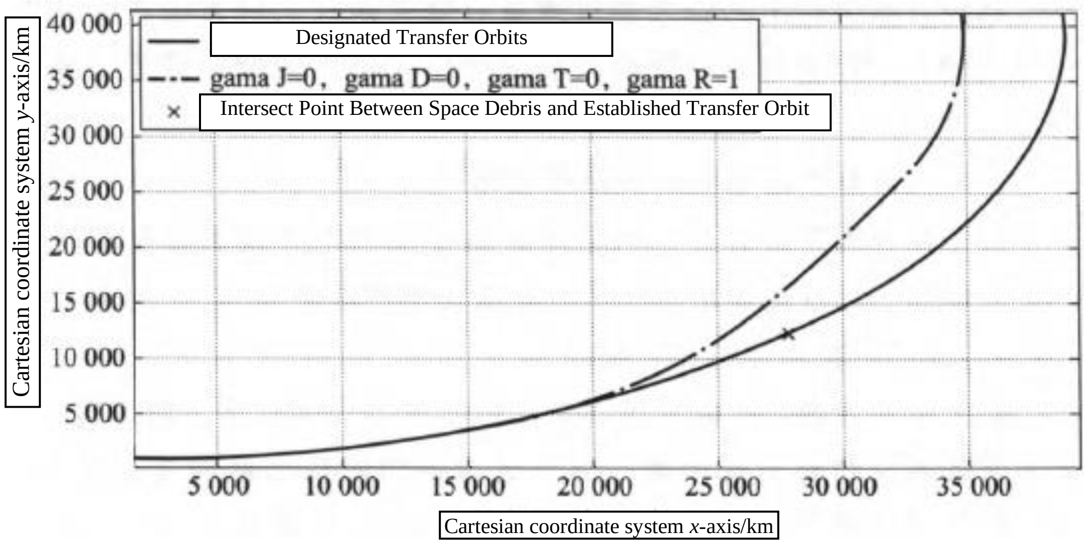


Figure 6-15: Optimal Avoidance Path with Fuel Consumption in Mind

Focusing on the avoidance path obtained by considering the braking aging factor, the braking time of spacecraft to avoid space debris can be better controlled, and the evasion behavior can be taken as soon as possible. The avoidance path obtained by focusing on the fuel consumption factor can better control the fuel consumption of the spacecraft to avoid maneuvering and obtain the path with the least fuel consumption.

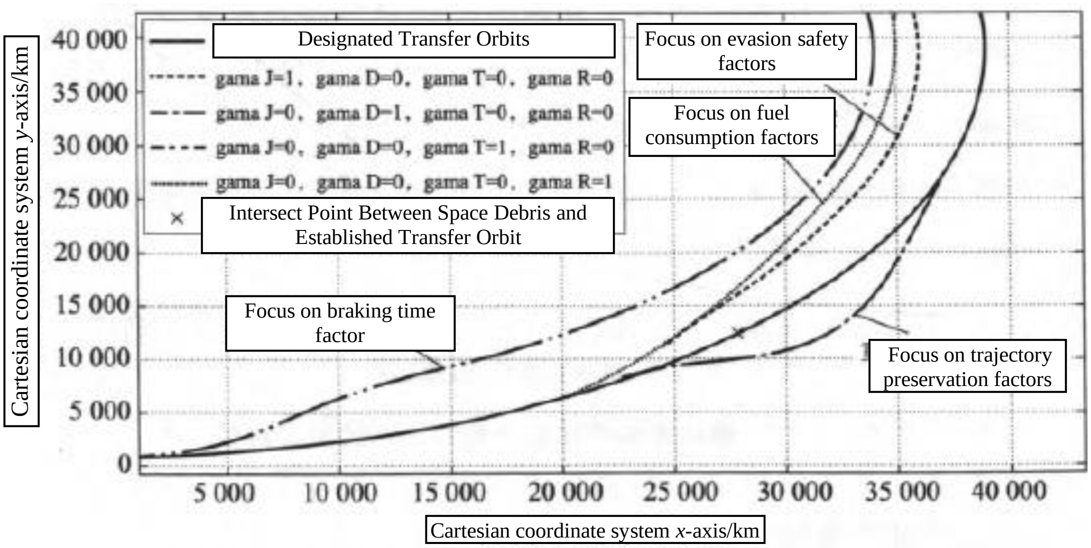


Figure 6-16: Optimal Avoidance Path of Spacecraft Under Single-Factor Preference (see color insert)

Conclusion: The spacecraft orbit temporary avoidance path planning method proposed in this book can plan the optimal avoidance path for spacecraft to meet different preferences according to the different avoidance preferences.

(2) Path planning to meet the preferences of composite factors

For the composite factor preference, the composite preference with different weight proportions is considered in the orbit avoidance path planning to obtain the optimal avoidance path that satisfies different preferences.

1) Compounding preference biased towards fuel consumption.

In the process of space debris avoidance, spacecraft should consider avoidance safety, orbit keeping, braking timeliness and fuel consumption. For the composite preference that is more biased towards fuel consumption, i.e., the weight of the preference for fuel consumption is relatively larger, the weight of the indicator will be $\gamma_J = 0.2$, $\gamma_D = 0.2$, $\gamma_T = 0.1$, $\gamma_R = 0.5$ to obtain the optimal avoidance path that favors the fuel consumption compound preference, as shown in Figure 6-17. The avoidance path takes into account the factors of avoidance safety, orbit maintenance and braking timeliness while focusing on fuel consumption and obtains an avoidance path with a better effect than the single preference factor avoidance path, with an estimated fuel consumption of 8.834 kg.

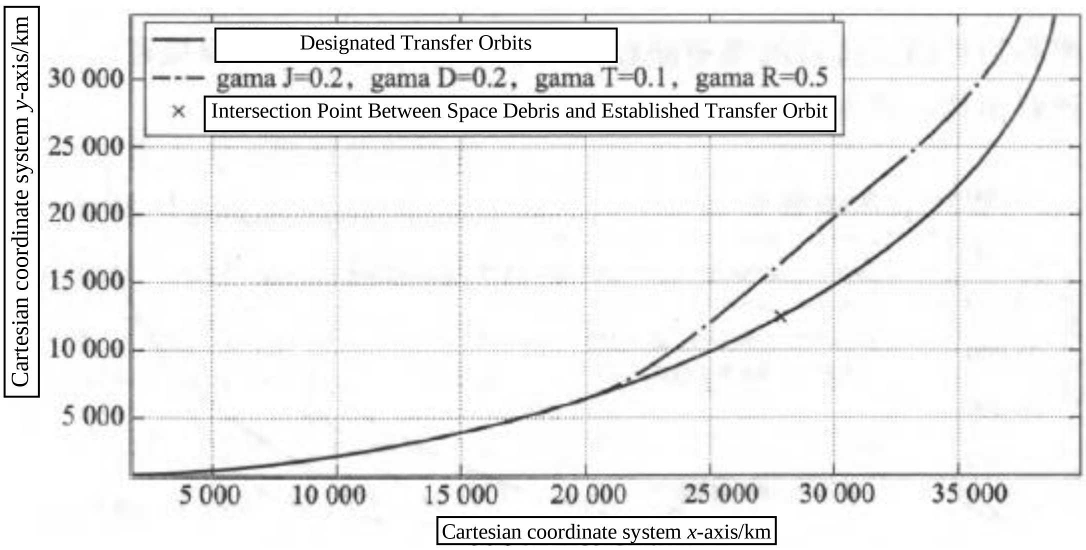


Figure 6-17: Optimal Avoidance Path for Fuel Consumption Compound Preference

2) Deviated orbit retention composite preference

For the more deviated orbit retention compound preference, that is, the weight of the orbit maintenance preference is relatively larger, the index weight will be $\gamma_J = 0.2$, $\gamma_D = 0.4$, $\gamma_T = 0.1$, $\gamma_R = 0.3$, so as to obtain the optimal avoidance path of the preferred orbit maintenance compound preference. Figure 6-18 shows the path.

The avoidance path takes into account the factors of avoidance safety, fuel consumption and braking timeliness while focusing on orbit maintenance, so that the evasion path can quickly return to the set transfer orbit after successfully avoiding the target, and the whole process is expected to consume 9.502 kg of fuel.

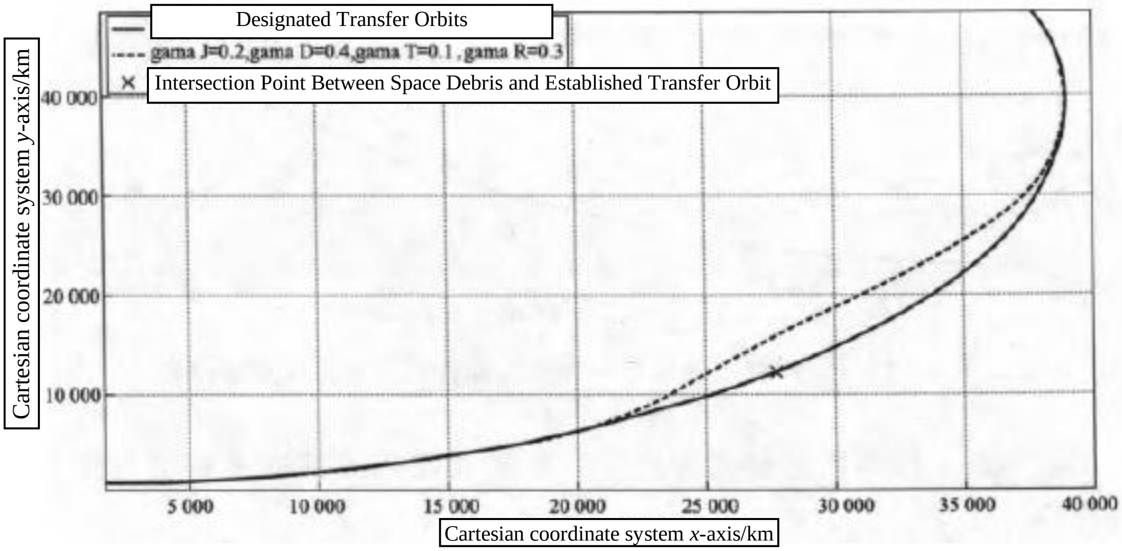


Figure 6-18: Optimal Avoidance Path for Deflected Orbit Retention Compound Preference

3) a composite preference for orbital retention and safety avoidance.

For the composite preference of orbit retention and safety avoidance, that is, the weight of orbit retention and safety avoidance preference is relatively larger, which will make the index weight $\gamma_J = 0.3$, $\gamma_D = 0.5$, $\gamma_T = 0.1$, $\gamma_R = 0.1$ to obtain the optimal avoidance path for the composite preference of biased orbit holding and avoidance safety, as shown in Figure 6-19. The avoidance path can pay more attention to orbit maintenance and avoidance safety factors and obtain an avoidance path with better avoidance path effect, and the fuel consumption is expected to be 10.996 kg.

As shown in Figure 6-20, the avoidance path that satisfies the preferences of each composite factor is placed in the same scenario. Comparing the avoidance paths, it can be seen that the avoidance paths obtained by biasing the fuel consumption compound preference can obtain the avoidance paths that save fuel and take into account the orbit maintenance and braking timeliness. The avoidance path obtained by the composite preference of biased orbit retention can better meet the comprehensive indexes of target avoidance, orbit keeping, braking flexibility and fuel saving. The avoidance path obtained by the composite preference of biased orbit retention and safety avoidance can further reduce the risk of uncertainty and achieve rapid recovery to the set transfer orbit, but this needs to be at the cost of higher fuel consumption.

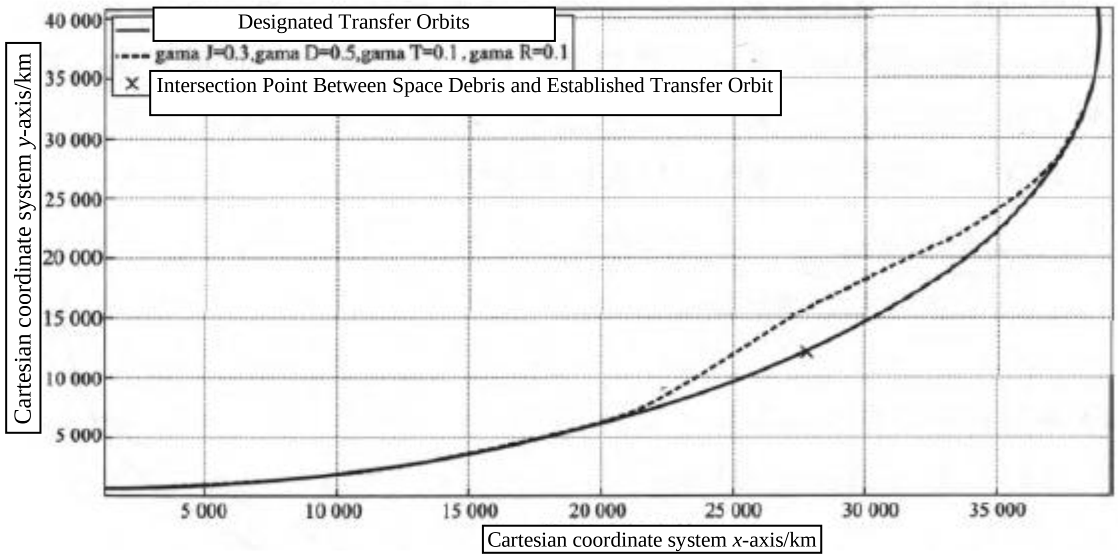


Figure 6-19: Optimal Avoidance Path for the Composite Preference of Biased Orbit Holding and Avoidance Safety

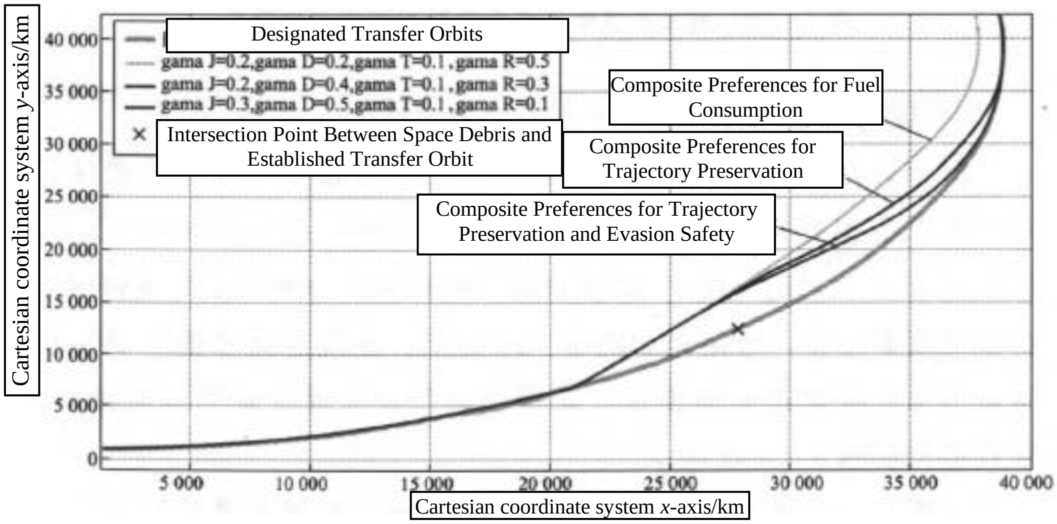


Figure 6-20: Optimal Avoidance Path of Spacecraft Under the preference of Composite Factors (see color insert)

Conclusion: The spacecraft orbit temporary avoidance path planning method proposed in this book can consider a variety of preference factors for the orbit avoidance problem at the same time and can obtain the optimal avoidance path to meet the needs of different preferences through the flexible setting of each preference weight.

Through the on-orbit servicing mission planning system, the spacecraft orbit temporary avoidance path can be planned, and the real-time avoidance dynamic effect of the spacecraft can be displayed. As shown in Figure 6-21, the yellow trajectory is the established space transfer orbit of the spacecraft, and the red curve is the optimal avoidance path of the spacecraft.

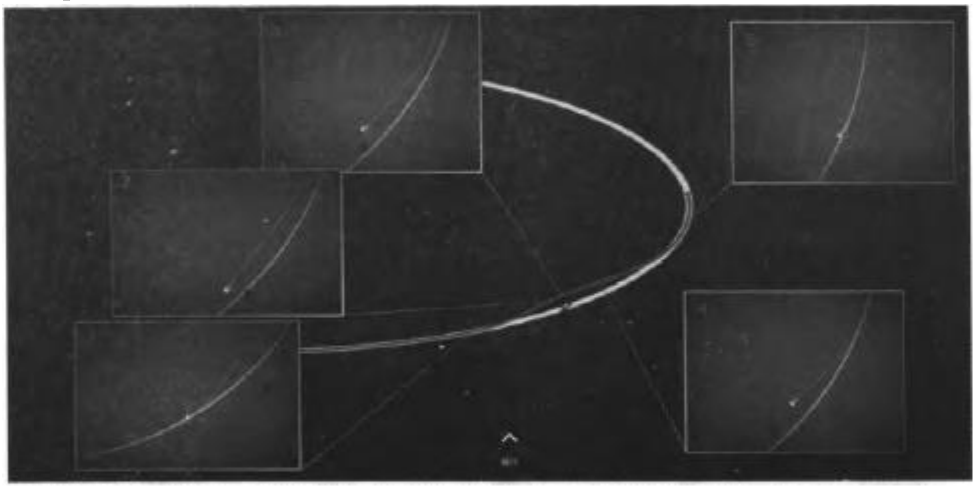


Figure 6-21: Example of Orbit Avoidance Effect in an On-Orbit Service Mission Planning System (see color insert)

6.3.2.3: Real-time planning of spacecraft orbit game strategy

In the process of performing on-orbit servicing tasks, spacecraft needs to use the real-time planning method of spacecraft orbit game strategy to plan the optimal tracking strategy for spacecraft in time. In order to effectively verify the method in this book, a variety of different initial states of spacecraft and space target are considered in the following article to test the effectiveness of the method in solving the orbital game problem.

(1) The target is located in the left front

When the space target and the spacecraft are in the same orbital plane and are located in front of the left side of the spacecraft, the initial state parameters of the spacecraft and the space target are shown in Table 6-12.

Table 6-12: Initial status of spacecraft and space target (target in front left)

	x/km	y/km	z/km	$\dot{x}$/(km/s)	$\dot{y}$/(km/s)	$\dot{z}$/(km/s)
P	0	0	0	-0.049 6	0.041 8	0
E	-69	70	0	-0.037 1	0.031 4	0

For the problem of non-complete information sequential game with non-cooperative targets, based on the Nash equilibrium strategy, the real-time planning method of spacecraft orbit game strategy is used to deal with the behavior of the target satellite, and the best game strategy can be obtained. In order to shorten the distance with the space target, the spacecraft has been trying to catch up in the y-axis direction, and at the beginning of the lateral distance to the target in the x-axis direction, it finally catches up with the target after 2,325 s. In this initial state, the entire orbital trajectory of the spacecraft against the space target is shown in Figure 6-22.

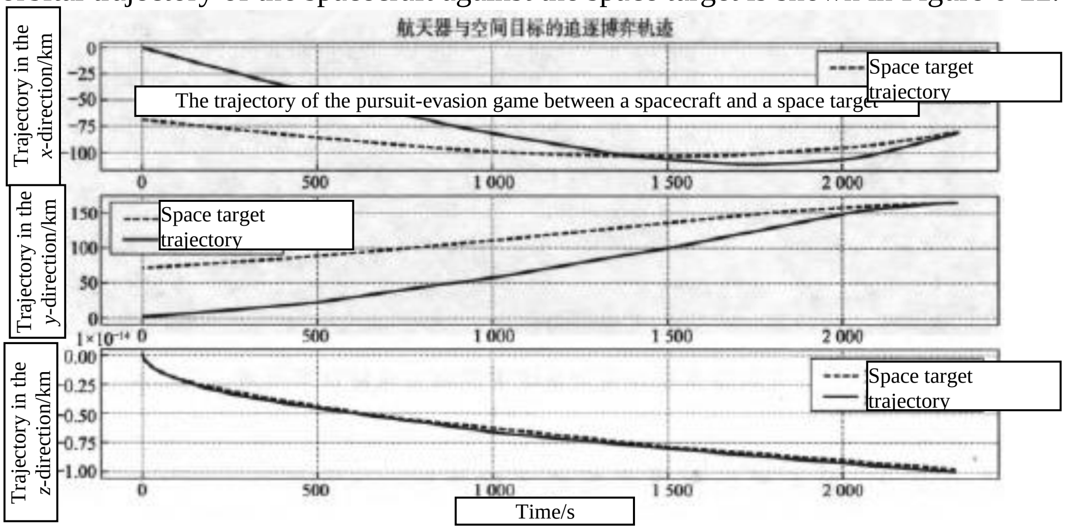


Figure 6-22: Orbital Game Between Spacecraft and Space Target when the Target is in Front Left

(2) The target is located in the front right

When the space target and the spacecraft are in the same orbital plane and are located in front of the right side of the spacecraft, the initial state parameters of the spacecraft and the space target are shown in Table 6-13.

Table 6-13: Initial status of spacecraft and space target (target in front right)

	x/km	y/km	z/km	$\dot{x}$/(km/s)	$\dot{y}$/(km/s)	$\dot{z}$/(km/s)
P	0	0	0	0.049 6	0.041 8	0
E	72	70	0	0.037 1	0.031 4	0

In order to shorten the distance between the spacecraft and the space target, the spacecraft has been trying to catch up in the y-axis direction and also in the x-axis direction, during which the game trajectory has interactive fluctuations due to the non-cooperative behavior of the space target, and the spacecraft successfully catches up with the target after 1,862 s.

In this initial state, the entire orbital trajectory of the spacecraft against the space target is shown in Figure 6-23.

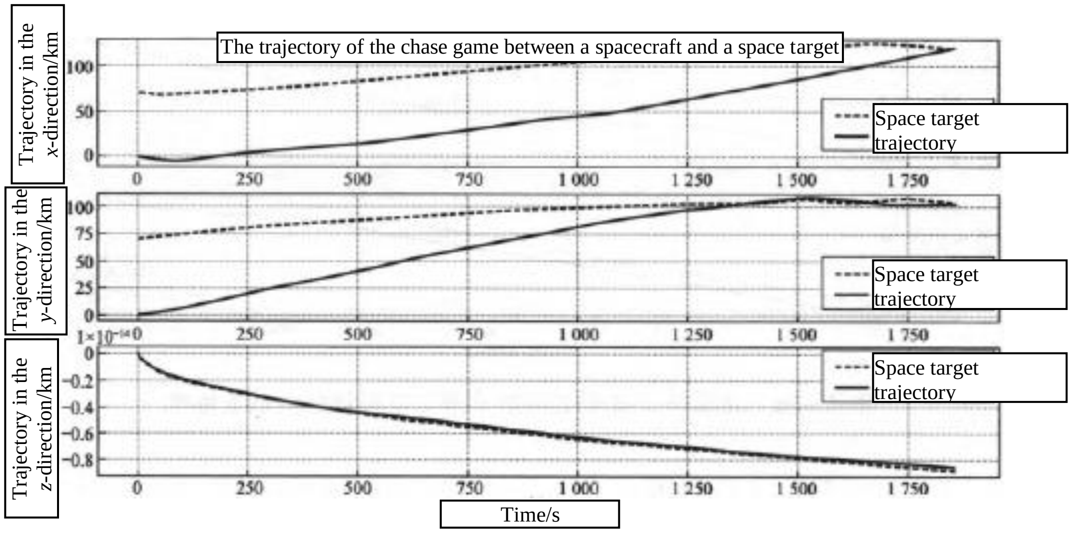


Figure 6-23: Orbital Game Between Spacecraft and Space Target When the Target is in Front Right

(3) The target is located in a high orbit

When the space target and spacecraft are located in different orbital planes, and the orbital position of the space target is higher than that of the spacecraft, the initial state parameters of the spacecraft and the space target are shown in Table 6-14.

Table 6-14: Initial status of spacecraft and space targets (targets in high orbit)

	x/km	y/km	z/km	$\dot{x}$/(km/s)	$\dot{y}$/(km/s)	$\dot{z}$/(km/s)
P	0	0	0	0.049 6	0.041 8	0.027 8
E	50	60	22	0.037 1	0.031 4	0

In this initial state, the spacecraft and the space target play a game with each other, during which the spacecraft has been trying to catch up in the axis direction in order to shorten the distance with the space target, reduce the lateral distance with the target as much as possible in the direction of the working axis, and reach the orbital altitude close to the space target by ascending the orbit, and finally it took 1986 s for the spacecraft to catch up with the target. In this initial state, the entire orbital trajectory of the spacecraft against the space target is shown in Figure 6-24.

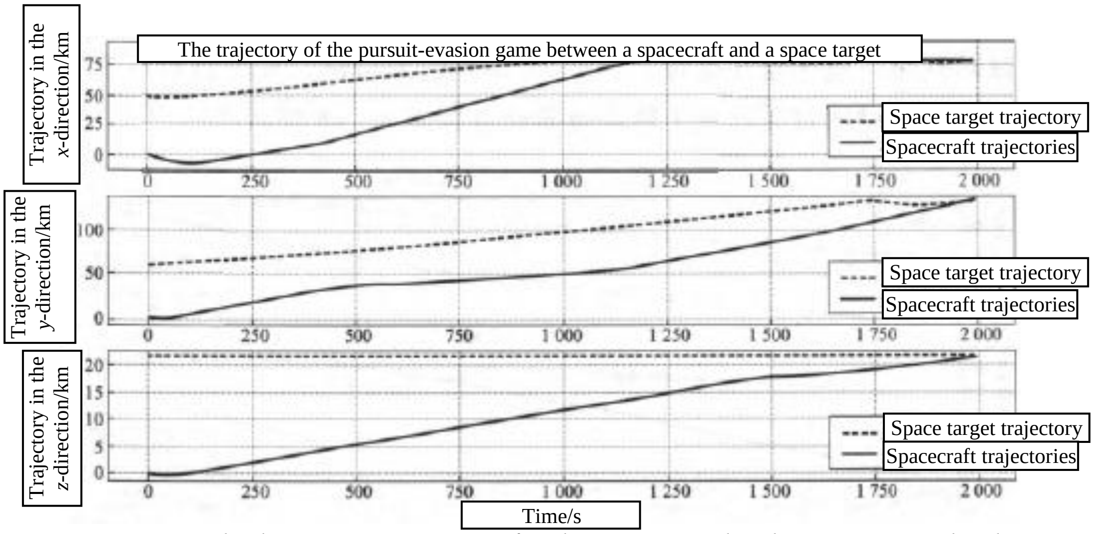


Figure 6-24: Orbital Game Between Spacecraft and Space Target When the Target is in High Orbit

(4) The target is located in low orbit

When the space target and the spacecraft are located in different orbital planes, and the orbital position of the space target is lower than that of the spacecraft, the initial state parameters of the spacecraft and the space target are shown in Table 6-15.

Table 6-15: Initial status of spacecraft and space targets (targets in low orbit)

	x/km	y/km	z/km	$\dot{x}$/(km/s)	$\dot{y}$/(km/s)	$\dot{z}$/(km/s)
P	0	0	0	-0.049 6	0.041 8	-0.027 8
E	-67	69	-30	-0.037 1	0.031 4	0

In order to shorten the distance from the space target, the spacecraft has been trying to approach the target in the x-axis direction and catch up in the y-axis direction, and finally caught up with the target after 2, 853 s by reducing the orbital altitude. In this initial state, the entire orbital trajectory of the spacecraft against the space target is shown in Figure 6-25.

Conclusion: In the face of the sequential game process of different initial states, completely unknown behaviors and incomplete information of space targets, the Nash equilibrium strategy of orbit game can be obtained independently through the branched deep reinforcement learning architecture, and finally the tracking rendezvous of space targets can be realized.

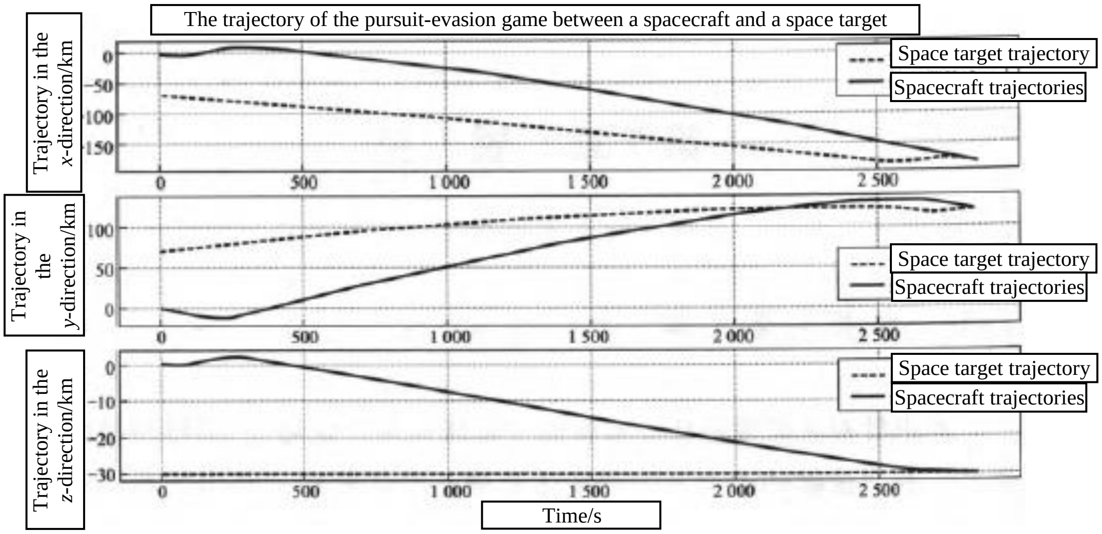


Figure 6-25: Orbital Game Between Spacecraft and Space Target When the Target is in Low Orbit

6.4: Chapter Summary

Guided by the application requirements of on-orbit servicing experiments and ground simulation training, this chapter introduces the foundation, requirements and key points of the construction of the mission planning system, clarifies the application requirements of the on-orbit servicing mission planning system, designs the architecture of the mission planning system, constructs the functional architecture of the mission planning system, and realizes the systematization, modularization, easy operation, easy expansion and easy maintenance characteristics of the system. Based on the case scenario, the on-orbit servicing task planning system is applied to test the feasibility and effectiveness of the intelligent planning method for on-orbit servicing tasks, meet the needs of on-orbit servicing task planning, and the visual simulation effect is good, which can provide strong support for on-orbit experiments or ground simulation training.

7.1: THE MAIN WORK OF THIS BOOK

In view of the academic problems faced by on-orbit servicing tasks, such as "nonlinear combinatorial optimization," "multi-limit shortest path" and "continuous dynamic interaction," this book focuses on the problems of "target allocation of composite service mode," "orbital avoidance path planning" and "orbital game strategy planning" in accordance with the research idea of "raising problems, analyzing problems, providing methods, and verifying and analyzing."

7.1.1: The Planning Requirements and Research Framework of On-Orbit Servicing Tasks Are Studied

In order to better deal with the problem of mission planning with a long duration and uneven types of situations, firstly, the on-orbit servicing task planning is summarized, the relevant concepts are defined, and the on-orbit servicing task planning process is analyzed. Secondly, from the aspects of timeliness, controllability and autonomy, the on-orbit servicing mission planning needs are analyzed, and it is pointed out that the mission planning will face the difficult problems of "uneven supply and demand," "space debris attack" and "non-cooperative target service."

7.1.2: The Method of On-orbit Target Allocation in the Composite Service Mode is Studied

In the face of the practical needs of the limited on-orbit force to adopt the composite service mode to serve multiple spatial targets with different orbit positions and different priorities, firstly, a nonlinear combinatorial optimization model was constructed to describe the problem of on-orbit target allocation under the composite service mode.

Secondly, in order to meet the characteristics of the on-orbit servicing target allocation problem and make up for the reward bias and overestimation problems of the classical method, the convergence and stability of the Deep Q Networks method are improved. Finally, in pursuit of the balanced development of execution efficiency and energy efficiency, a two-way training network for target allocation was built. The applicability and comparative advantages of the method are tested through case studies.

In order to solve the problem of "nonlinear combinatorial optimization" of spacecraft target allocation in orbit, a target allocation algorithm based on improved Deep Q Networks was proposed in order to solve the problem of "nonlinear combinatorial optimization" of spacecraft target allocation in orbit. The algorithm can effectively cope with the adverse effects of uneven supply and demand, give full play to the advantages of the composite service model, realize the balanced development of execution efficiency and energy efficiency, and achieve the comprehensive goal of less force investment and higher expected success rate, which can provide effective auxiliary decision-making for on-orbit services. Compared with other algorithms, the training time is shortened by 80% on average, the training error reduction rate is doubled, and the execution effect fluctuates less while increasing rapidly, which further improves the shortcomings of the conventional method that is difficult to directly apply to the composite service mode and the high computing time.

7.1.3: The Method of Temporary Avoidance Path Planning of Spacecraft Orbit is Studied

In the face of the spacecraft flying along the set transfer orbit, in order to avoid space debris, it is necessary to adopt temporary planning to obtain the optimal avoidance path, firstly, a multi-constraint shortest path model is constructed to describe the problem of orbital temporary avoidance path planning in the face of space debris. Secondly, a space motion coordinate system based on Frenet is constructed, which can take into account the absolute motion of flight along the transfer orbit and the relative motion of space debris avoidance, which solves the problem that space avoidance motion is not easy to represent. Finally, the artificial potential field function is improved, the action area of each potential field is adjusted, and a comprehensive potential field model with reference line traction, long-distance point repulsion neglect and obstacle point gravitational weakening is constructed, which avoids the premature trajectory deviation and local oscillation phenomenon of the traditional artificial potential field method. The applicability and comparative advantages of the method are tested through case studies.

In order to solve the outstanding risk of untimely avoidance due to temporary space debris inauguration, and in the face of the problem of "multi-restricted shortest path" presented by spacecraft orbit temporary avoidance, a path generation algorithm based on Frenet and improved artificial potential field was proposed. The algorithm further solves the problem that the relative position of the spacecraft and the set transfer orbit is not easy to express in path planning and realizes the simple representation of space avoidance motion. It avoids the premature trajectory deviation and local oscillation of the traditional artificial potential field method and realizes the autonomous avoidance of space debris.

It can consider avoidance safety, fuel consumption, minimum offset and braking aging factors at the same time to meet different avoidance needs and preferences. Compared with other algorithms, the average time is shortened by 47% and the distance is reduced by 11%, which further improves the shortcomings of conventional methods that are difficult to meet different avoidance preferences at the same time and that the orbit balance is weak.

7.1.4: The Real-Time Planning Method of Spacecraft Orbit Game Strategy is Studied

In the face of the situation handling needs of spacecraft approaching space targets, real-time planning is adopted to deal with non-cooperative goals, and the optimal strategy is quickly obtained, firstly, a sequential decision game model is constructed, and the real-time planning problem of spacecraft orbit game strategy is described. Secondly, the motion model between low-earth orbit spacecraft is constructed, and the Nash equilibrium strategy of orbital game is given, which translates the orbital game problem between orbital spacecraft into a differential countermeasure problem. Thirdly, a fuzzy inference model for continuous space solution is constructed to realize the mapping transformation of continuous state through fuzzy inference to continuous behavior output, which effectively avoids the problem of dimensionality disaster in continuous space in which traditional deep reinforcement learning copes with continuous space. Finally, a new branch deep reinforcement learning architecture is proposed, which realizes the branch training and sharing decision-making of behavior strategies and avoids the problem of the combination growth of the number of behaviors and mapping rules. The applicability and comparative advantages of the method are tested through case studies.

In view of the outstanding risk that the equilibrium strategy is difficult to obtain due to non-cooperative objectives, in the face of the problem of "continuous dynamic interaction" presented by the spacecraft orbit game, a game strategy solving algorithm based on branched deep reinforcement learning was proposed. The algorithm realizes the combination of optimal control and game theory, improves the learning ability of deep reinforcement learning on discrete behaviors, and further solves the problem that the differential countermeasure model is highly nonlinear and difficult to solve by classical optimal control theory, and can obtain the Nash equilibrium strategy of spacecraft orbit game in real time. Compared with other algorithms, the training time is reduced by 90% on average, and the growth rate of reward and punishment value is accelerated by 17% and is more stable, which further improves the shortcomings of the conventional method that it is difficult to deal with the bilateral control problem and the convergence domain of the equilibrium strategy is small.

7.1.5: The Design and Application of the On-orbit Service Mission Planning System Are Carried Out

Firstly, the foundation, requirements and key points of the construction of the mission planning system were studied based on the application requirements of the on-orbit servicing experiment and the ground simulation training application, and the application requirements of the on-orbit servicing mission planning system were clarified.

Then, the architecture of the mission planning system was designed, and the functional architecture of the mission planning system was constructed, which realized the systematization, modularization, easy operation, easy expansion and easy maintenance characteristics of the system. Finally, based on the case scenario, the on-orbit servicing task planning system is applied to test the feasibility and effectiveness of the intelligent planning method for on-orbit servicing tasks, meet the needs of on-orbit servicing task planning, and the visual simulation effect is good, which can provide strong support for on-orbit experiments or ground simulation training.

7.2: RESEARCH PROSPECTS

Throughout the ages, innovation is not easy, and on-orbit servicing-related research has caused a "high threshold" for related research due to its wide coverage, difficult theoretical innovation, and high scientific and technological content, which is daunting and even invisibly slows down the pace of the motherland's entry into outer space. In view of some problems in the planning of on-orbit servicing tasks, this book introduces the target allocation method of composite service mode, the orbit avoidance path planning method and the orbit game strategy planning method, but due to the limitation of knowledge reserve, we feel that there are still many problems worthy of further research and exploration.

7.2.1: From Simple to Complex, Continue to Conduct In-Depth Research on Planning Methods

"The Tao produced One; One produced Two; Two produced Three; Three produced All things." The development of any science and technology follows a process from simple to complex. This book studies the intelligent planning method for the on-orbit servicing task planning problem, which further solves the academic problems such as "nonlinear combinatorial optimization," "multi-limited shortest path" and "continuous dynamic interaction." However, the issue of on-orbit servicing mission planning will continue to expand in breadth and depth, from the current complete information, one-to-one, and single-mission planning problems to the future incomplete information, many-to-many, collaborative multi-mission planning and other issues, all of which clearly reflect the upward trend of mission planning with the continuous development of service means and aerospace technology. This forward-looking and upward trend of problem development is bound to have an important impact on the application of planning methods, and it is necessary to continue to study on the basis of existing research results to deal with more complex situations, such as the planning problem that spacecraft will face multiple attacks at the same time;

consider more factors and conditions, such as many-to-one or many-to-many spacecraft collaborative service issues. In the next step, we can take the existing research results as the cornerstone, combined with the diverse needs of the actual planning tasks in the future, and continue to conduct in-depth research, and it is believed that the research on the planning of on-orbit servicing missions will continue to develop and improve step by step along this spiral road.

7.2.2: Promote the Coordinated Development of Related Technologies

In the research on the intelligent planning method of on-orbit servicing mission, the knowledge of celestial mechanics and orbit dynamics in the field of aerospace, the mathematical modeling and optimization methods in the field of operations research, and the simulation and software development technologies in the computer field are used. This situation not only determines that the research on intelligent planning of on-orbit servicing missions needs to constantly absorb nutrients from the new achievements of aerospace science, operations research and computer science, and transform those things that can be used directly or through transplantation and transformation into their own content; moreover, it also determines the new achievements in the research of intelligent planning methods for on-orbit servicing tasks, which can also be selected and used by other sciences. It can be seen that continuing to carry out in-depth research on intelligent planning of on-orbit servicing missions is not only a matter in the aerospace field but can also promote the coordinated development of multiple disciplines from point to area, and produce more extensive impact and benefits.

7.2.3: From Computing to Cognition, Explore the Road of Intelligent Development

On-orbit servicing task planning is a typical kind of knowledge processing process, which will involve a variety of constraints and numerous reasoning decisions, which are essentially suitable for solving problems by artificial intelligence technology. Although this book has carried out in-depth research and exploration on the intelligent planning method of on-orbit servicing missions, the proposed intelligent planning method mainly uses intelligent methods to solve several on-orbit servicing mission planning problems, which is an attempt to research intelligent and unmanned aerospace technology in the future. In the future, with the continuous breakthroughs in emerging technologies such as quantum computing, brain-like intelligence and brain-inspired chips, intelligent planning technology is bound to make significant progress in the direction of perceptual intelligence and cognitive intelligence, which will greatly promote the innovation and development of the academic field of intelligent planning for space missions, and will lay the foundation for further research.

References

[1] Yang Leping, On Space[M]. Beijing: World Knowledge Press, 2017.
[2] Clausewitz, On War[M]. Beijing: Military Science Press, 2004.
[3] Wu Guanqun. Research on rendezvous trajectory optimization and close proximity control of on-orbit servicing spacecraft[D]. Harbin: Harbin Institute of Technology, 2019.
[4] Hu Xiaofeng, Rong Ming, On several issues in joint combat planning system[M]. Journal of Command and Control, 2017, 3(4): 273-280.
[5] Xing Qinghua, Fan Haixiong. Anti-missile mission planning technology: case-based reasoning[M]. Beijing: Science Press, 2020.
[6] Dana Nav, Malik G, Paoio T. Automatic Programming: Theory and Practice[M]. Jiang Yunfei, Yang Qiang, Ling Yingbiao, et al. Trans. Beijing: Tsinghua University Press, 2008.
[7] Xu Weidi. Development of space debris removal capabilities, one of the current priorities of China's space industry: a brief discussion on space environment governance[J]. Space Debris Research, 2020, 20(1): 10-16.
[8] Shen L, Chen L, Yan N, et al. Development of space debris removal capabilities, one of the current priorities of China's space industry: a brief discussion on space environment governance[J]. Space Debris Research, 2020, 20(2): 1-6.
[9] Li Ming, Gong Zizheng, Liu Guoqing. Development of cutting-edge technology and system for space debris monitoring and removal[J]. Chinese Science Bulletin, 2018, 63(25): 2570-2591.
[10] https://www.nasa.gov/mission_ pages/station/news/orbital_debris.html.
[11] Liu Huawei, Liu Yongjian, Tan Chunlin, et al. Analysis and suggestions on key technologies for space debris removal[J]. Spacecraft Engineering, 2017, 26(2): 105-113.
[12] Wang Xueyao. The latest development of foreign on-orbit service systems (Part 1)[J]. Space International, 2017(10): 24-29.
[13] Su Fei, Liu Jing, Zhang Yao, et al. Optimal pulse analysis for spacecraft in-plane maneuver avoidance[J]. Systems Engineering and Electronics, 2018, 40(12): 2782-2789.
[14] Chen Haipeng. On-orbit refueling: let the spacecraft rejuvenate[N]. China Aerospace News, 2016-07-02(1).

[15] LU YU, WANG XIAOWEI, SACCOCCIA GIORGIO, ET AL. Conceptual study of on-orbit refueling station [J]. Missile and Space Launch Technology, 2015(1): 1-7.
[16] Tang Qiong, Yan Ning, Zhang Feng. Research on the configuration of foreign in-orbit fuel filling stations[J]. Space International, 2017(11): 60-64.
[17] Jia Ping A brief analysis of the development of foreign on-orbit assembly technology [J]. International Space, 2016(12): 61-64.
[18] Wang Xueyao. The latest development of foreign on-orbit servicing systems (II)[J]. Space International, 2017(11): 65-69.
[19] Huang Xuxing, Li Shuang, Sun Pan, et al. Survey on the application of artificial intelligence in spacecraft guidance and control[J]. Acta Aeronautical News. https://kns.cnki.net/kcms/detail/11.1929.V.2020-0708.1512.004.html.
[20] AMBROSE R, WILCOX B, REED B, et al. Robotics, tele-robotics and autonomous systems roadmap [R]. NASA, Technology Area 04, 2012.
[21] Chen Meng, Xiao Yuzhi, Zhang Tao. Artificial intelligence technology in space service and control[J]. Manned Spaceflight, 2018, 24(3): 285-291.
[22] HOLDREN J P, BRUCE A, FELTEN E, et al. The national artificial intelligence research and development strategic plan [R]. National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, 2016.
[23] TETHERS. SL awards TUI contract for in-space manufacturing demo [EB/OL]. http://www.tethers.com/index.html.
[24] TETHERS. DARPA awards TUI/firmamentum contract to develop “constructable” persistent GEO platform [EB/OL]. http://www.tethers.com/index.html.
[25] WEISMULLER T, LEINZ M. GNC demonstrated by the orbital express autonomous rendezvous and capture sensor system [C]. In: Proceedings of the29th Annual AAS Guidance and Control Conference. 2006: 6-16.
[26] Meng Bo, Huang Jianbin, Li Zhi, et al. The flight mission and implications of the US high-orbit approach operation satellite MiTEx [J]. Spacecraft Engineering, 2014(23): 112-118.
[27] MITCHELL I T, GORTON T B, TASKOV K, et al. GNC development of the XSS-11 micro-satellite for autonomous rendezvous and proximity operations [C]. In: Proceedings of the 29th Annual AAS Guidance and Control Conference, 2006: 6-14.

[28] RUMFORD T. Demonstration of autonomous rendezvous technology (DART) project summary [C]. In: Proceedings Volume 5088, Space Systems Technology and Operations. Orlando, 2003: 10-19.
[29] Ouyang Qi, Zhao Yong, Chen Xiaoqian. On-orbit servicing mission planning for coplanar circular orbit spacecraft[J]. China Space Science and Technology, 2010, 30(1): 34-40.
[30] Yuan Li, Huang Huang. Current status and development of intelligent autonomous control technology for spacecraft[J]. Space Control Technology and Application, 2019, 45(4): 7-18.
[31] Xu Yingjie. Research on mission planning method of on-orbit servicing spacecraft for satellite maintenance[D]. Changsha: National University of Defense Technology, 2018.
[32] Li Yongjun, Huang Zhuo, Guo Bo. A review of weapon-target assignment problem[J]. Ordnance Automation, 2009, 28(11): 1-4, 9.
[33] Cao Qiying, He Zhangbing. Research on genetic algorithm for WTA problem[J]. Control Theory and Applications, 2001, 18(1): 76-79.
[34] Chen Zhen. Research on satellite collision probability estimation and avoidance strategy design[D]. Harbin: Harbin Institute of Technology, 2020.
[35] CHEN L, ZHOU B Z, HAN L. An analytic method of collision detection for active spacecrafts [C]. The 55th International Astronautical Congress, Vancouver, Canada, 2004.
[36] Bai Xianzong. Research on orbit prediction error and collision probability of space target[D]. Changsha: National University of Defense Technology, 2013.
[37] LAPORTE F, SASOT E. Operational management of collision risks for LEO satellites at CNES [C]. AIAA 2008-3409. SpaceOps 2008 Conference, Heidelberg, Germany, 2008, 5.
[38] AIDA S, KIRSCHNER M, WERMUTH M, et al. Collision avoidance operations for LEO satellites controlled by GSOC [C]. AIAA 2010-2298. SpaceOp s2010 Conference, Huntsville, Alabama, 2010, 4.
[39] MATSUDA I, HIROSE C, KUDO N. The JAXA conjunction assessment process [C]. AIAA 2010-2039. SpaceOps 2010 Conference, Huntsville, Alabama, 2010, 4.
[40] WEISMULLER T, LEINZ M. GNC demonstrated by the orbital express autonomous rendezvous and capture sensor system [C]. In: Proceedings of the29th Annual AAS Guidance and Control Conference. 2006: 6-16.

[41] DESOUKY S F, SCHWARTZ H M. Self- learning fuzzy logic controllers for pursuit-evasion differential games [J]. Robotics and Autonomous Systems, 2011(59): 22-33.
[42] Zhu L, Fang Shengliang, Sun Xiaojun, et al. Analysis and simulation of satellite data link efficiency under transponder jamming[J]. Electronic Information Countermeasure Technology, 2014, 29(3): 29-32.
[43] RICHARD A P. Modern Communications Jamming Principles and Techniques [M]. New York: Artech House Inc, 2011: 490-491.
[44] HESSEL M, MODAYIL J, VAN H H, et al. Rainbow: Combining improvements in deep reinforcement learning [J]. Association for the Advancement of Artificial Intelligence, 2017, 10(6): 3215-3222.
[45] Wang Yuejiao, Ma Zhong, Yang Yidai, et al. Satellite attitude control method based on deep reinforcement learning[J/OL]. China Space Science and Technology: 1-7 [2019-09-21]. http://kns.cnki.net/kcms/detail/11.1859.V.20190428.1714.020.html.
[46] Luo Haibo, Xu Lingyun, Hui Bin, et al. Research status and prospect of object tracking method based on deep learning[J]. Infrared and Laser Engineering, 2017, 46(5); 14-20.
[47] Liu Zhao, Xu Ke. Dynamic channel resource allocation algorithm for multi-beam satellite[J]. Mobile Communications, 2019, 43(5): 27-32.
[48] Wang Wei, Cheng Shuchang, Zhang Yuzhi. Research on a class of weapons-target assignment method based on genetic algorithm[J]. Systems Engineering and Electronic Technology, 2008, 30(9): 1708-1711.
[49] Tian Wen, Yang Fan, Yin Jianan, et al. Multi-objective optimization method for route spatiotemporal resource allocation[J]. Journal of Traffic and Transportation Engineering, 2020, 20(6): 218-226.
[50] Yang Shanliang, Huang Jian, Liu Yang, et al. Research on joint firepower WTA problem based on genetic algorithm[J]. Computer Simulation, 2012, 29(3): 61-64.
[51] Chang Tianqing, Bai Fan, Wang Qinzhao. Niche genetic algorithm for solving the weapon-target allocation problem of tank detachment[J]. Journal of Armored Force Engineering College, 2012, 26(1): 44-49.
[52] Wang Lei, Ni Mingfang, Yang Kuansi, et al. Combinatorial chaotic genetic algorithm based on direct comparison strategy to solve weapon target allocation problem[J]. Journal of System Simulation, 2014, 26(1): 125-131.
[53] Chen Si, Hu Tao. Weapon-target assignment based on multi-objective optimization genetic algorithm[J]. Ship Electronic Engineering, 2015(7): 54-57.
[54] Liu Zhen, Liu Wenbiao. Dynamic decision model and solution algorithm for island-ship integrated coordinated air defense[J]. Command, Control and Simulation, 2019, 41(1): 26-31.
[55] Yan Yuduo. Application of improved genetic algorithm to solve the problem of weapon target allocation[J]. Digital technology and application,

2018, 36(10): 138-141.
[56] Huang Guorui, Li Penghui, Ding Junxiang, et al. Weapon-target allocation model and solution algorithm based on fuzzy optimization theory[J]. Firepower and Command Control, 2013, 38(1): 56-60.
[57] Wu Congmeng, Wang Gongbao. Research on the application of genetic-ant colony algorithm in target allocation problem[J]. Ordnance Industry Automation, 2014, 33(4): 8-12.
[58] Yuan Mei. Solving WTA problem based on ant colony algorithm[D]. Harbin: Harbin Institute of Technology, 2007.
[59] Xiong Yu. Research on the application of improved ant colony algorithm in weapon target allocation[J]. Computer and Digital Engineering, 2014, 42(3): 399-402.
[60] Cui Lili. Research and application of improved ant colony algorithm in WTA problem[D]. Shanghai: Shanghai Jiao Tong University, 2011.
[61] Su Miao, Qian Hai, Wang Xufa. WTA problem solving based on ant colony algorithm based on immune memory[J]. Computer Engineering, 2008, 34(4): 215-217.
[62] Gao Shang, Yang Jingyu. Particle swarm optimization algorithm for weapon-target assignment problem[J]. Systems Engineering and Electronic Technology, 2005, 27(7): 1250-1252.
[63] Yang Qisong, Wang Shunhong, Wang Ranhui, et al. Modeling and decision-making of phased ground strike weapon-target allocation[J]. Journal of Ballistics, 2017, 29(2): 90-96.
[64] CHEN H, LIU Z, SUN Y, et al. Particle Swarm Optimization Based on Genetic Operators for Sensor-Weapon-Target Assignment [C] // Fifth International Symposium on Computational Intelligence & Design. IEEE Computer Society, 2012.
[65] Li Xinran, Fan Yongsheng. Improved particle swarm optimization for solving weapon target allocation problem[J]. Firepower and Command Control, 2014, 39(12): 58-61.
[66] Zhang Jiao, Wang Zhongxu, Chen Li, et al. Research on modeling and optimization method of air defense firepower distribution with multiple interception timing[J]. Acta Armamentarii, 2014, 35(10): 1644-1650.
[67] Fan Chengli, Xing Qinghua, Zheng Mingfa, et al. Weapon-target allocation optimization algorithm based on IDPSO[J]. Industrial Systems Engineering and Electronic Technology, 2015, 37(2): 336-342.
[68] Liu Shuangying, Han Xie. A quantum particle swarm algorithm for solving weapon-target allocation problem[J]. Computer Science, 2013, 40(2): 235-237.
[69] NG A, COATES A, DIEL M, et al. Autonomous inverted helicopter flight via reinforcement learning [C]. International Symposium on Experimental Robotic, 2006: 799-806.

[70] Zhu Xiaoyu, Qiao Bing, Zhang Qingzhan, et al. A round-trip on-orbit refueling task scheduling model and optimization algorithm based on fuel station[J]. Engineering Science and Technology, 2017, 49(S2): 186-194.
[71] Xiao Hai, Liu Xinxue, Shu Jiansheng, et al. Research on target allocation problem of multiple on-orbit servicing spacecraft[J]. Computer Simulation, 2017, 34(1): 90-93, 128.
[72] ZHOU H, YAN Y, HUANG X, et al. Multi- objective planning of a multiple geostationary spacecraft refuelling mission [J]. Engineering Optimization, 2017, 49(3): 531-548.
[73] Liu Bingyan, Ye Xiongbing, Zhou Chifei, et al. Resource allocation of on-orbit servicing based on improved DQN[J]. Acta Aeronautics and Astronautics, 2020, 41(5): 261-269.
[74] Tan Yinglong, Qiao Bing, Zhu Xiaoyu, et al. A spacecraft on-orbit refueling operation scheduling with fuel consumption as the optimization target[J]. Manned Spacecraft, 2018, 24(2): 143-149.
[75] Tao Sunjie, Song Zhu. A resource scheduling algorithm for station network with integrated measurement and control data transmission[J]. Telecommunication Technology, 2018, 58(7): 760-767.
[76] Li Xiamiao, Chen Xinjiang, Wu Guohua, et al. Relay satellite scheduling model and heuristic algorithm considering breakpoint resumable transmission[J/OL]. Acta Aeronautica Sinica:1-15 [2019-09-21]. http://kns.cnki.net/kcms/detail/11.1929.V.20190812.1640.002.html.
[77] Ma Shuo, Ma Yaping. Cluster UUV target allocation method based on hierarchical cluster auction[J]. Engineering and Ship Science and Technology, 2019, 41(9): 70-75.
[78] Li Yuan. Research on resource allocation and formation trajectory optimization method of multi-UAV cooperative mission[D]. Changsha: National University of Defense Technology, 2011.
[79] Zhang Huang. Research on obstacle avoidance navigation of unmanned vehicles based on multi-sensor data fusion[D]. Xi'an: Xi'an University of Technology, 2015.
[80] LIKHACHEV M, FERGUSON D. Planning long dynamically feasible maneuvers for autonomous vehicles [J]. The International Journal of Robotics Research, 2009, 28(8): 933-945.
[81] Wang Fengbo, Cui Xianguo, Ding Lin, et al. Research and implementation of driving navigation path based on Dijkstra algorithm[J]. Western China Science and Technology, 2011, 10(34): 26-28.
[82] Chen Yalin, Zhuang Liyang, Zhu Longbiao, et al. Research on path planning of parking system based on improved Dijkstra algorithm[J]. Modern Manufacturing Engineering, 2017, 2(8): 63-67.
[83] Ma Jing, Wang Jiabin, Zhang Xue. Application of A* algorithm in path planning of unmanned vehicles[J]. Computer Technology and Development, 2016, 26(11): 153-156.
[84] Xin Huang. Research on obstacle detection, prediction and collision avoidance method of unmanned vehicle moving obstacles[D]. Hefei:

University of Science and Technology of China, 2014.
[85] GUTJAHR B, LUTZ G, WERLING M. Lateral vehicle trajectory optimization using constrained linear time-varying MPC [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(6): 1586-1595.
[86] Wu Lingyu, Bai Guozhen, Guan Dengshi. Trajectory planning for automatic parking[J]. Agricultural Equipment and Vehicle Engineering, 2017, 55(9): 85-89.
[87] Chen Cheng, He Yuqing, Bu Chunguang, et al. Feasible trajectory planning of unmanned vehicle based on fourth-order Bezier curve[J]. Acta Automatica Sinica, 2015, 41(3): 486-496.
[88] Yu Zhuoping, Li Yishan, Xiong Lu. A review of motion planning algorithms for unmanned vehicles[J]. Journal of Tongji University (Natural Science Edition), 2017, 45(8): 1150-1159.
[89] Feng Laichun. Research on motion planning algorithm of parametric RRT unmanned vehicle based on guidance domain[D]. Hefei: University of Science and Technology of China, 2017.
[90] Ding Jiaru, Du Changping, Zhao Yao, et al. UAV path planning algorithm based on improved artificial potential field method[J]. Journal of Computer Applications, 2016, 36(1): 287-290.
[91] HE Renke, WEI Ruixuan, ZHANG Qirui, et al. Aircraft route planning method based on mimic electric potential energy[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1543-1549.
[92] Yang Lichun, Gu Yingyan, Bai Yu. Unmanned aerial vehicle (UAV) online route planning algorithm based on improved artificial potential field method[J]. Application of Electronic Technology, 2018, 44(4): 5-9, 13.
[93] Shang Pu. Research on path planning and autonomous obstacle avoidance control system of rotor UAV[D]. Xi'an: Xi'an University of Science and Technology, 2019.
[94] Cheng Zhi, Zhang Zhi'an, Li Jinzhi, et al. Path planning of mobile robot with improved artificial potential field method[J/0L]. Computer Engineering and Applications: 1-6 [2019-11-26]. http://kns.cnki.net/kcms/detail/11.2127.tp.2019-0624.1809.022.html.
[95] Jia Zhengrong, Wang Hangyu, Lu Faxing. Improved APF circulation path planning based on obstacle convexity[J]. Acta Aeronautica Sinica, 2019, 40(10): 254-268.
[96] Xu Xiaoqiang, Wang Mingyong, Mao Yan. Path Planning of Mobile Robot Based on Improved Artificial Potential Field Method[J/OL]. Computer Applications. https://kns.cnki.net/kcms/detail/51.1307.TP.20200731.0944.002.html.
[97] HOWARD T, KELLY A. Optimal rough terrain trajectory generation for wheeled mobile robots [J]. The International Journal of Robotics Research, 2007, 26(2): 141-166.

[98] MONTEMERLO M, BECKER J, BHAT S, et al. Junior: The Stanford entry in the Urban Challenge [J]. Journal of Field Robotics, 2008, 25(9): 569-597 (in Chinese).
[99] WERLING M, KAMMEL S, ZIEGLER J, et al. Optimal trajectories for time-critical street scenarios using discretized terminal manifolds [J]. The International Journal of Robotics Research, 2012, 31(3): 346-359.
[100] Wang Wei, Chen Huiyan, Ma Jianhao, et al. Intelligent vehicle path tracking based on Frenet coordinate system and control delay compensation[J]. Acta Armamentarii, 2019, 40(11): 2336-2351.
[101] Long Xiang, Gao Jianbo, Kui Hanbing. Development and test verification of an autonomous driving vehicle system architecture[J]. Journal of Chongqing University of Technology (Natural Science), 2019, 33(12): 45-54.
[102] Wang Shajing. Research on trajectory planning algorithm for autonomous driving based on Frenet coordinate system sampling[D]. Lanzhou: Lanzhou University of Technology, 2019.
[103] Zhao Ningning, Xu Demin, Gao Jian, et al. Multi-AUV formation path tracking control based on Serret-Frenet coordinate system[J]. Torpedo Technology, 2015, 23(1): 35-39.
[104] DESOUKY S, SCHWARTZ H. Q(λ)-learning fuzzy logic controller for a multi-robot system [J]. In IEEE International Conference on Systems, Man and Cybernetics, 2010(10): 4075-4080.
[105] Yan Yongsuo. Path tracking control of unmanned surface vehicle based on feedback-sliding mode strategy[D]. Harbi: Harbin Engineering University, 2019.
[106] Guo Yanning, Li Chuanjiang, Ma Guangfu. Autonomous attitude maneuver control of spacecraft based on potential function method[J]. Acta Aeronautics and Astronautics, 2011, 32(3): 457-464.
[107] Qian Y, Xu M, Guo D, et al. Research on short-term avoidance path planning for spacecraft[J]. Flight Mechanics, 2010, 28(5): 67-70.
[108] Yao Dangnai, Wang Zhenguo. Autonomous avoidance strategy for spacecraft on-orbit collision avoidance[J]. Journal of National University of Defense Technology, 2012, 34(6): 100-103, 110.
[109] SANG CHERL LEE, HAE DONG KIM. Collision avoidance maneuver planning using GA for LEO and GEO satellite maintained in keeping area [J]. International J. of Aeronautical & Space Sci, 2012, 13(4): 474-483.
[110] Zhang Jingrui, Luo Yang, Zhang Yao. Attitude maneuver control of constrained spacecraft based on improved potential function method[J]. Journal of Engineering and Tsinghua University (Natural Science Edition), 2013, 53(8): 1160-1165.
[111] LEE U, MESBAHI M. Feedback control for spacecraft reorientation under attitude constraints via convex potentials [J]. Aerospace and Electronic Systems, IEEE Transactions on,

2014, 50(4): 2578-2592.
[112] Miao Yuanming, Pan Teng. Spacecraft attitude maneuver path planning to avoid attitude forbidden zone[J]. Spacecraft Engineering, 2015, 24(4): 33-37.
[113] HU Q, DONG H, ZHANG Y, et al. Tracking control of spacecraft formation flying with collision avoidance [J]. Aerospace Science and Technology, 2015(42): 353-364.
[114] Yu Dateng, Wang Hua, Sun Fuyu. Optimal avoidance maneuver strategy for spacecraft considering potential threat area[J]. Acta Aeronautical News, 2017, 38(1): 286-294.
[115] DONG H, HU Q, AKELLA M R. Safety control for spacecraft autonomous rendezvous and docking under motion constraints [J]. Journal of Guidance, Control, and Dynamics, 2017, 40(7): 1-14.
[116] Su Fei, Liu Jing, Zhang Yao, et al. In-plane maneuver avoidance strategy based on optimal pulse [J]. Space Debris Research, 2018, 18(3): 34-41.
[117] Fan Shipeng, Wu Guang, Wang Liang, et al. Aircraft trajectory planning based on improved artificial potential field method[J]. Aerospace Control, 2018, 36(1): 50-54.
[118] SHEN Q, YUE C, GOH C H, et al. Rigid-body attitude stabilization with attitude and angular rate constraints [J]. Automatica, 2018(90): 157-163.
[119] Li Haohao, Zhang Jin, Luo Yazhong. Active avoidance strategy of spacecraft based on maneuvering target filter estimation[J]. Chinese Journal of Theoretical and Applied Mechanics, 2020, 52(6): 1560-1568.
[120] Yun Chaoming, Hu Min, Song Qinglei, et al. A review of the safety research of giant low-orbit constellation and its evasion maneuver strategy[J]. Space Debris Research, 2020, 20(3): 17-23.
[121] Wang Guogang, Dai Lu, Xu Kai, et al. Energy-optimal active collision avoidance control of small satellites under multiple constraints[J]. Space Control, 2020, 38(3): 17-24.
[122] CARMEN PARDINI, LUCIANO ANSELMO. Environmental sustainability of large satellite constellations in low earth orbit [J]. Acta Astronautica, 2020(170): 27-36.
[123] Qian Xingfang, Lin Ruixiong, Zhao Yanan. Missile Flight Mechanics[M]. Beijing: Beijing Institute of Technology Press, 2006.
[124] Li Chaoyong. Research on some problems of guidance and control of TBM interceptor[D]. Harbin: Harbin Institute of Technology, 2008.
[125] Liu Yanfang. Research on terminal guidance law of interceptor missile based on differential countermeasure theory[D]. Harbin:

Harbin Institute of Technology, 2014.
[126] Luo Yazhong, Li Zhenyu, Zhu Hai. A review of differential countermeasures for spacecraft orbital pursuit[J/OL]. Science China: Technological Sciences: 1-13 [2020-07-26]. http://kns.cnki.net/kcms/detail/11.5844.TH.20200511.1021.002.html.
[127] Sun Songtao. Research on the pursuit and escape countermeasures and numerical solution methods of two spacecraft in low earth orbit[D]. Harbin: Harbin Institute of Technology, 2015.
[128] Zhang Hao, Zhang Yiqun, Zhang Pengfei. Guidance law of differential countermeasures for intercepting active defense targets[J/OL]. Systems Engineering and Electronic Technology: 1-14 [2021-02-03]. http://kns.cnki.net/kcms/detail/11.2422.TN.20210112.1311.004.html.
[129] Guo Zhiqiang, Sun Qilong, Zhou Shaolei, et al. Guidance law of norm-type differential countermeasures for active defense aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1787-1796.
[130] Guo Zhiqiang, Zhou Shaolei, Yu Yunzhi. Guidance law of norm-based cooperative differential countermeasures for intercepting maneuvering targets[J]. Computer Simulation, 2020, 37(3): 23-26.
[131] Li Yifu, Song Guibao, Jia Runa, et al. Firepower optimization model of multi-type anti-ship missiles based on differential game[J]. Firepower and Command and Control, 2019, 44(11): 121-126.
[132] Qu Xiangju. Application of direct multiple target method in trajectory optimization[J]. Flight Mechanics, 1992, 10(1): 1321.
[133] Wang Hua, Tang Guojin, Lei Yongjun. Research on direct target method for finite thrust trajectory optimization problem[J]. China Space Science and Technology, 2003, 23(5): 51-56.
[134] Feng Haoyang, Yue Xiaokui, Wang Xuechuan. A new method for solving the perturbation Lambert problem with large-scale convergence: pseudo-linearization-local variational iterative method[J]. Acta Aeronautica Sinica, 2021, 42(X): 324699.
[135] Peng Kun, Peng Rui, Huang Zhen, et al. Implicit target method for solving the optimal lunar soft landing orbit[J]. Acta Aeronautica Sinica, 2019, 40(7): 159-167.
[136] Zhang Qiuhua, Sun Songtao, Mou Ying, et al. Pursuit and escape strategy and numerical solution of two spacecraft with fixed time[J]. Journal of Astronautics, 2014, 35(5): 537-544.
[137] Peng Qiqing, Li Haiyang, Shen Hongxin, et al. Optimization of lunar fixed-point landing orbit based on the combination of Gauss pseudospectroscopy method and direct target method[J]. Journal of National University of Defense Technology, 2012, 34(2): 119-124.
[138] Liao Yihuan, Li Daokui, Tang Guojin. Research on motion planning of space manipulator based on hybrid programming strategy[J]. Journal of Astronautics, 2011, 32(1): 98-103.
[139] Zhang Zixiong. Research and simulation of low and low orbit flight planning method of low-thrust spacecraft[D]. Taiyuan: North University of China, 2020.

[140] Li Longyue, Liu Fuxian, Shi Xiangfeng, et al. Pursuit and escape countermeasure model and coordination point solution method in missile attack and defense confrontation[J]. Systems Engineering and Electronic Technology, 2016, 38(5): 1067-1073.
[141] Li Longyue, Liu Fuxian, Shi Xiangfeng, et al. Modeling and solving differential countermeasures for missile chase and escape game[J]. Systems Engineering - Theory & Practice, 2016, 36(8): 2161-2168.
[142] SHIMA T. Optimal cooperative pursuit and evasion strategies against a homing missile [J]. Journal of Guidance, Control and Dymamics, 2011, 34(2): 414-425.
[143] Yu Jianglong, Dong Xiwang, Li Qingdong, et al. Cooperative guidance method for intercepting maneuvering targets based on differential countermeasures[J]. Journal of Command and Control, 2020, 6(3): 217-222.
[144] Zhao Jisong. Research on sparsity of local coordination method for solving trajectory optimization problems[J]. Journal of Astronautics, 2017, 38(12): 1263-1272.
[145] Xue Guohao, Chen Qiaoyu, Xu Yinsheng, et al. Research on obstacle avoidance path planning of intelligent vehicles based on direct matching point method[J]. Mechanics & Electronics, 2020, 38(8): 7-11, 16.
[146] Zhu Yingxuan, Zhao Kegang, Ye Jie, et al. Research on real-time trajectory tracking of intelligent vehicles based on matching point method[J]. Machine Tool & Hydraulics, 2020, 48(11): 150-154.
[147] RAUWOLF G A, COVERSTONE C. Near-optimal low-thrust orbit transfers generated by a genetic algorithm [J]. Journal of Spacecraft and Rockets, 1996, 33(6): 859-862.
[148] COVERSTONE CARROLL, VICTORIA. Near-optimal low-thrust trajectories via micro-genetic algorithms [J]. Journal of Guidance, Control, and Dynamics, 1971, 20(1): 196-198.
[149] LARRY D D, MENOM K K. Low-thrust orbit transfer optimization using genetic search [C]. AIAA Paper, 1999: 1999-4151.
[150] CHEN G, WAN Z, HU Y, et al. Genetic algorithm optimization of RLV reentry trajectory [J]. Systems Engineering and Electronics, 2006, 28(8): 1240-1243.
[151] KUMAR G N, PENCHALAIAH D, SARKAR A K, et al. Hypersonic boost glide vehicle trajectory optimization using genetic algorithm [J]. IFAC-Papers OnLine, 2018, 51(1): 118-123.
[152] Zeng Guoqiang, Du Xiaoning, Ren Xuan. Research on optimal small-thrust orbit change of lunar satellite[J]. Acta Astronomical Sinica, 2000, 41(3): 289-299.
[153] Wang X, Li J, Cui N, et al. Genetic algorithm optimization of soft landing orbit of lunar landing vehicle[J]. Journal of Tsinghua University

(Natural Science Edition), 2003, 43(8): 1056-1059.
[154] QINGZHEN Z, CUNJIA L, BO Y, et al. Reentry trajectory planning optimization based on ant colony algorithm [J]. IEEE International Conference Robotics and Biomimetics, 2008.
[155] Xie Lei, Zhang Hongbo, Zhou Xiang, et al. Trajectory planning of a sub-stage of launch vehicle based on combination algorithm[J]. Control and Information Technology, 2019(4): 79-84.
[156] OZAN, TEKINALP, MUGE, et al. Simulated annealing for missile optimization: developing method and formulation techniques [J]. Journal of Guidance, Control, and Dynamics, 2004.
[157] Wang Yin, Yao Keming, Lu Yuping. Re-entry trajectory design based on dynamic optimization of neural network[J]. Systems Engineering and Electronics, 2015, 37(10): 2347-2351.
[158] Cheng Lin, Jiang Fanghua, Li Junfeng. A review of deep learning applications in aircraft dynamics and control[J]. Engineering Mechanics and Engineering, 2020, 2(3): 267-276.
[159] BARRON E N, EVANS C L, JENSEN R. Viscosity Solutions of Isaacs' Equations and Differential Games with Lipschitz Controls [J]. Journal of Differential Equations, 1984, 53(2): 213-233.
[160] TAKAGI T, SUGENO M. Fuzzy identification of systems and its applications to modelling ad control [J]. IEEE Transactions on Systems, Man and Cybernetics, 1985, 15: 116-132.
[161] Li Xuehui. Research on spacecraft orbit configuration and autonomous rendezvous and docking control method[D]. Harbin: Harbin Institute of Technology, 2018.
[162] JANG J S R, SUN C T. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence [M]. New Jersey: Prentice-Hall, Inc., 1997.
[163] Gao Peng, Luo Jianjun. Adaptive artificial potential function guidance for spacecraft to avoid dynamic obstacles[J]. China Space Science and Technology, 2012, 32(5): 1-8.
[164] Zeng Xiangxin. Research on path planning and control method of free floating space robot[D]. Harbin: Harbin Institute of Technology, 2018.
[165] Chang Yan, Chen Yun, Xian Yong, et al. Guidance method for differential countermeasures for space rendezvous of maneuvering target[J]. Journal of Astronautics, 2016, 37(7): 795-801.
[166] Liu Bingyan, Ye Xiongbing, Gao Yong, et al. Solving non-cooperative target pursuit-and-escape game strategy based on branched deep reinforcement learning[J]. Acta Aeronautica Sinica, 2020, 41(10): 348-358.

[167] FRIEDMAN A. Differential games [M]. Rhode Island: American Mathematical Society, 1974.
[168] DICKMANNS E, WELL K. Approximate solution of optimal control problems using third order hermite polynomial functions [C]. Optimization Techniques IFIP Technical Conference, Novosibirsk, July1-7, 1974.
[169] Wang Qiang, Ye Dong, Fan Ningjun, et al. Satellite terminal pursuit control method based on zero miss distance[J]. Transactions of Beijing Institute of Technology, 2016, 36(11): 1171-1176.
[170] MATIGNON L, LAURENT G J, Le Fort Piat. Independent reinforcement learners in cooperative Markov games: a survey regarding coordination problems [J]. The Knowledge Engineering Review, 2012, 27(1): 1-31.
[171] Zhu Hai, Luo Yazhong, Zhang Jin. Analytical construction method of spacecraft chase and escape differential countermeasure boundary grid[C]. The 5th National Conference on Aerospace Dynamics and Control, 2017:1.
[172] Zhu Hai. Optimal control strategy for spacecraft orbital pursuit based on differential countermeasures[D]. Changsha: National University of Defense Technology, 2017.
[173] Wu Qichang, Zhang Hongbo. Spacecraft pursuit strategy and numerical solution based on survival differential countermeasures[J]. Control and Information Technology, 2019(4): 39-43.
[174] Sun Songtao, Zhu Qiangjun, Song Bin. Semi-direct solution method for spacecraft chase and escape problem based on CW equation[J]. Shanghai Aerospace, 2019, 36(3): 69-75.
[175] SUN S T, ZHANG Q H. Numerical solution of a pursuit-evasion differential game involving two spacecraft in low earth orbit [J]. Journal of Industrial and Management Optimization, 2015, 11(4): 1127-1147.
[176] SUN S T, ZHANG Q H, CHEN Y. Numerical solution for a class of pursuit-evasion problem in low earth orbit [C]. 2013, 9th Asian Control Conference, Istanbul, Turkey.
[177] Hao Zhiwei, Sun Songtao, Zhang Qiuhua, et al. Application of semi-direct matching method in solving spacecraft chase and escape problem[J]. Journal of Astronautics, 2019, 40(6): 628-635.
[178] Zhao L, Zhou J, Liu Y, et al. Three-dimensional space "pursuit-escape-defense" tripartite differential countermeasure method[J]. Systems Engineering and Electronic Technology, 2019, 41(2): 322-335.
[179] Wang Chunbao, Ye Dong, Sun Zhaowei, et al. Adaptive game strategy for spacecraft terminal interception[J]. Journal of Astronautics, 2020, 41(3): 309-318.

[180] SCHWARTZ H M. Multi-agent machine learning: a reinforcement approach [M]. Canada: John Wiley & Sons, Inc, 2014.
[181] Wang Yuqi, Ning Guodong, Wang Xiaofeng, et al. Maneuvering penetration strategy of near space vehicle based on differential countermeasures[J]. Acta Aeronautical Sinica, 2020, 41(S2): 69-78.
[182] CHRISTOPHER J C H, Watkins, Peter Dayan. Qlearning [J]. Machine Learning, 1992(3): 279-292.
[183] Chen Yanni. Research on finite-time adaptive dynamic programming guidance based on differential countermeasures[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019.
[184] Cao Lei. Key technologies of intelligent game confrontation based on deep reinforcement learning[J]. Command Information System and Technology, 2019, 10(5): 1-7.
[185] CHENG Y, SUN Z J, HUANG Y X, et al. Fuzzy categorical deep reinforcement learning of a defensive game for an unmanned surface vessel [J]. International Journal o£ Fuzzy Systems, 2019, 21(2): 592-606.
[186] LIU B Y, YE X B, GAO Y, et al. Forward-looking imaginative planning framework combined with prioritized replay double DQN [C]. International Conference on Control, Automation and Robotics, 2019, 4: 336-341.
[187] Wu Xiaoguang, Liu Shaowei, Yang Lei, et al. Ramp gait control method for bipedal robot based on deep reinforcement learning[J/OL]. Acta Automatica Sinica:1-13[2020-02-281]. https://doi.org/10.16383/j.aas.cl90547.
[188] Zhang Shuqin, Wang Zhonggui, Ran Long Tunnel. Space rendezvous and docking measurement technology and engineering application[M]. Beijing: China Aerospace Press, 2005.
[189] Li Gefei. Rendezvous and docking orbit control planning[M]. Beijing: National Defense Industry Press, 2016.
[190] Luo Yazhong, Zhang Jin, Zhu Yue. Space station operation mission planning[M]. Beijing: National Defense Industry Press, 2020.
[191] Yuan Jianping, Hou Jianwen, An Xiaomin, et al. Spacecraft Countermeasures against Orbital Dynamics[M]. Beijing: China Aerospace Press, 2014.
[192] Blockley, S. Aerospace Science and Technology Publishing Engineering (5): Dynamics and Control[M]. Jiang Ju, Zhou Jianjiang, Han Chao, et al. Beijing: Beijing Institute of Technology Press, 2016.
[193] Li Hao, Fan Xiangyu, Jin Hongbin, et al. Research on UAV swarm combat mission planning based on swarm intelligence[M]. Beijing: National Defense Industry Press, 2019.
[194] Richard S Sutton, Andrew G Barto. Reinforcement Learning[M]. 2nd edition. Translated by Yu Kai et al. Beijing: Publishing House of Electronics Industry, 2019.

[195] Zhao Rui'an. Orbital design of space weapons[M]. 2nd edition. Yu Kai, et al. Beijing: China Aerospace Press, 2008.
[196] Liu Lin. Algorithm of satellite orbit mechanics[M]. Nanjing: Nanjing University Press, 2019.
[197] Cheng Guocai. Theory and method of optimal control of space vehicle[M]. Beijing: National Defense Industry Press, 1999.
[198] Yuan Jianping, Zhao Yushan, Tang Geshi. Spacecraft deep space flight orbit design[M]. Beijing: China Aerospace Press, 2014.
[199] Tang Jinguo, Luo Yazhong, Yong Enmi. Theory, method and application of spacecraft trajectory optimization[M]. Beijing: Science Press, 2012.
[200] Liu Bingyan, Yu Hongyuan, Ma Xinyi, et al. Research on maneuver planning method of space vehicle in coplanar multi-orbit[J]. Journal of Dynamics and Control, 2021, 19(4): 81-88.
[201] Zhang Hongbo. Theory and method of spacecraft orbital mechanics[M]. Beijing: National Defense Industry Press, 2015.
[202] Wang Daheng, Wang Ganchang, Yang Jiaqi. High Technology Dictionary[M]. Beijing: Tsinghua University Press, Science Press, 2000.
[203] Yuan Jianping, He Xingsuo. Dynamics of Spacecraft Orbital Maneuvers[M]. Beijing: China Aerospace Press, 2010.
[204] Chen Chao, Xu Yuanzi, Cheng Guangquan, et al. Modeling and optimization method of military mission planning[M]. Beijing: National Defense Industry Press, 2019.
[205] Nash, Shapley, Haisani, et al. Game Theory Classics[M]. Beijing: Chinese Renmin University Press, 2013.
[206] Meng Yazhe. Extension solution of direct/indirect hybrid method for spacecraft fuel consumption optimum[J]. Acta Aeronautica Sinica, 2017, 38(1): 264-285.
[207] ZHU HAO, TIAN HUI, CAI G B, et al. Uncertainty analysis and design optimization of hybrid rocket motor powered vehicle for suborbital flight [J]. Chinese Journal of Aeronautics, 2015, 28(3): 676-686.
[208] JEAN-MARC SALOTTI. Robust, affordable, semi-direct Mars mission [J]. Acta Astronautica, 2016, 127(2016): 235-248.
[209] Zhang Yulin. Theory and application of distributed satellite system[M]. Beijing: Science Press, 2008:28-32.
[210] WANG H Y, ZHAO G W, HUANG H, et al. Equilibrium and station-keeping efficiency of cross-track multi-satellite arrays using a micro- electromagnetic formation flight [J].

Proc. IMechE Part G: J Aerospace Engineering, 2014, 228(11): 1937-1953.
[211] Wang Chaoyu, Hou Xiaogeng. Control method and hovering burnup analysis of space debris near on-site hovering[J]. China Space Science and Technology, 2020, 40(1): 27-36.
[212] Chen Xiaogang, Wu Zhilin, Gao Yu. Design and optimization of space-based interception orbit based on energy[J]. Computer Simulation, 2012, 29(5): 59-63.
[213] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [C]. In International Conference on Learning Representations, 2016.
[214] Zhang Xianchao. Deep Learning[M]. Beijing: Science Press, 2019.
[215] WANG Z, SCHAUL T, HESSEL M, et al. Dueling network architectures for deep reinforcement learning [J]. Association for the Advancement of Artificial Intelligence, 2016, 4(5): 1998-2003.
[216] Feng Chao. Reinforcement learning essence core algorithm and TensorFlow implementation[M]. Beijing: Publishing House of Electronics Industry, 2018.
[217] VAN H H, GUEZ A, SILVER D. Deep reinforcement learning with Double Q-learning [C]. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016, 2094-2100.
[218] RICHARD S S, ANDREW G B. Reinforcement Learning: An Introduction [M]. Second edition. London: The MIT Press, 2018.
[219] BA J, MNIH V, KAVUKCUOGLU K. Multiple object recognition with visual attention [C]. In ICLR, 2015.
[220] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518(7540): 529-533.
[221] WANG L X. A Course in Fuzzy Systems and Control [M]. New Jersey: Prentice-Hall, Inc., 1997.
[222] Ding Fan, Han Wei. Review of optimization algorithm for weapon target allocation problem[C]. Aviation Equipment Service Guarantee and Maintenance Technology Forum and Annual Conference of China Aviation Industry Technology and Equipment Engineering Association, 2019.
[223] Li Jinzhe, Wang Lei. Research on mathematical modeling and scheduling algorithm of spacecraft distributed intelligent computing system[J]. Space Control Technology and Application, 2019, 45(6): 38-46.
[224] Gu Zhongshun. Research on modeling and optimization technology of dynamic scheduling problem of relay satellite[D]. Changsha:

National University of Defense Technology, 2008.
[225] SHEN H. Optimal scheduling for satellite refuelling in circular orbits [D]. Georgia: Georgia Institute of Technology, 2003.
[226] LOXTON R, TEO K L, REHBOCK V, et al. Optimal Control Problems with a Continuous Inequality Constraint on the State and the Control [J]. Automatica, 2009, 45(10): 2250-2257.
[227] Zhu Xiaoyu. Scheduling algorithm for in-orbit refueling service of circular orbit spacecraft based on space fuel station[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017.
[228] Yu Jing. Research on spacecraft on-orbit servicing mission planning technology[D]. Changsha: National University of Defense Technology, 2015.
[229] Zhao L, Wang Shuo, Hao Y, et al. In-orbit mission planning of agile remote sensing satellites based on energy optimization[J]. Journal of Aeronautics and Astronautics, 2017, 38(6): 207-225.
[230] Liu Bingyan, Ye Xiongbing, Fang Shengliang, et al. Autonomous planning of on-orbit avoidance path based on Frenet and improved artificial potential field[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(4): 731-741.
[231] ENGWERDA J. Algorithms for computing Nash equilibria indeterministic LQ games[J]. Computational Management Science, 2007, 4(2): 113-140.
[232] DAI X, LI C, RAD A. An approach to tune fuzzy controllers based on reinforcement learning for autonomous vehicle control [J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(3): 285-293.
[233] Yu Da Teng. Research on spacecraft anti-rendezvous evasion maneuver method[D]. Changsha: National University of Defense Technology, 2013.
[234] JURE BAJIC, MARU AN HERCEG, IVAN RESETAR, et al. Trajectory planning for autonomous vehicle using digital map [J]. 2019 Zooming Innovation in Consumer Technologies Conference (ZINC), 2019, 7.25.
[235] FASSBENDER D, HEINRICH B C, WUENSCHE H T. Motion Planning for Autonomous Vehicles in Highly Constrained Urban Environments [J]. International Conference on Intelligent Robots and Systems (IROS), 2016, 4708-4713.
[236] DINEEN S. Multivariate calculus and geometry [M]. New York: Springer, 2001.
[237] Zhang Qing. Spatial characteristics of highway geometry under Frenet framework[D]. Changsha:

Central South University, 2014.
[238] ROSS T J. Fuzzy Logic with Engineering Applications [M]. The United States of America: John Wiley & Sons, Ltd, 2010.
[239] WERLING M, ZIEGLER J, SOREN K, et al. Optimal trajectory generation for dynamic street scenarios in a Frenet Frame [J]. IEEE International Conference on Robotics & Automation, 2010, 4(5): 5-6.
[240] Zhang Yu. Research on hybrid motion planning of autonomous vehicles[D]. Beijing: Beijing Institute of Technology, 2018.
[241] Wang Shajing. CHEN Jianye. Research on motion planning of intelligent vehicles based on Frenet coordinate system[J]. Mobile Power and Vehicle, 2019(1); 22-29.
[242] Zhang Xiangyu. Research on optimal control of spacecraft orbit under specific thrust direction constraints[D]. Harbin: Harbin Institute of Technology, 2016.
[243] Cui Hongzheng, Liu Wenling, Tang Geshi, et al. Orbital maneuver detection of non-cooperative space targets under different thrusts[J]. Journal of Astronautics, 2016, 37(3): 253-261.
[244] Zhang Chen, Zhao Yushan. Design method of hybrid propulsion fuel-saving track[J]. Journal of Astronautics, 2015, 36(8): 869-876.
[245] Zhou Ting. Research on joint maneuvering and fuel optimization in the approach segment of space rendezvous[D]. Beijing: Tsinghua University, 2010.
[246] Liu Bingyan, Ye Xiongbing, Wang Xinbo, et al. Path avoidance algorithm for unmanned ground vehicle based on improved artificial potential field[J]. Chinese Journal of Inertial Technology, 2020, 28(6): 769-777.
[247] KHATIB O. Real-time obstacle avoidance for manipulators and mobile robots [J]. The international journal of robotics research, 1986, 5(1): 90-98.
[248] Fan Shipeng, Qi Qi, Lu Kunfeng, et al. Autonomous obstacle avoidance technology of cruise missile based on improved artificial potential field method[J]. Transactions of Beijing institute of Technology, 2018, 38(8): 828-834.
[249] Wang Wenbin, Qin Xiaolin, Zhang Lige, et al. Dynamic track planning of UAV based on rolling time domain[J]. Journal of Intelligent Systems, 2018, 13(4): 524-533.
[250] HUANG Z C, WU Q, MA J, et al. An APF and MPC combined collaborative driving controller using vehicular communication technologies [J]. Chaos, Solitons & Fractals, 2016(89): 232-242.
[251] JURGEN FRANKE, WOLFGANG KARL HARDLE, HAFNER C M. Neural Networks and Deep Learning [M]. Verlag: Springer, Statistics of Financial Markets, 2019.

[252] TAKAHASHI A, HONGO T, NINOMIYA Y, et al. Local path planning and motion control for AGV in positioning [C]. The Autonomous Mobile Robots and Its Applications. IROS' 89. 1989, 392-397.
[253] DIJKSTRA E W. A note on two problems in connexion with graphs [J]. Numerische Mathematik, 1959, 1(1): 269-271.
[254] Wu Hongbo, Wang Yingjie, Yang Xiaoxiao. Urban traffic path analysis based on Dijkstra algorithm optimization[J]. Journal of Beijing Jiaotong University, 2019, 43(4): 116-121, 130.
[255] CHEN Y, HE Z, LI S L. Horizon-based lazy optimal RRT for fast, efficient replanning in dynamic environment [J]. Autonomous Robots, 2019, 43(8): 2271-2292.
[256] Zhang Jing, He You, Peng Yingning, et al. Cooperative game path planning based on neural network and artificial potential field[J]. Acta Aeronautics Sinica, 2019, 40(3): 322493.
[257] Yu Dayeng, Wang Hua, Zhou Wanmeng. Anti-rendezvous avoidance maneuver considering spatial geometric relationship[J]. Journal of National University of Defense Technology, 2016(38): 89-94.
[258] Zhang Weiying. Game and Society[M]. Beijing: Peking University Press, 2018.
[259] HAO Z W, SUN S T, ZHANG Q H, et al. Application of Semi-Direct Collocation method for solving pursuit-evasion problems of spacecraft [J]. Journal of Astronautics, 2019, 40(6): 628-635.
[260] ISAACS R. Differential games [M]. New York: Wiley, 1965.
[261] Ma Zhonggui. Game theory and its application in wireless communication network[M]. Beijing: National Defense Industry Press, 2015.
[262] Liu Bingyan, Ye Xiongbing, Yue Zhihong, et al. Continuous Space Chase and Escape Game Algorithm Based on Multiple Parallel Deep Q Networks[J]. Acta Armamentarii, 2021, 42(3): 663-672.
[263] Chai Yuan, Luo Jianjun, Wang Mingming, et al. Space target proximity control based on pursuit-and-escape game[J]. Aerospace Technology, 2020, 4(1): 30-38.
[264] CRANDALL M G, EVANS L C, LIONS P L. Some Properties of Viscosity Solutions of Hamilton- Jacobi Equations [J]. Transactions o£ the American Mathematical Society, 1984, 282(2): 487-502.
[265] MARTIN L P. Markov Decision Processes: Discrete Stochastic Dynamic Programming [M]. New York: John Wiley & Sons, 1994.
[266] EVANS L C. Partial Differential Equations [M]. Providence: American Mathematical Society, 1998.

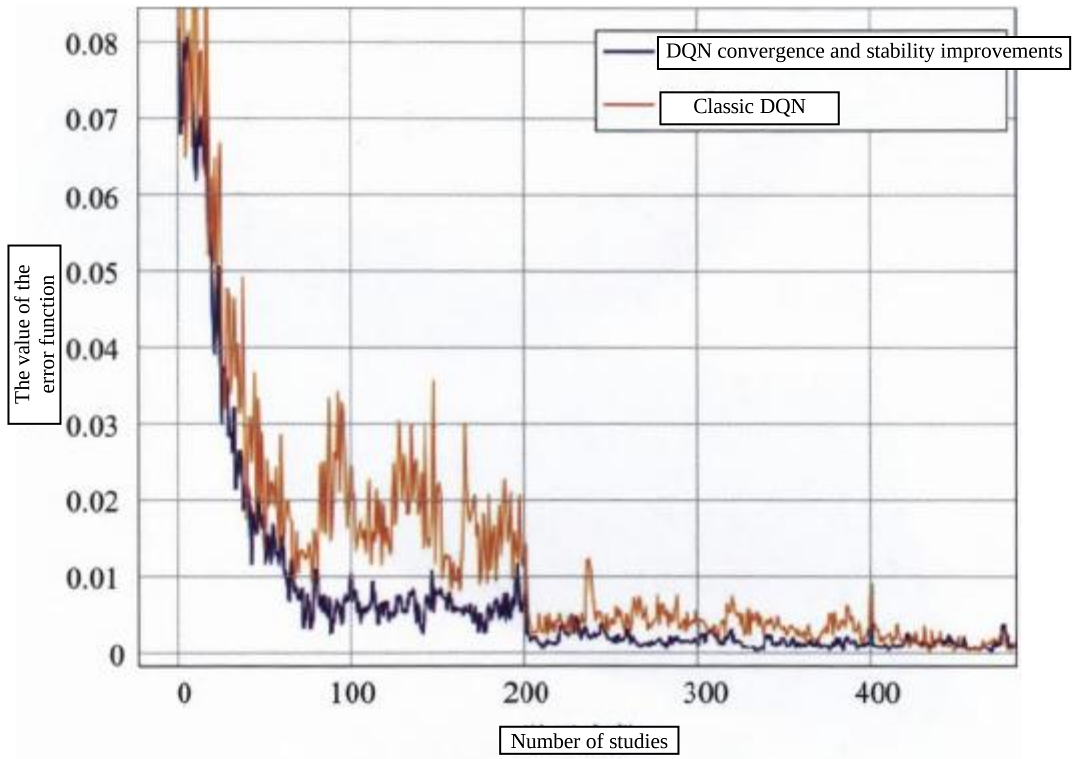


Figure 3-19: Comparison of Error Function Values Between the Two Methods (p108)

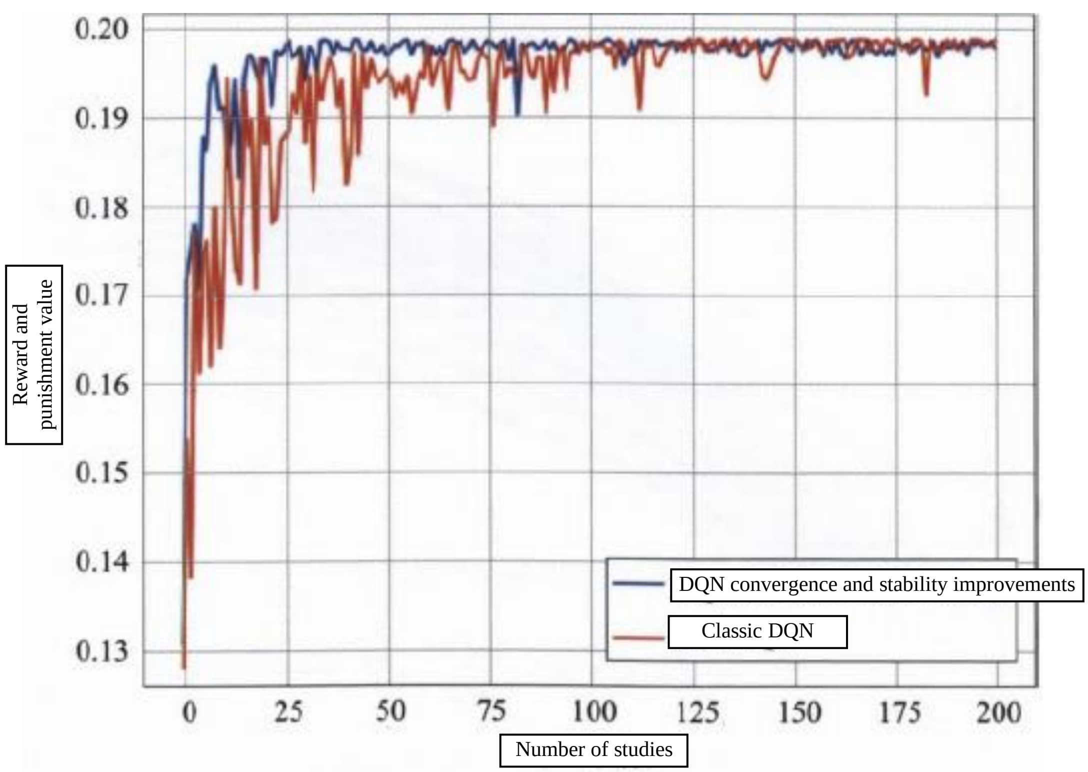


Figure 3-20: Comparison of Reward Values of the Two Methods (p108)

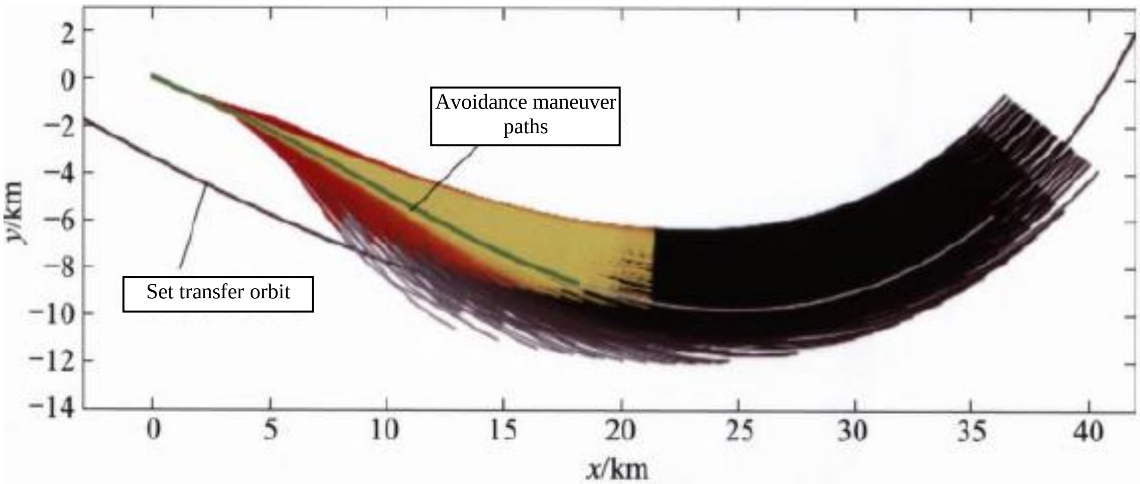


Figure 4-4: Acceleration Rate of Change Curve of Orbital Avoidance Path (p124)

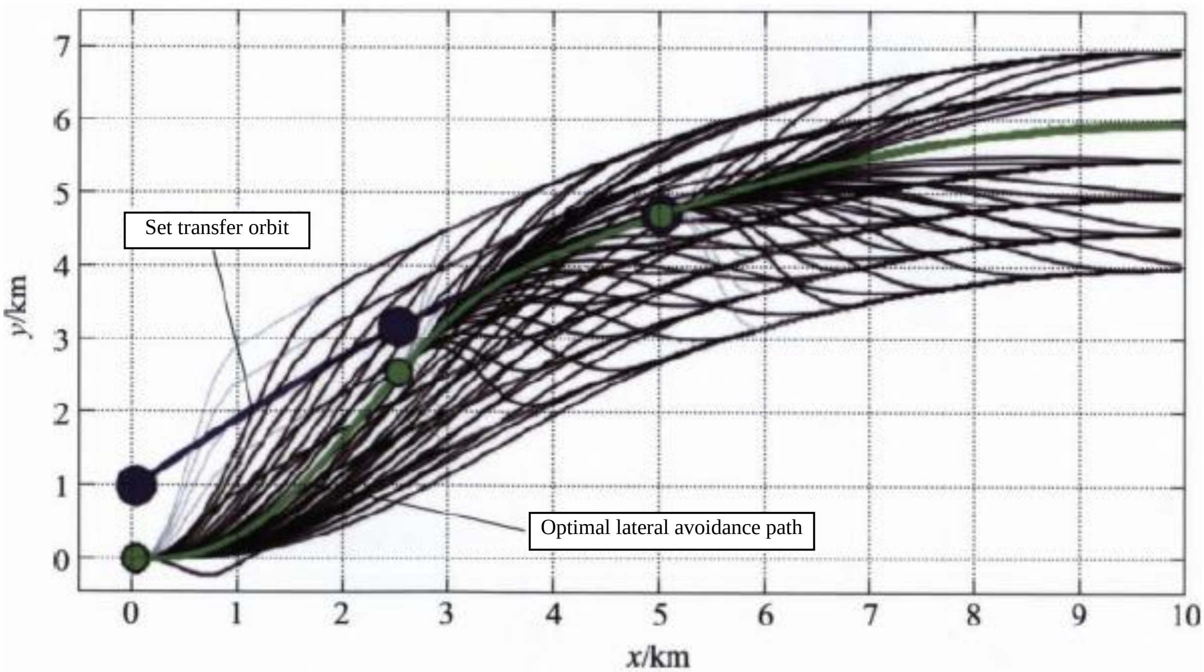


Figure 4-5: Lateral Offset Curve of Orbital Avoidance Path (p125)

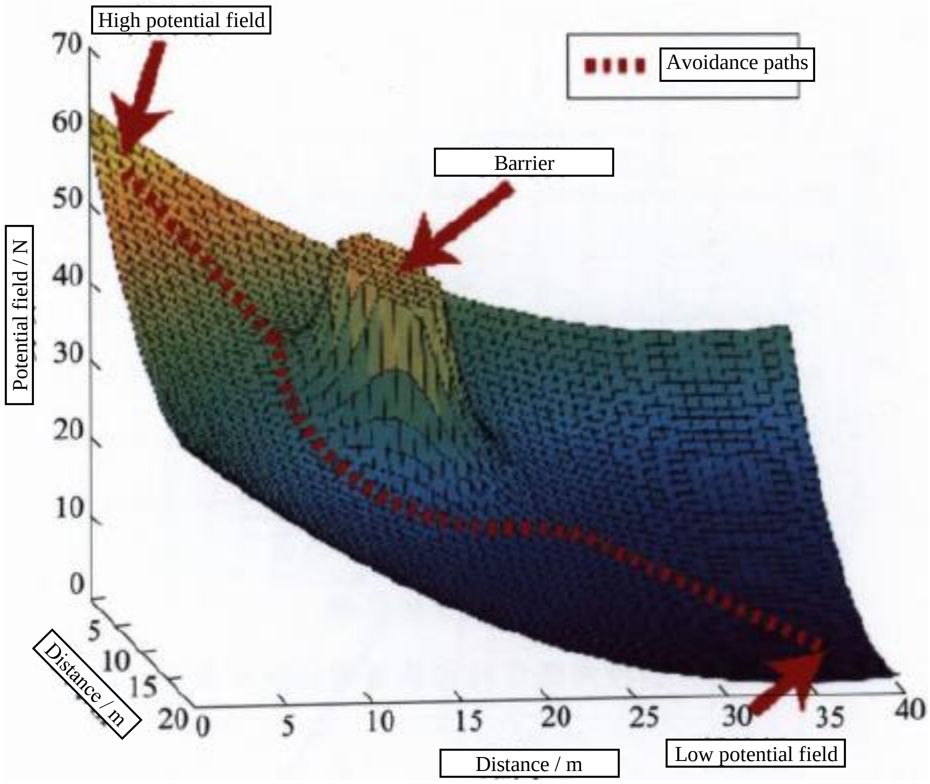


Figure 4-7: Schematic Diagram of Artificial Potential Field Avoidance Path (p129)

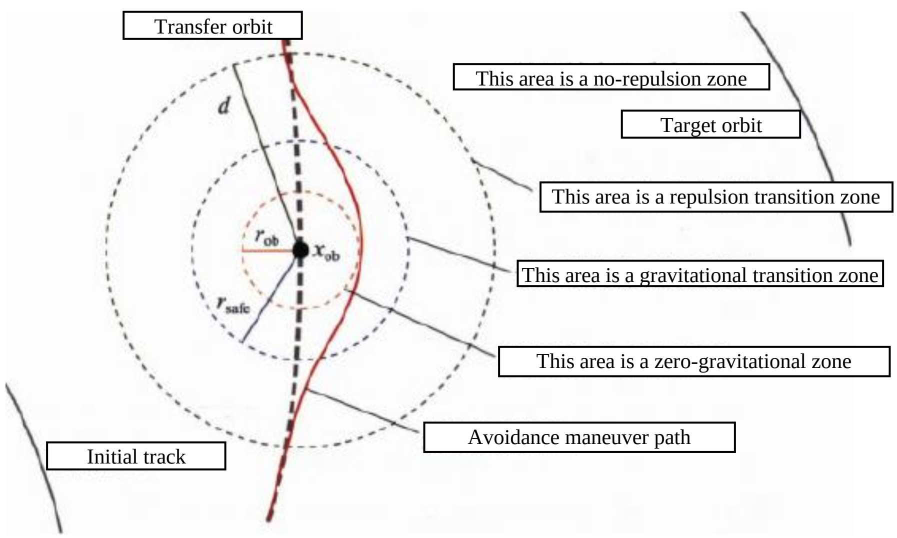


Figure 4-12: Schematic Diagram of Spacecraft Avoiding Maneuvering Artificial Potential Field (p135)

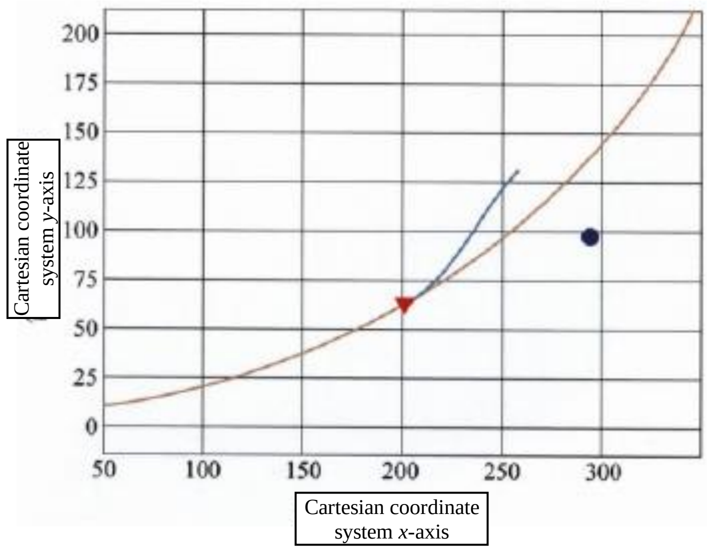


Figure 4-13: Simulation Diagram of a Spacecraft Leaving a Given Orbit (p143)

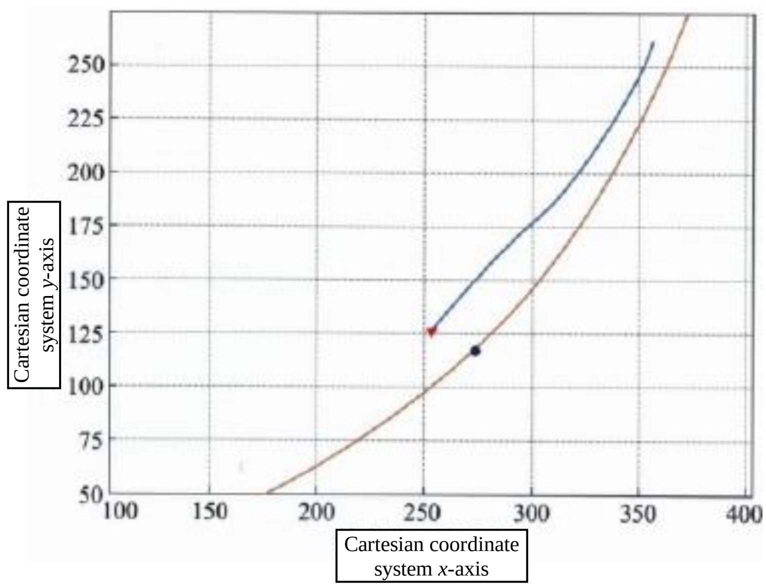


Figure 4-14: Simulation Diagram of Spacecraft Successfully Avoiding Space Debris (p143)

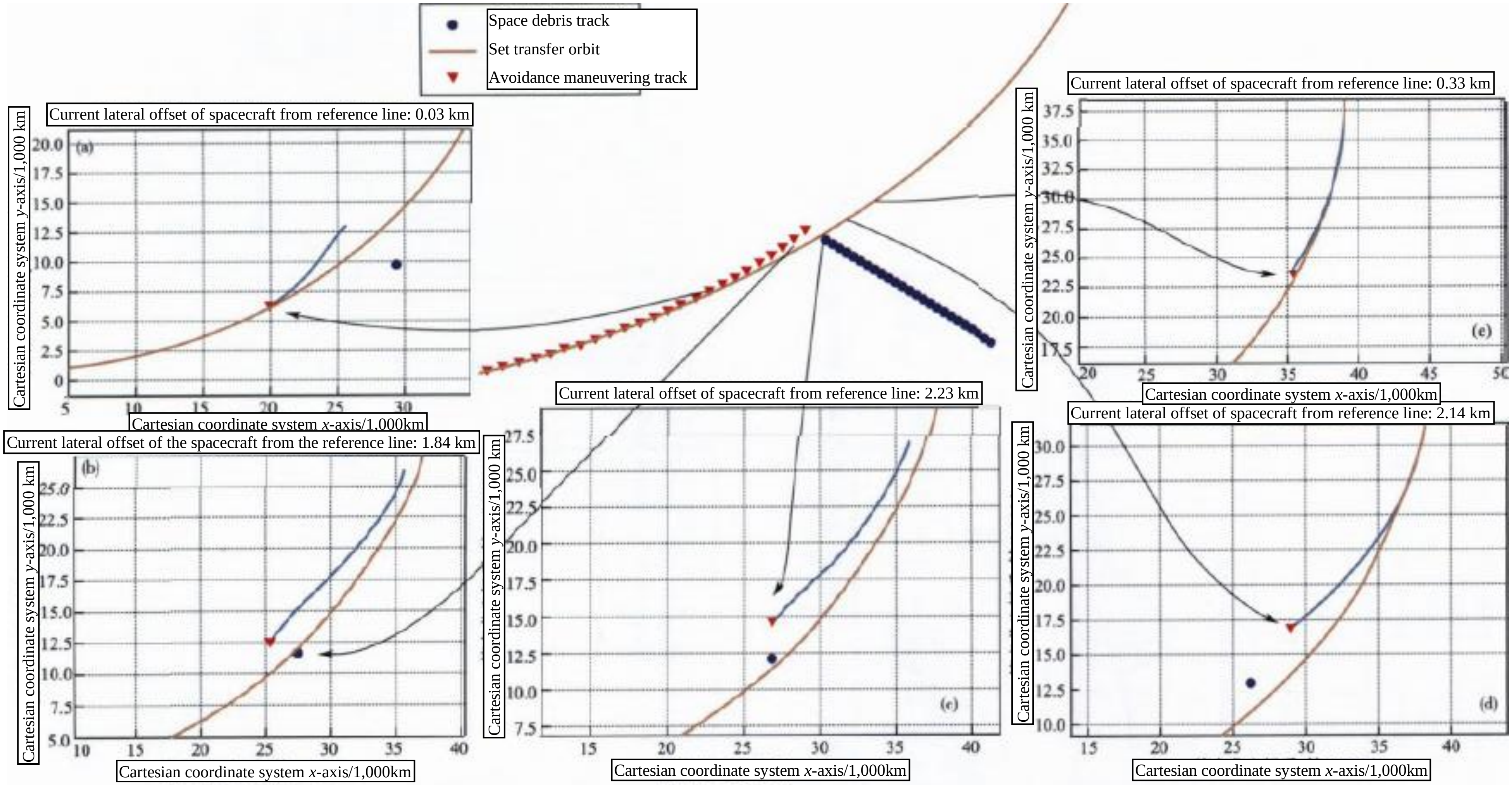


Figure 4-15: Overall Simulation Rendering of Spacecraft Space Debris Avoidance (p144)

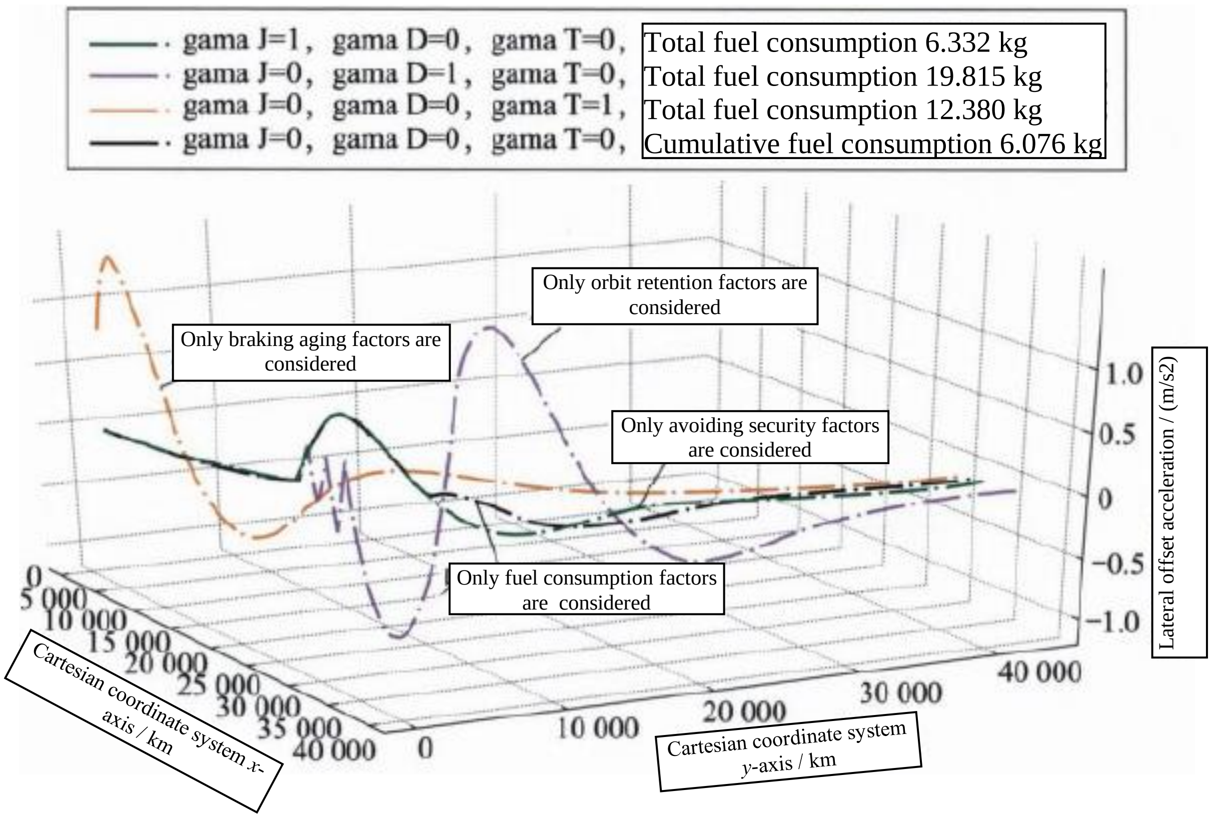


Figure 4-16: Lateral Offset Acceleration in Single-Objective Optimization Mode (p4145)

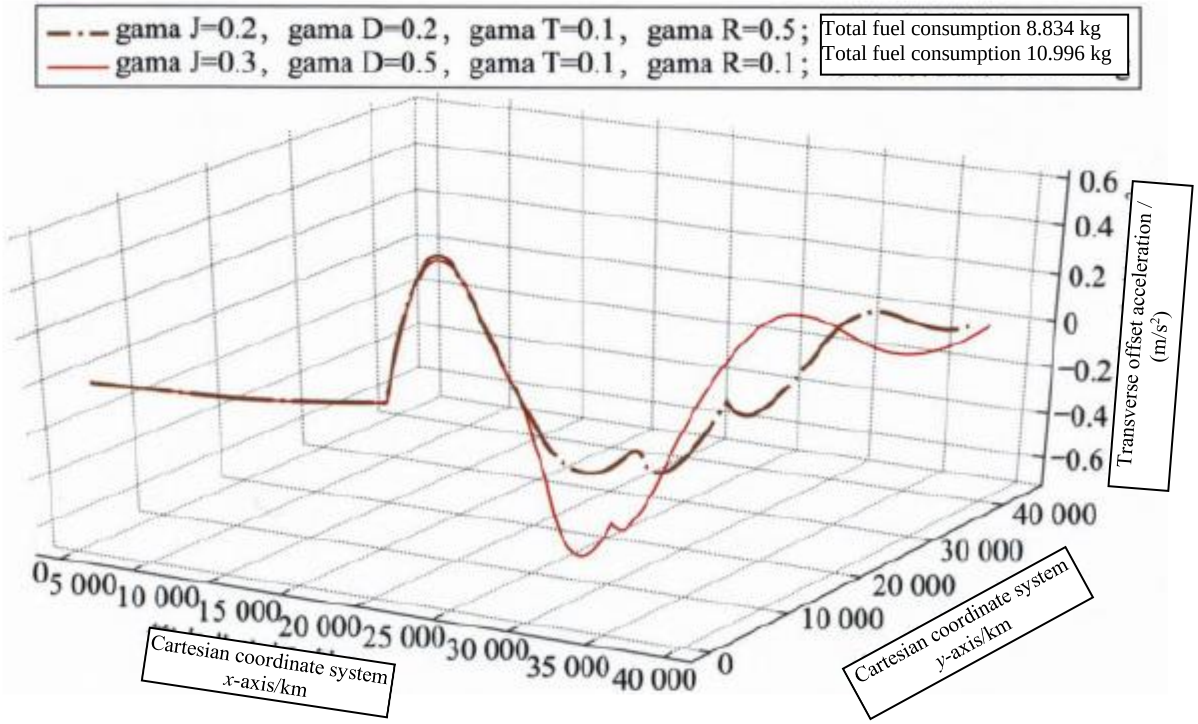


Figure 4-17: Lateral Offset Acceleration in Multi-Objective Optimization Mode (p147)

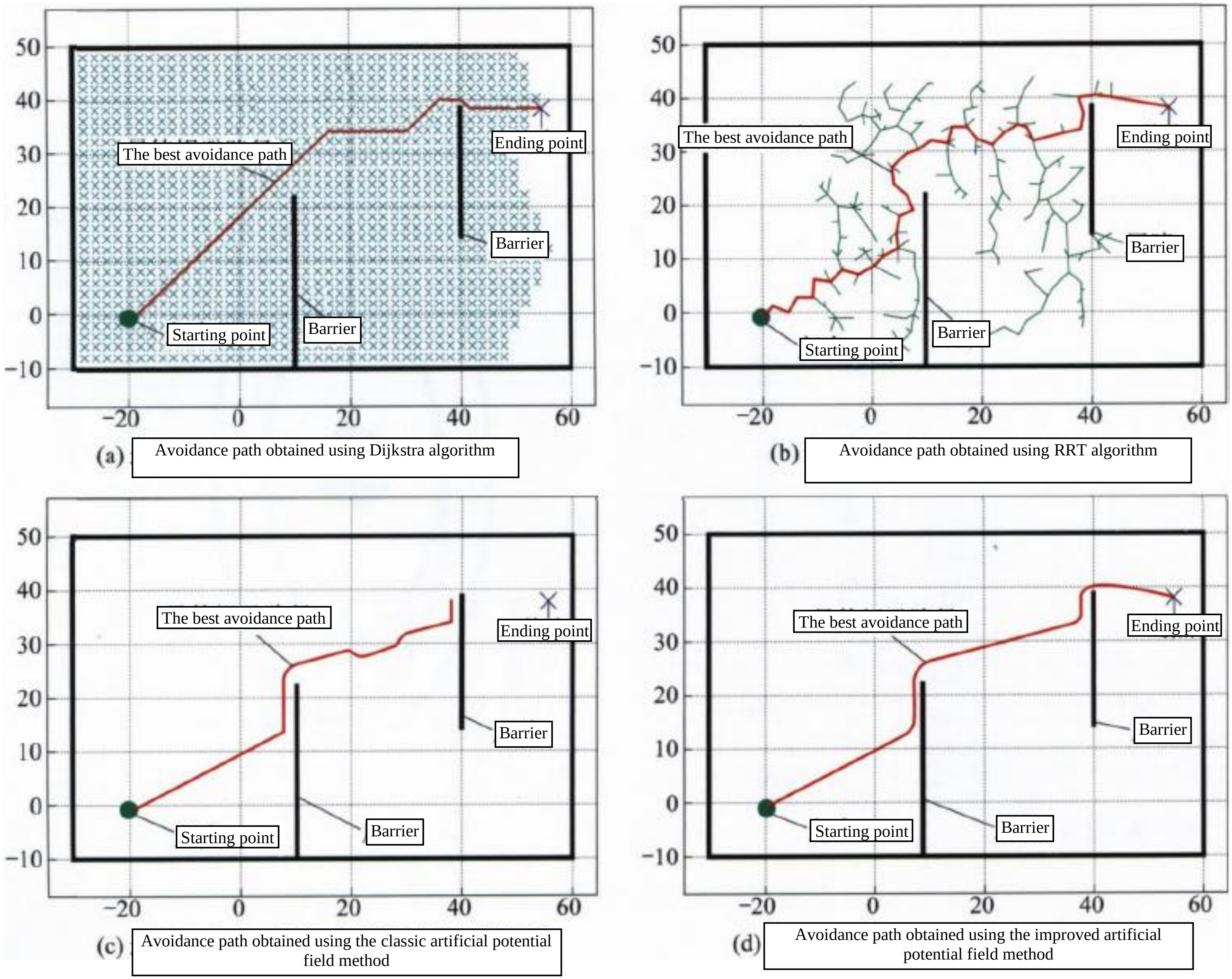


Figure 4-18: Solution Effects of Different Algorithms for Barrier Avoidance Problem (p150)

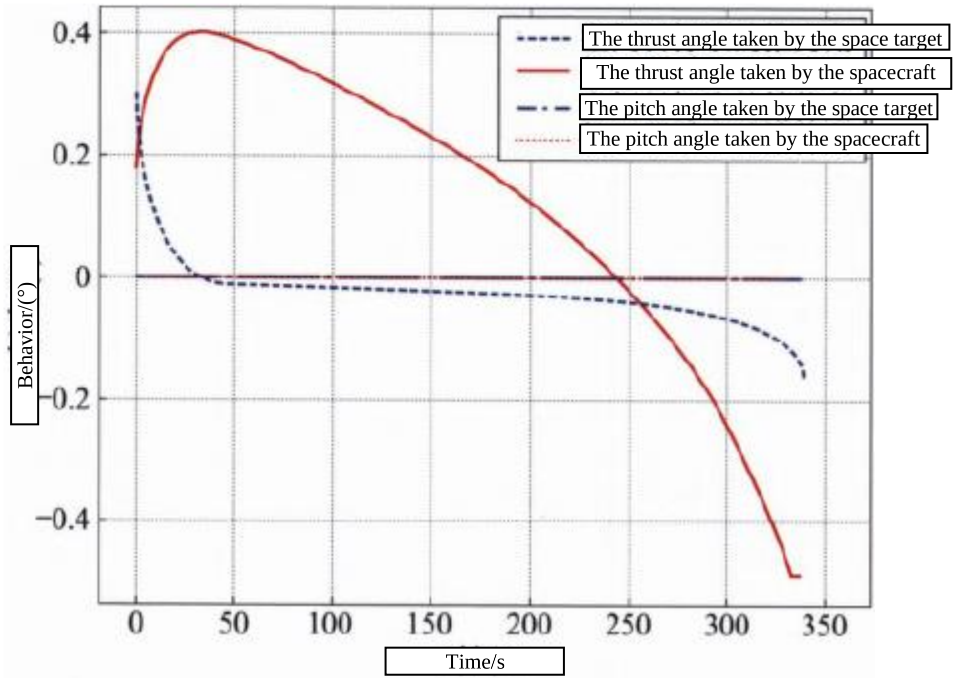


Figure 5-11: Behavior Control after 1,000 Sessions (p183)

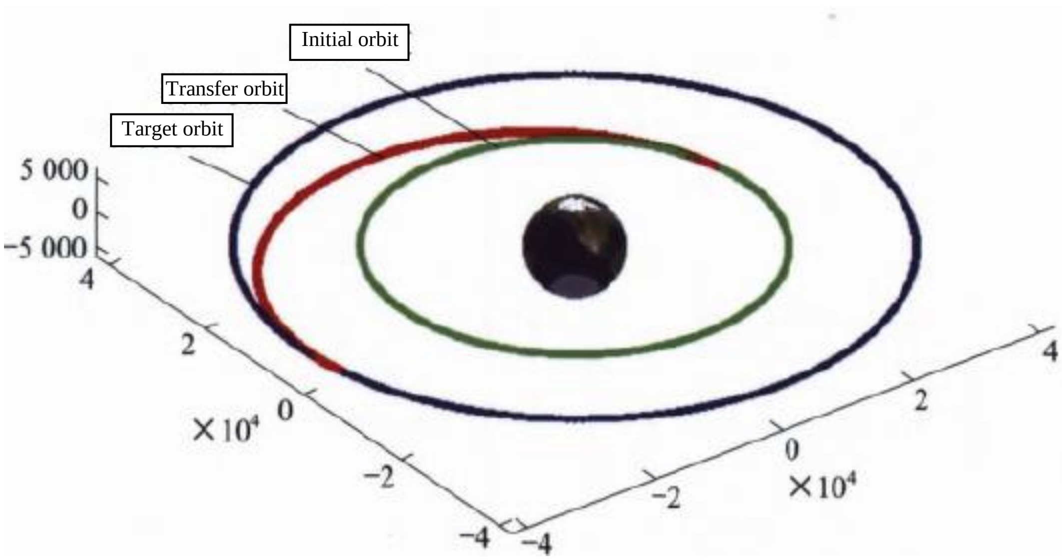


Figure 6-7: Coplanar Transfer Orbit from Spacecraft 1 to Target (p205)

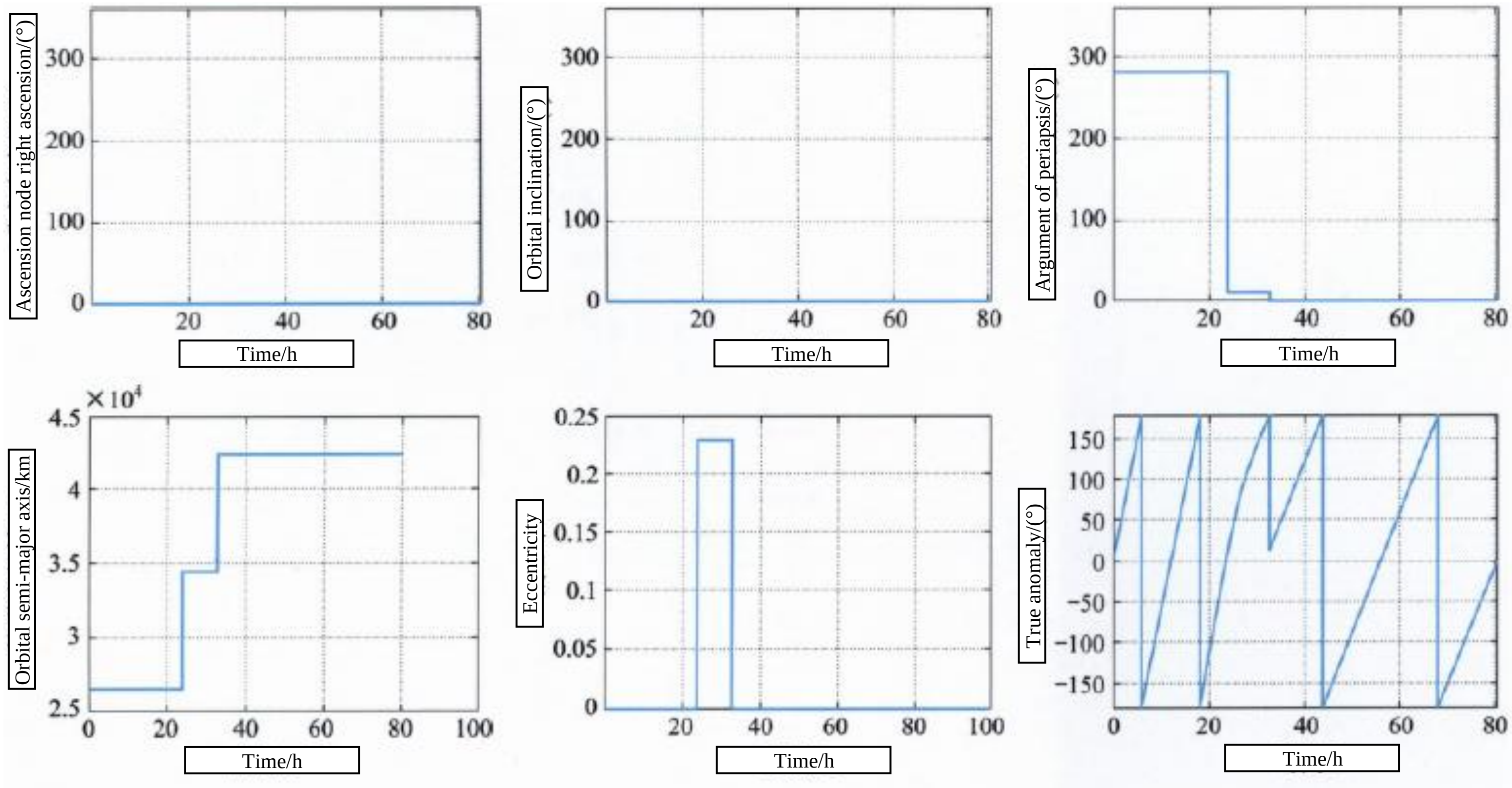


Figure 6-8: Variation of Orbital Root Number of Coplanar Transfer Orbitals (p206)

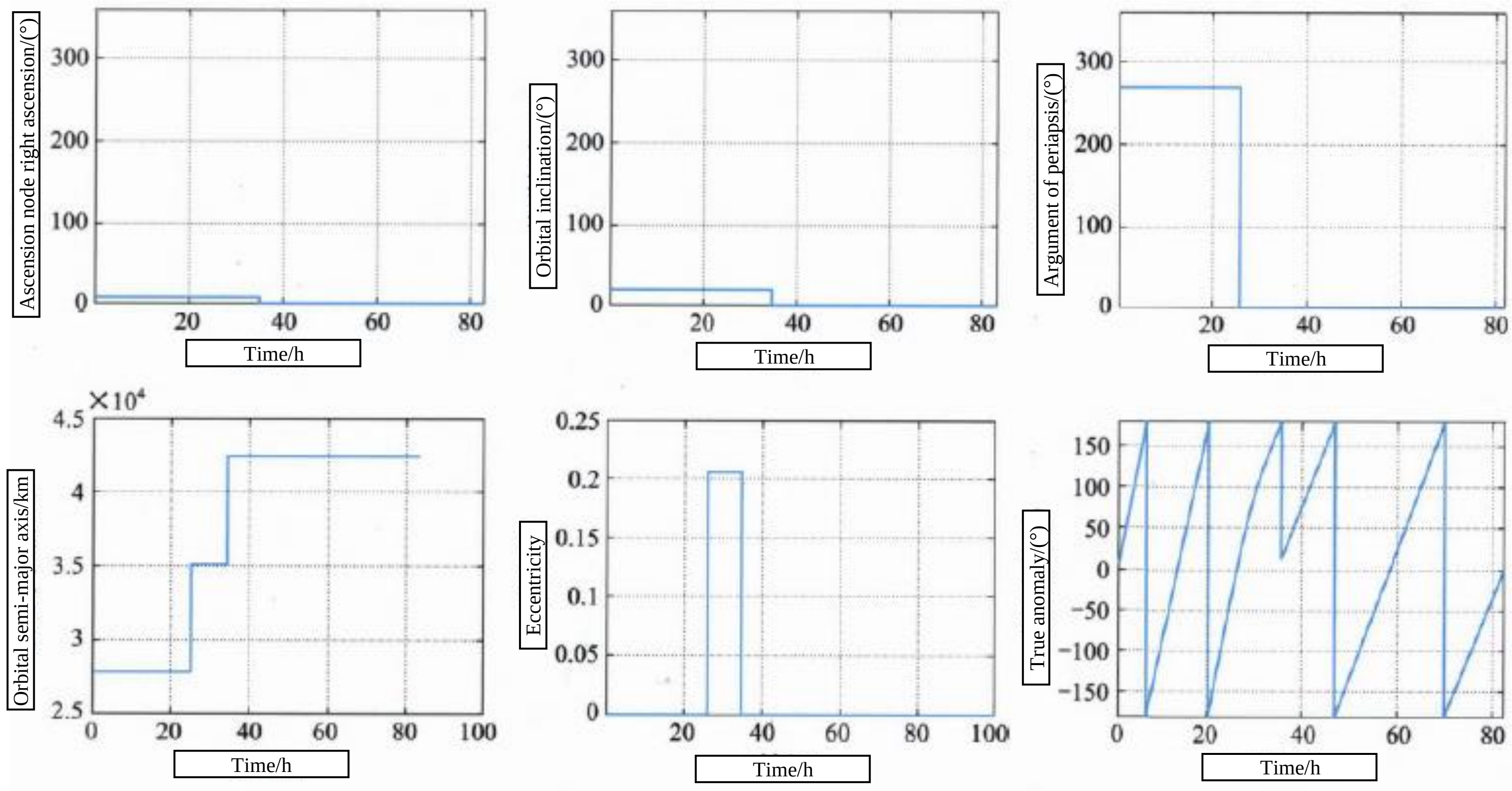


Figure 6-10: Changes in the Number of Orbital Roots in Non-Coplanar Transfer Orbitals (p210)

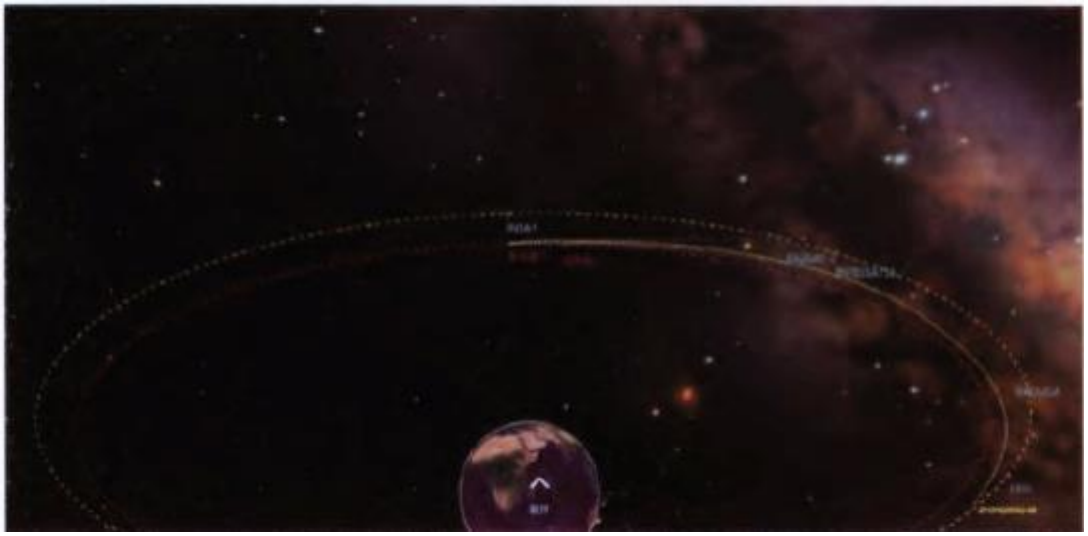


Figure 6-11: Example of Target Allocation Effect in an On-Orbit Servicing Task Planning System (p212)

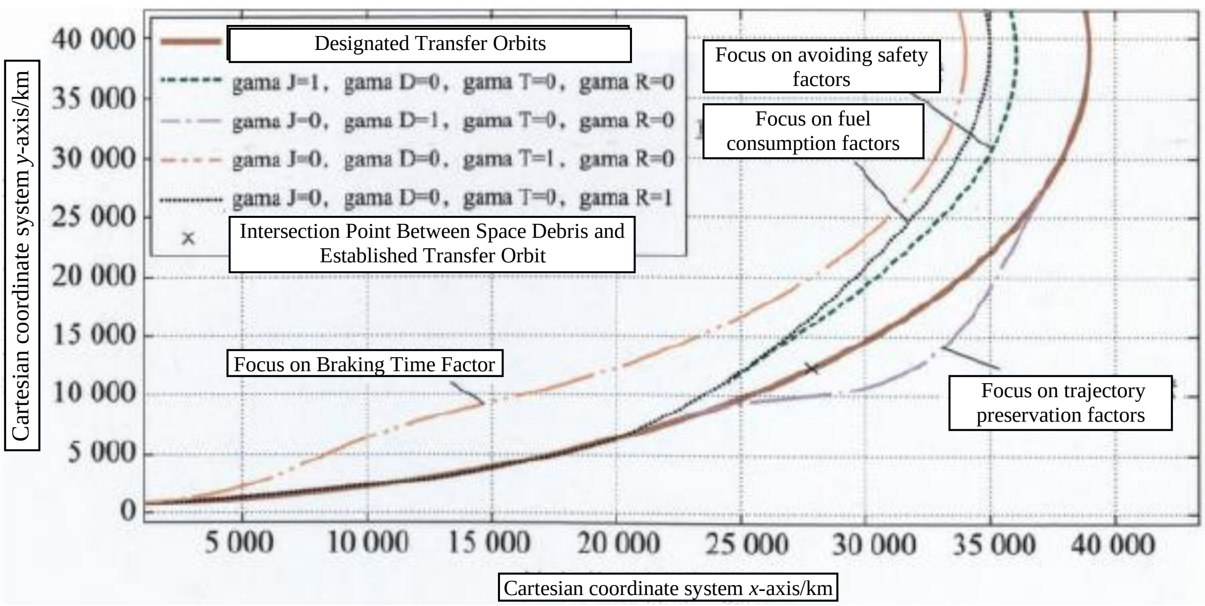


Figure 6-16: Optimal Spacecraft Avoidance Path under Univariate Preference (p215)

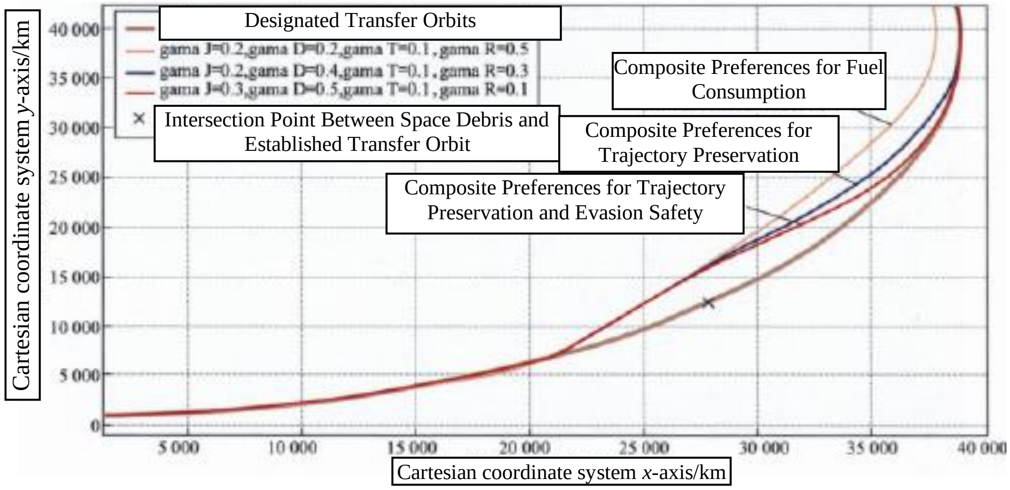


Figure 6-20: Optimal Avoidance Path of Spacecraft Under Composite Factor Preference (p218)

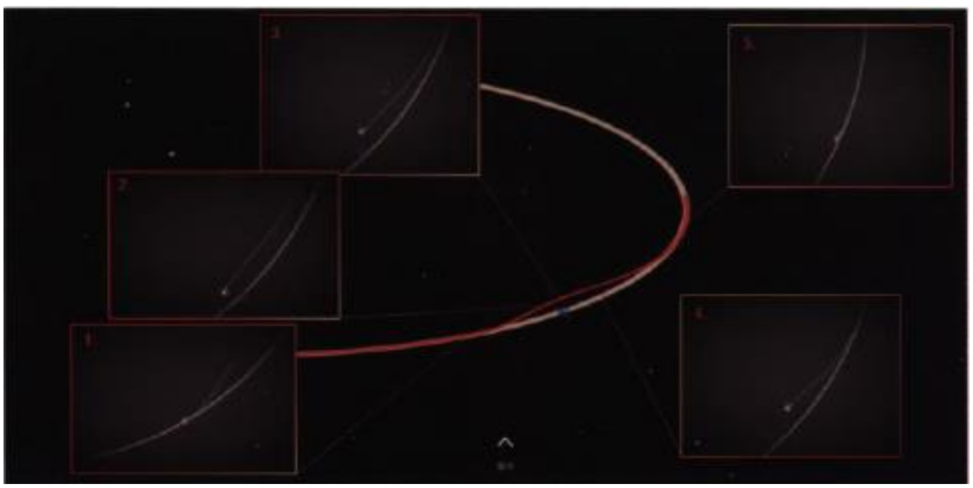


Figure 6-21: Example of Orbit Avoidance Effect in an On-Orbit Servicing Mission Planning System (p219)

As an important means to ensure the long-lasting and stable operation of spacecraft in the space environment, on-orbit servicing is an important development direction in the field of aerospace science and technology. In order to further meet the actual needs of mission planning throughout the whole process of on-orbit servicing, and better deal with such a long-lasting period and uneven type of mission planning, this book combines the development of on-orbit servicing technology and highlights the needs of risk disposal, explores the ways and means of applying advanced artificial intelligence technology to on-orbit servicing task planning, explores the research framework of using artificial intelligence methods to realize on-orbit servicing task planning, and focuses on an on-orbit target allocation method under a composite service model. A spacecraft orbit temporary avoidance path planning method and a spacecraft orbit game real-time planning method. The system design is carried out for practical applications, and through the combination of theoretical research and technical practice, it provides scientific planning means for on-orbit services and can provide reference for solving mission planning problems in other fields.