Editor in Chief

Professor Yaping Lei
President of Xi'an Technological University, Xi'an, China

Associate Editor-in-Chief

Professor Wei Xiang
Electronic Systems and Internet of Things Engineering
College of Science and Engineering
James Cook University, Australia（AUSTRALIA）

Dr. Chance M. Glenn, Sr.
Professor and Dean
College of Engineering, Technology, and Physical Sciences
Alabama A&M University,
4900 Meridian Street North Normal, Alabama 35762, USA

Professor Zhijie Xu
University of Huddersfield, UK
Queensgate Huddersfield HD1 3DH, UK

Professor Jianguo Wang
Vice Director and Dean
State and Provincial Joint Engineering Lab. of Advanced Network and Monitoring Control, CHINA
School of Computer Science and Engineering, Xi'an Technological University, Xi'an, China

Administrator

Dr. & Prof. George Yang
Department of Engineering Technology
Missouri Western State University, St. Joseph, MO 64507, USA

Professor Zhongsheng Wang
Xi'an Technological University, China
Vice Director
State and Provincial Joint Engineering Lab. of Advanced Network and Monitoring Control, CHINA

Associate Editors

Prof. Yuri Shebzukhov

International Relations Department, Belarusian State University of Transport, Republic of Belarus.

Dr. & Prof. Changyuan Yu
Dept. of Electrical and Computer Engineering, National Univ. of Singapore (NUS)

Dr. Omar Zia
Professor and Director of Graduate Program
Department of Electrical and Computer Engineering Technology
Southern Polytechnic State University
Marietta, Ga 30060, USA

Dr. Liu Baolong
School of Computer Science and Engineering
Xi'an Technological University, CHINA

Dr. Mei Li
China university of Geosciences (Beijing)
29 Xueyuan Road, Haidian, Beijing 100083, P. R. CHINA

Dr. Ahmed Nabih Zaki Rashed
Professor, Electronics and Electrical Engineering
Menoufia University, Egypt

Dr. Rungun R Nathan
Assistant Professor in the Division of Engineering, Business and Computing
Penn State University - Berks, Reading, PA 19610, USA

Dr. Taohong Zhang
School of Computer & Communication Engineering
University of Science and Technology Beijing, CHINA

Dr. Haifa El-Sadi.
Assistant professor
Mechanical Engineering and Technology
Wentworth Institute of Technology, Boston, MA, USA

Huaping Yu
College of Computer Science
Yangtze University, Jingzhou, Hubei, CHINA

Ph. D Wang Yubian

Department of Railway Transportation Control
Belarusian State University of Transport, Republic of Belarus

Prof. Xiao Mansheng
School of Computer Science
Hunan University of Technology, Zhuzhou, Hunan, CHINA

Qichuan Tian
School of Electric & Information Engineering
Beijing University of Civil Engineering & Architecture, Beijing, CHINA

Language Editor

Professor Gailin Liu
Xi'an Technological University, CHINA

Dr. H.Y. Huang
Assistant Professor
Department of Foreign Language, The United States Military Academy, West Point, NY 10996, USA

Table of Contents

The Research of Tai'an Big Data Ecological Status Based on IPV9

Gao Yalin
Department of General Endocrinology
The First Affiliated Hospital of Xi'an Medical University
E-mail: 979382497@qq.com

Ma Yihua
Department of General Endocrinology
The First Affiliated Hospital of Xi'an Medical University
E-mail: 910107571@qq.com

***Abstract*—In recent years, medical problems have always been a topic of concern. With the improvement of people's life quality, the attention to diseases has also reached a new height. It has also become a livelihood issue for the whole nation to enjoy low-cost and high-quality medical services. This paper researches the current situation of China's medical treatment, summarizes the existing problems in China's medical treatment, analyzes the development situation and opportunities of China's big data in medical treatment, and finally introduces and analyzes the Tai 'an health ecosystem based on the independent and controllable new network IPV9 technology, and forecasts the brand-new digital medical network. In information age, new IT technology will inevitably become the key to the development of medical care.**

Keywords-Health Platform; Medical Big Data; IPV9

I. PROJECT BACKGROUND

In the "13th Five-Year Plan for Deepening the Reform of the Medical and Health System" issued by the State Council in December 2016, it was proposed that during the "13th Five-Year Plan" period, it is necessary to have conduct New breakthroughs on grading diagnosis and treatment, modern hospital management, universal medical insurance, drug supply security and New breakthroughs have been made in the construction of 5 systems including supervision, and promote reforms in related fields at the same time[1].

The“13th Five-Year Plan for National Population Health Informatization Development Plan” issued by the National Health and Family Planning Commission in January 2017 requires: strengthen the construction of population health informatization and health medical big data service system vigorously, and promote the government health medical information system and Public health and medical data interconnection , open sharing, eliminate information barriers and islands, focus on improving the ability of population health information governance, vigorously promote the development of medical big data applications, and explore new models of "Internet medical treatment" services.

The “ ‘Thirteenth Five-Year Plan’ for Shandong Province’s Population Health Informatization Construction” issued by Shandong Province in August 2016 clearly pointed out that it is necessary to accelerate the promotion of population health informatization construction and improve the level of health and family planning services. By 2017, all provincial and municipal platforms will be completed and interconnected. By 2020, the health and family

DOI: 10.21307/ijanmc-2020-011

planning information network and grassroots information management system will expand to cover non-governmental health and family planning institutions. It is proposed to improve the construction of regional population health data centers and health and family planning information networks, strengthen the standardization of provincial and municipal population health information platforms, and strengthen the application of public health management information, family planning management information, medical services and security information, and drug procurement management. Information application, comprehensive management information application, administrative office information service, improve the construction of information security system, promote the "one-card use" of population health information, and promote the "Internet + health care" services to facilitate people and benefit the people.

II. PROBLEMS IN THE MEDICAL INDUSTRY

Population growth and aging, the expansion of the medical market in developing countries, the advancement of medical technology and the continuous increase in labor costs will drive the growth of expenditure. Global medical expenditures in 2017-2021 are expected to grow at a rate of 4.1% per year, while the growth rate in 2012-2016 is only 1.3%. Increasing incidence of chronic diseases, changing dietary habits and increasing obesity have exacerbated the upward trend of chronic diseases, especially cancer, heart disease and diabetes. There are currently about 114 million diabetic patients in China, and the number of patients in the world is expected to increase from 415 million at present to 642 million in 2040. Traditional research and development (R&D) costs have risen, and the time to market has been slow. From 2004 to 2014, drug development costs increased by 145%. The labor force is insufficient. With the rapid changes in the population structure and the rapid development of technology, skilled and semi-skilled healthcare workers will be greatly reduced.

A. Supply-demand structure imbalance

In 2016, there were 2.31 practicing (assistant) physicians per 1,000 population in my country. In 2015, the number of physicians per thousand population in my country ranked between 25-30 among the countries counted by OECD. In addition, the poor practice environment of doctors in my country has made fewer outstanding talents enter the medical system year by year.

B. Unbalanced development of medical resources

From 2010 to 2016, the compound growth rate of the number of diagnosis and treatment and the number of institutions in tertiary hospitals were 10.7% and 8.3%, respectively, while that of primary hospitals was only 1.5% and 0.4%. The traditional mode of medical treatment makes the tertiary hospitals overcrowded, resulting in poor medical experience and serious waste of high-quality medical resources. Therefore, my country still regards the implementation of graded diagnosis and treatment as the current primary task.

C. Medicare Overdraft

"The China Medical and Health Care Development Report 2017" predicts that by 2024, a cumulative deficit of 735.3 billion yuan will be in deficit. Improving the ability of medical insurance to control fees and exploring innovative payment mechanisms are imminent. Therefore, it is imperative to improve the existing medical treatment model and implement graded diagnosis and treatment.

In 2015, the National Health and Family Planning Commission proposed that a tiered diagnosis and treatment system will be fully established in 2020, including primary consultation at the grass-roots level, two-way referral, swift triage, and up-down linkage triage. During the construction of the new model, there are three main problems.

1) Information does not circulate, and most medical institutions are isolated islands of information, and patient information cannot be quickly shared and circulated;

2) Resources are not circulating, high-quality doctors are mostly concentrated in the top hospitals of major provincial capitals, and the top three hospitals have limited energy, and the primary medical care they can support every year is limited;

3) Interests are not interlinked, and there is no effective mechanism of interest binding between hospitals to promote the circulation of patients between hospitals.

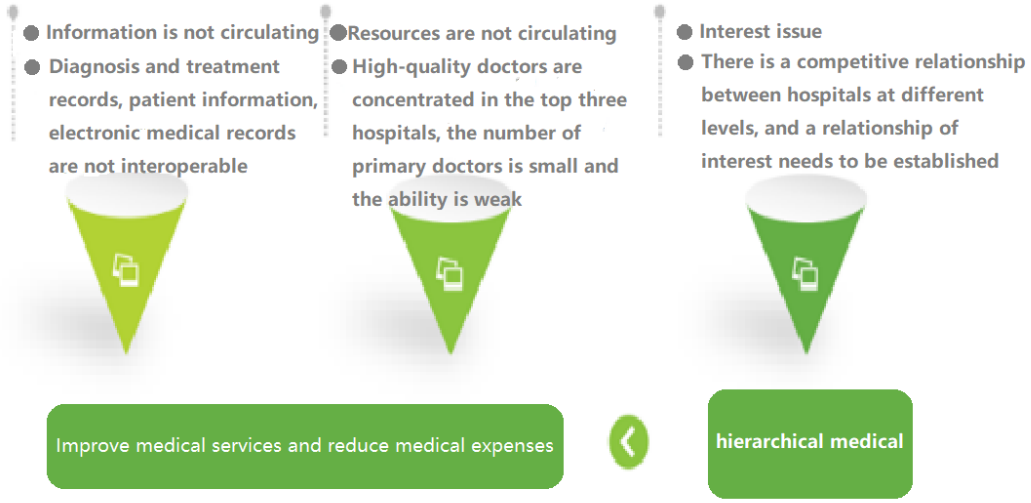


Figure 1. Problems encountered in the implementation of graded diagnosis and treatment in 2018

D. Aging enters a serious stage

According to the China Industry Information report, in 2017, my country's population over 65 years old accounted for 11.4%, the birth rate was 12.43% (2016 birth rate 12.95%), and the population structure showed an aging trend. Based on 2013, the over-two-week visit rate of people over the age of 65 is 26.4%, and under the influence of modern living habits, the incidence of chronic diseases such as diabetes and hypertension is also increasing[2]. The number of diagnosis and treatment will continue to rise, and the medical system will also face great pressure. As shown in Figure 2, Figure 3.

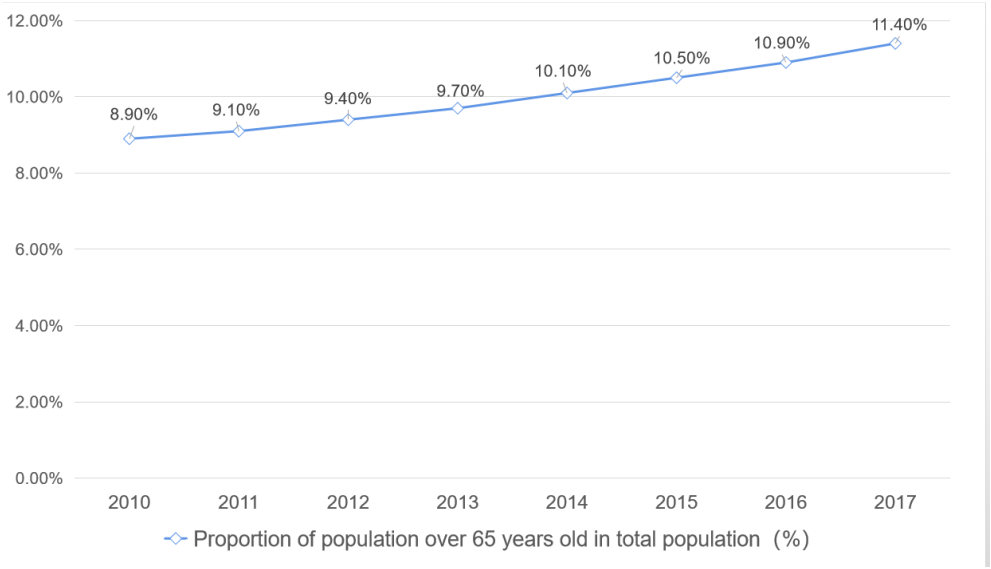


Figure 2. Proportion of the population aged 65 and over in China from 2010 to 2017

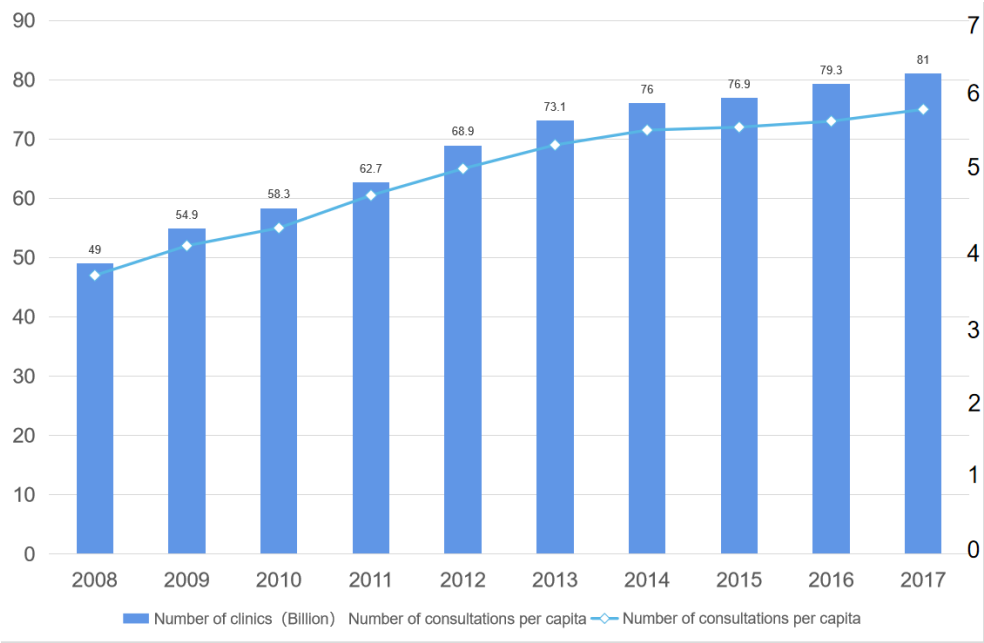


Figure 3. Number of consultations and treatments per capita in my country from 2008 to 2017

E. High proportion of chronic diseases, sub-health, and lung cancer

There are 170 million people with hypertension and more than 100 million people with high blood fat in my country. There are 92.4 million people with diabetes, and an average of one person has diabetes every 30 seconds. At present, the sub-health ratio of white-collar workers in mainstream cities in my country is as high as 76%, and the number of white-collar workers who are overworked is close to 60%. The proportion of healthy people in the true sense is less than 3%. my country has more than 800,000 newly diagnosed lung cancer patients each year, and the death rate ranks first among all malignant tumors.

Chronic diseases in my country have the characteristics of "large number of patients, long duration of illness, high cost of medical treatment, and large demand for services", which has formed the predicament of the Chinese medical industry.

III. DEVELOPMENT STATUS OF CHINA'S MEDICAL MARKET

The development of the medical big data industry is driven by valued medical care (that is, the win-win of medical service quality and medical cost), and its potential value space is huge, and it is generated in specific application scenarios. The service targets of medical big data can be residents, medical service institutions, scientific research institutions, medical insurance management institutions and commercial insurance companies and public health management departments.

McKinsey's 2013 report predicted that in the United States alone, the application of medical big data is expected to reduce medical expenses by 300 billion to 450 billion US dollars per year. China has huge population base, serious waste of medical resources, shortage of medical resources and unreasonable allocation, excessive growth of medical expenses, and weak development of commercial insurance. The application of medical big data is rich and can be deeply explored. The market size is at least 100 billion.

At present, the domestic medical big data industry is still in its infancy, and it is moving towards the growth stage, and the market scale is constantly rising. According to the analysis report of the Global Health and Medical Big Data Industry Development Prospects and Investment Strategic Planning Analysis 2018-2023 released by the Prospective Industry Research Institute, the market size of China's medical big data industry in 2015 was about 46.6 billion, and there is still room for future growth. bigger. The market size is expected to reach 91 billion yuan by 2022. As shown in Figure 4.

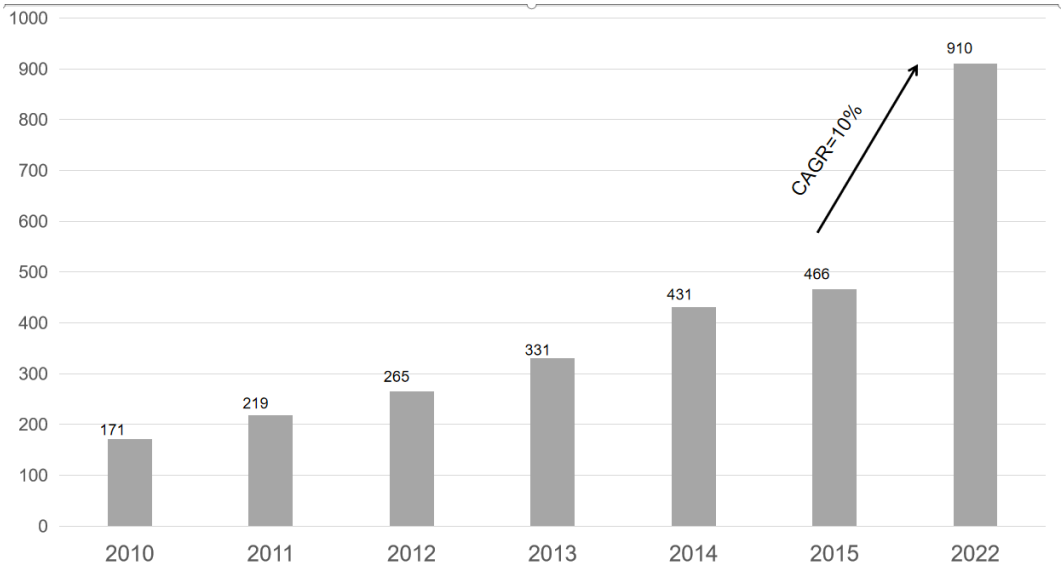


Figure 4. Market scale and forecast of China's health and medical big data industry (unit: 100 million yuan, %)

IV. TAI'AN HEALTH BIG DATA PLATFORM

A. Future Network IPV9

IPV9 officially applied for a patent certificate in December 2011. The main content refers to the 0-9 Arabic digital network as a virtual IP address, and the decimal is used as a text representation method, that is, a method of use that is easy to find online users; Efficiency and convenience for end users, some of which can be used directly as domain names; it has an endless number of assignable IP addresses and has a maximum of 2 × 2048-bit addresses, which is the cornerstone of the future digital world[3]. At the same time, because it uses the classification and coding of the original computer network, cable broadcast TV network and telecommunications network services, it is also known as "the new generation of safe and reliable information integrated network protocol."

Compared with IPv4 and IPv6, IPV9 has more obvious advantages, its address space is larger, the routing table forwards packets faster, it can provide better and cheaper network platform services, and has better security And compatibility, the most important thing is that IPV9 has independent intellectual property rights, and China has mastered the core technology, based on IPV9 technology to establish a safe and controllable and compatible traditional network Internet system, further promote the formation of ISO future network standards, thus breaking the United States Internet hegemony puts cybersecurity in your own hands. The formation of network standards in the future will further promote the independent establishment of Internet systems in various countries to achieve safe and controllable cyberspace.

B. China's Big Data Ecosystem

1) Opportunities for The Construction of China's Big Data Ecosystem

Taking the opportunity of urban transformation as an opportunity to actively plan for the construction of new smart cities and create a benchmark for the construction of new smart cities in my country has become a development goal in the process of urbanization everywhere. Among them, smart medical is not only an important part of the construction of smart cities, but also an important part of promoting the reform of the medical and health system.

The maturity of IT technology, the emergence of SOA technology, SaaS application mode, and the rapid development of high-tech technologies such as wireless networks, the price of IT equipment is getting lower and lower, making this construction practically feasible from a technical and economic perspective. At the same time, with the continuous application of cloud computing technology in the practice of medical informatization, regional medical informatization construction can achieve better results on this basis.

The national health industry development realizes major projects for the conversion of old and new kinetic energy, and provides great opportunities for the development of big data ecological domains throughout the health area. Continue to integrate and promote the leapfrog development of the city's health service industry, and promote economic development with the adjustment of the health service supply side.

In August 2007, the Ministry of Information Industry formally defined IPv9 as the next-generation Internet to distinguish IPv6 from the next-generation Internet. On February 23, 2013, the State Council issued a notice on the national mid-term and long-term plan for the construction of major scientific and technological infrastructure (2012-2030), which stated in the construction of key future network test facilities during the "The Twelfth Five-Year Plan" period: Relying on increased bandwidth and progressive improvement can no longer meet the needs of future development. To break through the future network basic theory and support the new generation of Internet experiments, build future network test facilities, mainly including: original network equipment system, resource monitoring and management system, covering cloud Open services such as computing services, IoT applications, spatial information network simulation, network information security, high-performance integrated circuit verification, and quantum communication networks.

For the future network based on IPv9, the document requires accelerating the construction of major scientific research infrastructure such as the national future network test facility, and actively carries out network new technology, new application test verification and application demonstration, and requires significant enhancement of the network information technology independent innovation ability Form the first-mover advantage of future network technology.

2) The Situation Facing China's Big Data Ecological Domain Construction

The environment at home and abroad has undergone profound changes, medical service capabilities and residents' family health awareness have been continuously improved, and the "Healthy China" strategy has been implemented in depth. The Party and the country attach great importance to the health work of the whole people. The Fifth Plenary Session of the Eighteenth Central Committee of the Party will raise "healthy China" into a national strategy. The country has successively issued policy documents such as “Outline of “Healthy China 2030””, “Outline of Strategic Planning for Traditional Chinese Medicine Development (2016-2030)”, and “National Nutrition Plan (2017-2030)”. The development of the healthcare industry has created a good environment.

The State Council approved the "Overall Plan for the Construction of Comprehensive Experimental

Zones for Conversion of New and Old Kinetic Energy in Shandong". Shandong Province is vigorously implementing major projects for the conversion of old and new kinetic energy, accelerating the promotion of "four new" to promote "four modernizations", and actively creating a national demonstration province with integrated medical and nursing care As one of the "top ten" industries in the province's conversion of old and new kinetic energy, the healthcare industry clearly proposes the development goal of building a trillion-level healthcare industry by 2022. As an important meeting point for the improvement of people's livelihood and the conversion of old and new kinetic energy, Shandong Province will continue to increase its support for the medical and health industry.

A new round of scientific and technological revolution and industrial transformation has accelerated. Life sciences and technology continue to make new breakthroughs, accelerated application of major technologies such as genetic engineering, molecular diagnostics, stem cell therapy, 3D printing, and the depth of the next generation of information, biology, engineering technology and medical health fields such as big data, cloud computing, Internet, artificial intelligence Integration is getting closer and closer, and technologies such as telemedicine, mobile medicine, precision medicine, and smart medicine are developing vigorously, and new forms and models of health industries such as health management, health care for the elderly, health tourism, leisure health, and "Internet + health" are flourishing.

The consumption structure of residents continued to upgrade. With the substantial improvement of living standards and the rapid transformation of life concepts, the consumption structure of residents has accelerated to a developmental and enjoyable type, the general public's health awareness has increased, and health needs have diversified from a single medical service to disease prevention, health promotion, health care and rehabilitation, etc. Services are changing, and people's demand for health products and services is increasing. At the same time, the continuous improvement of the social security system and the rapid development of the medical insurance business will certainly further stimulate the market demand for medical care and health.

The problem of population aging is becoming increasingly prominent. In recent years, the aging population in Shandong has shown a large base, rapid growth, aging, disability, and empty nesting, and the aging population has been deepening. On the one hand, the needs of elderly people's life care and medical health are doubled, the consumer demand in the field of medical care and health is strong, and the development space of related industries is huge. On the other hand, the medical care and health industry in Shandong is still in its infancy, with relatively insufficient supply-side capabilities, structural contradictions and policy barriers, lack of high-quality resources, narrow coverage of medical and nursing integration, and insufficient professionals to meet the needs of the elderly Demand for health care services at different levels.

C. Health Tai'an Big Data Platform

On November 20, 2018, at the 9th Dacheng Road, Beijing, organized by the Chinese People's Liberation Army General Staff of the IPv9 Technology Project Application Demonstration Seminar, the meeting discussed and demonstrated the healthy Tai'an big data ecological domain as an IPv9 technology application case. It is required to accelerate the construction of healthy Tai'an big data ecological domain and rapidly increase the scale of IPv9 network, and strive to build Tai'an into an IPv9 network technology demonstration area through the construction and promotion of healthy Tai'an big data ecological domain[4].

Health Tai'an big data ecological domain can provide residents with personalized health management and health care methods to improve residents' satisfaction, so as to achieve minor illnesses at the

grassroots level, serious illnesses into hospitals, and rehabilitation in the community; can provide residents with life-cycle health Information to provide residents with networked and information-based health services and health management, residents can obtain continuous, comprehensive and high-quality health care services; can improve the efficiency of health services, reduce residents' waiting time for medical treatment; can support high-quality regional health The rational use of resources effectively solves the rational division of labor and resource allocation among grassroots and second-level large hospitals and preventive and health care institutions, and alleviates the problem of difficulty and expensive medical treatment for residents.

By building a healthy Tai'an big data ecological domain system, the demand for "health services in the same city" will be realized. The construction process realizes the above-mentioned "smart medical" development in Tai'an City by establishing four standards: unified data standards and norms for health information in Tai'an City, sharing of public health information resources in the same city, electronic two-way referral and inspection results, mutual recognition in the same city and application of health cards in the same city content. Taking the healthy Tai'an big data ecological domain as the core, it realizes the interconnection and sharing of information from horizontal to side and vertical to the end, as well as comprehensive business collaboration. Form a full life cycle of medical and health services, intelligent medical care, refined management and scientific decision-making, promote the development of large health industries, achieve the purpose of more scientific management, smarter business, and more beneficial to residents, and promote the openness of health and family planning in Tai'an City. The development by leaps and bounds strongly supports the construction of the medical highlands in Shentai'an City and the health of Tai'an. Through the construction of this platform, the information construction of health and family planning in Tai'an has reached the national first-class level. Since the system was built, it has been tested in medical institutions across the city and achieved good results.

D. 5G Communication Technology Pushes Medical Industry to New Heights

5G is the abbreviation of the fifth generation mobile communication technology. However, unlike 4G, 3G, and 2G, 5G is not an independent, new wireless access technology, but a Including 2G, 3G, 4G and WiFi) technology evolution, as well as some new supplementary wireless access technology after the integration of the general term[5].

Compared with the previous generations of networks, most of which focus on the characteristics of communication technology itself, the advantages of 5G networks have brought the high-speed, safe and free communication between people, people and things, and things. Smart development in industries with higher requirements, such as the medical industry, offers the possibility. This is mainly to apply 5G's high-speed, low-latency and high-capacity features to improve first aid, remote diagnosis and treatment, and personalized medical treatment.

Facing the opportunities brought by 5G and Internet of Things technology, the Chinese government is actively advancing and continuously formulating relevant policies. 5G represents a brand-new digital medical network, and is likely to enhance the patient's medical experience. It can help users maintain health through three major capabilities: medical Internet of Things (IoMT), enhanced mobile broadband (eMBB) and mission-critical services. When these three are brought together, they can provide users with comprehensive and personalized services anytime, anywhere.

V. CONCLUSION

The medical problems of the masses are always issues that the country attaches importance to. The

medical level of a country determines the political foundation and ruling foundation of the country, and reflects the comprehensive national strength and the security of people's livelihood. The development of China's medical industry is at a critical stage. With the promotion of relevant policies and the support of the new future network IPV9 technology, Internet technology is bound to become the core of medical development. This paper analyzes the problems, characteristics and development status of China's medical industry and introduce of the Tai'an Health big data platform based on IPV9. IPV9 is independently developed by our country, has independent intellectual property rights, and can achieve low cost and high efficiency. It is the core key technology of the next generation Internet. It is believed that with the support of these technologies, the new Internet technology will establish a set of targeted systems that meets my country's national conditions, allowing my country's medical treatment to develop on the right path at high speed and high quality.

REFERENCES

[1] Akemi Matsuzawa, Yuko Shiroki, Junichi Arai, Akemi Hirasawa. Care coordination for children with medical complexity in Japan: Caregivers' perspectives[J]. Child: Care, Health and Development, 2020, 46(4).

[2] T Jencsik, I Boncz,A Pakai,M Gyuró,T Csákvári, D Endrei,Z Horváthné Kövés, R Vajda. Influenza-Related Attitudes of Healthcare Workers At Institutions for Acute and Chronic Diseases[J]. Value in Health, 2017, 20(9).

[3] Xie Jianpingetc.A method of assigning addresses to network computers using the full decimal algorithm[P]. CN: ZL00135182.6, 2004.2.6.

[4] Husam Barham, Tugrul Daim. The use of readiness assessment for big data projects[J]. Sustainable Cities and Society, 2020, 60.

[5] Jae H. Jahng,Seung K. Park. Simulation-based prediction for 5G mobile adoption[J]. ICT Express, 2020, 6(2).

Research and Implementation of Future Network Router

Sun Huai
[1.] School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
[2.] State and Provincial Joint Engineering Lab. of Advanced Network, Monitoring and Control
Xi'an, China
E-mail: sh1227467868@163.com

Wang Zhongsheng
[1.] School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
[2.] State and Provincial Joint Engineering Lab. of Advanced Network, Monitoring and Control
Xi'an, China
E-mail: wzhsh1681@163.com

***Abstract*—This paper summarizes the process and characteristics of IPV9, and introduces the digital domain name system, route management and security. The tunneling and dual protocol stack technologies are described. Both the transition from IPv4 to IPV9 and from IPV9 to IPv4 will be based on mature network conversion services that support protocol compatibility. The IPV9 router is currently available in two models, 100-megabit and gigabit. Both routers have good ability to meet the network needs of IPV9. The film transmission system of Beijing University of Posts and Telecommunications uses IPV9 and 5G technology, which is a relatively advanced transmission application of IPV9 at present. It summarizes the convenience and security brought by the new network address, and provides a new idea for the development of the network.**

Keywords-Future Network; Router; Network Transformation

I. THE NEW NETWORK IPV9

Internet Protocol (IP) is a communication protocol designed for computers to communicate with each other in the network. IP provides a common rule for computers to access the Internet. The Internet has become the largest open network in the world. With the rapid development of the global economy, the advancement of communication technology and network technology, the penetration rate of computers and mobile terminals is getting higher and higher. The problems with IPv4 are also exposed. For example, in the address space, performance, network security and routing bottlenecks, IPv4 makes it difficult to meet the needs of the Internet in the future. To solve the IPv4 many problems, IPv6, IPV9 and other Internet protocols have been born.

A. The Production of IPV9

The new network covers three new technologies: address coding design, new addressing mechanism and new address architecture design[1]. It aims to build a core technology system based on the underlying IP network. On this basis, a new framework can be formed. Connected and compatible with a network system that covers existing networks (Internet with IPv4 and IPv6 technologies). 2011 US government agency has the authority of the professional and technical confirmation from the law, my country has IP framework with the United States Internet network to the prior art, proprietary technology core network sovereignty. This is the patented technology of IPV9 (Method of using whole digital code to assign address for computer). The official patent name is “the method of allocating addresses to computers using full digital coding”.

The IPV9 protocol refers to the 0-9 Arabic digital network as the virtual IP address, and uses decimal as the text representation method, which is a convenient way to find online users[2]. In order to improve efficiency and facilitate end users, some of the addresses can be directly used for domain name. At the same time, it is also called “new generation security and reliable information integrated network protocol”. It uses the classification and coding of the original computer network, cable radio and television network and telecommunication network.

B. The Characteristics of IPV9

By using IPV9 routers, clients, protocol conversion routers and other devices to build a pure IPV9 network, IPV9/IPv4 hybrid network to achieve a new generation of Internet systems with independent and secure intellectual property rights. Including the domestically controllable IPV9 future network root domain name system, promote technology convergence, service integration, data convergence, and achieve cross-level, cross-regional, cross-system, cross-department, cross-business collaborative management and services. With the data concentration and sharing as the way, we will build a national integrated national big data center, accelerate the promotion of domestically-controlled independent control alternative plans, and build a safe and controllable information technology system.

In the existing TCP/IP protocol, conventional packet switching cannot support true real-time applications and circuit switching, and supports applications such as transmitting sound or images in circuits in a four-layer protocol. With the demand for voice, image and data triple play, the incompatibility of human-machine interface and the environmental protection requirements for redundant links, especially the original security mechanism is unreasonable, it is imperative to establish a new network theory foundation. So in 2001 , China established the Decimal Network Standard Working Group to study and implement security-based first-come-authentication communication rules, address encryption, as short as 16 bits up to 2048 bits of address space, resource reservation, virtual real circuit The communication network transmission mode, such as character direct addressing and three-layer four-layer hybrid network architecture.

The existing TCP/IP protocol is a unreliable packet protocol with a maximum packet length of 1514 bytes. The TCP/IP/M protocol of IPV9, which is led by China, not only inherits the unreliable packet protocol of the existing TCP/IP protocol, but also develops absolute code stream and long stream code[3]. The data packet can reach tens of megabytes or more. After three can be transmitted directly by telephone and cable television data link is established without affecting the existing transmission network until four transmission new transmission theory until they have finished the removal of three of four transport protocol.

1) Digital Domain Name System

In the digital domain name system, IPv4 and IPv6 are domain name resolutions through the United States, while IPV9 is set by countries, which avoids the limitation of IP addresses and reduces the use of domain names by the state[4]. IPV9 is a "decimal network" with independent intellectual property rights developed according to the invention patent "Method of Allocating Addresses for Computers Using All Digital Encoding". Its decimal network introduces a digital domain name system, which can be used to convert the original binary through a decimal network. The address is converted into decimal text, allowing the computers on the network to connect to each other, to communicate and transmit data to each other, and to be compatible with Chinese and English domain names.

The digital domain name technology used by the IPV9 decimal network reduces the difficulty of network management, the vast address space and the newly added security mechanism, and solves many problems faced by the existing IPv4. The advantages of

other aspects can also meet the different levels of demand for various devices in the future.

2) Routing

In terms of routing, the increase in the size of the Internet has caused the IPv4 routing table to swell, making the efficiency of network routing declining. The emergence of IPV9 solves this problem, and the optimization of routing improves the efficiency of the network. IPV9 establishes an IPV9 tunnel between the mobile unit and the proxy , and then relays the data packet sent to the mobile unit's home address received by the "proxy" used as the mobile unit to the current location of the mobile unit through the tunnel, thereby implementing Network terminal mobility support.

IPv6 has a smaller routing table than IPv4. IPv6 addresses are assigned according to the clustering principle, which enables the router to represent a piece of network with one record in the table, reduces the length of routing table in the router, and improves the speed of forwarding packets from the routing table. The address allocation of IPV9 follows the principle of spatial clustering, which enables the IPV9 router to represent a national network and an application network with one record, greatly reducing the length of routing table in the router and improving the speed of forwarding packets by routing table. At the same time, this network can express a specific geographical location[5]. According to this logic, only one route is needed between countries. For example, the route to China is 86. The IPv4 routing table is large and irregular, and the IPv6 routing table is smaller than the IPv4 routing table, but the IPv6 routing table contains no geographic information and the routing is messy.

3) Security

IPV9 encryption technology and authentication technology have significantly improved than IPv4, and the encryption technology proposed by IPV9 is difficult to decipher at the physical level, and the confidential performance has been significantly improved. However, at the level of network information security, there are still many factors that cause insecure network information in China. The fundamental reason is that the root servers of IPv4 and IPv6 are in the United States. Many patents related to the network are in the hands of the United States. At the same time, the risk of information exposure may also be introduced. The IPV9 is to have independent intellectual property rights of Internet Protocol, can bring a lot of protection to the information security of the country[6]. IPV9's address space enables end-to-end secure transmission, making it possible for people to use devices to directly assign addresses. Both IPv4 and IPv6 do not have the concept of national geographic location. Most of their domain name resolution servers are in the United States, and IPV9 proposes the concept of "sovereign equality", which enables each country to have its own root domain name system.

II. Introduction to Network Access Technology

IPV9 has a huge IP address resource space, which not only completely solves the current situation of IPv4 address resource shortage, but also far superior to IPv4 network in terms of the number of IP addresses it can use. Due to the large scale of the current IPv4 protocol, no matter which protocol is in use, it is impossible to fully replace IPv4 in a short period of time. It must go through a cyclic and gradual replacement process. Therefore, the problem of transition mechanism should be considered.

In order to successfully complete the process of IPV9 protocol replacing IPv4 protocol, the first consideration is to deal with the relationship between the existing IPv4 network and the future IPV9 network. The problem to be solved is how to achieve a smooth transition of IPV9 network, so that it can solve the problem of IP address shortage in a short time. In fact, the solution of the transition mechanism problem can promote the application of IPV9, which is of great

significance to whether it can become the next generation Internet protocol or just a LAN protocol.

A. Tunnel Technology

Tunneling technology is a technology that converts the data gram in IPV9 format into the data gram in IPv4 format and finally transmits it in the routing system of IPv4 network. A tunnel has a tunnel entrance and a tunnel exit[7]. There is only one tunnel entrance and one tunnel exit. First, at the tunnel entrance, the IPV9 data gram is transformed and processed, and the data information is parsed and encapsulated into a data gram in IPv4 format. The processed data gram is then transmitted along the virtual link identified by the tunnel marker. When the data gram arrives at the exit of the tunnel, it is handed over to the IPv4 protocol. According to the corresponding protocol, the data gram is processed, along with the field value, which is the value of the next header field of the article. After the processing, if the tunnel protocol value can still be detected, the IPv4 header will be discarded and the data gram will be unsealed to obtain the destination address in the original IPV9 message, and then the message will be sent to the original address according to the destination address. At the same time, the transmitted data gram is processed using the IPV9 protocol. During the transition from IPv4 to IPV9, tunneling enabled IPV9 node communication by using existing IPv4 networks. Figure 1 is a schematic diagram of IPV9 data gram encapsulation.

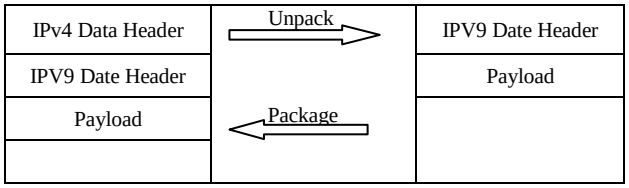


Figure 1. Schematic diagram of IPV9 data gram packaging

B. Double Stack Technique

Dual stack refers to the fact that a single node supports both IPv4 and IPV9 protocol stacks at the same time. Such a node can directly communicate with IPv4 nodes based on IPv4 protocol, or with IPV9 nodes based on IPV9 protocol. Therefore, it can serve as a connection point between IPv4 network and IPV9 network, and such a node is IPV9/IPv4 node mentioned earlier. Since the new IPV9 protocol stack is mainly aimed at the original IPv4 protocol stack network layer part of the major changes, for the transport layer and other layers above the basic changes, IPV9/IPv4 nodes are usually implemented using a double IP layer structure. This is shown in figure 2.

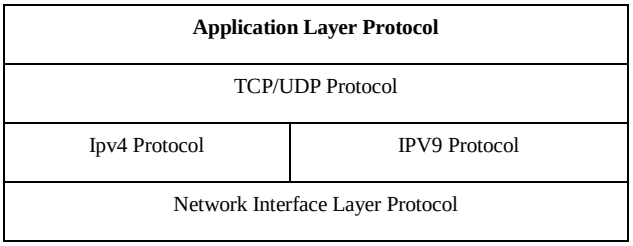


Figure 2. Double IP layer structure diagram

For the router, however, because of the large changes have taken place in IP protocol, and IP routing protocol that is close to the corresponding changes have taken place goes, so "double stack" router is refers to in a router equipment maintenance able and IPv4 two routines by the protocol stack, so that half of the router can also can communicate with host able with IPv4 hosts, respectively support independent able and IPv4 routing protocol, IPv4 and able routing information routing protocols to calculate, according to their different routing table maintenance. The routing table obtained by IPV9 data gram is forwarded according to the routing protocol of IPV9, and the routing table obtained by IPv4 data gram is forwarded according to the routing protocol of IPv4. The router protocol structure that supports IPV9 and IPv4 dual protocol stacks is shown in figure 3.

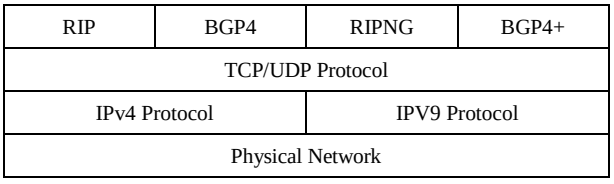


Figure 3. Protocol structure of dual stack router

The network using the dual-stack technology does not have the problem of inter working, so it has certain convenience. However, it needs to assign an IPv4 address to each IPV9 node, which will lead to the problem of IPv4 address resource strain. In addition, every IPV9/IPv4 nodes to run at the same time IPv4 and able two kinds of protocol stack, at the same time save two sets of command set, at the same time calculation, maintenance and storage of two list items, for gateway device also need to two message transformation and encapsulation protocol stack, this undoubtedly and increase the load of each node, higher demands on the performance of these nodes. In addition, DNS servers must support the mapping of host domain names to IPV9 addresses on a dual-stack network.

III. INTRODUCTION TO IPV9 ROUTING EQUIPMENT

With the expansion of the new generation Internet routing system in China, the new generation Internet routing system has been deployed in Beijing, Shanghai, Chongqing, northeast China, Sichuan, Xinjiang, Shandong, Guangdong, Zhejiang, Hong Kong, Macao and other places. Manual configuration of address and route has not been able to meet the current development form, in the configuration, inefficient, complex and other defects have been gradually exposed. In order to cope with the huge routing node management and configuration, an efficient and automated configuration system is needed to handle the allocation and configuration management of the new generation of Internet addresses.

A. 100-megabit Router

The whole machine of the megabit router is shown in figure 4, including two signal antennas and the main body of the machine.

Figure 4. Whole figure

The panel on the gigabit router is shown in figure 5. The panel includes network connection indicator light, WIFI signal light, USB connection signal light and power signal light.

Figure 5. Panel figure

The warning light of the gigabit router is as follows. When the red light of the power indicator is on, the power supply is normal. When the blue light of WAN indicator flashes, WAN is normal. The blue light of the LAN indicator flashes when the LAN is normal. When the blue light of WIFI indicator flashes, WIFI is normal; USB is normal when the blue light of USB indicator flashes. The prompt light diagram is shown in table 1.

The front panel is shown in figure 6. The front panel has USB interface and TF card slot. USB interface can be used to connect USB mouse, keyboard, USB disk and other devices. The TF card slot is used to insert the TF card.

TABLE I. LIGHTS TABLE

Indicator Light	Description	Features
	Power Supply	Red light on : Normal
	WAN	Blue light flashing : Normal
	LAN	Blue light flashing : Normal
	WIFI	Blue light flashing : Normal
	USB	Blue light flashing : Normal

Figure 6. Front panel figure

The back panel is shown in the figure 7. The back panel has a DC port, which is used to connect the power supply to the router. Note that using a mismatched power supply can cause damage to the router. The RESET button is the restore button. Press to restart the device. Long press for about 5 seconds to restore the device to the factory default Settings. When the system status indicator changes from slow flicker to fast flicker, it means that the router has successfully resumed the factory setting. At this time, release the RESET button and the router will restart. The WAN port is the WAN port jack. This port is used to connect Ethernet cables. LAN4, LAN3, LAN2 and LAN1 ports are: LAN port jack. This port is used to connect to a hub, switch, or computer with a network card installed on the LAN.

Figure 7. Back panel

B. Gigabit Router

The gigabit router's whole machine is shown in the figure 8, including six signal antennas and the main body of the machine. The six signal antennas include four 2.4g antennas and two 5G antennas.

Figure 8. Whole figure

The front panel is shown in the figure 9, and the back panel has a power port, which is used to connect the power supply to power the router. Note that using a mismatched power supply can cause damage to the router. The RESET button is the restore button. Press to restart the device. Long press for about 5 seconds to restore the device to the factory default Settings. When the system status indicator changes from slow flicker to fast flicker, it means that the router has successfully resumed the factory setting. At this time, release the RESET button and the router will restart. The WAN port is the WAN port jack. This port is used to connect Ethernet cables. LAN4, LAN3, LAN2 and LAN1 ports are: LAN port jack. This port is used to connect to a hub, switch, or computer with a network card installed on the LAN. USB interface can be used to connect USB mouse, keyboard, USB disk and other devices. The TF card slot is used to insert the TF card.

IV. IPV9 Routing Setup Procedures and Methods

A. Login Router Configuration Interface

Enter 192.168.1.1 in the browser address bar as shown in the browser login interface below, and press enter to enter the system login interface as shown in figure 10.

Enter the default user name root and the default password shsjzwlxxkjyxgs in the system login interface as shown in figure 11. Click the login button and you will see the overview interface as shown in figure 12.

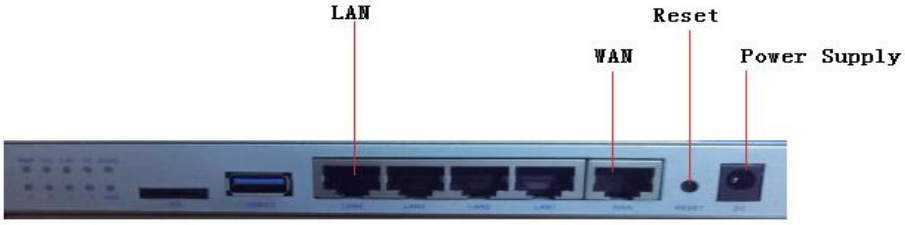


Figure 9. Front panel figure

Figure 10. Browser login

Figure 11. System login interface

Figure 12. IPV9 overview interface

B. IPv9 User Registration

Users who use the system for the first time should register before using the router function. User registration is shown in figure 13 .The user types are divided into individuals and enterprises. After the personal router registers the device, the address is automatically assigned without manual allocation. The enterprise router manually assigns the address, and one enterprise account can register multiple devices. Before

you register, you need to send the SMS verification code and enter the SMS verification code.

1) Click the menu IPv9 drop-down key to select the user registration function.

2) Enter the user name, password, confirmation password, registration type, real name, certificate type, certificate number, E-mail, mobile phone, enterprise name, address, postal code and remarks according to the text prompt of user registration.

3) Enter your phone number and click the send button on the right.

4) After the phone receives the verification code, fill in the verification code.

5) Click the user registration button below.

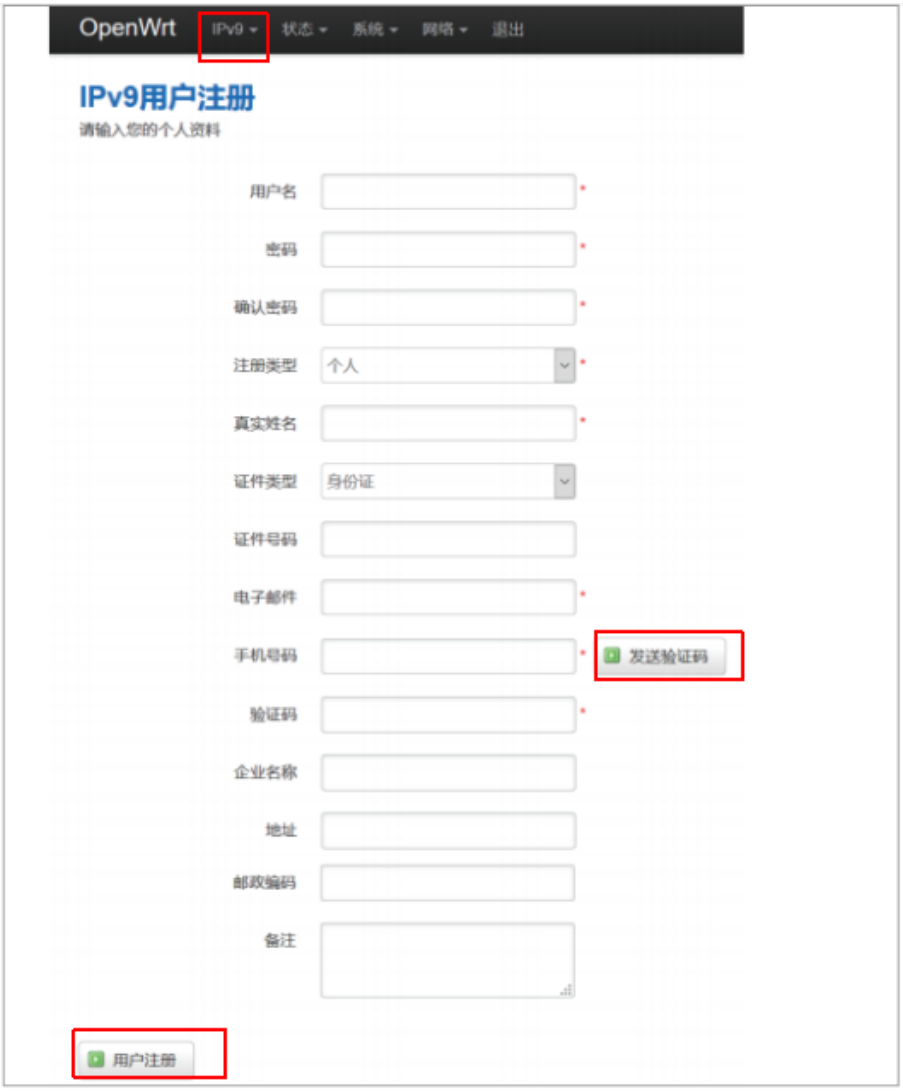


Figure 13. IPV9 overview interface

C. Equipment registration

When the individual user registers the device, not only the device information is registered, but also the IPv4/IPv9 address is automatically assigned to the device, while the enterprise user only registers the device information. The device registration is shown in figure 14.

1) Click the menu IPv9 drop-down button to select the device registration function.

2) Choose your city.

3) Click the device registration button below.

D. Configuration of WIFI

Click the network to select WIFI and enter the WIFI configuration interface.

In figure 15, you can see the working state of WIFI. If you need to modify the configuration of WIFI, click modify.

Modify the name of WIFI as shown in figure 16.

Figure 14. IPV9 user device registration

Figure 15. Configuration of WIFI

Figure 16. IPV9 user device registration

Configure the device WIFI encryption method. Generally, it is recommended to use the encryption mode shown（WPA_PSK/WPA2-Psk Mixed Mode） above to enter the new password and click "save & apply" button. This is the normal router setup, after which a new router is configured. The encryption configuration is shown in figure 17.

Figure 17. Encryption configuration

V. IPV9 AND 5G MOVIE TRANSMISSION SYSTEM

Able movie network issuance application of Beijing Unicom 5G network and China mobile company in Suzhou 5G network has already passed the Beijing university of posts and telecommunications able fiber routing backbone nodes and able backbone fiber optic cable connected directly across the country, and on May 21, 2019, the first in the world at 500 MBPS end-to-end to 1000 MBPS speed, in local access able the national backbone network, successfully carried out the digital film distribution network (each film data capacity in hundreds of GB or so). The national online distribution of Chinese films was the first in the world to enter the new era of "one hour", and the intelligent cinema based on IPV9 address was realized.

The IPV9/IPv4 router is a dual protocol stack router. By building this dual protocol stack router, the IPV9 network can be realized and the IPv4 network can be compatible with IPV9, thus achieving a good transition from IPv4 to IPV9. Through the configuration of different interfaces of routers, the conversion between different protocols can be realized to realize the data transmission of pure IPV9 protocol packet, IPV9 over IPv4 and IPv4 over IPV9, while IPV9 over IPv4 and IPV9 over IPV9 are realized by tunnel technology, which is an important way to realize the intercommunication between IPv4 and IPV9 networks.

A. IPV9 Private Network Transmission

Figure 18 shows the transmission of pure IPV9 private network. An IPV9 and IPV4 router is directly set up in the state administration of film and the cinema, which is connected through the IPV9 private network. The router interface is configured as the IPV9 transmission mode, so the line between the state administration of film and the cinema is the pure IPV9 protocol transmission. The network between the two is full of IPV9 protocol packets, which can guarantee absolute security.

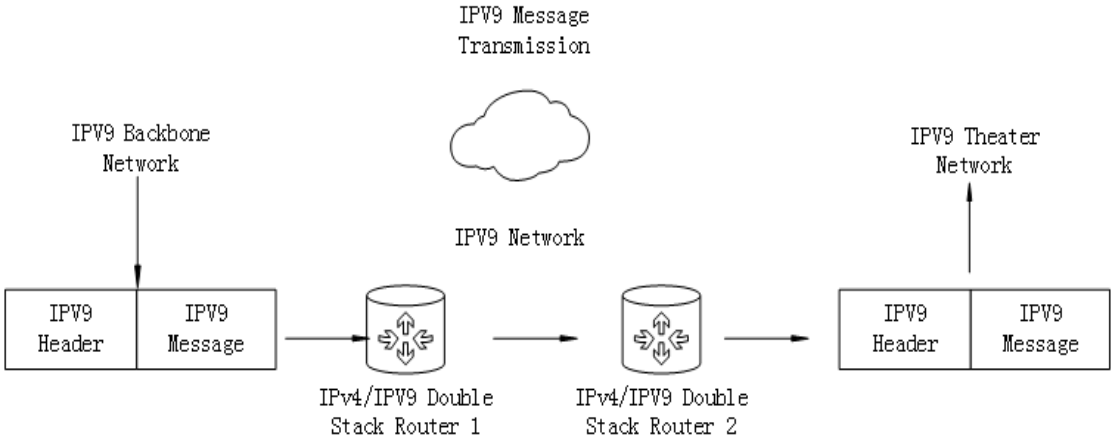


Figure 18. IPV9 private network transmission

B. IPV9 Private Network Tunnel Transmission

IPv4 over IPV9, IPV9/IPv4 router can be compatible with IPv4 network, namely IPv4 packets can be transmitted on the able private network, realize the IPv4 over able, to design a scheme of transmission as shown in figure 19. Both ends set up double protocol stack router, the router for private network between. The advantage is the backbone adopts able transmission protocol, security can get reliable guarantee.

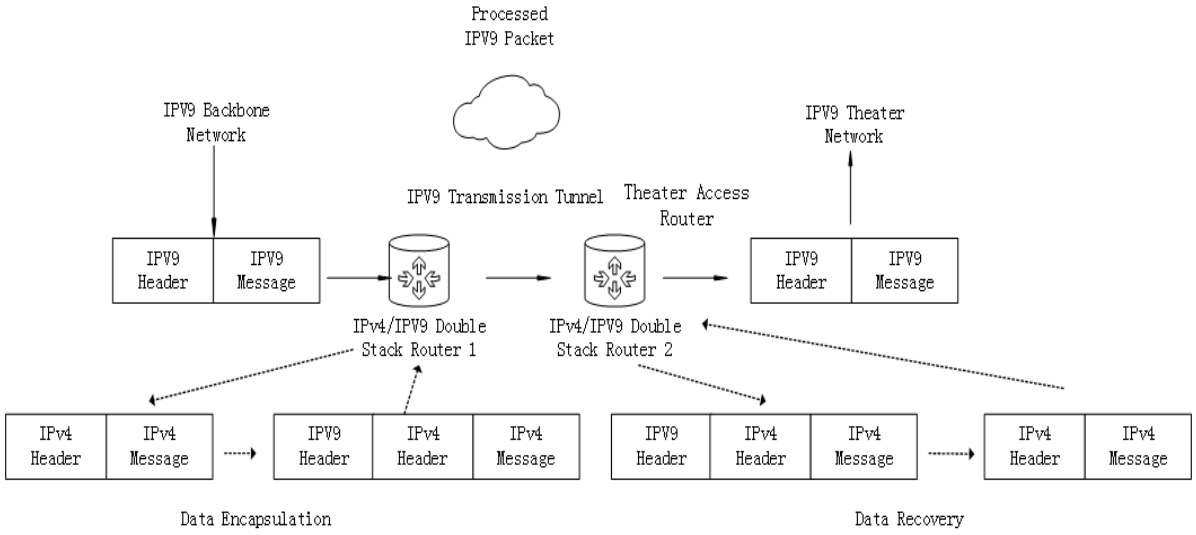


Figure 19. Private network tunnel transmission

The transmission process of IPV9 based IPv4 packets is as follows:

1) Router I received IPv4 packets from the IPv4 network.

2) *Router 1 encapsulates IPv4 packets in IPV9 message.* The source address and destination address of IPv9 message correspond to the entrance address and exit address of the tunnel respectively, that is, the IPV9 address of router 1 and router 2.

3) The encapsulated IPV9 packets are transmitted along and through the marked tunnel link, routed to the IPV9 address, and arrive at the destination router 2 to complete the transmission of IPV9 private network line.

4) Router 2 receives the IPV9 packet from router 1, unseals the IPV9 packet to obtain the original IPv4 packet, and then sends it out.

VI. CONCLUSION

With the development of the Internet and the increasing number of Internet users, the shortage of IP address resources has become a bottleneck restricting its development. The application of IPV9 is spreading in China, especially in the government, Banks and other sectors. IPV9 has a large address capacity, is compatible with IPv4 and IPv6, and uses special encryption mechanisms to make the network environment more secure. This paper discusses the IPV9 address architecture and the digital domain name system, discusses the different network access technologies in detail, introduces the megabit and gigabit routers that support the use of new networks, and discusses the configuration methods of routers in detail. This paper analyzes the film transmission system of Beijing university of posts and telecommunications at the network level, which is very important for the future deployment of IPV9 network.

REFERENCE

[1] Xie Jianping etc. Method of using whole digital code to assign address for computer [P]. US: 8082365, 2011.12.

[2] Xie Jianping etc. A method of assigning addresses to network computers using the full decimal algorithm [P]. CN: ZL00135182.6, 2004.2.6.

[3] V. Fuller, T. Li,Network Working Group. Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy, RFC-1519, 1993.9.

[4] Kohler E, Li J, Paxson V, et al. Observed Structure of Addresses in IP Traffic [J]. IEEE/ACM Transactions on Networking, 2006, 14(6):1207-1218.

[5] Information technology-Future Network- Problem statement and requirement-Part 2: Naming and addressing, ISO/IEC DTR 29181-2, 2014, 12.

[6] Xie Jianping etc. Digital domain name specification, SJ/T11271-2002, 2002.07.

[7] M. Crawford. Network Working Group.Transmission of IPv6 Packets over Ethernet Networks.RFC-2464, 1998.12.

Image Inpainting Research Based on Deep Learning

Zhao Ruixia
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 1364343954@qq.com

Zhao Li
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 332099732@qq.com

***Abstract*—With the rapid development of computer technology, image inpainting has become a research hotspot in the field of deep learning. Image inpainting belongs to the intersection of computer vision and computer graphics, and is an image processing technology between image editing and image generation. The proposal of generative adversarial network effectively improves the problems of poor image inpainting effect and large difference between the inpainting image and the target image, and promotes the development of image inpainting technology. In this paper, the image inpainting is based on the generation of confrontation networks. Its network structure establishes two repair paths, namely the reconstruction path and the generation path, and the two paths correspond to two groups of networks. The encoder and generator in the network respectively complete the encoding and decoding tasks based on the residual network. The discriminator also uses the patch block discriminator on the basis of the residual network to discriminate the authenticity of the image. This paper uses Places2 data set to verify the algorithm, and uses PSNR and SSIM two objective evaluation methods to evaluate the quality of the repaired image. Experiments show that the algorithm inpainting effect is better.**

Keywords-Image Inpainting; Generation Adversarial Networks; Residual Network; Patch

I. INTRODUCTION

With the development and popularization of computer technology, Internet technology and multimedia technology, digital image processing technology has also developed rapidly. In the process of storage, transmission and use of digital image information, the phenomenon of image information damaged and loss will occur. These damaged areas affect the visual effect of the picture and the integrity of the information, and have a certain impact on the application of the picture. People urgently need a technology and method that can automatically inpainting damaged digital images, so digital image inpainting technology is born.

Image inpainting is one of the most popular areas of deep learning. Its basic principle is to give an image of a damaged or corroded area, and try to use the intact information of the known area of the damaged image to inpainting the damaged area of the image[1-2]. Digital image inpainting methods can be divided into two major categories: traditional image inpainting methods and deep learning-based image repair methods. Traditional image repair methods can be divided into: structure-based image repair technology and texture synthesis-based image inpainting technology. Both image inpainting algorithms based on structure and texture can inpainting the loss of small areas such as folds. With the expansion of the missing areas, the inpainting effect gradually deteriorates. There are problems such as incomplete semantic information and blurred images in the inpainting results, which makes the image inpainting effect ineffective, ideal. The emergence of deep neural networks allows the model to obtain the understanding of image semantic information through multi-level feature extraction, and to a certain extent improves the repair effect of large-area damaged images.

As deep learning shows exciting prospects in the fields of image semantic inpainting and situational awareness, and image inpainting algorithms based on deep learning can capture more advanced features of images than traditional inpainting algorithms based on structure and texture, so often used for image inpainting. At present, image inpainting based on generative adversarial networks is a major research hotspot in the field of deep learning image inpainting, which lays a solid foundation for the development of image inpainting technology.

DOI: 10.21307/ijanmc-2020-013

A. The basic idea of generating adversarial networks

Generative adversarial network is undoubtedly one of the popular artificial intelligence technologies, and was rated as the "Top Ten Global Breakthrough Technologies" in 2018 by the MIT Technology Review. The generative adversarial network is composed of a generative network and a discriminant network. The purpose of the generative network is to estimate the distribution of data samples from a given noise and generate synthetic data. The purpose of the discriminant network is to distinguish the input data from the generated data or the real data. The generative network and the discriminant network are a set of confrontational relationships. The source of the confrontational ideas comes from the zero-sum game in game theory. The two sides of the game use each other's strategy to change their confrontation strategy in an equal game, so as to achieve the goal of winning[3]. It is extended to the generative antagonistic network, that is, the generative network and the discriminant network are the two sides of the game. The optimization goal is to achieve Nash equilibrium[4], the generative network tries to produce closer to real data. Accordingly, the discriminant network tries to distinguish more perfectly between real data and data generated by generators. As a result, the two networks progressed in confrontation, and continued to confront each other after the progress, the data obtained by the generating network became more and more perfect, approaching the real data.

B. Development of deep learning models

GeneratingSince the input of the GAN generation model is random noise data, in actual applications, there are generally clear variables to control the category or other information for the data to be generated, such as generating specific numbers from 1 to 9 numbers. In order to solve the problem of generating labeled data, Conditional Generative Adversarial Networks are proposed, and information such as category labels and pictures are added to the input to make the image more in line with the target[5]. The foundation of image inpainting technology based on deep learning is the convolutional neural network, which uses the convolutional neural network to extract high-dimensional features and information prediction, which makes the image inpainting technology develop rapidly[6-7]. Because the network of generating model and discriminating model in GAN is too simple, there will be image blur when generating large-size images. In order to generate clear images, Radford A et al.[8] proposed deep convolutional generation adversarial networks. With the emergence of several unsupervised image conversion models, such as CycleGAN[9], DualGAN[10], DiscoGAN[11], it provides better ideas for image inpainting technology.

II. NETWORK STRUCTURE

Image inpainting not only requires that the results conform to human visual habits, making it difficult for the human eye to detect the traces of inpainting (undetected)[12], meanwhile inpainting the information contained in the missing pictures as much as possible, so that the restored image can be as much as possible Same as the image before the damage. Based on this restoration goal, this paper builds an image inpainting network framework suitable for this article by studying and analyzing the structure principles of GAN.

Using the neural network's ability to extract high-dimensional features of images, the structural framework of this paper is built. In this paper, a parallel dual-path framework based on GAN is used: one is to reconstruct the path, and use the given real image and masked image to obtain its complementary image to reconstruct the original image; the other is to generate the path and use the given masked image to inpainting. The input image of the generated path and the input image of the reconstructed path are complementary images of each other. The network structure is built on the basis of the residual network. Its structure includes three parts: encoder, generating network and discriminating network. The image inpainting process in this paper is: (1)Input the masked image and the complement image (the masked image and the supplementary image are the real image) into the encoders E1 and E2 of the reconstruction path and the generation path to encode; (2)The extracted two image features were fused and input into generator G1 and G2; (3)The generator reconstructed image and the real image are input into the discriminator D1 for discrimination; (4)The generated image, the fused image and the real image are input into the discriminator D2 for discrimination; (5)The discriminators D1 and D2 output the discriminant results and feed them back to the encoder, generator and discriminator through the back propagation algorithm to update the network parameters and train the network. The overall structure of the network is shown in Figure 1.

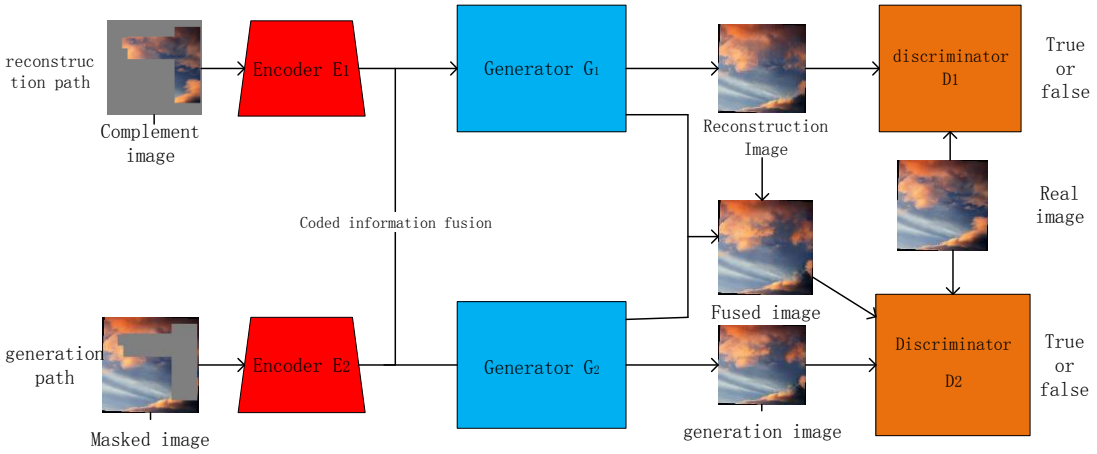


Figure 1. Data flow diagram of GAN

A. *Encoder*

The encoder extracts the features of the image based on the residual network. The inputs of encoders E1 and E2 are three-channel images of 256×256 pixels. The residual block is composed of two layers of convolution and one layer of skip link. The first layer uses a convolution kernel of size 3×3. The length is 1 and the padding size is 1. The second layer uses a 3×3 convolution kernel with a sliding step size of 1 and no padding. The residual structure of the encoder is shown in Figure 2.

In this paper, there are two parallel paths for image inpainting: reconstruction path and generation path. The network structure of the encoder is the same, and the combination of residual modules is used. The network structure contains 7 residual modules. The network structure of the encoder is shown in Figure 3.

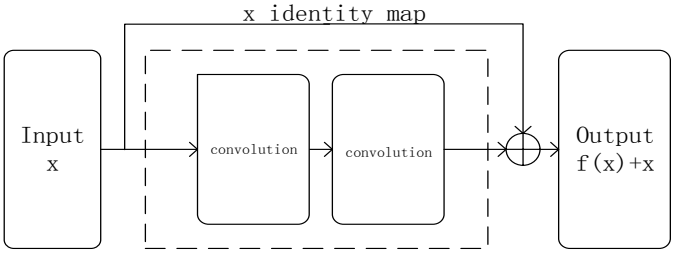


Figure 2. Residual structure of the encoder

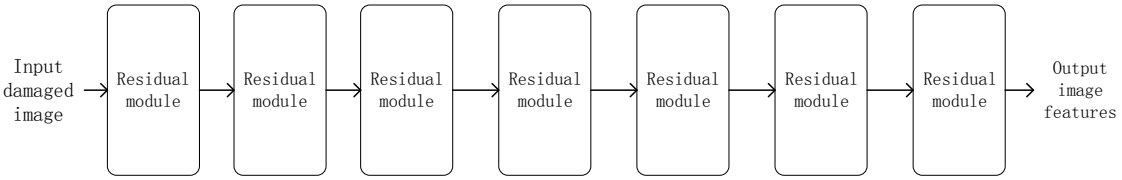


Figure 3. Encoder network structure

B. *Generate network*

The generating network adopts Res-Net network structure, and uses the residual decoding block to decode the features extracted in the encoding stage. In the generation network, the residual block is used in the decoding stage. The residual block in the decoding stage is composed of three parts: a convolution layer, a deconvolution layer, and a skip link layer. The convolutional layer uses a convolution kernel with a

size of 3×3, a sliding step size of 1, and a padding of 1. The deconvolution layer uses a 3×3 convolution kernel with a sliding step size of 2 and a padding of 1. After the deconvolution operation, the padding of the output image is 1. The skip link layer performs a deconvolution operation, using a convolution kernel with a size of 3×3, a sliding step size of 2, and a fill of 1. After the deconvolution operation, the output image has a fill of 1. The generated network uses the Spectral Normalization method to normalize the output data. The network structure of the residual block in the decoding stage is shown in Figure 4.

A self-attention mechanism has been added to the network. The self-attention mechanism uses residual blocks and uses Short+Long Term to ensure the consistency of the appearance of the generated image. The network structure of the generated network is shown in Figure 5.

C. The training principleDiscrimination Network

The discrimination network adopts the structure of PatchGAN. The difference between PatchGAN and ordinary GAN is that the output of ordinary GAN is the evaluation of the entire image, and the output of PatchGAN is an N×N matrix. Each element of the N×N matrix represents the original image. The larger receptive field in the map corresponds to a patch in the original picture. This paper runs a patch discriminator on the image in a convolution mode. The discriminator outputs a patch block of 70×70 size, and each element represents the probability value of the real image. This paper judges that the input of the network is a picture, the target picture is used as a positive example, and the inpainting picture is used as a negative example, so as to judge whether the inpainting picture is true. The discriminators D1 and D2 in this paper have the same network structure and use five-layer convolution. The first three layers use a 4×4 convolution kernel with a sliding step size of 1 and a padding of 1; the last two layers use a 4×4 convolution kernel with a sliding step size of 2 and a padding of 1. The discriminant network first extracts the features of the input image, and then analyzes and compares the extracted features. The network structure of the discrimination network is shown in Figure 6.

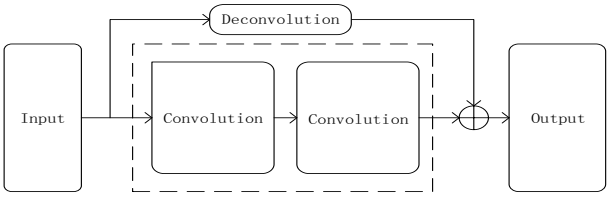


Figure 4. Decoding residual block network structure

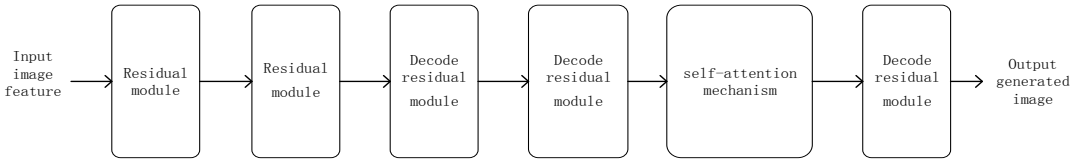


Figure 5. Generate network structure diagram

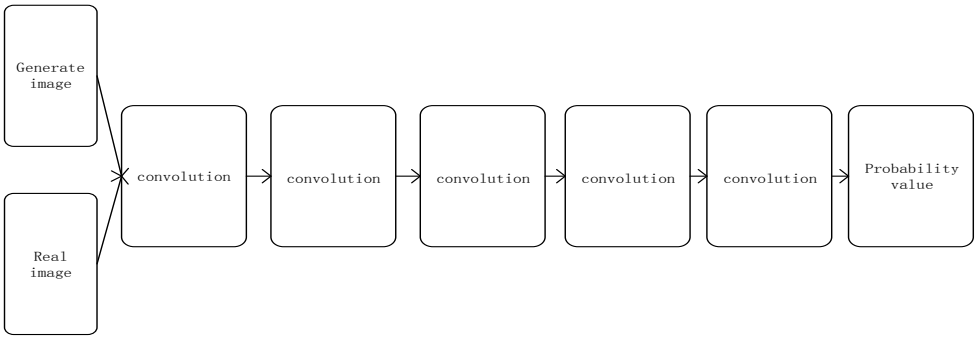


Figure 6. Discriminant network structure diagram

III. NETWORK TRAINING

In this paper, WGAN-GP loss is used to optimize the network structure. WGAN-GP is an improvement of WGAN. A gradient penalty method is proposed to improve the continuity constraint conditions, making GAN convergence more stable. The loss function of WGAN-GP is composed of the loss LG of the generator and the loss LD of the discriminator. The calculation formula of generator loss can be written as

$$\begin{aligned} L_D^{WGAN} &= E\left[D(x)\right] - E\left[D\left(G(z)\right)\right] + L_{gp} \\ L_{gp} &= \lambda E\left[\left(\left|\nabla D\left(\alpha x - \left(1-\alpha G(z)\right)\right)\right| - 1\right)^2\right] \\ L_D &= L_D^{WGAN} + L_{gp} \end{aligned} \tag{1}$$

Where x represents a randomly selected sample in the data set and $D(x)$ represents the result output when the input of the discriminant model is a real sample. L_D^{WGAN} Represents the loss function corresponding to the WGAN discriminator, Lgp represents the gradient penalty loss function newly added in WGAN-GP, and λ represents the penalty coefficient.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experimental environment

In order to verify the effectiveness of the algorithm proposed in this article, on the Ubuntu platform, the Python language and the PyTorch deep learning framework are used. Experiment with 5000 images of Place2, a public data set. The image size is 256×256 pixels, and the ratio of 8: 2 is used for training and testing.

B. Experimental results

Since the image inpainting task is to repair the incomplete part of the image, the data set should be mask processed before the inpainting task. In this paper, the image preprocessing is divided into two methods: random masked and intermediate masked. After the data processing is completed, the image inpainting task is performed.

The inpainting result of occlusion in the image is shown in Figure 7. Where (a) represents the damaged image, (b) represents the inpainting image, and (c) represents the real image.

The inpainting result of random masked in the image is shown in Figure 8. Where (a) represents the damaged image, (b) represents the inpainting image, and (c) represents the real image.

C. Experimental analysis

At this stage, there are mainly two kinds of image evaluation methods: subjective evaluation method and objective evaluation method. This article combines the subjective evaluation method and the objective evaluation method to evaluate the repaired image.

1) Subjective evaluation

From the experimental results of 4.2, it can be seen that the content of the image inpainting by this method is basically the same as the target image, the color is very similar to the target image, and direct visual observation of the image is real and natural. The inpainting of texture is natural and continuous.

2) Objective evaluation

The objective evaluation method uses peak signal-to-noise ratio measurement (PSNR) and structural similarity (SSIM) to evaluate the repaired image. The higher the PSNR, the less distortion in the picture inpainting process, and the better the inpainting picture. SSIM measures the similarity of the two images. A higher value indicates that the two images are more similar. The maximum value is 1. The definition of peak signal-to-noise ratio, the expression is:

$$\begin{aligned} \mathrm{MSE} &= \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}(I_0(i,j) - I(i,j))^2}{M \times N} \\ \mathrm{PSNR} &= 10\log\left(\frac{G_f^{\ 2}}{MSE}\right) \end{aligned} \tag{2}$$

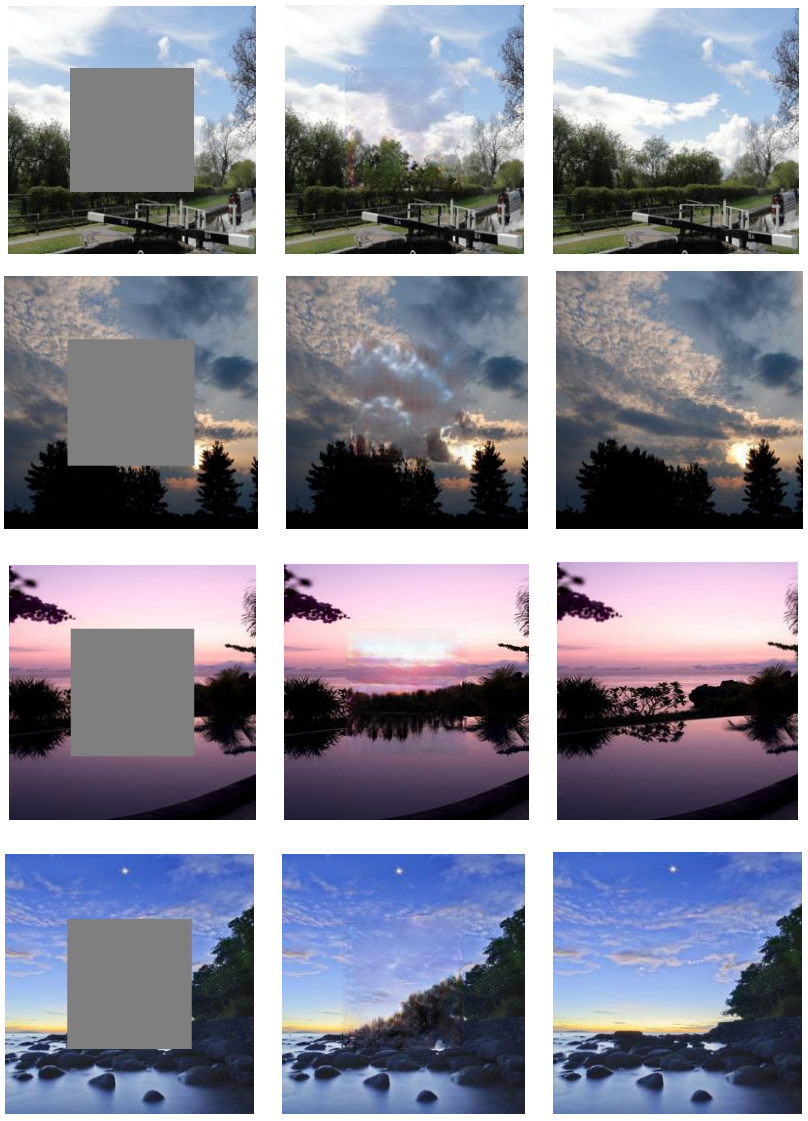
MSE is the mean square error. The default value is 255, $I_0(i,j)$ represents the pixel value at (i,j) in the real image, $I(i,j)$ represents the pixel value at (i,j) in the inpainting image, and M * N represents the area size of the inpainting image.

The definition of structural similaritycan be written as

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^{\ 2} + \mu_y^{\ 2} + C_1)(\sigma_x^{\ 2} + \sigma_y^{\ 2} + C_2)} \tag{3}$$

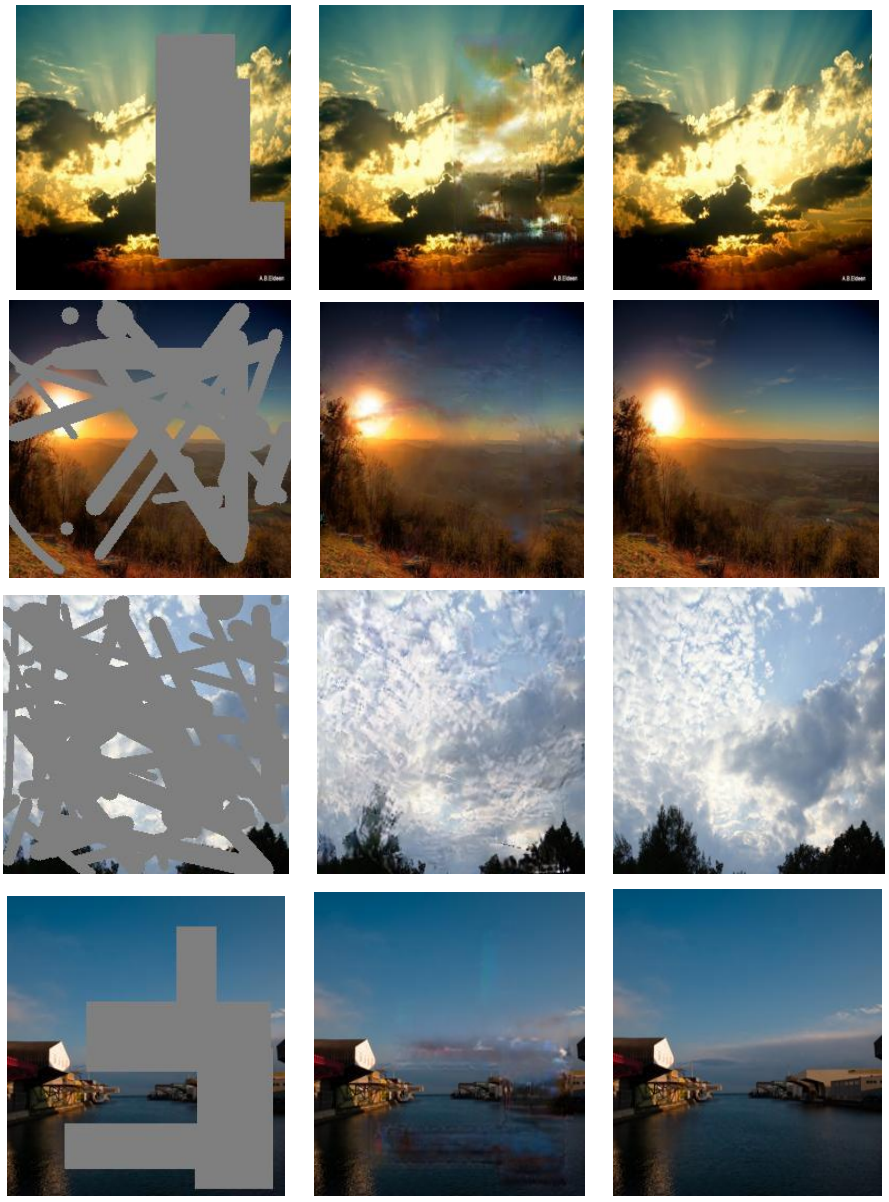
x and y represent the two input images, where μ_x is the average of x, μ_y is the average of y, σ_y^2 is the variance of x, σ_y^2 is the variance of y, σ_{xy} is the covariance of x and y, and C_1, C_2 are Used to maintain a stable constant. L is the dynamic range of pixel values, generally taken as 255.

This paper compares four different image inpainting models, using PSNR and SSIM methods to evaluate.



(a)damaged image (b)inpainting image (c)real image

Figure 7. Inpainting result of intermediate masked



(a)damaged image (b)inpainting image (c)real image

Figure 8. Inpainting result of random masked

TABLE I. EVALUATION RESULTS OF PSNR AND SSIM METHODS

Image inpainting model	PSNR	SSIM
CE[13]	18.72	0.843
GL[14]	19.90	0.836
GntIpt[15]	20.38	0.855
GMCNN[16]	20.62	0.851
Ours	24.06	0.857

V. CONCLUDE AND PROSPECT

In this paper, the image inpainting network structure is built based on GAN. The residual network is used in the encoding and decoding process to reduce the gradient disappearance and gradient explosion problems. Using the loss function of WGAN-GP to update the network parameters to inpainting the image, not only the similarity of the inpainting image structure, but also the matching degree of the image texture. The Place2 dataset is used for network training and testing. The subjective evaluation method and the objective evaluation method are used to evaluate the inpainting image. The objective evaluation method selects SSIM and PSNR to make an objective evaluation of the inpainting image. The comparison between the image inpainting model and the inpainting model of other papers verifies the effectiveness of the algorithm in this paper.

REFERENCES

[1] Bertalmio M, Sapiro G, Caselles V, et al. Image inpainting[C]. international conference on computer graphics and interactive techniques, 2000: 417-424.

[2] Guillemot C, Meur O L. Image Inpainting : Overview and Recent Advances[J]. IEEE Signal Processing Magazine, 2014, 31(1): 127-144.

[3] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.

[4] Ratliff L J, Burden S A, Sastry S, et al. Characterization and computation of local Nash equilibria in continuous games[C]. allerton conference on communication, control, and computing, 2013: 917-924.

[5] Mirza M, Osindero S. Conditional Generative Adversarial Nets[J]. Computer Science, 2014:2672-2680.

[6] Pathak D, Krahenbuhl P, Donahue J, et al. Context Encoders: Feature Learning by Inpainting[C]. computer vision and pattern recognition, 2016: 2536-2544.

[7] Yang C, Lu X, Lin Z, et al. High-Resolution Image Inpainting Using Multi-scale Neural Patch Synthesis[C]. computer vision and pattern recognition, 2017: 4076-4084.

[8] Radford A, Metz L, Chintala S, et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. arXiv: Learning, 2015.

[9] Zhu J, Park T, Isola P, et al. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks[C]. international conference on computer vision, 2017: 2242-2251.

[10] Kim T, Cha M, Kim H, et al. Learning to Discover Cross-Domain Relations with Generative Adversarial Networks[J]. arXiv: Computer Vision and Pattern Recognition, 2017.

[11] Yi Z, Zhang H, Tan P, et al. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation[C]. international conference on computer vision, 2017: 2868-2876.

[12] Efros A A, Freeman W T. Image quilting for texture synthesis and transfer[C]. international conference on computer graphics and interactive techniques, 2001: 341-346.

[13] Pathak D, Krahenbuhl P, Donahue J, et al. Context Encoders: Feature Learning by Inpainting[C]. computer vision and pattern recognition, 2016: 2536-2544.

[14] Iizuka S, Simoserra E, Ishikawa H, et al. Globally and locally consistent image completion[J]. ACM Transactions on Graphics, 2017, 36(4).

[15] Yu J, Lin Z, Yang J, et al. Generative Image Inpainting with Contextual Attention[C]. computer vision and pattern recognition, 2018: 5505-5514.

[16] Wang Y, Tao X, Qi X, et al. Image Inpainting via Generative Multi-column Convolutional Neural Networks[C]. neural information processing systems, 2018: 329-338.

Research on the Application of Convolutional Neural Networks in the Image Recognition

Gu Hongxian
[1.]School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
[2.]State and Provincial Joint Engineering Lab of Advanced Network, Monitoring and Control
Xi'an, 710021, China
E-mail:15757118020@163.com

Gao Zhiyu
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail:516750613@qq.com

Liu Bailin
[1.]School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
[2.]State and Provincial Joint Engineering Lab of Advanced Network, Monitoring and Control
Xi'an, 710021, China
E-mail:498194312@qq.com

Mu Jing
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail:55897293@qq.com

***Abstract*—Over the past few years, benefits from the strong feature extraction advantages of Convolutional Neural Networks (CNN)themselves and the efforts and applicationby researchers making, research work on CNN in the field of image recognition has yielded many results and achieved the best performance in classification and regression tasks. This paper focuses on the improvement and application history ofCNNand summarizes the direction of improvement and optimization of CNN in recent years from the perspective of the structure of CNN themselves and their applications in various fields. Finally, this review is summarized with a further outlook on the development direction of CNN.**

Keywords-Image Recognition; Deep Learning; Machine Learning; Convolutional Neural Networks

I. INTRODUCTION

Since the concept of deep learning was proposed by Hinton et al[1]. In 2006, during more than a decade of development, machine learning is closer to the original goal of "artificial intelligence". Deep learning is a hierarchical machine learning approach that involves multiple levels of nonlinear transformations that learn the inherent laws and representation levels of sample data, and the feature information obtained in the process of learning can help the machine achieve analytical judgments about the data.

Compared to traditional machine learning methods, it has achieved good results in search technology, image recognition, machine translation, natural

DOI: 10.21307/ijanmc-2020-014

language processing, multimedia learning, speech, recommendation and personalization technologies.

With the practice of researchers in various fields, many network models have been proposed, such as DBN (Deep Belief Network), CNN (Convolutional Neural Network), RNN (Recursive Neural Network), etc. The introduction of CNN into the field of image recognition has taken researchers a long time to explore and practice. Image recognition technology originated in the 1940s, when it was not rapidly developed due to inadequate technology and inadequate hardware facilities. It was not until the 1990s that artificial neural networks combined with support vector machines(SVM) facilitated the development of image recognition technology, which was widely used. However, traditional image recognition techniques are based on shallow structural models, which require human pre-processing of the image, resulting in reduced accuracy of image recognition. As computer hardware levels and GPU evolved, researchers began to work on deeper models of network structure, and in 2012, Krizhevsky et al. reduced the error rate of the tested Top-5 to 15.3% in ImageNet's large-scale visual recognition challenge competition based on CNN, 10.9% lower than the error rate of the second-place team's Top-5, showing the great potential of deep models. In the following years, CNN have made leaps and bounds in digital image recognition and processing with their powerful feature extraction capabilities.

II. OVERVIEW OF CNN

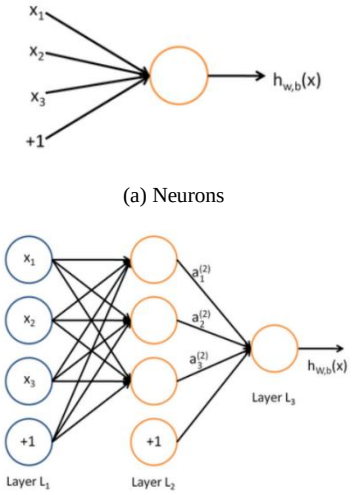
CNN, compared to other network models, are better able to adapt their structures to image structures while extracting features and classifying them, with outstanding performance in image processing. In addition, its weight sharing feature educes the training parameters of the network, which makes the network structure simple and more generalizable, and has become a current research hotspot.

A Development history

The prototype of the CNN is the BP algorithm proposed by Rumelhart in 1986[2].In the 1990s, Lecun proposed the LeNet-5 model[3], which was mainly applied to image classification of handwritten numbers, used the stochastic gradient descent method and reverse propagation method for supervised training of the CNN, and achieved the best recognition results on the MNIST dataset[4], laying the foundation of modern CNN In 2006, Hinton proposed the concept of deep learning in his paper and pointed out that multi-cryptic neural networks have better feature learning capabilities and their complexity in training can be effectively mitigated by layer-by-layer initialization[1]. In the next few years, the development of CNN has also had some achievements, thanks to the substantial update of computer hardware devices and the rapid development of GPU.In 2012, the ImageNet competition, the model based on CNN took the first place with a 10% accuracy rate higher than the second place, and was once again pushed to the deep research boom by scholars.In 2014, the Computer Vision Group of Oxford University and Google DeepMind jointly developed VGGNet[5], and won the first and second place in the ImageNet competition respectively.In 2015, Kaiming He et al. proposed the residual neural network ResNet[6], which solves the problem of deep networks being difficult to train by fitting the residual term with cross-layer connections.Although the number of network layers reaches 152, the complexity is lower and the Top-5 error rate on ImageNet is only 3.57%.

B Basic structure and working principle

The basic building blocks of a CNN are also neurons one by one, containing weights with learning abilities and paranoid term constants. When multiple neurons are combined with a hierarchical structure, a neural network model is formed. A figurative representation of both is shown in Figure 1.



(a) Neurons

(b) Neural networks

Figure 1. Neurons and neural networks.

CNNmodels, are neural network modelsthat contain feature extractors consisting of convolutional and pooled layers. A typical network structure is shown in Figure 2, which includes five parts: input layer, convolution layer, pooling layer, full connection layer, and output layer. Among them, the convolutional layer and the pooling layer are the core modules to realize the feature extraction function of convolutional neural networks.

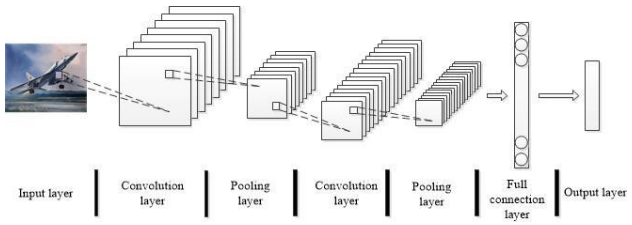


Figure 2. Typical structure of a CNN

CNN is a multilayered supervised learning neural network structure with the following workflow: A series of pre-processing operations are performed on the data at the input level, such as data normalization, de-normalization, etc. Entering the convolutional layer, the image of the input layer is convolved with a convolutional nucleus, and then the activation function outputs a feature extraction diagram of the layer, which is expressed as (1).

$$X_j^{i-1} = \sum_{i \in M_j} X_j^l * W_{ij}^l + b_j^l \quad (1)$$

Wherein, f(·) represents theactivation function,$M_j$represents the set of images participating in the current convolutional layer operation, X_i^{l-1}represents the value of a certain pixel input to the current layer image, $*$represents the convolutional operation, W_{ij}^l represents the weight vector of the l layer convolutional nucleus, and b_j^l represents the paranoid term of the l layer. Currently the commonly used activation functions are: Sigmod function, Tanh function, ReLU function, etc[7].

Connecting the pooled layer after the convolutional layer allows a certain degree of feature invariance and can reduce the amount of data. The input feature map is split into non-overlapping regions in the layer, and for each subregion features are further extracted by pooling operations, the common pooling operations being average pooling and maximum pooling.

After a number of alternating convolutional and pooling layers, multiple sets of highly abstracted feature maps are obtained; then the fully connected layers are entered and the multiple sets of feature maps are combined into one feature map. Then based on business needs, the final output of classification or identification results.

The goal of training is to minimize the loss function $L(W, b)$of the network, so the weights and biases of each layer need to be constantly updated during the training.

Common loss functions are Mean Squared Error (MSE) function, Negative Log Likelihood (Negative LogLikelihood, NLL) function, etc. In practice, in

order to reduce the occurrence of overfitting, the loss function increases the L2 parametre to control the overfitting of the weights and the parameters to control the strength of the overfitting effect.

$$J(W,b) = L(W,b) + \frac{\lambda}{2} W^T W \quad (2)$$

The weights and biases were updated as (3) and (4).

$$W_{ij}^{l} = W_{ij}^{l} - \alpha \frac{\partial}{\partial W_{ij}^{l}} J(W,b) \quad (3)$$

$$b_{i}^{l} = b_{i}^{l} - \alpha \frac{\partial}{\partial b_{ij}^{l}} J(W,b) \quad (4)$$

C Significance of the study

CNNhave been so successful in a number of applications and researchers have moved on to other areas, which brings more challenges. In terms of the CNN itself, that's where the research comes in.

1) Refinement of the theoretical system through domain application effects.

CNN have undergone more than 70 years of bumpy development, from MP models, BP neural networks, to various deep learning networks that are popular nowadays, all of them are judged directly by experimental effects, and there has been no complete set of theories for mathematical verification of these methods. Thus, as the field of application of CNN expands, it will also promote theoretical research in the field of CNN to a certain extent.

2) Facilitate the optimization of neural network structures and extend their application value.

At present,CNN have a place in natural language processing, image recognition, speech processing and other fields, and theirtrend is positive. And the network still has the problem of gradient disappearance, training sample size limit, computing power limit, is the short board of its development. For the problem that networks are difficult to train, an analysis of the problems related to network training is also given in the literature[8], but the solutions given have not become mainstream. Therefore, there is an urgent need to improve the learning capabilities of deep neural networks, so that the networks have better generalization capabilities and can be adapted to more complex application scenarios.

III. IMPROVEMENT AND OPTIMIZATION OF CNN

In recent years, improvements in CNN have been driven primarily by factors such as final detection effectiveness, network operating efficiency, and computational complexity. As of now, the improvement and optimization of CNN are mainly considered in three aspects: network depth, network structure, and network training methods.

A Network depth

Lecun et al. designed the Lenetnetwork[9], which uses alternately connected convolutional and pooled layers, and eventually passes through full-connected layers.There is 5 layers in Lenet and Lenet became the originator of CNN.In 2012, Krizhevsky proposed AlexNet[10], a network model with five convolutional layers, some of which are followed by a pooled layer for downsampling and finally two fully connected layers. The last layer is the softmax output layer, with 1000 nodes, corresponding to 1000 image classifications in the ImageNet graph set and using the Dropout mechanism and ReLu function, which has improved the accuracy and training time. Subsequently, the VGGnet proposed by Karen et al. uses almost all 3 $\times$3 convolutional nuclei, while adding pooled layers after several convolutional layers, instead of pooling immediately after each convolutional layer, to guarantee the depth of the network in many ways.VGGnet[11]demonstrated that increasing the number of network layers is beneficial for improving the accuracy of image classification. This increase is not unlimited and too many layers can create network degradation problems. The number of layers that affect the test results VGGnet was finally determined in two

versions, 16 and 19 layers. In the following years, there were different teams of researchers who proposed GoogleNet (22 layers), ResNet (152 layers) and they all deepened in terms of network depth and got better and better.

In terms of network depth alone, increasing the network level has an effect on the learning effect of convolutional neural networks, which also confirms the need for deeper learning.

B Improvements of the network structure

Improvements to the network structure mainly revolve around the idea of reducing network complexity. In 2014, Google proposed GoogLeNet[12], whose main innovation is the Inception mechanism, which sets up different convolutional cores in the same layer, i.e., multi-scale processing of images, and adding 1*1 convolutional core before 3*3, 5*5 for dimensionality reduction, which reduces parameters and improves the accuracy of image recognition on ImageNet datasets by about 10%.In 2015, Springenberg J T et al. proposed a full convolutional structure in literature[13]. Instead of the classical pairing of alternating convolutional and pooled layers in a classical convolutional neural network, the stride convolutional layer is used instead of the feature extraction layer in this structure. It was found that the error rate of this new network structure was reduced by 10 percentage points compared to traditional convolutional neural networks, and it was found that in some cases, the addition of apooling layer to this network structure resulted in weakened performance. Lin M et al. proposed a network-in-network structure in the literature, Network In Network[14]. is a subversion of the traditional structure of convolutional neural networks. The mesh structure replaces the full connection layer with a global averaging pooling layer, allowing the input feature map to be classified directly at the output, improving performance. However, this structure makes the convergence process longer and, on the whole, is not a very effective network structure.

The Google DeepMind team proposed a self-contained transformation module to flexibly and efficiently extract image invariant features[15], the specific structure of the module is shown in Figure 3, the model named Spatial Transformer is mainly composed of three parts: Localization Network, Grid generator and Sam-pler. Localization Network uses input feature mapping and outputs spatial transformation parameters through multiple implicit layers; Grid generator uses predictive transformation parameters to create sampling grids; Sampler uses feature mapping and sampling grid as inputs to generate output mapping from grid points.

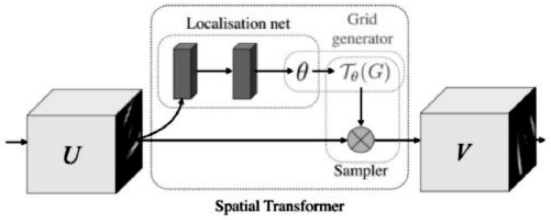


Figure 3. Invariant image feature

This model is used to good effect: it is a self-contained module that can be added anywhere in the network structure (not just the CNN), and there are no limits. It is easy to differentiate and can be used directly for end-to-end training. Its easy-to-differentiate and fast nature allows it to be added to the structure without slowing down the original network. The features extracted by this method are more common than the conventional structures.

Shanshan Xu[16] proposed an algorithm for progressively extending the network structure of a convolutional neural network to optimally adjust the network structure to make the network suitable for real-world problems, and at the same time proposed an improved feature extraction method to realize that feature extraction does not occur on its own in the network, and applied this method to handwritten number recognition, and the experimental results

showed that the recognition accuracy and recognition efficiency were higher than other algorithms using the classification method of convolutional neural networks.

C Methods of network training

1) The dropout mechanism[17] was added to the eliminate overfitting.

The Dropout proposed by Hinton et al. effectively improves the generalization performance of the network by randomly ignoring a certain percentage of node responses during training compared to traditional fully connected neural networks. However, the performance improvement of Dropout for convolutional neural networks is not significant, mainly because convolutional neural networks greatly reduce the number of training parameters compared to fully connected networks due to the weight sharing properties of convolutional nuclei, which itself avoids the more severe overfitting phenomenon. Based on the idea of Dropout, Wan et al[18]proposed the Drop Connect approach. Unlike Dropout which ignores some of the node responses of the full connection layer, Drop Connect randomly disconnects a certain percentage of the neural network convolutional layer. Forconvolutional neural networks, Drop Connect, which acts on the convolutional layer, has a much stronger past-fitting capability than Dropout, which acts on the full-connect layer.

Although the functional layer of Dropout and Drop Connect are different, the underlying principle is to increase the sparsity or randomness of network connections in network training in order to eliminate overfitting and improve network generalization capabilities.

2) Training methods using knowledge transfer

Overfitting and gradient dispersion are prone to occur when training against conventional convolutional neural networks. A training strategy using knowledge migration was proposed by Rocco et al. in the literature[19]. Pre-training (Pre-training with soft target, PST) is first performed on the soft target (soft target is a class distribution containing information between sample classes), and the migration of the soft target from the source model to the target model in the same domain allows more supervisory information to be obtained from a limited sample than from a single tag, solving the problem of missing samples. Then the target model adjacent convolutional layer is divided into a module to learn the low-level features of the source model in a modular way, similar to DBN's layer-by-layer pre-training strategy, and the combination of MMT and PST, the sample class information and low-level features of the two knowledge migration at the same time, so that the model convergence to a better position, and then use the SGD algorithm fine-tuning, so that the generalization performance is greatly improved.

Ⅳ. Application of CNN in Image Recognition

Image classification is an image processing technique that identifies different things by the characteristic information given by the image. With the rise of machine learning, automatic image classification techniques have been applied in various development fields. Cao used the classical convolutional neural network VGG-16 as a prototype in the literature[20], and added a multi-scale sampling layer at the end of the convolutional part, so that the model can input any size of images for training and testing, while reducing the number of neurons in the full connection layer, which improves the training speed of the model while ensuring accuracy, and applies it to the problem of multi-attribute classification of human faces. In the literature[21], Wenxu Shi proposed a CNN-based multiscale approach combined with a feature extraction algorithm for reverse convolutional networks (MSDCNN) and classified adenocarcinoma pathology images. Classification experiments performed on adenocarcinoma pathology cell images showed that the MSDCNN algorithm improved the classification

accuracy of the final convolutional feature scale by about 14% over the conventional CNN algorithm and about 1.2% over the classification accuracy of the fusion network model approach, which is also based on multi-scale features. In the literature[22], Chunlei Zhang proposed a parallel network model based on convolutional neural networks for military target image classification. The method uses two edge detection operators to extract the target image features separately and then input them into the convolutional neural network for deep feature extraction, which improves the classification recognition rate by 1.2% and reaches 97% compared to theconventional convolutional neural network. The theoretical analysis and experimental data illustrate that the model enables the effective differentiation of military target image data and is important for the accurate provision of military operational information.

Target detection is a fundamental problem in the field of machine vision as well as artificial intelligence, whose main goal is to pinpoint the category and location border information of various targets in an image. Target detection algorithms based on convolutional neural networks, such as RCNN, Fast RCNN, Faster RCNN, Mask RCNN, etc., are widely used in security monitoring, intelligent transportation and image retrieval and other fields. A target detection algorithm based on multi-scale feature extraction was proposed by Jianghao Rao in the literature[23] and applied to the detection of infrared pedestrian small targets with better results than conventional networks. A Mask-RCNN-based method for building target detection was proposed by Dajun Li et al. in the literature[24]. In the literature[25], Ding Peng "fine-tuned" a variety of mainstream depth convolutional neural networks based on Faster RCNN detectors on two classical data sets for target detection of optical remote sensing images. In response to the problem of target detection in traffic roads, Zhang Qi et al. proposed a traffic target detection method based on anchor point clustering, all-anchor point training strategy and reinforced intersection and merger ratio (SIoU) in the literature[26].

Ⅴ. CONCLUSION

This paper describes the basic structure of classical CNN from the development of CNN, briefly analyzes the features of CNN that have been improved and optimized in the development process, and finally elaborates on the wide application of CNN in the field of image classification and target detection. CNN have been developed to date and occupy an important position in the field of image recognition.

In terms of current trends, CNN will continue to evolve and makeCNNsuitable for various application scenarios, such as 3D CNN facing video understanding. There are also challenges, such as limited data sets, network generalization performance, robustness to be improved, and high training costs. The aforementioned issues will be a direct driver for the future development of convolutional neural networks and will directly contribute to the further deepening of artificial intelligence.

ACKNOWLEDGMENT

This work is supported by the Natural science foundation of Shaanxi province(2019JM-603).

REFERENCES

[1] Hinton G E, Salakhutdinov R R , Reducing the dimensionality of data with neural networks[J].Science, 2006,313(5786):504-507.

[2] Rumelhart D E, Hinton G E, Williams RJ. Learning representations by back-propagating errors[J].Nature, 1986, 323:533-538.

[3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, november 1998.

[4] LeCun Y, Cortea C. MNIST handwritten digit database[EB/OL].http://yann.lecun.com/exdb/mnist, 2010.

[5] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition.

[6] He K, Zhang x, Ren S, et al. Deep residual learning for image recognition[EB/OL].http://arxiv.org/abs/1512.03385,2015.

[7] Gu Jiu-Xiang, Wang Zhen-Hua, Jason Kuen,etal. Recentadvances in convolutional neural networks. arXiv:1512.07180v5,2017.

[8] X. Glorot, Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. AISTATS, 2010

[9] Lecun Y, Boser B, Denker J S, etal. Back propagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4):541-551.

[10] Krizhevsky A, Sutskever I,Hinton G E.Imagenet classifi-cation with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems, 2012 : 1097-1105.

[11] Bengio Y, Simard P, Frasconi P.Learning long-term depen-dencies with gradient descent is difficult[J].IEEE Transac-tions on Neural Networks, 1994, 5(2):157-166.

[12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, Going Deeper with Convolutions, Arxiv Link: http://arxiv.org/abs/1409.4842.

[13] Rougui J E, Istrate D, Souidene W, et al. Audio Based Sur-veillance for Cognitive Assitance Using a CMT Microphonewithin Socially Assistive Living [J]. 2009, 2009 (2009):2547-2550.]

[14] Lin, Min, Qiang Chen, and Shuicheng Yan. Network in network. arXiv preprint arXiv:1312.4400

[15] Zang D, Chai Z, Zhang J, et al. Vehicle license plate recog-nition using visual attention model and deep learning[J].Journal of Electronic Imaging, 2015, 24(3):033001.

[16] Xv Shanshan. Research and application of convolutional neural network [D]. Nanjing forestry university, 2013.

[17] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaption of feature detectors[R /OL].[2015-10-26]. http: / / arxiv. org / pdf /1207. 0580v1.

[18] Wan L, Zeiler M, Zhang S, et al. Regularization of neural networks using drop connect [C] / / Proceedings of the 2013 Inter-national Conference on Machine Learning. New York: ACM Press, 2013: 1058－1066.

[19] Luo Ke, Zhou Anzhong,Luo Xiao. A Convolutional neural network training strategy using knowledge transfer[J]. Control and decision-making, 2019, 034 (003) : 511-518.

[20] Cao Ge. Research on face image classification application based on deep convolutional neural network [D]. Jilin university, 2019.

[21] Shi Wenxu, Jiang Jinhong, Bao Shengli. Application of improved convolutional neural network in pathological image classification of adenocarcinoma[J]. Science, technology and engineering, 209,19 (35) : 279-285.

[22] Zhang Chun-lei. Military target image classification technology based on parallel convolutional neural network[J]. Electronic design engineering, 209,27 (08) : 76-80.

[23] Rao Jianghao. Target identification, positioning and detection based on CNN integration [D]. University of Chinese academy of sciences (institute of optoelectronic technology, Chinese academy of sciences), 2018.

[24] Li Dajun, He Weilong, Guo Bingxuan,Li Maosen, Chen Minqiang. Building target detection algorithm based on Mask - RCNN[J]. Science of surveying and mapping, 202,44 (10) : 172-180.

[25] DingPeng. Research on optical remote sensing target detection technology based on deep convolutional neural network [D]. University of Chinese academy of sciences (changchun institute of optics, precision mechanics and physics, cas), 2019.

[26] Zhang Qi, Ding Xintao, Wang Wanjun, SZhou Wen. Traffic target detection method based on Faster RCNN [J]. Journal of west anhui university, 202,35 (05) : 50-55.

Design and Research of Security System in University Laboratory

Ding Li
School of Health Services Management
Xi'an Medical University,
Xi'an, Shaanxi, China
E-mail: dingli@xiyi.edu.cn

Zhang Dujuan
School of Health Services Management
Xi'an Medical University,
Xi'an, Shaanxi, China
E-mail: ZhangDj1981sx@162.com

Shi Xuerong
School of Health Services Management
Xi'an Medical University,
Xi'an, Shaanxi, China
E-mail: 39480919@qq.com

Su Yulei
School of Health Services Management
Xi'an Medical University,
Xi'an, Shaanxi, China
E-mail: 541952030@qq.com

***Abstract*—The laboratory is an important place for teaching and scientific research in major universities, and plays an important role in the work of universities. The resources equipped by the laboratory occupy a large part of the resources of the entire school. The laboratory security system plays an important role in ensuring the safety of the laboratory and preventing the loss of equipment. Today's laboratory security systems have factors such as low automation, insufficient management, and inadequate safety in the experimental environment. Moreover, most security systems are built using wired methods, which have problems such as cumbersome wiring, aging lines, and difficult maintenance. In view of the many problems in the current laboratory management, this paper proposes a security system designed and implemented by the ZigBee technology of the Internet of Things. The system uses a wireless network to monitor laboratory access control and the environment, realizing the intelligence of the laboratory security system.**

Keyword-Internet of Things (IoT); Sensor; Security

I. INTRODUCTION

In the field of security, with the diversification of networking methods, hardware and software platforms and application technologies, the implementation of security systems has become more and more diverse. At present, the laboratory security system has a large number of alarm constraints and redundancy, and poor real-time monitoring. In response to this series of problems, this paper uses the Internet of Things and other related technologies to carry out research work on the laboratory security system design, and proposes a set Networked laboratory security system. The system uses sensor design nodes to form a ZigBee wireless sensor network to realize the collection of environmental information in the laboratory, thereby automatically monitoring the occurrence of some security risks.

II. IOT TECHNOLOGY

Internet of Things (abbreviation: IoT) originated in the media field and is the third revolution of the

DOI: 10.21307/ijanmc-2020-015

information technology industry[1]. The Internet of Things refers to connecting any object with the network through information sensing equipment according to the agreed protocol, and the objects exchange information and communicates, through the information transmission medium to achieve intelligent identification, positioning, tracking, supervision and other functions. There are two key technologies in IoT applications, namely sensor technology and embedded technology[2]. Unlike traditional networks, the terminal of the Internet of Things technology is no longer a PC, and its terminal is an embedded computer system and its supporting sensors. The Internet of Things technology is widely used in industrial, medical, transportation and other fields.

III. NETWORK TOPOLOGY

A typical ZigBee network supports three topologies, namely a star network topology, a tree network topology, and a mesh network topology. The star network topology is composed of a network main coordinator and multiple terminal equipment nodes. The main coordinator node is an FFD device. In a star topology, each terminal node can only communicate with the coordinator node, and each terminal node cannot transmit data. The tree topology is composed of a network coordinator and multiple terminal equipment nodes and routing nodes. The main coordinator and the routing node can have their own child nodes, and the terminal device node can autonomously choose to join the coordinator or the router node. The tree structure follows the order of parent node and child node, that is, each terminal node has its fixed parent node, which must be passed through the parent node layer by layer. The advantage of this topology is that the network coverage is large, but if one of the parent nodes is damaged, all the child nodes connected to it will be unable to communicate. The mesh topology is also composed of a coordinator node, multiple terminal device nodes and router nodes. This structure is different from the tree network structure in that all routing nodes can communicate with each other, and there is no strict communication sequence. When a routing node fails, data can be transmitted through other routing nodes. This topology not only reduces the delay of information transmission, but also improves the routing efficiency and reliability. Through the corresponding routing algorithm, the best path can be found.

According to the current status of the laboratory, the range of experiments that need to be monitored is large, and the types of data monitored are many. From a practical point of view, the wireless network that needs to be established can dynamically delete and add nodes at any time. According to these characteristics, choosing a mesh topology is more suitable for the current laboratory system. The mesh topology has a strong stability and has a strong advantage in a large-scale laboratory monitoring system.

IV. STRUCTURE DESIGN OF LABORATORY SECURITY SYSTEM

The laboratory security system mainly relies on the monitoring of the laboratory environment to determine potential safety hazards or the presence of suspicious persons[3]. Various types of sensors are arranged in the monitoring range. These sensors are connected to the terminal nodes and placed in different locations in the laboratory. The system uses infrared sensors to monitor whether there are people in the room, door magnetic sensors to detect whether someone breaks into the window, smoke and temperature sensors to detect the presence of fire in the laboratory, and harmful gas sensors to detect the leakage of pharmaceutical reagents. When an abnormality occurs, the sensor in the detection unit will send out an alarm signal, and the alarm signal will be transmitted through the communication unit to the monitoring center software via the ZigBee network. The structure of the laboratory security system is shown in Figure 1.

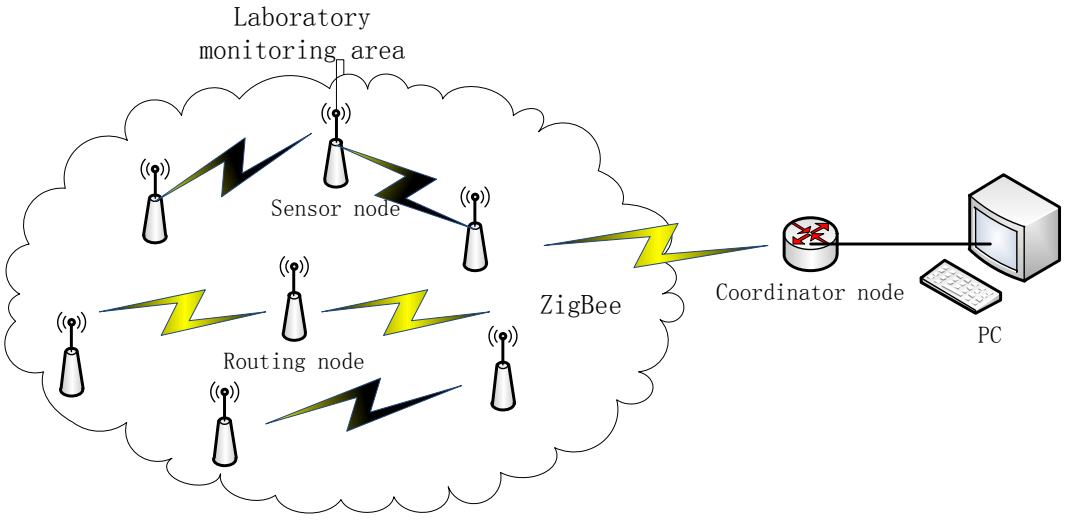


Figure 1. Laboratory security system structure

V. SYSTEM HARDWARE DESIGN

A. Monitoring unit design

The detection unit is mainly composed of an infrared sensor, a door magnetic sensor, a smoke sensor, a temperature sensor, a harmful gas and a combustible gas sensor module. The infrared sensor is mainly used to detect whether there are outsiders in the laboratory. The system uses a pyro electric infrared sensor, which mainly detects the infrared radiation radiated by the human body in a non-contact manner to determine whether someone is in the surveillance area.

This system uses the human body infrared sensor module HC-SR501. This module is fully automatic. When someone enters the sensing range, the module will automatically output high level. If it can always feel the presence of the person, it will always output high level. ， After detecting the existence of human, it will switch from output high level to output low level. The wireless door magnetic sensor is a device commonly used in security systems, mainly used to detect whether the doors and windows are opened or closed illegally. When the wireless door sensor collects the signal that the door sensor is open, it sends an alarm signal. The combustible gas sensor is a detector used to detect the combustible gas content in the air. The system uses the MQ-9 sensor to measure the combustible gas concentration[4]. When the sensor detects that the combustible gas content in the air reaches or exceeds the threshold, it sends an alarm signal. The temperature sensor adopts DS18B20 digital sensor, which is small in size and easy to use. The interface mode is single wire, which can realize the networking function of multiple sensors. The harmful gas sensor module minds the effective detection of many gases such as ammonia, sulfide, methane, and carbon monoxide. The MQ135 sensor is used for the detection of harmful gas content[5].

B. Communication unit design

The communication unit is mainly responsible for data transmission in the wireless sensor network environment monitoring system, which is the core part of the entire system. Data is transmitted through a large number of wireless sensor network nodes arranged in the network. The mutual communication between the various nodes realizes the communication of the entire network and forms the communication unit of the system. The node structure designed in the wireless

sensor network is shown in Figure 2. It is mainly composed of three parts: wireless node module, sensor module and power intelligent main board.

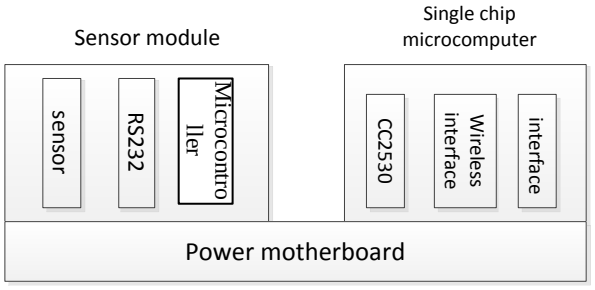


Figure 2. Wireless sensor network node

The microcontroller in the sensor module selects STC89C52RC produced by STC. The controller has the characteristics of high speed and low power consumption. It uses an enhanced 51 core, the calculation speed is higher than the ordinary 8051. The controller can wake up the idle mode at any time to reduce power consumption and ensure that the end node power supply is used for a longer time.

The wireless node module selects CC2530 chip as the microcontroller of the module. The CC2530 chip is compatible with various ZigBee standards, and has the characteristics of stable performance, low power consumption, and strong anti-interference ability[6]. The sensor module is arranged in the monitoring environment and is responsible for data collection and transmission in different areas, and transmits the collected data information to the coordinator. If it cannot be directly transmitted, it is transmitted to the coordinator in a multi-hop manner.

The design part of the power module chooses to use an input voltage of 9v, and the microcontroller and various sensors on the front-end measurement version use a 5v power supply[7]. The ZigBee module CC2530 uses a 3.3V power supply, so the system uses a 7805 voltage regulator to convert the 9V voltage It is 5V voltage. At the same time, the 5V voltage is reduced to 3.3V through the 1117 voltage regulator tube, which is used to provide power for the ZigBee module.

VI. SYSTEM SOFTWARE DESIGN

ZigBee is a wireless transmission technology based on the IE802.15.4 standard specification[8]. It has the characteristics of self-organizing network, low cost, low power consumption, and enemy complexity [9]. It does not need to apply for authorization for the working frequency band, which is convenient and cost-effective to use. The system uses ZigBee to develop a set of wireless sensor networks according to the actual monitoring needs in the laboratory, and the data transmission between the nodes in the network is transmitted using the ZigBee protocol.

The laboratory monitoring system mainly includes terminal node coordinator, terminal node and routing node. The coordinator is responsible for forming a network, collecting data, and providing an interface with a computer to realize the formation of a sensor network and the establishment of a data transmission channel. The terminal node is responsible for collecting data information in the laboratory, including temperature, smoke, and harmful gases. These data are processed by the microcontroller, transmitted to the ZigBee network via the wireless network module, and finally transmitted to the coordinator, which is then transmitted to the monitoring center to realize the collection and transmission of laboratory environmental data. The routing node is responsible for the signal enhancement and forwarding of the data. Because the monitoring area is large, some data may be lost during the transmission process. To avoid this situation, a routing node is designed in the network to enhance and forward the data signal. , To ensure the stability and accuracy of data transmission throughout the network..

A. Terminal node design

1) Design of sensor data acquisition module

The system data acquisition module uses 51-microcontroller, through C language programming, to achieve sensor status and data acquisition, judgment and so on. These data are sent to the wireless network through the ZigBee module and transmitted to the coordinator. The flow chart of the sensor data acquisition module software design is shown in Figure 3.

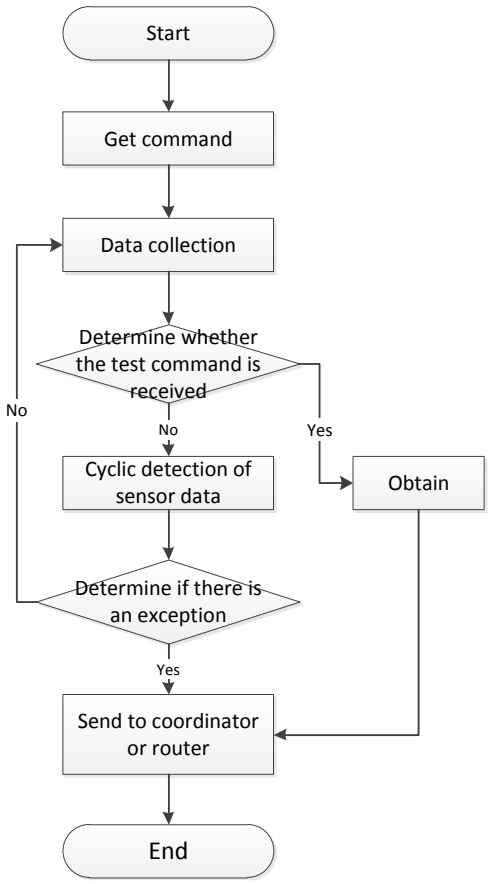


Figure 3. Software design flow chart of information collection module

After the data acquisition module software is running, it reads the background from the serial port of the ZigBee module to obtain background commands, collects the status of the sensors connected to the module, and sends the data to the network. If there is no background acquisition command, the program cyclically executes the acquisition sensor status command, and judges whether there is an abnormal situation in the laboratory through the sensor status data. If abnormal data appears, send laboratory status data to the monitoring center, and if there is no abnormality, continue to collect laboratory sensor data.

2) Wireless network module design

The information collected by each terminal node in this system is different. According to different sensors, multiple data will be collected within the scope of laboratory monitoring. In order to realize the data transmission in the system, it is necessary to establish a network and add these terminal nodes to it. In this system, the terminal node sends a request to join the network to the coordinator. The coordinator determines whether to allow joining according to the situation, responds to the request and sends it to the terminal node, so as to join the network and perform data transmission. The flow chart is shown in Figure 4.

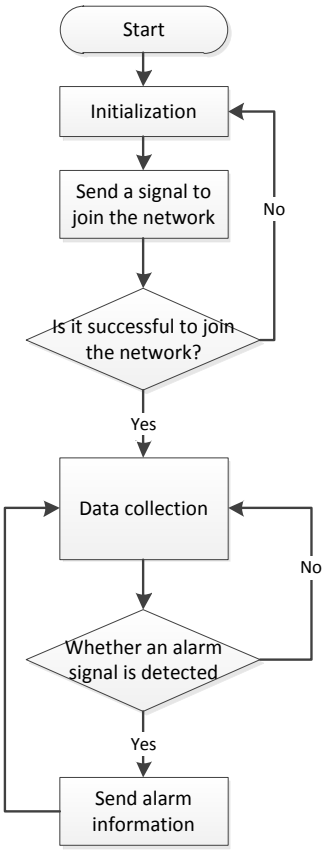


Figure 4. End node flow chart

The terminal node first finds the coordinator node in the network, and scans again if it is not found until it is found. Then send the association request command, wait for the coordinator to process, if you agree, then join the network, and get a short 16-bit address assigned, the terminal node sends data to the coordinator through the network.

B. Coordinator node design

The main role of the coordinator is to create the entire network and serve as a bridge between the terminal nodes and the control center. It needs to receive all kinds of data collected by terminal nodes and routing nodes through ZigBee wireless communication protocol, and then send it to the monitoring center. In the entire system, the coordinator must be a full-function device, and there can only be one coordinator node in a network, so the coordinator initialization setting is required at the time of design. The flow chart is shown in Figure 5.

The coordinator in the network is responsible for constructing the entire network, and adopts the method of ad hoc network to construct.

1) Determine network coordinator

The process is to determine whether the node is an FFD node, and then to determine whether the FFD node is in another network or whether a coordinator already exists in the network[9]. Through active scanning, send a beacon request command, and then set a T_scan_duration, if no beacon is detected within the scanning period, then it is considered that FFD does not have a coordinator in the entire network, it can build your own ZigBee network, and as this The network coordinator continuously generates beacons and broadcasts them.

2) Carry out the channel scanning process

It includes two processes: energy scanning and active scanning. First, perform energy detection on the specified channel or the default channel to avoid possible interference. Then carry out active scanning to search for network information within the communication radius of the node. This information is broadcast on the network in the form of beacon frames. The node obtains these beacon frames through active channel scanning, and then selects a suitable channel based on this information.

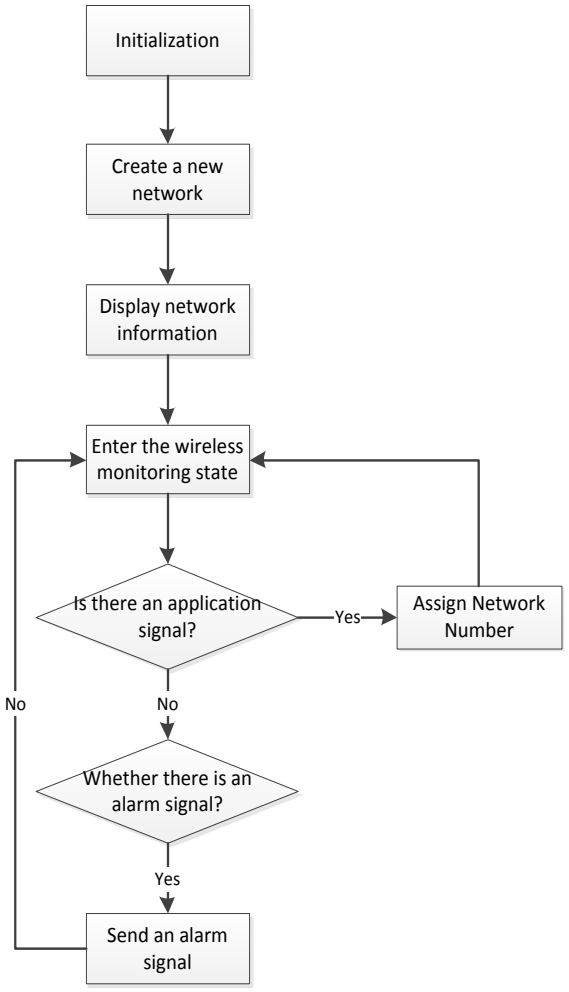


Figure 5. Coordinator node program flow chart

3) Set the network ID

After finding a suitable channel, the coordinator will select a network identifier (PAN ID) for the network. There are two address modes in the ZigBee network: extended address (64-bit) and short address (16-bit), where the extended address is assigned by the IEEE organization and is used for unique device identification.

After the above steps are completed, the ZigBee mesh network is successfully initialized, and then waits for other nodes to join. After the node successfully joins the network, it will get a short network address and send and receive data through this address. The flow chart of ad hoc network is shown in Figure 6.

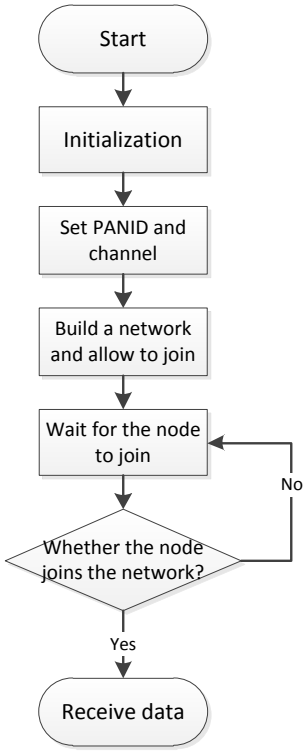


Figure 6. Coordinator ad hoc network flow chart

C. Design of routing nodes

As a relay node of the entire detection system, routing nodes are suitable for environmental monitoring in a large area. In wireless sensor networks, it may not be possible to connect and communicate because of the long distance between nodes. At this time, the routing node acts as a relay station to connect the terminal node and the coordinator. When setting up a routing node, you need to first initialize the CC2530 device and protocol stack, send a signal to join the network, the front-end coordinator will respond accordingly, agree to join the network, and assign a network address. After the routing node joins the network, it starts the function of data forwarding, thereby ensuring the transmission of data throughout the network. The program flow Figure 7 is shown.

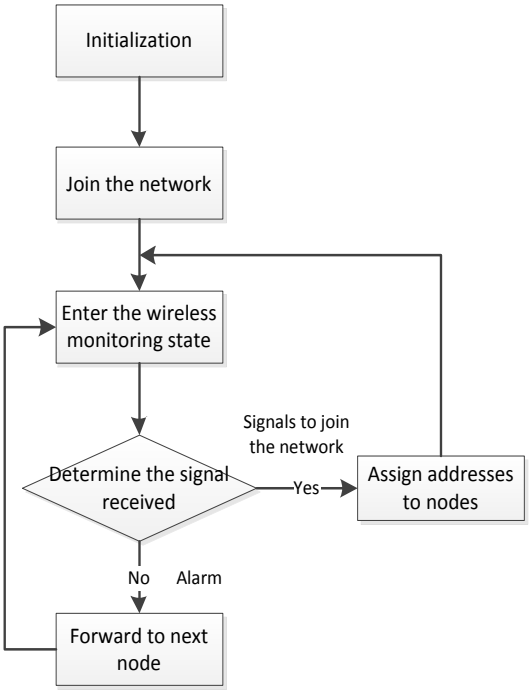


Figure 7. Routing node flow chart

VII. COMMUNICATION PROTOCOL DESIGN

Communication protocol refers to the rules and conventions that both entities must follow to complete communication or services. The protocol defines the format used by the data unit, the information and meaning that the information unit should contain, the connection method, and the timing of information transmission and reception, so as to ensure that the data in the network is smoothly transmitted to the determined place.

This system uses wireless communication, a specific communication rule designed for communication in the formulation of protocols. In the formulation of the transmission module protocol, according to the ZigBee communication protocol and the needs of the actual system, the communication protocol of this system was formulated. The single transmission format is shown in Table 1.

TABLE I. TRANSMISSION DATA FORMAT

FD	length	target address	Date 1	Date 2	Date 3	……

FD indicates that ZigBee sends data point to point, and the length indicates the total length of the piece of data. This length can determine how many bits of data need to be read, which is convenient for taking out the data. The target address stores the ZigBee address of the received information. The following data 1 represents the laboratory number, data 2, 3, 4 and so on represent the data collected by the sensor. Each node sets different data according to its sensor.

The single data receiving format is shown in Table 2.

TABLE II. TRANSMISSION DATA FORMAT

FD	length	target address	Date 1	Date 2	Date 3	……	Source address

FD means ZigBee point-to-point transmission. The length represents the total length of the data, and the target address represents the ZigBee address of the received message, that is, the node address to which the data needs to be sent. Data 1 represents the laboratory number. Data 2 represents the data information received by the sensor. Data 3 and data 4 are also the information received by the sensor. The amount of this information depends on the number of sensors at the terminal node. The source address represents the ZigBee address of the sent message, that is, from which node the message was received.

VIII. CONCLUSION

The laboratory security monitoring system mainly monitors the environment in the laboratory to alert the abnormal situation in time to avoid the economic loss caused by aging equipment or human negligence. The development of such a system has a positive effect on the laboratory construction of colleges and universities, and has certain application prospects.

REFERENCES

[1] ZHANG Hui YUAN Yong-zhe WANG Wen-jun “Security and intelligent system for buses based on mobile internet of things “Electronic Design Engineering. 2020, 28(9) 37-40.

[2] SUN Qi-bo LIU Jie LI Shan FAN Chun-xiao SUN Juan-juan” Internet of Things:Summarize on Concepts,Architecture and Key Technology Problem” JOURNAL OF BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS. 2010, 33(3).1-2.

[3] LIU Qiang CUI Li CHEN Hai-ming “Key Technologies and Applications of Internet of Things” COMPUTER SCIENCE. 2010, 37(6).1-4.

[4] HAN Tuanjun YIN Jiwu ZHANG Panfei ZHENG Xu ZHENG Zhengbing.” Design of remote laboratory safety management system based on GPRS and Zigbee” Experimental Technology and Management. 2019, 36(11) 240-244.

[5] YAN Yaling LI Bo LIU Weijie.” ZigBee-based Laboratory Fireproof Remote Monitoring System”. 2019, 38(5) 282-285.

[6] Jiang Yunchen.” Monitoring and managing system of laboratory based on campus LAN and ZigBee technology” EXPERIMENTAL TECHNOLOGY AND MANAGEMENT. 2010, 27(7).104-106.

[7] FAN Yan YU Yang LI Yong-yi DING Yi-xing “Research on Remote Monitoring System Based on ZigBee Wireless Sensor Network” Research and Exploration in Laboratory. 2016, 35(1).80-84.

[8] ZHANG Pan HUO Lian-song “Design of An Wireless Temperature Monitoring System Based on ZigBee” Sensor World. 2018, 24(8).23-27.

[9] WANG Sihua ZHENG Shuqiang DING Yihua “Discussion on ZigBee Sensor Network Technology and Its Application” Radio Engineering. 2020, 50(5).415-417.

Research and Implementation of Future Network IPV9

Liu Zhang
[1.]School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
[2.]State and Provincial Joint Engineering Lab. of Advanced Network, Monitoring and Control
Xi'an, China
E-mail: 604463203@qq.com

Wang Yubian
Department of Railway Transportation Control
Belarusian State University of Transport
34, Kirova street, Gomel, 246653
Republic of Belarus
E-mail: alika_wang@mail.ru

***Abstract*—Nowadays, IPv4 has been difficult to meet the needs of the Internet in terms of performance, address space, security, etc. In order to solve the relevant needs of IPv4, protocols such as IPv6 and IPV9 have been born. This article introduces the current status and characteristics of IPv4 and IPv6, compares with IPV9, summarizes the relevant characteristics of IPV9, and introduces the production process of IPV9, its protocol composition, system architecture and related application introduction. IPV9 is controlled by my Chinese core technology and has independent intellectual property rights, which is the foundation of my country's future network.**

Keywords-Future Network; Decimal System; IPV9

I. IP

A. The Introduction of IP

IP(Internet Protocol), is the network layer Protocol in the TCP/IP architecture.

When we use the Internet, the most important question is whether my messages and actions can be successfully sent and whether I can receive messages from the outside. Today, our needs are fundamentally assured through IP. Sending and receiving is actually a kind of information transmission, our various operations will be various applications in the form of packets for transmission. The problem is getting from the beginning to the end, and it's not a direct highway, but a ladder of different routes that takes multiple hops to get there. The purpose of IP is to solve the problems of network connection and computer communication. Each IP address consists of a network address (NetID) and a host address (HostID). A network address represents which network in the Internet it belongs to, and a host address represents which host in that network it belongs to.

B. The Introduction of IP Address

IP Address(Internet Protocol Address), is a unified address format that assigns a logical address to each network and host on the Internet, just like our mobile phone number, which can be used to mask the physical address differences while making communication more convenient.

All IP addresses consist of network ID and host ID. Depending on the network ID and host ID, the Internet commission has defined five IP address types to suit networks of different capacities, namely class A to class E. Among them, A, B and C are the basic classes, while D and E are used as multicast and reserved. This is shown in figure 1.

DOI: 10.21307/ijanmc-2020-016

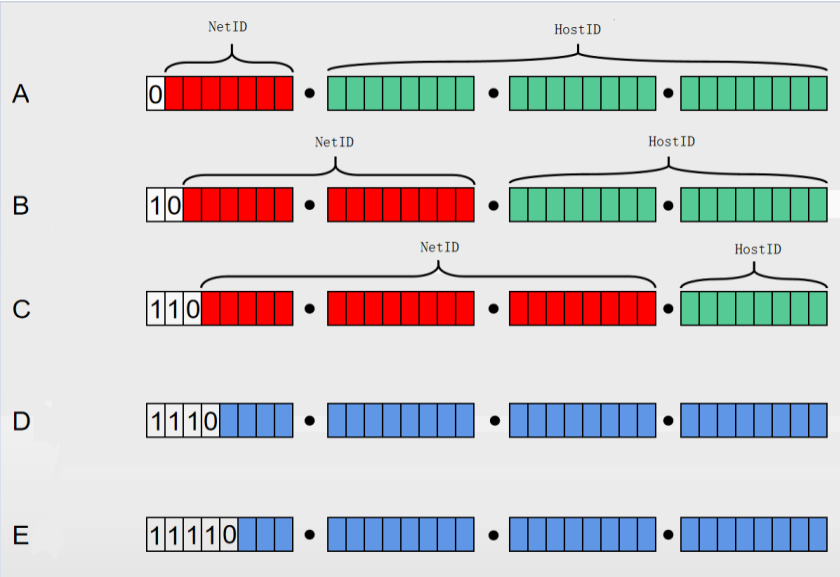


Figure 1. IP address

II. The Features and Problems of IPv4 and IPv6

A. *The Present Situation of IPv4*

IPv4 has played a key role in the development of networks, but with the expanding of network size, it can not meet the demand of network development, the first is the address resources are exhausted, lead directly to address the crisis, although no classification of addressing CIDR technology, network address translation NAT technology to alleviate the crisis, but still can't solve the problem.

The second problem is the expansion of routing table. The topology structure of address space directly leads to the fact that the form of address allocation is irrelevant to the network topology. With the growth of the number of networks and routers, the over-expansion of routing table increases the cost of searching and storage and becomes the bottleneck of the Internet. At the same time, the length of packet head is not fixed, so it is very inconvenient to extract, analyze and select the route by hardware, so it is difficult to improve the throughput rate of route data. Then there is the uneven distribution of IP addresses, because of the origin of the United States, more than half of all addresses are owned by the United States, resulting in a serious imbalance in the distribution of IP addresses.

There is also a lack of QoS (Quality of Service) support. IPv4 did not want to be open to the public at the beginning of the design, so it is very lacking in security, and it is difficult to provide rich QoS functions for real-time multimedia, mobile IP and other commercial services, although later The developed protocols such as RSVP provide QoS support, but the cost of planning and constructing an IP network is relatively high.

B. *The Features and Problems of IPv6*

IPv4 is a widely deployed Internet protocol. The IPv4 protocol is simple, easy to implement and interoperable. However, with the rapid development of the Internet, the deficiencies of IPv4 design have become increasingly obvious. IPv4 address space is insufficient, and the number of routing table entries that need to be maintained is too large. To solve these problems, the IETF designed IPv6. Compared with IPv4, IPv6 has the following features：

- IPv6 has a larger address space. In IPv4, the length of the IP address is 32 bits, that is, there

are 2^32-1 addresses; In IPv6, the length of an IP address is 128 bits, or 2^128-1 addresses. Compared with the 32-bit address space, the address space is greatly increased.

- IPv6 uses smaller routing tables. IPv6 address assignment follows the principle of Aggregation at the beginning, which enables the router to represent a subnet with an Entry in the routing table, greatly reducing the length of routing table in the router, and improving the router's forwarding speed of packets.
- IPv6 adds enhanced Multicast support and Flow Control, which allows for the development of multimedia applications on the network and provides a good network platform for QoS Control.
- IPv6 adds support for Auto Configuration. This is the improvement and extension of DHCP protocol, making network management more convenient and fast.
- IPv6 has Better header formatting, it uses a new header format with options that are separated from the base header and can be inserted between the base header and the upper data if desired. This simplifies and speeds up the routing process because most options do not need to be routed.

Despite the obvious advantages of IPv6, the number of IPv4 routers is huge, and the transition from IPv4 to IPv6 is a gradual process, with IPv6 being backward compatible. Therefore, IPv6 and IPv4 will coexist for a long time to come. In addition, IPv6 has a big flaw in the design idea of its address structure. IPv6 confuses the network hierarchy in design. The interface ID embeds the physical address into the logical address layer, which on the one hand leads to the limitation of the physical address space to the empty IP address. Security does not belong to the content of the IP layer, so it is inappropriate to design security technology in the IP layer. Because with the development of security technology, security methods and key length will constantly change, so the development of security technology will eventually lead to the requirements of IP address redesign. Due to the chaotic logic of the network hierarchy, IPv6 creates far more new problems than it solves.

III. The Introduction of IPV9

A. The Production of IPV9

In 1998, Chinese researcher Xie Jianping proposed IPV9, which means "Method of using whole digital code to assign address for computer." IPV9 is a "nickname" borrowed from the American concept of IP. In order to distinguish China's IPV9 from America's IPv4 and IPv6, the V in China's IPV9 is uppercase, not lowercase. The patent covers the new address coding design, the new addressing mechanism and new address three technical architecture design, form a new system of IP network at the bottom of the core technology, on the basis of the design of the new framework, to form a network system that is connected and compatible to cover the existing network (the Internet using IPv4 and IPv6 technologies).

In 2011, the authoritative professional agencies of the us government confirmed legally and technically that China owns the core technologies of the sovereign network under the IP framework, which are different from the existing technologies of the us Internet and have independent intellectual property rights. This is the IPV9 patented technology, the official name of the patent is "the method of assigning addresses to computers in full numeric code."

China's IPV9 was approved in 2001 (CN98 1 22785), and has been granted patents in more than 10 countries and regions including South Africa, Turkey, kazakhstan, Russia, the republic of Korea, the democratic People's Republic of Korea, Hong Kong, Canada, Singapore, Australia, Mexico and Norway.

In 2004, IPV9 applied for the us patent, which was successively issued by the us Patent Office seven times with "non-final rejection opinions" and six times with final rejection letters. During this period, IT was repeatedly criticized by senior members of the US IETF and famous IT companies in the US.

In December 2011, The United States Patent and Trademark Office (PTO) officially issued the patent certificate No. and US 8,082,365, and stated in its notification of approval that the applicant's verification report was "very convincing".

IPV9 protocol refers to the 0-9 Arabic digital network as virtual IP address, and the decimal system as the text of the representation method, that is, a convenient way to find the use of the Internet users; For efficiency and end-user convenience, some of the addresses can be used as domain directly; It has an infinite number of allocatable IP addresses, with a maximum of 2 by 2048 bits, and is the cornerstone of the future digital world. At the same time, due to the use of the original computer network, cable broadcast television network and telecommunications network business classification code, therefore, also known as the "new generation of secure and reliable information integrated network protocol."

B. The Characteristics of IPV9

Compared with IPv4 and IPv6, IPV9 has more obvious features and advantages, mainly reflected in the following points:

1) Address space is huge

IPV9 has a larger address space than IPv4/IPv6. IPv4 defines the bit length of IP address is 32, that is, there are 232-1 addresses; While the length of IPv6 is 128, that is, 2128-1 addresses, the standard length of an IPV9 address is 2256-1, with 42 layers address structure design will be 10256-1 (21024-1). To put it mildly, if IPv6 were widely used, every grain of sand in the world would have an IP address. Then after IPV9 is widely used, the smallest molecule of bright matter in the whole universe will have a corresponding address. It is no exaggeration to say that if IPV9 is fully applied, every cell and living gene in the world can be assigned to an IPV9 address. Layer 42 is the asset management address (including legal digital currency space) compatible with ean-ucc128 barcode length.

2) Route tables are smaller

IPv6 has a smaller routing table than IPv4. The address allocation of IPv6 follows the principle of Aggregation at the beginning, which enables the router to represent a subnet with an Entry in the table, this greatly reducing the length of routing table in the router, and improving the speed of forwarding packets in the routing table.

The routing table of IPV9 is very small, and the address allocation of IPV9 follows the principle of Geo-spatial clustering from the beginning, which enables IPV9 router to represent a country subnet and an application subnet with a single record, it greatly reducing the length and cleanliness of routing table in the router, and improving the speed of forwarding packets by routing table. At the same time, this subnet can express a specific geographical location, for example, we assign the IPV9 address segment of Shanghai as 86[21[5]/96, then in other routers of the same level, only one route pointing to the address segment of 86[21[5]/96 can realize the IPv9 address routing of Shanghai. According to this logic, only one route is needed from country to country. For example, the route to China is 86/64. The IPv4 routing table is large and irregular, and the IPv6 routing table is smaller than IPv4, but the IPv6 routing table contains no geographic information and the routing is messy.

3) Automatic configuration support

IPV9 adds support for automatic configuration of variable length addresses, which is an improvement and extension of DHCP protocol of IPV9, making network management more convenient. IPV9 supports multicast, and supports the ISO/IEC C6 future network << naming and addressing >>TCP/IP/M model, and

supports long packet code streams for virtual and real circuits. This allows multimedia applications on the web to ensure video quality and reduce overhead, provide faster and faster applications such as industrial controls and unmanned vehicles, and provide better and cheaper service over the Internet than IPv6.

4) Address length could be select

IPV9 address length has a variety of options, which can realize the change of 16, 32, 64, 128, 256, 512 and 1024 bit address length, and select the most appropriate address length according to different usage scenarios to reduce the routing overhead.

5) Dual encryption

The address length of IPV9 is long enough to realize dual encryption from the transmission of source and target addresses, which plays an important role in some specific network transmission fields. IPV9 network makes use of logical isolation features to make network information transmission more secure and effective.

6) Add location information to the address

IPV9 addresses can be embedded with geo-location information, as well as personal and industry ID information, this making IP addresses uniquely tied to personal information.

7) Compatible with previous addresses

IPV9 address is backward compatible with IPv4/IPv6 address. In order to absorb the upgrade difficulty of IPv6 incompatibility with IPv4, IPV9 protocol remains and unchanged, so that IPv4/IPv6 upgrade to the new version of IPV9, the upgrade cost is very low.

8) Sovereignty is different

IPv4/IPv6 addresses Spaces and copyright ownership: United States.

IPV9 address space and copyright ownership: China.

IPV9 has its own intellectual property rights and was proposed by The Internet Assigned Numbers Authority (IANA), but it is China that has succeeded in developing and mastering the core technology. Compared with IPv4/v6, China has the core patent digital domain system of IPV9 technology, which is of great significance for the future development of China's network and the mastery of the security of cyberspace.

C. The Construction of IPV9 Protocol

The IPV9 protocol includes message protocol, address protocol, transition protocol, mobile communication protocol, etc, as shown in figure 2.

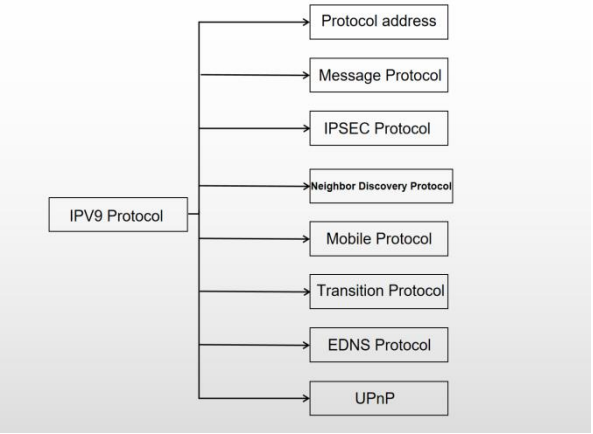


Figure 2. IPV9 protocol

1) Address Protocol

The IPV9 network expands the number of address bits to 256 bits, realizing a huge addressing space. And according to different data transmission methods, IPV9 addresses are divided into three types: unicast, anycast and multicast. In summary, it is the difference between one-to-one, one-to-one recent and one-to-many. Unicast type Each interface is configured with an identifier, and the packet identifies the identifier to reach the specified interface; an identifier in any on-demand type represents a group of interfaces of different nodes, and the shortest path interface is selected through a routing protocol And transmit the data packet to the interface; multicast is to use the multicast address to send the data packet to each

interface indicated by the identifier, and the shortest path interface will not be selected. IPV9 uses the "decimalization and brackets" approach in two forms:

a) Use the complete brackets to represent 2048 bits. In this way, the brackets can be ignored when entering a web address in the browser's address bar.

b) Divide the 256-bit address into 8 segments, each segment being 32 bits, "a [b] [c] [d] [e] [f] g [h". IPV9 addresses are very compatible with IPv4 and IPv6. The mapping relationship is shown in table 1 and table 2. The addresses of IPv4 and IPv6 are kept intact in the last bit address segment, and the value of the first address is used as an identifier to point to IPv4 or IPv6.

TABLE I. MAPPING RELATIONSHIP FROM IPV4 TO IPV9

Address number	1-96	97-128	129-160	161-256
Length (bits)）	96	32	96	32
Mapping	0[0[0	0	0	IPv4 address

TABLE II. MAPPING RELATIONSHIP FROM IPV4 TO IPV9

Address number	1-96	97-192	193-256
Length (bits)）	96	32	96
Mapping	1[0[0	0	IPv6 address

For IPV9 nodes in tunneling technology, they need to be assigned IPv4/IPv6 compatible addresses to communicate with other nodes in the corresponding network. The mapping strategy for this situation is shown in table 3, where the first prefix is 1000000000, the 000 and 001 values in the token bit correspond to IPv4 and IPv6 respectively, and the rest are reserved for future function expansion.

TABLE III. IPV9 COMPATIBILITY WITH IPV4/IPV6

Address number	1-10	11-29	30-32	33-96	97-128	129-214	215-256
Length (bits)）	10	19	3	64	32	96	32
Content	prefix	keep	mark	0	scope	IPv6	IPv4 address

2) Message Protocol

TABLE IV. IPV9 MESSAGE PROTOCOL

Version number	**Traffic Flow Type**				**Flow Label**
	Address length	***Priority traffic class***	***Address authentication***	***Absolute Traffic***	
Payload Length	Next Header				Limit jump times
Source Address（256bit）					
Destination Address（256bit）					
Time					
Identification Code					

The total header length of IPV9 is 72 bytes, which is more than that of IPv4 but more concise. The format is shown in table 4, which consists of ten parts: protocol version number, communication flow type, payload length, stream label, next header, hop limit, source address, destination address, time, identification code, etc.

In IPV9, the optional information of other layers is placed in the extended header between the high-level protocol header and the IPV9 header, and its structure is shown in table 5. An IPV9 packet can carry one or more or even no extension headers, and each subsequent extension header location is marked in its previous header.

TABLE V. EXTENSION HEADERS

IPV9 header **Next header=TCP**	**TCP header + data**		
IPV9 header Next header=route	IPV9 header Next header=TCP	TCP header + data	
IPV9 header Next header=route	IPV9 header Next header=data segment	IPV9 header Next header=TCP	TCP data segment Header + data

3) Transition Protocol

The IPV9 transition protocol specifies the IPV9 transition header format and the definition of the address text representation, addressing model, and node address, including a detailed description of the currently defined transition header and address format.

The header in the transition period uses the original IPv4 header, and only changes the version number to 9 to distinguish it from the original IPv4 header. The last two segments of the IPV9 address are adopted for the interim address, which is 64 bits in total.

D. The System Architecture of IPV9

IPV9/Future Network root domain name server system, consisting of a parent root server, a master root server, 13 equal-name root domain name servers named by 13 English N-Z, Top-level domain name servers of 239 countries and regions like .CHN,.USA,.HKG,.MAC, routing management systems, application servers and 10 Gigabit backbone routers. Its working principle is that 13 root domain name servers read the main root server first, then read the parent root server, and after obtaining the data, they will spread to the whole network. Only 13 root domain name servers can access this hidden distribution host. This hidden publishing host is maintained. 13 root domain name servers read its data, which is read by the mirror server, and then spread to the entire network. The IPV9 root domain name server system is shown in Figure 3.

The root name server is the highest-level domain name server in the Internet domain name system (DNS). It is mainly used to manage the Internet's home directory, and is responsible for providing authorized domain name server addresses for Top Level Domain TLD resolution. It is the necessary infrastructure for constructing the Internet. Many computer scientists refer to the root domain name server as "truth", which shows its importance. Currently, the Internet's root domain name server, gTLD, and ccTLD are all managed and controlled by ICANN (Internet Corporation for Assigned Names and Numbers) authorized by the US government. Attacking the root domain name server is the most direct and deadly method of attacking the Internet. In the existing Internet, the root server is completely controlled by the United States, which poses a great risk to other

countries. The 13 IPV9 root DNS that can adapt to IPv4 networks, IPv6 networks, and IPV9 networks, use decimal network technology to organize, build, secure, controllable, face global users, and serve Chinese, English, digital and other languages , And can provide personalized broadband multimedia communication services on the communication network to provide English, digital, Chinese domain name resolution function. The IPV9 resolution system can ensure that the domain used by online users are resolved by the domain server to obtain the IP address of the corresponding access object, which is compatible with the current various domain services.

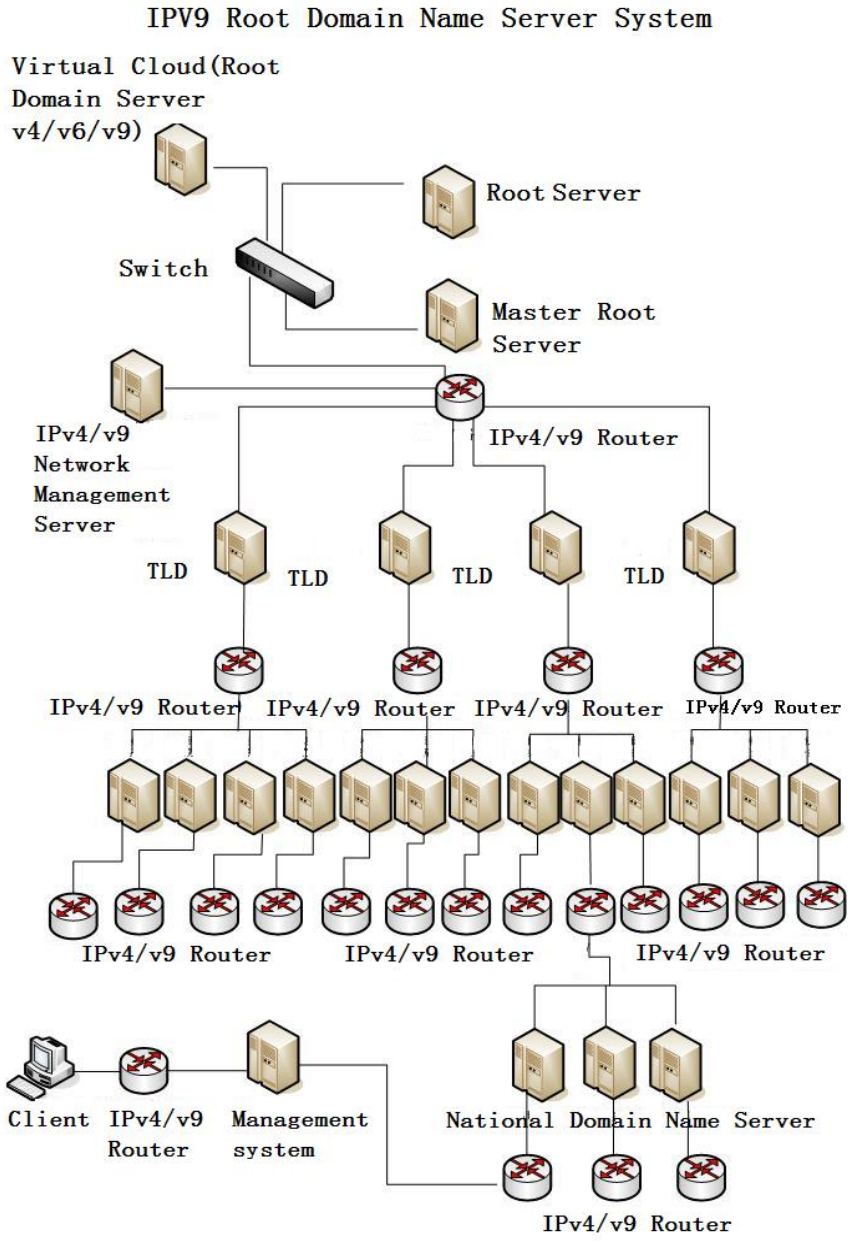


Figure 3. The System of IPV9 Root Name Server

This 13 root domain name resolution systems based on IPV9, able to adapt to IPv4 network, IPv6 network, IPV9 network, through the organization and construction of decimal network technology, with a safe and controllable appearance for global users, and can provide services and personality in various languages Provide broadband, multimedia communication service communication network to provide English, digital, Chinese domain name resolution function. The IPV9 resolution system can ensure that the domain names used by online users are resolved by the domain name server to obtain the IP address of the corresponding access object, and can also send requests for non-numeric domain names to the corresponding English domain name server or Chinese domain name server, as well as various Language domain name servers, while providing digital domain name resolution functions, are also compatible with providing Chinese and English domain name resolution services.

E. The Architecture Design of IPV9

The conventional data packet exchange of the current TCP / IP protocol cannot support true real-time applications and circuit switching, and the application of circuit transmission of sound or images in the four-layer protocol. In addition, the existing TCP / IP protocol is a connectionless and unreliable data packet protocol with a maximum packet length of 1514 bytes. With the integration of voice, image and data, the establishment of a new network theoretical foundation has become an urgent task. The design purpose of IPV9 is to avoid large-scale changes of the existing IP protocol, leading to the next-generation Internet can be backward compatible. The main idea of the design is to merge the IP protocol of TCP/IP with circuit switching. Using a router compatible with the two protocols, the designer envisions that through a series of protocols, the addresses of the three protocols (Ipv4/Ipv6/IpV9) Simultaneous use in the Internet, gradually replacing the current Internet structure without excessively affecting the current Internet. Due to the rational design of IPV9, it has received the attention of ISO and the International Internet Association.

1) The level system of IPV9

The IPV9 system uses a three-layer circuit / four-layer packet hybrid network architecture, and adopts the communication network transmission mode of authentication before communication rules. It was first proposed by China and has formed a demonstration project. The architecture is shown in Figure 4.

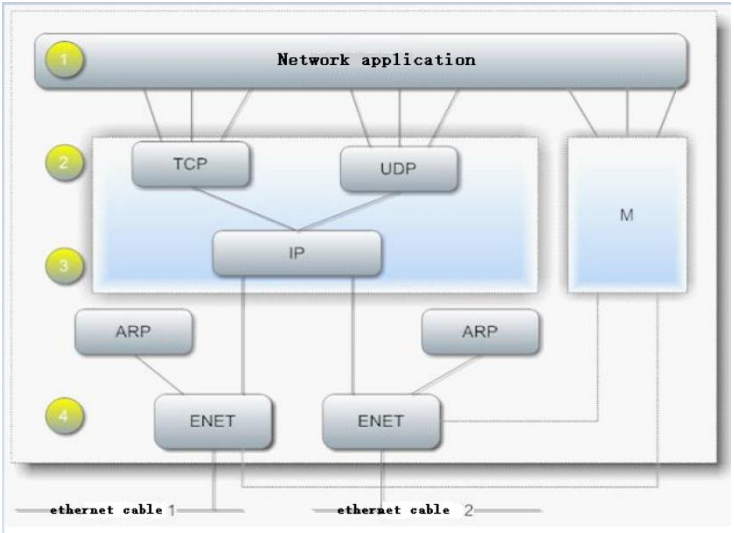


Figure 4. The level system of IPV9

2) The Connection of IPV9

IPV9's TCP / IP / M protocol, in addition to inheriting the existing TCP / IP protocol connectionless and unreliable data packet protocol, also develops absolute code streams and long stream code classes. Long packets can reach more than tens of megabytes. Can use three layers to directly transmit telephone and cable TV data to establish a four-layer three-layer transmission protocol with the new transmission theory. The connection method is shown in Figure 5.

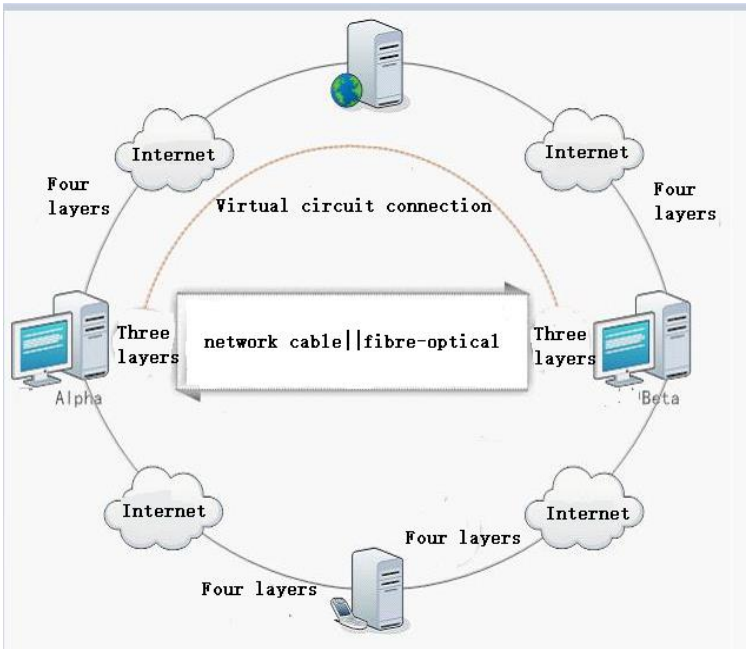


Figure 5. The Connection of IPV9

The IPV9 network management system is a comprehensive network management system that provides network monitoring and other functions based on a web interface. It can monitor various network parameters and server parameters to ensure the safe operation of the server system; it also supports IPV4 and IPV9 protocols, and provides a flexible notification mechanism to allow system administrators to quickly locate and solve various problems.

Through the use of IPV9 design routers, clients, protocol conversion routers and other equipment to build a pure IPV9 network, IPV9/IPv4 hybrid network to achieve a new generation of independent intellectual property security and control of the Internet system. Including the domestically-controlled and self-controllable IPV9/future network root domain name system, promoting technology integration, business integration, and data integration to achieve cross-layer, cross-region, cross-system, cross-department, and cross-business collaborative management and services. Take data centralization and sharing as a means to build a nationally integrated national big data center and gateway bureau, speed up the promotion of domestically made independent and controllable alternative plans, and build a safe and controllable information technology system. Be independent of the control of the US domain name system and realize an independent domain name system.

F. The Application Examples of IPV9

1) The application of 5G-future network/IPV9 movie network release application

Now the 5G network of China Unicom Beijing and China Mobile Suzhou have been directly connected through the IPV9 fiber routing backbone node of

Beijing University of Posts and Telecommunications and the IPV9 national backbone optical cable network, and achieved the world's first time End-to-end 500Mbps to 1000Mbps speed on May 21 this year. On the IPV9 national backbone network +5G local access/5G core network, the digital film program network distribution work was successfully carried out, and the national network distribution of Chinese movies was first entered in the new era of "one hour".

2) "Health Tai'an " IPV9 Big Data Platform

"Health Tai'an "IPV9 big data platform project relies on the existing backbone optical cable and user transmission access network of Shandong Broadcast Network Co. Ltd. Tai'an Branch, using IPV9 network technology to upgrading and construction, cover the medical and health institutions of the city, county, township and village levels and the medical insurance bureau, the administrative department and the finance bureau of Tai'an, and further expand to families and individuals. The bandwidth meets the requirements of healthy Tai'an big data business and can be sustainable. The expansion realizes compatible security operation between IPV9 network and IPv4 network (also realizes logical security isolation between IPV9 and IPv4 and IPv6 networks).

IV. CONCLUSION

Nowadays, the lack of IP addresses has become the main reason restricting its development. IPV9 has a huge address capacity, and it is better than IPv6 in terms of security, compatibility, efficiency, and cost savings. It is more suitable in China Development. This article introduces the characteristics, production process, protocol and composition of IPV9. IPV9 is independently developed by Chinese and has independent intellectual property rights. At the same time, it can solve the remaining problems of IPv4 and can be the core key technology of the next generation Internet. The new network should not be an upgrade of the old network, but a new network system structure. If it can be promoted, it will definitely promote the great development of the Internet.

REFERENCE

[1] Information technology-FutureNetwork-Problem statement and requirement-Part 5: Security, ISO/IEC DTR 29181-5, 2014, 12.

[2] HeYudan, Zhu Lian, etc. Commercial radio-frequency identification tag data format. SB/T 10530-2009, 2010, 03.

[3] Xie Jianping, Kong Ning, Etc. Domain name specification based on RFID Technology used for products and service. SJ/T11605-2016, 2016, 06.

[4] Wang Wenfeng, Xie Jianping, etc. Product and service digital identification format for information procession.SJ/T11603-2016, 2016.06.

[5] Xie Jianping etc. Digital domain name specification, SJ/T11271-2002, 2002.07.

[6] Xie Jianpingetc. A method of assigning addresses to network computers using the full decimal algorithm[P]. CN: ZL00135182.6, 2004.2.6.

[7] Tang Xiaodan etc. Computer Operating System (third edition) [M]. Xi'an: Xidian university press, 2010

[8] Xie Jianping, Xu Dongmei, etc. Digital domain name specification.SJ/T11271-2002, 2002.07.

A Study of Edge Computing Offloading Based on Security

Luo Pei
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail: 1985815398@qq.com

Lei Juchao
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail: 274766943@qq.com

***Abstract*—With the widespread use of IOT scenario and the development of mobile network technology, the access to many devices at the edge of the network has generated an exponential increase in the volume of data, bringing higher than ever before the demand for data transmission bandwidth, traditional centralized cloud computing can no longer meet the demand, the need to adopt edge computing distributed collaborative approach to data processing. In this paper, we study a mobile edge computing model based on energy consumption optimization management under a certain delay constraint, focus on an edge computing offload scheme based on security management, and design an offload decision scheme based on a multi-objective optimization hybrid quantum evolution algorithm (MHQEA) to ensure the security of user equipments during computing offload in the edge network and reduce total system energy consumption.**

Keywords-Edge Computing; Computing Offloading; Resource Allocation

I. INTRODUCTION

In recent years, with the rapid emergence of compute-intensive applications such as high-definition video, augmented reality/virtual reality, Internet of Things, Internet of Vehicles, and industrial Internet, people have put forward higher demands on the transmission rate and service quality of the network, resulting in an explosive growth in network traffic. At the same time, mobile devices are gradually emerging, due to the increasingly powerful design features, mobile devices have limited computing power, poor real-time and energy consumption limitations, it is difficult to bear the needs of computation-intensive, latency-sensitive and high energy consumption applications, In order to meet these challenges, the first solution proposed is to apply compute offload technology in mobile cloud computing (Mobile Cloud Computing (MCC)) architecture. By offloading the compute tasks of the terminal to a cloud data center with sufficient compute and storage resources, the delay and power consumption problems caused by the lack of computing power of the smart terminal can be alleviated to some extent.

In order to solve the latency problem of MCC, researchers have focused on mobile edge computing, which is a core technology of 5G in 2014[2], it was proposed by the European Telecommunications Standardization Institute (ETSI) to provide computing, communication and storage capabilities closer to users by sinking cloud data centers to the wireless network edge. In mobile edge computing (MEC) systems, network edge devices such as base stations, access points and routers are given cloud-like computing and storage capabilities to serve users as an alternative to the cloud. At the same time, because it is placed close to mobile device terminals and data sources, it can significantly reduce mobile device load and reduce transmission latency. Among them, compute offloading, as a cutting-edge technology for edge computing, can be achieved by offloading compute tasks to a well-resourced edge cloud. The problem of computational offloading is mainly focused on offloading decisions and resource allocation, and many scholars have done relevant research on these issues. in the literature[3], a one-dimensional search algorithm is used to minimize the latency target and optimize the user task execution latency based on the application queue buffer queue state. Experimental results show that the optimal strategy proposed in this scheme can reduce the latency by up to 80% compared to the local execution strategy and up to 44% compared to the cloud execution strategy, which can effectively respond to the dynamic arrival of intensive applications; the user device in the literature[4] adopts energy harvesting technique to optimize the energy consumption of the execution task, and proposes that the dynamic optimization algorithm can reduce the execution time by up to 64% by offloading the task to the MEC, while preventing the offloader from being dropped. The shortcomings of both papers are that they do not consider the energy

consumption of the user device, in the literature[4], it is assumed that the user device utilizes energy harvesting techniques and the energy consumption of the user device is ignored in the decision making process, energy harvesting techniques do not completely solve the energy consumption problem. The literature[5] investigates offload decisions during single node computation, the goal of which is not only to minimize execution latency but also to minimize the power consumption of edge computation, considering dense user devices capable of accessing multiple edge servers through nearby small nodes, and then proposes an optimal solution using an equivalent discrete framework, however, as the number of edge servers increases, this approach leads to high communication overhead and high computational complexity. Therefore, the authors solve this problem through an application assignment indexing scheme, which broadcasts through the node's own indexing policy and selects the most appropriate edge server to minimize execution latency and energy consumption, the downside of this scheme is that it does not consider scenarios with multiple compute nodes. The literature[6] considers the problem of joint optimization of energy consumption and latency of devices in a multi-user, multi-channel environment, where devices can optimize performance by adjusting the weighting parameters, however, this literature assumes that the computational resources of the MEC are sufficient, and the problem of insufficient edge computing resources exists in realistic production environments with multi-user conditions. However, the above literature does not consider the impact of security issues on system performance. Security issues such as privacy breaches may be encountered during the offloading process, so security issues are equally important in the study of edge computing offloading.

Unlike the aforementioned literature, this paper designs a secure computational offloading method, the main contents of which are as follows: considering the impact of time delay and energy consumption on the system, the resource allocation in the task transfer process and the offloading decision in the task processing process are jointly optimized to achieve the cost optimization of mobile devices, and security is added on top of this. The specific work is as follows:

1) Design the MEC network architecture and ensure security.

2) Considering minimum energy consumption under time constraints and joint optimization to improve QoE.

3) An offloading decision solving method based on the Multi-Objective Optimization Hybrid Quantum Evolution Algorithm (MHQEA) is proposed.

4) Through simulation and comparison with conventional algorithms, it is proved that the energy saving cost proposed in this paper is lower, the total energy consumption of the system is lower, and the safety of the system is guaranteed.

II. SECURITY-BASED MEC CALCULATION OFFLOADING SCHEME

The compute offload technology offloads the compute tasks of the user terminal to the cloud service, solving the bottleneck in compute performance, resource storage, etc. of the user terminal. Computing offload technology was first proposed in the cloud, and the emergence of MEC has provided a new direction for the development of computing offload technology. Performing computational offload tasks in MEC not only solves the problem of high network resource utilization, but also solves the problem of latency. Therefore, the convergence of EMC and compute offload is an important direction for network development in the future.

In mobile edge computing, the edge cloud needs to process tasks from a variety of mobile terminals in real time, and needs to coordinate the distribution of edge server resources and task loads to meet the heterogeneous requirements of different tasks for processing delay, execution energy consumption and reliability. Offload decision and resource allocation will directly affect the performance of computing offload in mobile edge computing, which has great research significance.

A. Security-based system model design

Although edge computing networks reduce request latency and core network pressure and improve network performance, but some problems in network security have been exposed, especially its distributed deployment mode leads to network security reduction, making DOS attacks easier, therefore, how to ensure the security of edge computing becomes one of the problems facing edge computing networks.

The system establishes a model of multi-user multi-edge computing services in a wireless network. It is assumed that each edge server has complete control over the channel gain, local calculation energy, and input data size of all end users. In daily life, the number of users is greater than the number of base stations set in the edge cloud. We assume that the MEC system

consists of m mobile users and n base stations that can be connected to the edge cloud.

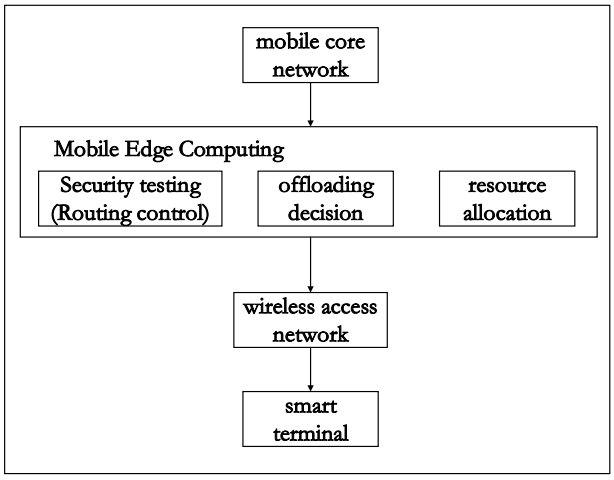


Figure 1. Security-based mobile edge computing network architecture

B. Design of Unloading Scheme for Edge Computing Based on Security

The computational offloading technique in edge computing involves the offloading decision and resource allocation, etc. The offloading decision is optimized for the performance of different services, and in this paper, the offloading decision is carried out with minimum energy consumption under the delay constraint, and the resource allocation is considered after the decision is offloaded.

1) Network model

We use C={c1,c2,... ,cn} to represent the different edge nodes, and the MEC set of each edge node is represented as M={m1,m2,... .mp}, which provides computational offload for mobile terminals, and the UE set is denoted as N={n1,n2,... .nq}.

2) Offload decision model

The offload decision model, as the core issue of computational offload techniques, is largely dependent on the computational task through the computing power of the device itself and the time delay and energy consumption that results when the offloading is completed to the edge. Therefore, it is necessary to calculate and analyze the delay and energy consumed by local computing and edge computing to complete the computing task.

The unloading decisions within each time slot are expressed in matrix form, i.e., the unloading decision matrix is represented by A, as in (1).

$$A=\begin{bmatrix} a_{11} & \dots & a_{1,m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,m} \end{bmatrix} \tag{1}$$

Where, $a_{n,m} \in \{0, 1\}$ denotes whether device n offloads the task to the edge server, $a_{n,m}$ =0 denotes local execution, and anyway denotes the task is offloaded to the edge server. Each edge server can only handle one task at a time. Thus the constraints are as in (2).

$$\sum_{n=1}^{N} \mathrm{a}_{n,m} \le 1, \sum_{m=1}^{M} \mathrm{a}_{n,m} \le 1 \tag{2}$$

3) Local execution

Local execution means that user hands over task t to the mobile device for direct execution and the local execution of the task consumes mainly computational delay, as in (3).

$$T_{\mathrm{n}_i}^{l} = \frac{C_{n_i}}{f_{n_i}^{l}} \tag{3}$$

where C_{n_i} represents the CPU required to complete the task and $f_{n_i}^{l}$ represents the computing power of user n's mobile device.

The energy consumption for the local execution of task t is mainly the energy consumption for processing computational tasks, as in (4).

$$E_{\mathrm{n}_i}^{l} = \alpha_{\mathrm{n}_i} \cdot C_{n_i} \tag{4}$$

In the formula, α_{n_i} indicates the energy consumption factor of the mobile device.

According to formula (3) and formula (4), the locally executed valuation function is obtained as in (5).

$$\begin{aligned} & D_{\mathrm{n}_i}^{l} = \beta_{n_i}^{l} T_{n_i}^{l} + \beta_{\mathrm{n}_i}^{e} E_{n_i}^{l} \\ & s.t. \quad \beta_{n_i}^{l}, \beta_{n_i}^{e} \in [0,1], \beta_{n_i}^{l} + \beta_{n_i}^{e} = 1 \end{aligned} \tag{5}$$

In the formula, $\beta_{n_i}^{l}, \beta_{n_i}^{e}$ used as the weight of the local execution time and energy consumption of the

task, respectively, in order to determine the energy consumption of each user.

4) MEC server implementation

The execution mode of the MEC server refers to that the user n migrates the task t to a virtual machine on the MEC server through a wireless channel for execution, and allocates computing and communication resources to the user equipment through the trust detection server. The delay of the entire calculation task is mainly transmission delay and calculation delay, as in (6).

$$T_{\mathrm{n}_i}^{c} = \frac{B_{\mathrm{i}}}{R_i} + \frac{H_{\mathrm{i}}}{F} \tag{6}$$

The data volume of the calculation result is much lower than the data input volume, so the transmission consumption of the calculation result is ignored. The energy consumption of the task t server mainly includes transmission energy consumption, MEC server calculation energy consumption and server monitoring energy consumption, as in (7).

$$E_{n_i}^{c} = P_{n_i} \cdot \frac{B_{\mathrm{i}}}{R_i} + \gamma H_{\mathrm{i}} + E_{monitor} \tag{7}$$

According to formula (6) and formula (7), the locally executed valuation function is obtained as (8):

$$D_{n_{\mathrm{i}}}^{c} = \beta_{\mathrm{n}_i}^{c} T_{n_i}^{c} + \left(1 - \beta_{n_i}^{c}\right) E_{n_i}^{c} \tag{8}$$

5) Security inspection

Introduce a security module in the MEC server to perform security checks on the offloading process. Because the routing module is responsible for the traffic, and because edge nodes have limited capacity compared to cloud computing centers, they are vulnerable to traffic attacks, and although individual edge nodes are compromised and nearby networks quickly find replaceable nodes to adjust to, a malicious attack can bring down a server. Therefore, traffic forwarding is divided into traffic types and firewalls are set up between the center and branches to enhance DDoS protection. It also detects virtual machine operation in real time to prevent malicious virtual machine migration behavior.

The security detection module mainly detects the offloaded computational tasks through entropy detection algorithm, which can detect anomalies sensitive to the information entropy of changes in network parameters and UE parameters. In the case of potentially malicious offloads, the MEC controller transfers them to the security monitoring server, updates their trust values, labels them as trust violations, and verifies the user's access to the network through a combination of trust violation values and information entropy, increasing detection accuracy and reducing system overhead by unifying the monitoring of potentially malicious users. If there is any error in security detection, the defense policy module will continue to secure the network through reallocation of resources or security transfer.

The entropy detection algorithm can accurately sense changes in network parameters and calculate the corresponding information entropy to detect whether a user is maliciously unloading. Energy is consumed during the detection of offloading computational tasks, aiming at the minimum energy consumption under time constraints and based on a safe offloading scheme, where the attributes involved in the entropy detection algorithm are user trust, offloading frequency, network environment, CPU and memory utilization, etc. The distribution of the property z in G belongs to a polynomial distribution, and the probability equation is as in (9).

$$P(G_{\mathrm{z}}) = \frac{|G|}{\Pi_{\mathrm{z}=1}^{|G|} z!} \Pi_{z=1}^{|G|} G_{\mathrm{z}} \tag{9}$$

Where, the set of attributes G={g1,g2, ... ,g3}($1 \leqslant z \leqslant 5$), A represents the proportion of users with attribute G_{z} to the entire system users, $R_{n,i} = \sum_{G} P$ can be calculated, we use the maximum threshold method to determine if there is an abnormal packet unloading, if it is greater than the set threshold, then it is a malicious unloading task, unloading is not allowed during the unloading decision. The entropy detection algorithm is as in (10).

$$R_{\mathrm{n}} = -\sum_{i=1}^{5} R_{n,i} \log(R_{n,i}) \tag{10}$$

Where R_{n}^{H} is the detection threshold and unloading is determined by the detection result, as in (11).

$$\delta_n = \begin{cases} 1, & R_n \geq R_n^H \\ 0, & R_n < R_n^H \end{cases} \quad (11)$$

Detection of controlled energy consumption as in (12).

$$E_{\text{monitor}} = \frac{M_n R_n}{M_c} \quad (12)$$

E_{monitor} represents the server's monitored energy consumption, M_n represents the memory resources provided by the MEC server for one user device, and M_c represents the resources available for the entire server.

The overall optimization function is expressed as in (13):

$$\min[\sum_{n\in N}\left(\sum_{c\in C}\left(\delta_n E_{n_i}^l\right)+\left(1-\delta_n\right)E_{n_i}^c\right)]$$

$$s.t.\ \ \delta_n = \begin{cases} 1, & R_n \geq R_n^H \\ 0, & R_n < R_n^H \end{cases} \quad (13)$$

$$\sum_{n\in N}\delta_n F_{n,c} \leq F_c$$

In the constraint condition, F represents the total CPU computing resources of the MEC server. The computing resources allocated by the MEC cannot exceed the total computing resources.

III. Algorithm Designs

In the security-based computing offload scenario, the problem is an NP-hard problem because of the large data volume. To further solve this problem, this paper adopts a solution of MHQEA to find an optimal approximation of the model, and the process of finding an optimal approximation is represented as Algorithm 1.

IV. Simulation Performance Evaluation

Set the system time gap length to 20ms, calculate the local calculation energy consumption and unloading energy consumption by the total number of system user equipments 10, 20, 30, 40, 50, 60, respectively, and repeat each set of calculation 1000 times to take the average. As shown in Figure 2.

Algorithm 1: A computational offloading algorithm based on MHQEA

```
begin
    The current iteration number t = 0, and the maximum iteration
number is set to M
    Initialize the offload decision matrix An(t)
    Find the optimal solution matrix Pn(t) by observing the state of
An(t) through the make subroutine
    Correction of Pn(t) by reparation subroutine
    Evaluate the total energy consumption of the system and find the
optimal solution B(t)
      while(t<M)do
      begin
      Number of current iterations t plus 1
      An(t-1) is observed by the make subalgorithm to determine the
Pn(t)
      Evaluate Pn(t) for the minimum total system energy consumption
      Updating An(t) through the update subroutine
      Find the optimal solution in Pn(t) and B(t-1) to b
      if(Current number of iterations meets migration conditions)then
        Migration b to B(t)
        end if
    end
end
```

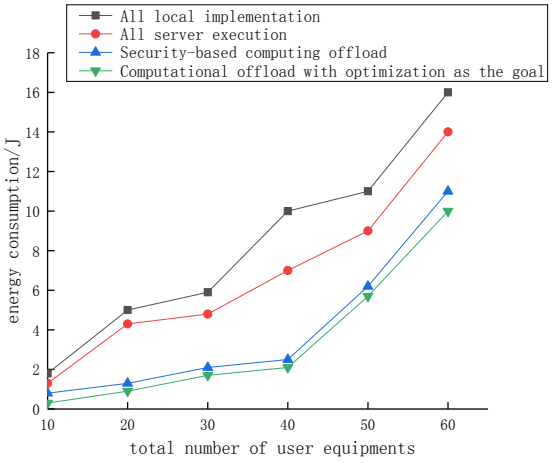


Figure 2. Relationship between offloading delay and total number of user equipments in different ways

As can be seen from Figure 2, the system energy consumption is higher for all local execution and all server execution, and the energy consumption is slightly higher for a security-based offload decision scheme than for an optimized energy-based offload

scheme. Because of the energy consumption required for security testing, it improves the overall security of the system and ensures the quality of service for the user's device unloading.

The optimal energy-based delay and security-based delay are calculated using the total number of user equipments in the system 10, 20, 30, 40, 50 and 60, respectively, and the average of each group is repeated 1000 times. As shown in Figure 3.

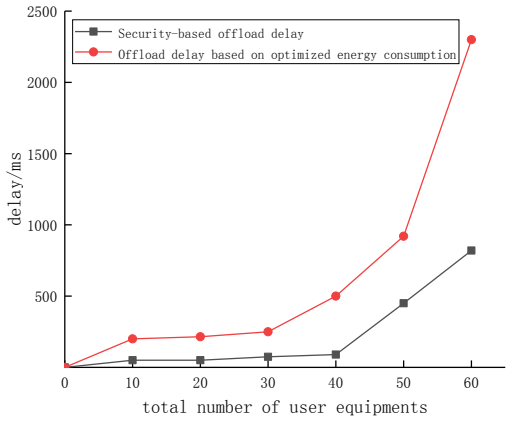


Figure 3. Relationship between different ways of offloading energy consumption and total number of user equipments

As can be seen from Figure 3, a security-based offloading decision scheme is better than an energy-optimized offloading scheme.

V. CONCLUSION

In the article, an offloading scheme for edge computing under security is proposed, which minimizes the total overhead of the system and improves the security of the offloading process, and we use an MHQEA-based algorithm to find the optimal offloading decision matrix. The simulation results show that the total system overhead for this option is lower than for other options, while maintaining safety.

Because MEC servers are small data centers, each server has far less energy than a cloud data center. With the development of the industrial Internet, MEC servers will be heavily deployed, which will cause the overall energy consumption of the system, how to deal with the greater energy consumption is also something we should study in the future.

REFERENCES

[1] Xie Renchao, Huang Tao, Yang Fan, Liu Yunjie. Principles and practice of edge computing [M]. Beijing: People's Publishing House, 2019.152-155

[2] Taleb T, Samdanis K, Mada B , et al. On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration[J]. Communications Surveys & Tutorials, IEEE, 2017, 19(3):1657-1681.

[3] Liu J, Mao Y, Zhang J, et al. Delay-Optimal Computation Task Scheduling for Mobile-Edge Computing Systems[J]. 2016.

[4] Mao Y, Zhang J, Letaief K B. Dynamic Computation Offloading for Mobile-Edge Computing with Energy Harvesting Devices[J]. IEEE Journal on Selected Areas in Communications, 2016:1-1.

[5] Guo X, Singh R, Zhao T, et al. An index based task assignment policy for achieving optimal power-delay tradeoff in edge cloud systems[C]//ICC 2016 - 2016 IEEE International Conference on Communications. IEEE, 2016.

[6] Chen X, Jiao L, Li W, et al. Efficient Multi-User Computation Offloading for Mobile-Edge Cloud Computing[J].IEEE/ACM Transactions on Networking, 2015:1-1.

[7] Chen Yanbing, Li Juan, Zhang Qi. Exploration of quantum evolution algorithms for solving constrained multi-objective optimization problems[J]. Electronic Technology and Software Engineering, 2015, 000(016):P.185-186,187.

Application Research of Crawler and Data Analysis Based on Python

Wu Hejing
East University of Heilongjiang
Heilongjiang, 150086
E-mail: 499917928@qq.com,

Zhao Long
East University of Heilongjiang
Heilongjiang, 150086

Shao Yabin
East University of Heilongjiang
Heilongjiang, 150086

Liu Fang
East University of Heilongjiang
Heilongjiang, 150086）

Cui Ran
East University of Heilongjiang
Heilongjiang, 150086

***Abstract*—Combined with the actual situation, this paper explores how to develop a crawler method based on the specific framework for the complete interface of steam manufacturers and stores, which should be able to automatically and efficiently crawl the data of specific targets, analyze the dynamic pages, and complete the data cleaning, downloading, saving and other operations, explore the methods of general data analysis, and Analyze the downloaded data, extract useful information from it, analyze and summarize the specific crawler method and data analysis method through practical application.**

Keywords-Python; Scrapy; Selenium; BeautifulSoup

I. INTRODUCTION

The 21st century is a book written by information. With the rapid development of information technology, today's society has become a huge information polymer, and there are various kinds of data in this huge polymer. Data is a kind of embodiment of information. In this era of information explosion, how to efficiently find the data we want from all kinds of miscellaneous data and extract them from the network in batches has become a key problem. However, sometimes the unprocessed data itself may be confusing for people. How to process the huge and complex data obtained through what kind of technical means, and finally become an intuitive number, or trend, and become the information that people can obtain intuitively is also a very important topic to be studied in this data age.

II. STATISTICAL INVESTIGATION ON THE PREFERENCE SALES VOLUME

In this project, the American Steam online game platform mall is selected as the research object of the crawler. By setting a specific game company as a search keyword in steam's online mall, the data of all works of the company in steam mall are crawled, and the useful information is extracted by analyzing the basic data of each manufacturer's preference for game production type, series sales volume, and praise In addition, the game manufacturers are comprehensively scored and evaluated.

DOI: 10.21307/ijanmc-2020-018

III. Relevant Technology and Framework

This project will use the scrapy framework based on Python language to crawl steam website. Python as a language has the advantages of lightweight, simplicity, wide range of application and so on. At present, various crawler frameworks and application libraries based on Python have been very mature, among which the crawler framework is very popular in the application of general web crawlers. Its first version was released in 2008, and now it is quite mature as a crawler framework. The basicprinciple of the scrapy framework is shown in Figure 1.

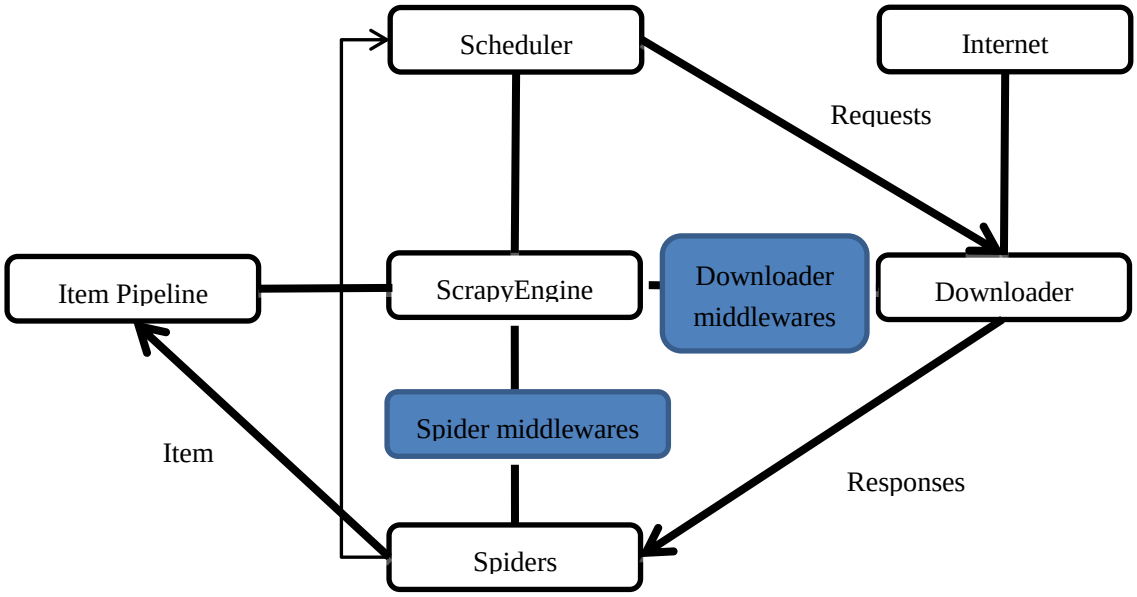


Figure 1. Basic principles of Scrapy frame

IV. Design of Crawler

A. General design idea

The process of crawler itself is actually to simulate the user's operation on the browser with a program. First of all, the starting point and range of crawling need to be specified. As the target of crawling is for manufacturers and their works, the interface of manufacturers is taken as the starting point. For example, the page of paradox, a manufacturer, first analyzes the entire manufacturer's page, and finds that the page links and information of all games or game related DLC downloads of the manufacturer are stored in the recommendation div framework of each sub recommendation of recommendations rows, as shown in Figure 2

B. Design and implementation of reptile functions

The crawler architecture is composed of items, spiders, piplings and middleware. Among them, items are mainly used to define the items to be crawled, spiders are responsible for defining the whole process of crawling, what means to crawl, pipes are responsible for the basic operations such as data cleaning and saving, middleware can be responsible for the bridge service of scratch and other plug-ins or architectures.

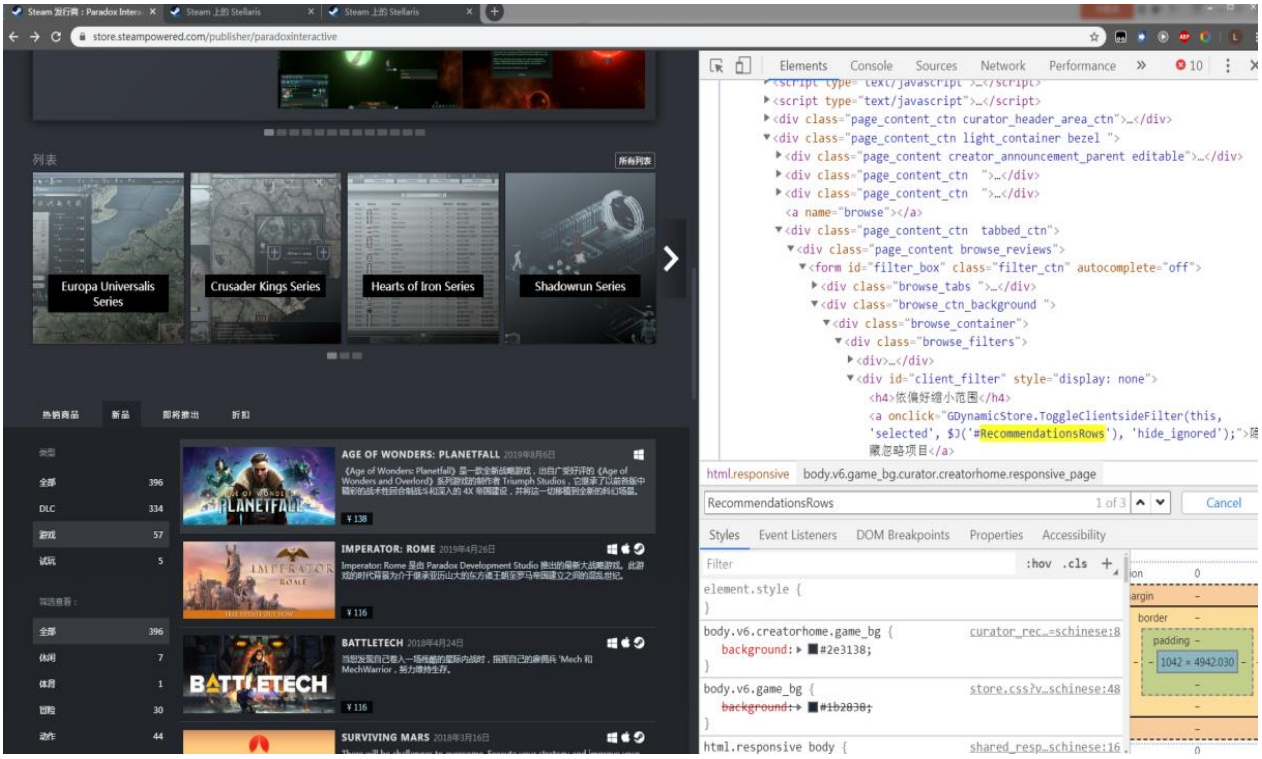


Figure 2. Investigation of HTML page structure of steam manufacturers by using viewers

First, the items to be crawled are defined in the items file. Finally, these items may be submitted to the analysis part for data analysis. The specific design and implementation code is:

```
import scrapy
class SteamDevItem(scrapy.Item):
    # define the fields for your item here like:
    # name = scrapy.Field()
    qry_nam = scrapy.Field()
    if_dev = scrapy.Field()
    pub_sum = scrapy.Field()
    pub_gam_sum = scrapy.Field()
    pub_dlc_sum = scrapy.Field()
    dev_nam = scrapy.Field()
    pub_nam = scrapy.Field()
    gam_title = scrapy.Field()
    res_date = scrapy.Field()
    gam_type = scrapy.Field()
    gam_tag = scrapy.Field()
    if_muti = scrapy.Field()
    gam_score = scrapy.Field()
    gam_score_sum = scrapy.Field()
    gam_score_ratio = scrapy.Field()
    pass
```

C. Spider design

The design of spider is the key point of this project. Whether the initial dynamic page connection or the last static page information crawling mode will be defined

in this file. In this project, spider will be named steam, and some key implementation codes will be pasted here, with running results and some notes attached. First, introduce start_ the design method of dynamic page crawling of selenium in requests method:

```
chrome_opt = webdriver.ChromeOptions()
    prefs = {

"profile.managed_default_content_settings.images": 2,
        'permissions.default.stylesheet': 2
    }
    chrome_opt.add_experimental_option("prefs", prefs)
    browser = webdriver.Chrome(options=chrome_opt)
    browser.get("https://store.steampowered.com/" + Qry_sta + "/" + Qry_Target)
    bs = BeautifulSoup(browser.page_source, 'html.parser')  #Beautiful Soup
```

The specific store connections of each product exist in the a anchor label of each entry, and these connections are read to the defined links using the loop_ In the list list, crawling of the list is completed, but sometimes the text and picture in the entry may contain a tag, and they all point to the same page. If direct application may cause repeated crawling, a loop is used here, and if not in statement is used to de duplicate the list.

After using the print statement to verify the function of the module, the verification results are shown in Figure 3.

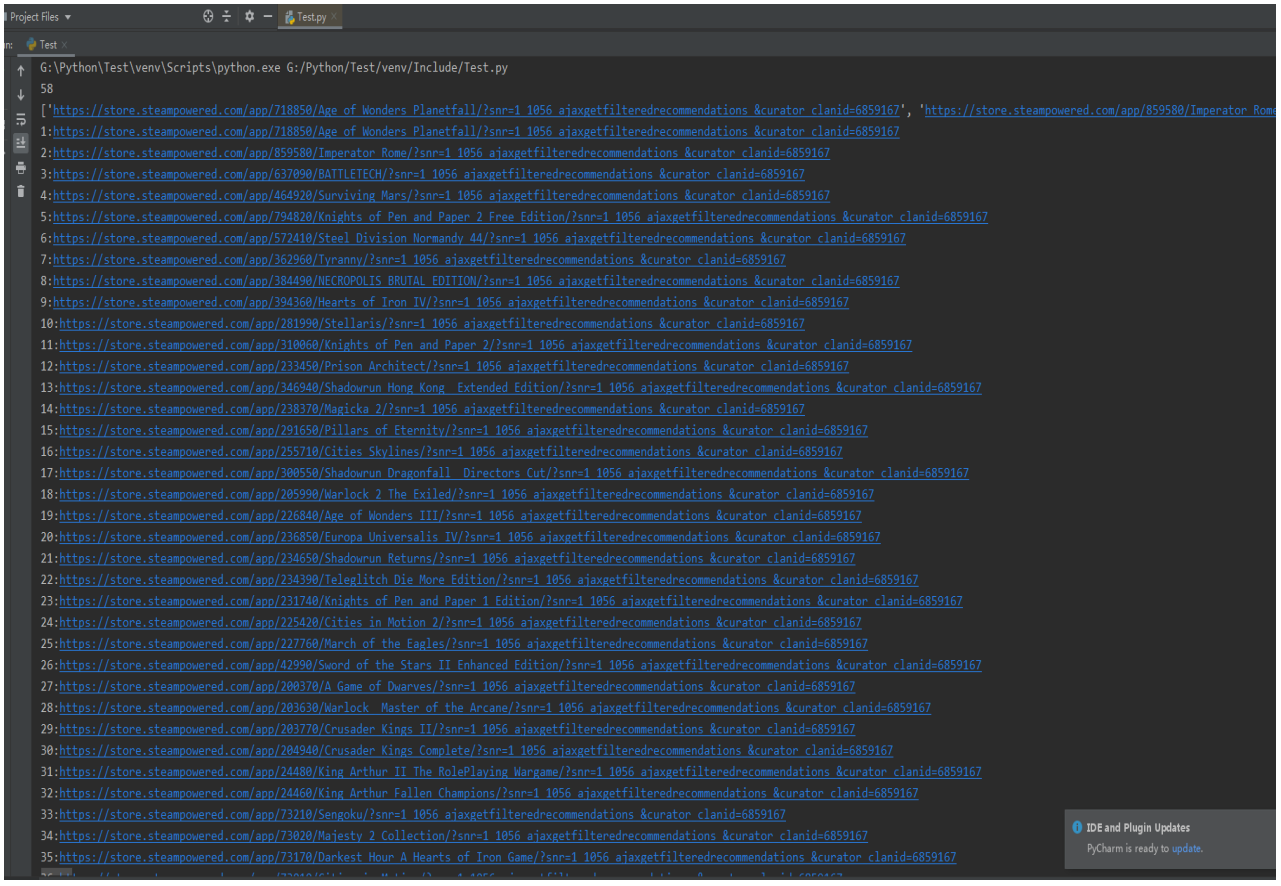


Figure 3. List of URLs obtained by selenium and beautiful soup

D. Start directional climbing

After designing and debugging the spider, run the CMD command window of the system, open the root directory of the crawler file, and input the crawler stream-o SteamDev.csv , crawl the target website. Input - O SteamDev.csv The purpose is to let the crawler save the last crawled data in the form of CSV table. The saved data appears in the project root. See Figure 4 for the climbing process.

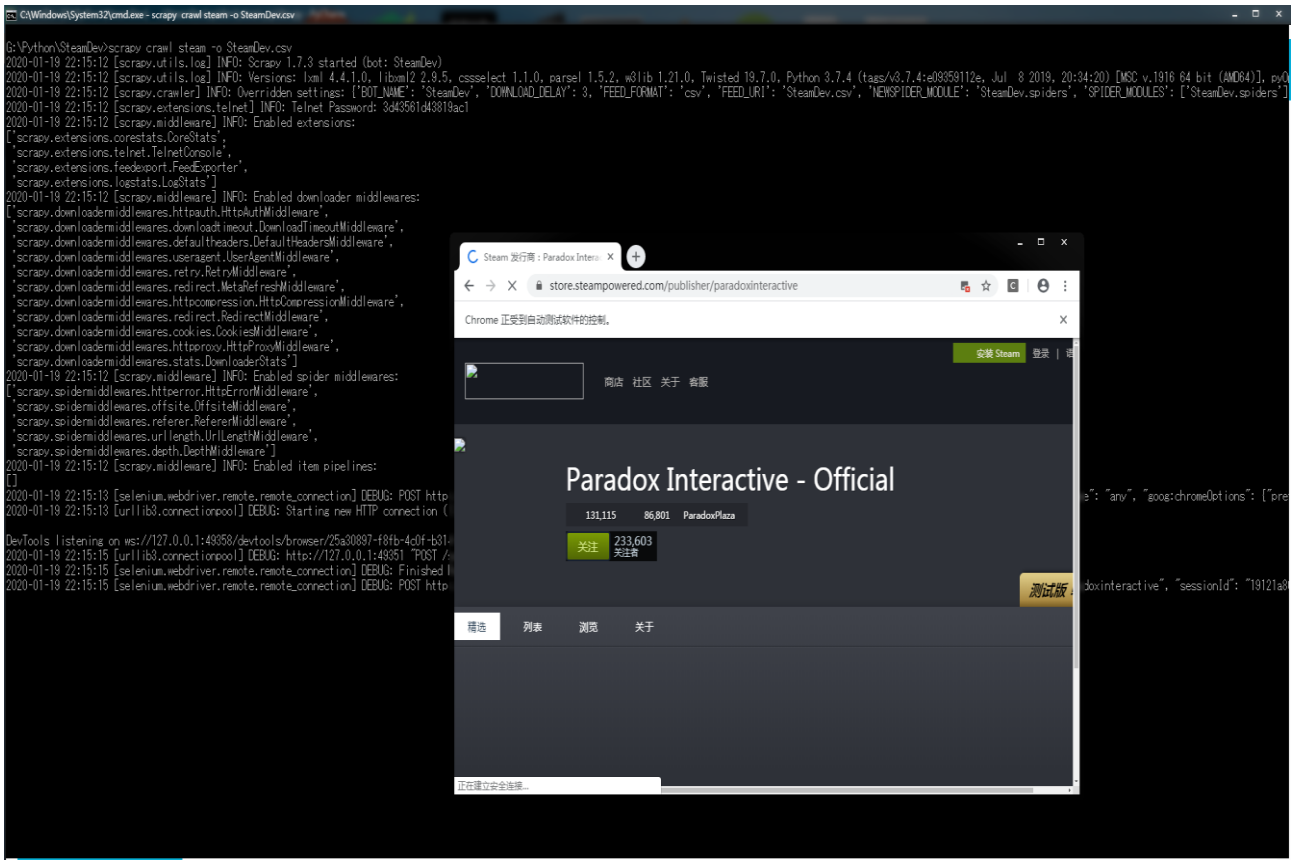


Figure 4. Executing the start request method selenium pop-up browser to crawl the dynamic page

V. DATA ANALYSIS

Next, we will perform basic visual operations on the crawled data in the form of operation tables. In the crawler project, we crawled for the Paradox Interactive publisher. The crawled data is presented in the form of CSV tables, as shown in Figure 5.

Through the use of spreadsheets and further collation of the crawled data, the following data are obtained: the publisher has published 396 works in steam platform, of which the majority of DLC has published 334 DLC, most of the games published are single player games, and each game published in its mall has an average of 6800 reviews, of which the proportion of favorable reviews is about 76.4 8%, see the chart below for detailed visual analysis.

dev_nam	gam_score	gam_score	gam_score	gam_tag	gam_title	gam_type	if_dev	if_muti	pub_dlc_s	pub_gam_s	pub_nam	pub_sum	qry_nam	res_date
							发行商		334	57		396	Paradox Interactive - Official	
Triumph S	1845	7	Very Posi	Strategy,	Age of W	Strategy		Single-player			Paradox Interactive			6 Aug, 2019
Paradox I	11187	6	Mostly Po	Strategy,	Imperator	Simulation, Strateg		Single-player			Paradox Interactive			25 Apr, 2019
Harebrair	10594	7	Very Posi	Mechs, Str	BATTLETEC	Action, Adventure, S		Single-player			Paradox Interactive			24 Apr, 2018
Haemimont	7060	9	Very Posi	Colony Si	Surviving	Simulation, Strateg		Single-player			Paradox Interactive			15 Mar, 2018
Kyy Games	116	6	Mixed	RPG, Indie	Knights o	Indie, RPG, Simulati		Single-player			Paradox Interactive			21 Feb, 2018
Eugen Sys	3661	7	Very Posi	World War	Steel Div	Action, Simulation,		Single-player			Paradox Interactive			23 May, 2017
Obsidian	5949	9	Very Posi	RPG, Story	Tyranny	Adventure, RPG		Single-player			Paradox Interactive			10 Nov, 2016
Harebrair	3384	6	Mostly Po	Souls-lik	NECROPOLI	Action, Adventure, I		Single-player			Paradox Interactive			16-Jul
Paradox I	47255	9	Very Posi	Space, Str	Stellaris	Simulation, Strateg		Single-player			Paradox Interactive			9 May, 2016
Kyy Games	1234	9	Very Posi	RPG, Adver	Knights o	Adventure, Indie, RF		Single-player			Paradox Interactive			20 Oct, 2015
Double El	34388	9	Very Posi	Simulatio	Prison Ar	Indie, Simulation, S		Single-player			Paradox Interactive			6 Oct, 2015
Harebrair	2106	9	Very Posi	RPG, Cyber	Shadowrur	Adventure, Indie, RF		Single-player			Paradox Interactive			20 Aug, 2015
Pieces Ir	5097	7	Very Posi	Magic, Co-	Magicka 2	Action, Adventure		Single-player			Paradox Interactive			26 May, 2015
Obsidian	9125	9	Very Posi	RPG, Fante	Pillars o	RPG		Single-player			Paradox Interactive			26 Mar, 2015
Colossal	74207	9	Very Posi	City Buil	Cities: S	Simulation, Strateg		Single-player			Paradox Interactive			10 Mar, 2015
Harebrair	3520	9	Very Posi	RPG, Cyber	Shadowrur	Adventure, Indie, RF		Single-player			Paradox Interactive			18 Sep, 2014
Ino-Co Pl	697	6	Very Posi	Strategy,	Warlock 2	Strategy		Single-player			Paradox Interactive			10 Apr, 2014
Triumph S	5144	9	Very Posi	Strategy,	Age of W	RPG, Strategy		Single-player			Paradox Interactive			31 Mar, 2014
Paradox I	47223	9	Very Posi	Grand Str	Europa Ur	Simulation, Strateg		Single-player			Paradox Interactive			13 Aug, 2013
Harebrair	7411	9	Very Posi	RPG, Cyber	Shadowrur	Adventure, Indie, RF		Single-player			Paradox Interactive			25 Jul, 2013
Test3 Pro	819	9	Very Posi	Action Ro	Teleglitc	Action, Indie		Single-player			Paradox Interactive			24 Jul, 2013
Behold St	1541	9	Very Posi	RPG, Turn-	Knights o	Indie, RPG		Single-player			Paradox Interactive			18 Jun, 2013
Colossal	1016	6	Mixed	Simulatio	Cities ir	Simulation, Strateg		Single-player			Paradox Interactive			2 Apr, 2013
Paradox I	158	6	Mixed	Strategy,	March of	Simulation, Strateg		Single-player			Paradox Interactive			18 Feb, 2013
Zeal Game	305	6	Mixed	Strategy,	A Game of	Casual, Strategy		Single-player			Paradox Interactive			23 Oct, 2012
Ino-Co Pl	738	9	Very Posi	Strategy,	Warlock -	Strategy		Single-player			Paradox Interactive			8 May, 2012
Paradox I	49512	9	Very Posi	Grand Str	Crusader	Free to Play, RPG, S		Single-player			Paradox Interactive			14 Feb, 2012
Paradox I	111	6	Mixed	Strategy,	Crusader	Strategy		Single-player			Paradox Interactive			14 Feb, 2012
NeocoreGa	233	6	Mixed	Strategy,	King Arth	RPG, Strategy		Single-player			Paradox Interactive			27 Jan, 2012
NeocoreGa	27	6	Mixed	Strategy,	King Arth	RPG, Strategy		Single-player			Paradox Interactive			16 Sep, 2011
Paradox I	268	6	Mixed	Strategy,	Sengoku	RPG, Simulation, Str		Single-player			Paradox Interactive			15 Sep, 2011
1C:InoCo	701	7	Very Posi	Strategy,	Majesty 2	Strategy		Single-player			Paradox Interactive			19 Apr, 2011

Figure 5. Crawled data list

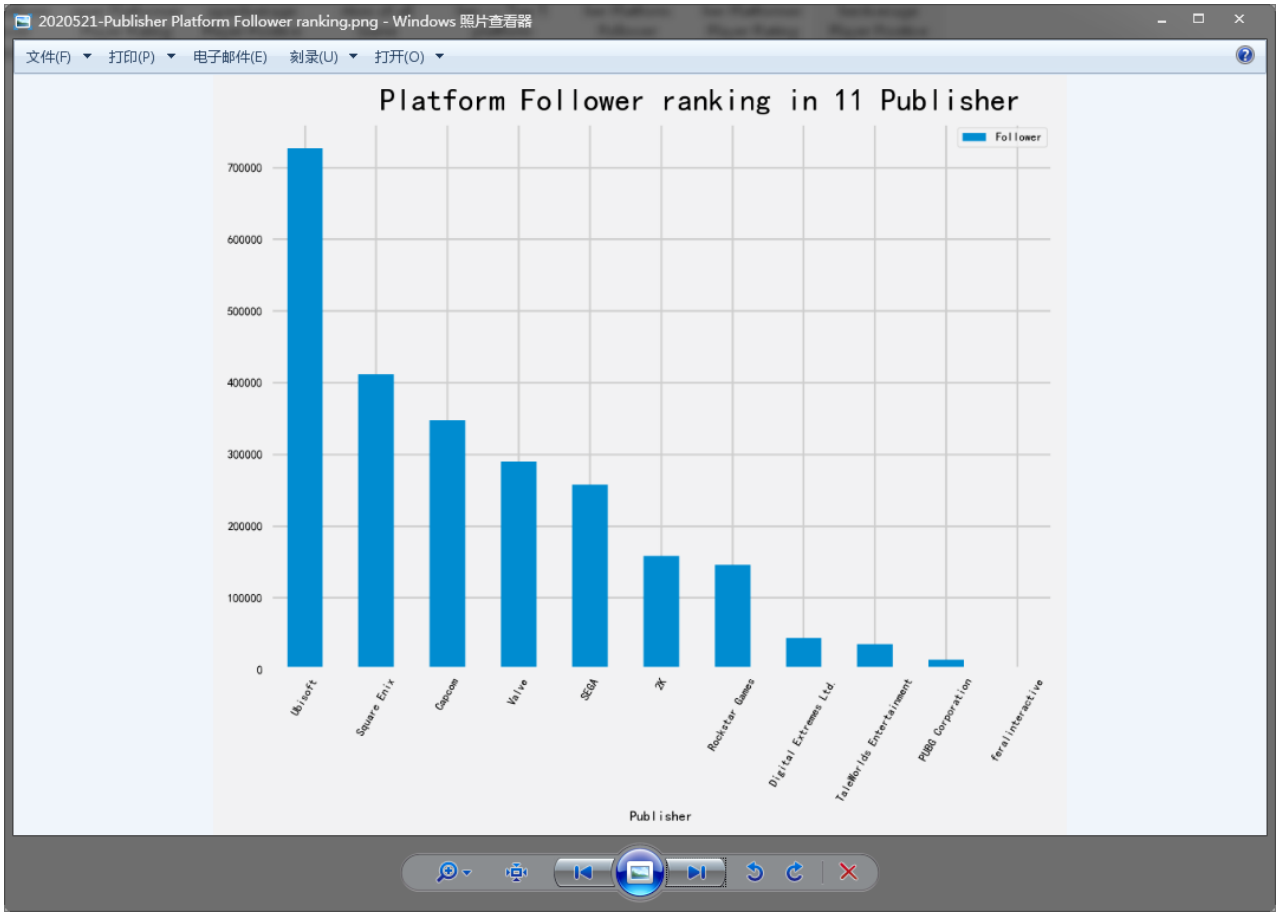


Figure 6. Output the publisher platform follower ranking chart

VI. CONCLUSION

Through demonstration and part of practice, this paper explores the process of data crawling and basic data analysis of dynamic pages by combining the general Python's story framework with selenium + beautiful soup through crawling the steam online game mall website.

The crawler has good scalability. For example, if you want to compare the crawling data of multiple game manufacturers, you can write a query manufacturer list to get the product URL list from the dynamic web page of the manufacturer list first. In terms of anti-crawler, selenium itself has a very good anti crawler ability. If you want to further anti crawler, you can also expand multiple cookies, and even establish a proxy IP pool.

ACKNOWLEDGMENT

This paper is about the scientific research project of Heilongjiang Oriental University in 2019, "Implementation of Crawler Based on Python Scrapy Framework", project number HDFKY190109

REFERENCE

[1] Yuhao Fan. Design and implementation of distributed crawler system based on scrapy[J].IOP Conference Series: Earth and Environmental Science,2018,108(4):2-8.

[2] Jing Wang, Yuchun Guo. Scrapy-based crawling and user-behavior characteristics analysis on taobao[P]. Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2012 International Conference on, 20120:1-5.

[3] Ryan Mitchell. Python web crawler authority Guide (Second Edition) [M]. Beijing: People's post and Telecommunications Press, 2019:57-70.

[4] Wei Chengcheng. Data information crawler technology based on Python [J]. Electronic world, 2018 (11): 208-209.

[5] Mark.Lutz . Python learning manual (Fifth Edition, Volume I) [M]. Beijing: Mechanical Industry Press, 2019:1-2.

[6] Fan Chuanhui. Python reptile development and project practice [M]. Beijing: Mechanical Industry Press, 2017 (3): 69-72.

[7] Song Yongsheng, Huang Rongmei, Wang Jun. research on Python based data analysis and visualization platform [J]. Modern information technology: 2019 (21): 1-4.

[8] Liu Yuke, Wang Ping. Statistics and graph output of student achievement data based on Python + pandas + Matplotlib [J]. Fujian computer. 2017 (11): 2-6.

[9] Liu Yuke, Wang Ping. Statistics and graph output of student achievement data based on Python + pandas + Matplotlib [J]. Fujian computer. 2017 (11): 2-6.

[10] Long Hu, Yang Hui. Data analysis and visualization in the context of big data [J]. Journal of Kaili University. 2016 (03): 1-3.

Design of Electric Power Line Drawing Algorithm

Sun Pengzhan
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail: 2994553975@qq.com

Lei Juchao
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, China
E-mail: leijuchao@139.com

***Abstract*—At present, the drawing of power lines is still in the state of semi-manual design, mainly completed by AutoCAD, which not only consumes human resources, but also has low efficiency. Therefore, this paper designs an automatic drawing software for power lines. It integrates drawing and data management. Based on the basic information of power lines, this software imports and automatically converts the basic power data (latitude and longitude data) containing GPS/ BeidDou positioning information into the distance, direction, angle and other information of poles and cables. Then, it realizes the function of automatically drawing the power line graph according to the scale by multiple tree traversal algorithm. This software is easy to operate and reliable, and has achieved good results in practice.**

Keywords-Power Lines; Automatic Drawing; The Database; Multiple Tree; Traversal Algorithm

I. INTRODUCTION

The power grid is directly facing the end users and is closely related to the production and life of the people. It is an important public infrastructure to serve the people's livelihood. In recent years, with the development of new-type urbanization, China has been committed to the construction and renovation of power grid, so as to realize reliable power supply, high-quality service and improve supply efficiency. Power grid construction and reconstruction has become an urgent task of the current power industry. After the completion of the power grid transformation, the power supply enterprises need to draw and archive the power line diagram, which is usually drawn semi-manually by AutoCAD. However, AutoCAD requires professionals to use different drawing tools to draw on the software when drawing, and if it is necessary to change the position of a pole or change the information of a power line, it needs to be redrawn. Therefore, this kind of design method's shortcoming is obvious, not only the design workload is big, time-consuming, laborious, the efficiency is low, moreover has not been able to adapt to the electronic age development demand[1-2].

In order to realize the convenient drawing function, this paper designs and implements an automatic drawing software for power lines. The system can calculate various parameter information in the line through GPS data (longitude and latitude data) and store it in the database. Then it will automatically generate a standard and beautiful power circuit diagram according to the information stored in the database. If some information needs to be changed after the completion of power grid construction and reconstruction, only the corresponding data need to be changed to automatically draw again. This software provides a strong guarantee for reducing labor intensity, improving work efficiency and plotting quality, and shortening plotting cycle.

II. POWER LINE STRUCTURE ANALYSIS

Power equipment mainly includes generator set, power distribution device, lighthouse bridge column, fuse, transformer, transformer table, automatic control device, watt-hour meter, high-voltage switch, high-voltage circuit breaker, capacitor, lightning arrester, current transformer, voltage transformer, cable line, power line, etc. Since the generator set is not fixed on the power line, the power line mentioned in this system mainly refers to the line composed of outdoor poles, cables and transformer stations (including box transformers).

Due to the wide coverage area of outdoor lines, the complex and changeable environment, and the diversified routing modes of the lines, the factors to be considered in drawing are also relatively complex. However, no matter how complicated and diverse the actual power lines are, all power lines are connected to the user end through transformer outgoing lines and several poles, so the power lines are essentially standard multi-branch tree structures. As shown in fig. 1, the power line in each area is essentially a

DOI: 10.21307/ijanmc-2020-019

multi-branch tree with transformer as root node (node 0), and each pole in the line is a node of the multi-branch tree, i.e. T = {0,1,2,3,4,5,6,7} is a tree with root 0. The other nodes except the root node 0 can be divided into n disjoint finite sets, each set is a subtree of the root node[3].

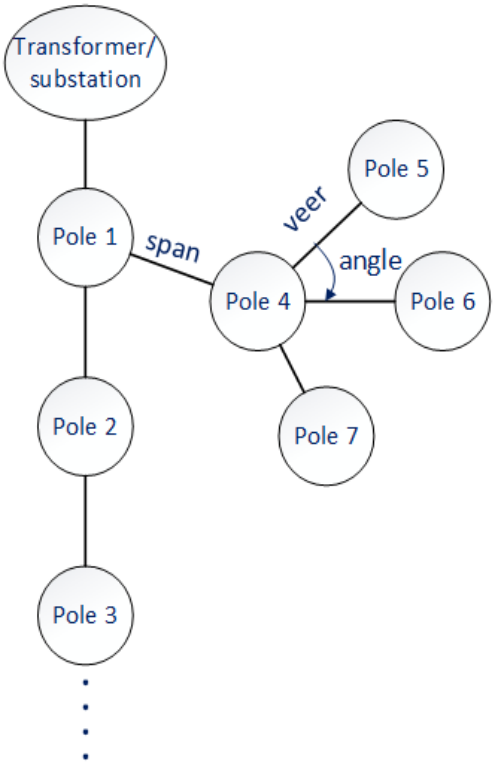


Figure 1. Schematic diagram of power line multi-branch tree structure

In the actual drawing process, whether manual drawing or automatic drawing, transformer is generally selected as the starting point. The process is: starting from the starting point to find the next pole connected to it through the relationship between the nodes, get the parameters of the node and the parent node, such as the distance, rotation angle, steering and other parameters, and then calculate the relative coordinates of the node. This drawing process is also very similar to the traversal process of a multi-fork tree, which provides the possibility for the realization of automatic power drawing. In the actual circuit drawing process, the four basic data sources of gear length, angle, steering, and pole number can be obtained during power grid reconstruction or power grid construction, or GPS can be used to obtain the latitude and longitude information of the pole and then converted to span, veer and angle. The automatic drawing of power lines can be realized by processing the four parameters in the database while traversing each node, calculating the relative coordinates of each node and converting them into absolute coordinates.

III. Power Line Automatic Drawing Function Design

A. Traversal of Multi-fork Trees

Since the relationship between transformer and pole is tree structure, when drawing circuit diagram, this multi-branch tree can be constructed by traversing power lines and all nodes in the tree can be processed one by one. Therefore, it has the possibility to realize the automatic drawing of power line graphics.

Common multi-tree traversal can be divided into two types: level-first traversal and depth-first traversal. Depth first traversal can be divided into: first order traversal, middle order traversal, and second order traversal. Its principle is to traverse the nodes of the tree along its depth and search the branches of the tree as deep as possible. Hierarchy priority traversal is to traverse the nodes of the tree according to hierarchy. Its principle is to first visit the root node and find all its children. These sub-nodes are also the root nodes of the sub-tree, so these nodes are accessed again, and so on and so forth, in order of first, second, third, etc[4].

B. Power Line Traversal Algorithm

Although the power line graph has a multi-branch tree structure, there are many problems in the actual traversal operation. In the past multi-tree applications, whether it is depth traversal or hierarchical traversal, the node stores the node's sub-node information when it is stored, so its sub-nodes can be directly found by accessing the node, and traversal can be completed from top to bottom. However, in power lines, because each pole node only knows its parent node but not its child node and has no other relationship information, it is difficult to determine the location of other nodes except the root node, so the traditional traversal algorithm is not applicable.

In order to solve this problem, this research has made some improvements on the basis of the hierarchical traversal algorithm. Because the information of each node's parent node is clear, we can easily find all the nodes with the same parent node. This research is based on this point to achieve top-down traversal. First, create an instance class object “Tree”, which represents that each row of data in the database is an instance object in java, and create an attribute “children” in the instance object Tree to save all child nodes of the current node. Second, get the data in the database to the instance object, and save all the

instance objects in the list collection. The specific traversal algorithm is to find the root node and put it into the result, and use the children attribute to find all the child nodes of the current layer from top to bottom through recursion, and save them in the result. Recursive traversal code is as follows.

```
//Store all node data information object collection
List<Tree> result = new ArrayList<Tree>();
//Query node data information
List<Tree> lists = new TestJDBC().queryAll();
//Traverse to determine whether it is the root node
    for (Tree tree : lists) {
        if(null == tree.getPid()) {
            result.add(tree);
        }
}
// Recursive call judgment function
for (Tree parent : result) {
        parent = recursiveTree(parent, lists);
    }
//Judgment function adds an object to the child attribute of the parent node object
public static Tree recursiveTree(Tree parent, List<Tree> list) {
        for (Tree tree : list) {
            if(parent.getId().equals(tree.getPid()))
{
                tree = recursiveTree(tree, list);
                parent.getChildren().add(tree);
            }
        }
        return parent;
}
```

C. Calculation of Plane Coordinates

The purpose of traversal is not only to construct the multi tree, but also to locate the coordinates of the traversal nodes. It is very important to convert the latitude and longitude data of the pole to the coordinate points in the plane rectangular coordinate system because it is impossible to draw the figure directly by GPS. In order to draw the figure accurately in the plane coordinate, first we must convert the longitude and latitude data into the plane coordinate data, and calculate the angle and the distance between the two points through the longitude and latitude data[5]. With poles A and B, the following calculation methods can be obtained through the derivation of mathematical geometric formula to realize data conversion.

(1) The calculation method of the angle formed by two points A and B is as follows:

$$\theta = \arctan(\frac{JA - JB}{WA - WB}) * \frac{180}{\pi} \quad (1)$$

Among them, θ is the angle (degree) between A and B, JA is the longitude of point a, JB is the longitude of point B, WA is the latitude of point A, WB is the latitude of point B.

(2) The calculation method of the distance between two points A and B is as follows:

$$L = 2R * \arcsin(\sqrt{\sin^2(\frac{WA - WB}{2}) + COS(WA) * \cos(WB) * \sin^2(\frac{JA - JB}{2})}) \quad (2)$$

Among them, L is the distance between points A and B (m), R is the earth radius (6371km), JA is the longitude of point A, JB is the longitude of point B, WA is the latitude of point A, WB is the latitude of point B.

According to the calculation results of formulas (1) and (2), the storage in the database is shown in Figure 2.

During traversal, the coordinates are located from the starting point, and then each child node of the node is found by traversal algorithm. The relative coordinates of the nodes are calculated by the span, rotation angle and steering parameters between the child node and the parent node. Due to the need to modify the relevant information of the pole during the traverse process, in order to ensure the integrity and reliability of the original database and improve the speed of database traversal, the information of relevant substation should be saved in the temporary database before the actual traversal[6-7]. The traversal process is as follows.

(1) Using SQL "select into temporary table from original table where change table name" command to create temporary data table and copy the change table data information to be drawn to the temporary table.

(2) Find the starting point "transformer" of the transformer in the temporary data table, and set the starting point coordinate of the drawing as $[x_0, y_0]$ according to the actual position of the transformer in the transformer.

num	The name of the line	Nature of the line	ID	Route	Height	Veer	Angle	Span	PID	Longitude	Latitude
1	Panjin north power distribution station	Station Area	1	Shenshan line	12	0	0	0	0	121.80306	41.33611
2	Panjin north power distribution station	Station Area	2	Shenshan line	12	East	10	10	1	122.1407	41.6107
3	Panjin north power distribution station	Station Area	3	Shenshan line	12	East by south	34	41	1	121.12083	41.01556
4	Panjin north power distribution station	Station Area	4	Shenshan line	12	East by north	18	21	1	121.13592	41.12497
5	Panjin north power distribution station	Station Area	5	Shenshan line	12	East by south	45	10	3	120.35314	40.32225
6	Panjin north power distribution station	Station Area	6	Shenshan line	12	East by north	66	13	3	122.76694	41.55722
7	Panjin north power distribution station	Station Area	7	Shenshan line	12	South by west	10	25	3	122.4255	41.4436
8	Panjin north power distribution station	Station Area	8	Shenshan line	10	North by east	33	18	2	120.315	40.333
9	Panjin north power distribution station	Station Area	8-1	Shenshan line	12	South	0	20	2	120.7575	40.74083
10	Panjin north power distribution station	Station Area	9	Shenshan line	12	East by south	30	19	2	119.76688	40.00035
11	Panjin north power distribution station	Station Area	10	Shenshan line	12	North by east	78	20	3	119.76492	39.99816
12	Panjin north power distribution station	Station Area	11	Shenshan line	12	East by north	21	22	6	121.14092	41.12663

Figure 2. Storage in a database

(3) Find the distribution board connected with the transformer, calculate the coordinates $[x_0, y_0]$ of the distribution board according to the distance between the distribution board and the transformer, the steering and the steering angle, mark in the temporary database that the relative position coordinates of the node and the parent node have been calculated, and mark in the parent node that the child node coordinates of the parent node have been calculated.

(4) Execute the cycle in the temporary database to select the relative coordinates of this node and the parent node have been calculated, but the poles of the coordinates of the child nodes have not been calculated.

(5) Repeat step (3) until all poles are traversed.

In the process of traversal, every record in the database must participate in multiple operations. If there are many poles in a substation, the cycle of traversal operation may be too long. Therefore, in the actual traversal operation, it may be faster to use the stored procedure of the database to complete this work.

IV. GENERATION OF POWER CIRCUIT DIAGRAM

After traversal calculation and plane coordinate positioning, the actual drawing results also need to display other relevant information for users. Therefore, the following work should be completed before automatically generating circuit drawing.

(1) Traverse from the starting point (transformer) to calculate the power supply radius of each pole.

(2) Draw title block, mark drawing scale, drawing date, type and quantity of various poles, model and length of conductor, number of meter boxes, maximum power supply radius, etc.

(3) Statistics of the types, specifications, quantities of various poles in the substation, the length of wires of various specifications and types, the number of various meter boxes (power meter boxes and lighting meter boxes), etc.

(4) According to the requirements of the power supply enterprise, the pole will be represented as "○", and the container will be represented as "□".

(5) Walk through the database again and draw the line diagram from the starting point.

A multi-fork tree structure with transformer as the root node line as the path and pole as the node is established by analyzing the demand of power supply enterprises for the drawing and archiving of power circuit diagrams and combining with the characteristics of power lines. After identifying the information stored in the database, a hierarchical traversal algorithm based on multi-tree is proposed. In the process of drawing, only four parameters of node need to be processed : span, veer, angle and Id number. and then the relative coordinates of each node are calculated and converted into absolute coordinates by using a multi-tree layer traversal algorithm, so that the automatic drawing of power lines can be realized. Then, the relative coordinates of each node are calculated and converted into absolute coordinates by using a multi-tree layer traversal algorithm, so that the automatic drawing of

power lines can be realized. The drawing method discussed in this paper has been successfully applied to many power supply enterprises in our country. Figure 3 shows the circuit diagram of a certain substation automatically generated.

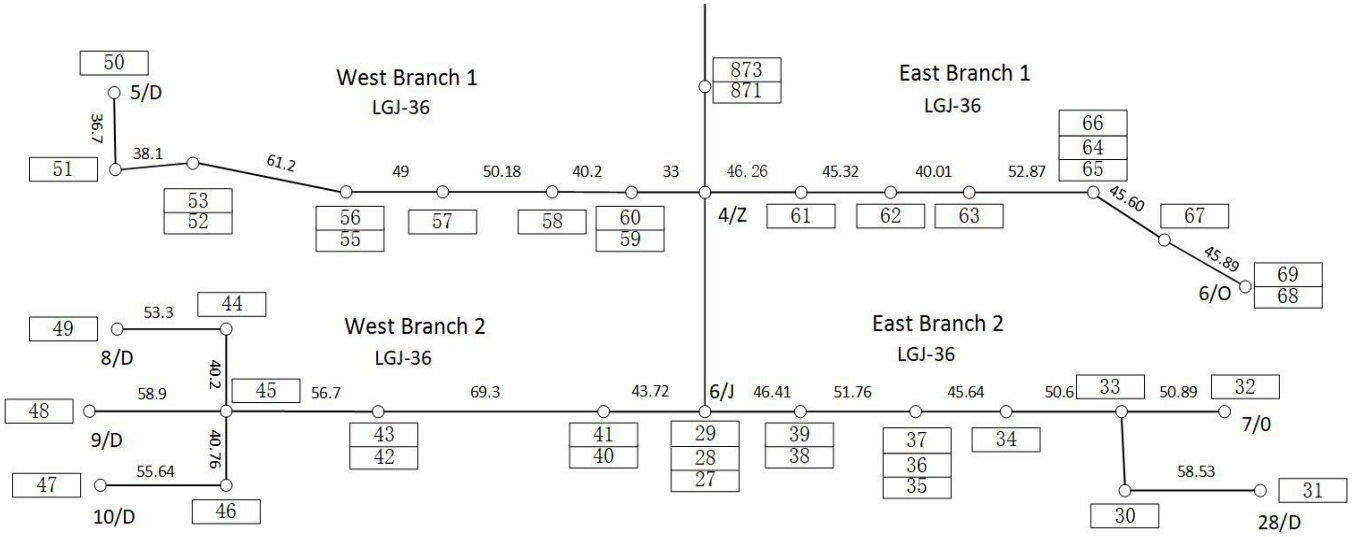


Figure 3. Circuit diagram of a transformer station automatically generated

V. CONCLUSION

The system manages power line parameter data through a database, and realizes the function of automatically drawing power graphs according to the proportion set by users, thus saving human resources. It can quickly change the information in the database when the information about a power line is changed. Therefore, it can well meet the needs of China's power grid reconstruction with short drawing cycle and high working efficiency. This system has achieved good results in the practical application of power supply enterprises in many places.

REFERENCES

[1] Peng Liangyong, Li Xiaodan. Realization of fast coordinate transformation method for CAD graphics [J]. Surveying and mapping and spatial geographic information, 2020,43(04):199-201.

[2] Li Huiping, He Lianfang. Discussion on the Teaching of Auxiliary Modules of AutoCAD Software [J]. Journal of Higher Education, 2020(14):90-92+96.

[3] Peng Qiwei, Zhang Hao. Adaptive Relay Algorithm for Medium Voltage Distribution Line Carrier Based on Multitree Traversal [J]. Power System Communication, 2009,30(03):81-84.

[4] Wang Fangxiu, Liu Chunhong. A New Algorithm for Constructing Binary Tree by Hierarchical Traversal and Other Traversal [J]. Journal of Wuhan University of Light Industry, 2016,35(04):67-72.

[5] Zhao Xiang, Chen Yan. A Coordinate Transformation Algorithm in Geographic Information System [J]. Computer and Network, 2007(11):40-41.

[6] Data Structure and Algorithm [M]. Beijing University of Aeronautics and Astronautics Press, Yu Xiaomin, 2010.

[7] Kou Chunpeng. Design and Implementation of SQLServer Database Information Acquisition System [D]. beijing university of chemical technology, 2008.

Improved Faster R-CNN Algorithm for Sea Object Detection Under Complex Sea Conditions

Liu Yabin
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shaanxi, China
E-mail: 1135616154@qq.com

Hu Zhiyi
Engineering Design Institute
Army Research Loboratory
Beijing, 100000, China
E-mail: 763757335@qq.com

Yu Jun
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shaanxi, China
E-mail: yujun@xatu.edu.cn

***Abstract*—In In the process of sea surface object detection by remote sensing images, this paper proposes an improved Faster R-CNN sea surface object detection algorithm for the problems of low accuracy and long detection time using deep neural networks. Firstly, the size of the bounding box in the region proposal networks is analyzed, and then the object in the image is clustered using the K-Means algorithm, and the clustering results are input into region proposal networks, thereby realizing the region proposal networks improvement of. Secondly, the Soft-NMS algorithm is used to screen the target candidate frame to obtain the detected sea objects. The experimental results show that the algorithm of this paper can detect the sea surface object of remote sensing images under complex sea conditions, and its mAP can reach 87.25%, which is an average increase of 3.75% compared with the commonly used detection methods.**

Keywords-Sea Surface Object Detection; Faster R-CNN Algorithm; K-Means Clustering Algorithm; Soft-NMS

I. INTRODUCTION

As an important field of computer vision research, marine object recognition is aimed at achieving automatic recognition of marine ships. Realizing the automatic identification of ships can help the construction of unmanned ports, ship search and rescue, monitoring and combating illegal activities such as illegal fishing, smuggling and piracy. At the same time, it also has certain application value in the military[1-2], such as supervising our country's territorial waters, monitoring illegal intrusion ships, and assisting in analyzing the deployment of the enemy's key ports and military warships. Therefore, it is of great practical significance to improve the detection accuracy of sea surface objects.

In recent years, with the sharp increase in the resolution of satellite remote sensing images in China, many trainable remote sensing image samples have been provided for deep learning. How to effectively use massive high-resolution remote sensing images is one of the main problems faced by current object detection[3-6]. Reference[7] studied the object recognition of remote sensing images based on Faster R-CNN deep network, and proved that the deep learning method can realize the rapid and accurate recognition of remote sensing image objects. Reference[8] discusses that when satellites take remote sensing images of the sea, they are greatly affected by weather conditions. Remote sensing images may have problems such as relatively small ship size, cloud cover, and background interference on land. Reference[9] proposed a ship detection algorithm based on the Mask R-CNN framework, but this algorithm will have problems such as missed detection and false detection for closely arranged multi-scale ships. Reference[10] cites the hollow convolution and global attention modules to extract more feature information, and finally constrains the salient features obscured by clouds. However, when there are small objects, cloud cover and land background interference, the detection accuracy of the above method is not high. In this paper, the Faster R-CNN[11-13] object detection algorithm is improved. The K-Means clustering algorithm[14] is

introduced to perform cluster analysis on the size of the object in the image, and the clustering results are directly input into the area recommendation network to achieve the improvement of the area recommendation network. Using Soft -The NMS algorithm replaces the NMS algorithm to reduce the miss detection probability of small objects. Finally, the average accuracy of object detection is improved.

II. IMPROVEMENT OF FASTER R-CNN

The object detection algorithm studied in this paper is mainly the Faster R-CNN object detection network model, and it has been improved.

A. Principle of Faster R-CNN

Faster R-CNN is a object detection network model proposed by Shaoqing Ren. It introduces RPN (Region Proposal Networks) to extract features from object candidate regions, and uses RPN to replace the original selective search algorithm. The detection process of Faster R-CNN model can be shown in Figure 1. First, input images of any size and corresponding annotation files into the network model. Second, they will go through the convolution layer to extract the features of the input image. Third, use RPN for region prediction, and use the ROI mapping operation to map the predicted candidate frame to the feature map. Finally, identify the target category and locate the bounding box. It can be seen that in the Faster R-CNN detection process, the recommendation of target candidate regions, feature extraction, and object detection process are all integrated into the network model.

In the Faster R-CNN network model, the main function of RPN is to generate target candidate boxes. The input of the RPN is a n×n size dimensional feature map, and the output is a series of target candidate boxes. Before entering the area generation network, a small network will choose sliding windows on the final convolutional feature map. For each window obtained, k target candidate regions are predicted, which are called anchors. Each anchor corresponds to a different scale, size, and center point on the convolutional feature map.

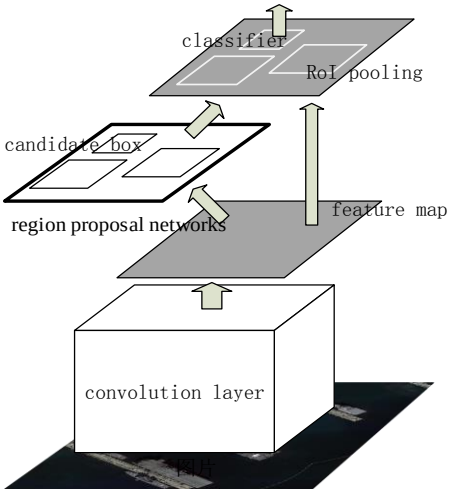


Figure 1. Fast R-CNN model framework

The Region Proposal Networks is shown in Figure 2. A w×h size convolutional feature map, corresponding to w×h anchors of different sizes, each window is mapped into a low-dimensional vector, and then this feature vector is input into two subnetworks: border classification network, border Back to the network. The function of the border regression network is to modify the size of the anchor to obtain an accurate candidate area. Its output is the translation zoom value of each anchor, and each window will have 4k outputs. The function of the boundary classification network is to determine whether the anchor belongs to the background or the object, and its output is the probability that each anchor belongs to the background or object, and each window corresponds to 2k values.

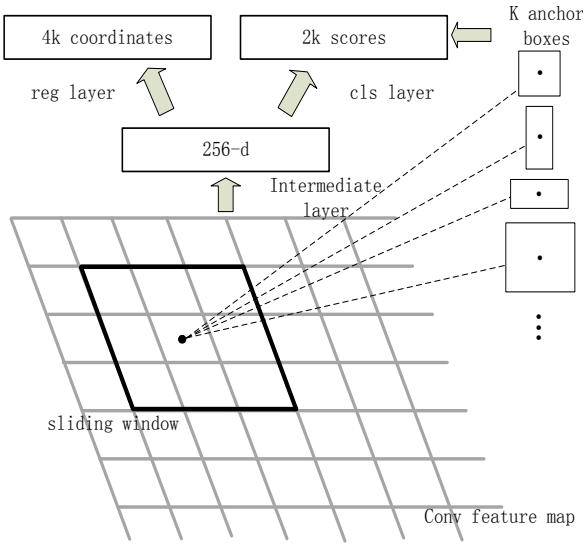


Figure 2. Region Proposal Networks

The Region Proposal Networks training process is end-to-end. The optimization methods used are stochastic gradient descent method and back propagation method. The loss function of the network model is a combined function of classification error and regression error. The formula is shown in formula (1).

$$L(\{P_i\}+\{t_i\})=\frac{1}{N_{cls}}\cdot\sum_i L_{cls}(p_i,p_i^*)+\frac{\lambda}{N_{reg}}\cdot\sum_i p_i^*\cdot L_{reg}(t_i,t_i^*) \quad (1)$$

Among them, i in the formula represents the i-th anchor point; pi represents the probability that the candidate box is the object, if pi=1, it means that the i-th anchor point is a positive sample; t* represents the deviation between the candidate box and the real box. During training, if the IoU value of the candidate area and any real frame is greater than 0.7, it is expressed as a positive sample; if the IoU value of the candidate area and any real frame is less than 0.3, it is expressed as a negative sample; if it does not belong to the above two cases, It means that the candidate area is not used in the training.

B. Faster R-CNN improved method

Although the Faster R-CNN network model integrates the RPN network and the Fast R-CNN network into a network model, a object detection algorithm framework based on deep convolutional networks that uses end-to-end training in a true sense is constructed. But Faster R-CNN also has shortcomings, mainly including:

a) The accuracy of Faster R-CNN on remote sensing data set is low. Especially when there is cloud and fog blocking and shore-based interference, the problem of missed detection of objects is prone to occur.

b) The threshold of Non-Maximum Suppression (NMS) is difficult to determine. If the threshold is set too small, it will cause the problem of missed detection of the network model; if the threshold is set too large, it will cause the model to misdetect.

Based on the above problems, this paper proposes an improved detection model of Faster R-CNN algorithm, which mainly includes:

a) Use the K-means algorithm to readjust the size and number of RATIOS.

b) Replace the non-maximum suppression (NMS) algorithm with the Soft-NMS algorithm.

Each will be described below. The following will be described in detail.

1) K-means algorithm readjusts the size and number of RATIOS

Because the existing public data sets (VOC, COCO) do not contain data related to remote sensing sea surface object images, if the original parameters of Faster R-CNN are used for training, it will have a certain impact on the training speed and detection effect. Therefore, it is necessary to re-cluster the labels of the remote sensing data sets to obtain the size and number of RATIOS that are favorable for detecting the remote sensing data sets.

TABLE I. ANCHOR_RATIOS FOR DIFFERENT K VALUES

number	K=1	K=2	K=3	K=4	K=5	K=6	K=7
	0.6	1.3	0.5	0.7	0.6	0.5	0.5
	-	1.4	1.2	1.2	1.2	0.7	0.6
	-	-	1.6	1.3	1.3	1.2	1.1
results	-	-		2.9	1.7	1.3	1.2
	-	-	-	-	2.6	2.4	1.8
	-	-	-	-	-	3	2.1
	-	-	-	-	-	-	2.9
mAP(%)	83.84	84.3	84.32	84.98	84.29	84.57	84.34
Time(s)	503	510	515	518	524	515	517

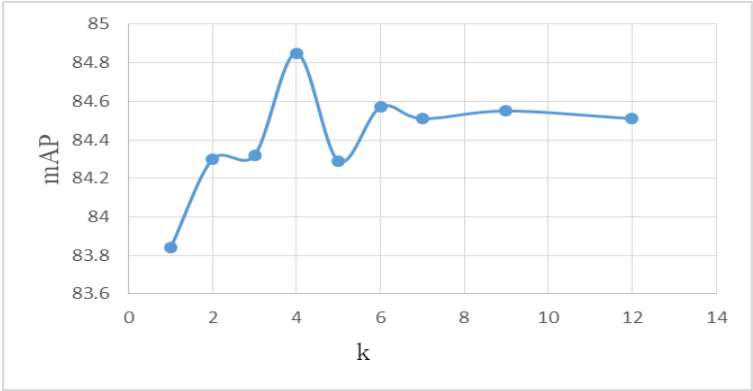


Figure 3. mAP of different K values

The K-means algorithm is used to perform dimensional clustering analysis on the labels of the remote sensing data set. With different values of K, the values of RATIOS are shown in Table 1. In the case where other parameters are unchanged, the resulting average accuracy change is shown in Figure 3.

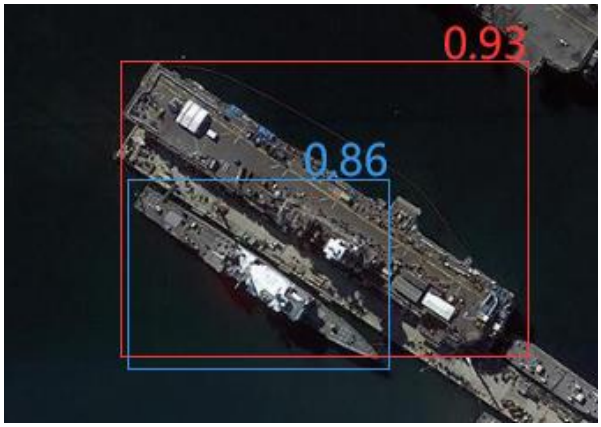
It can be seen from Figure 3 that when K=4, the average accuracy rate is the highest; as the K value increases, the average accuracy rate tends to be stable; combined with Table 1, it can be seen that with the increase of the K value, the detection time It is also increasing; therefore, the clustering result of K=4 is taken as the improved parameter of RATIOS, that is, RATIOS =[0.7, 1.2, 1.3, 2.9].

2) Replace NMS algorithm with Soft-NMS algorithm

Non-maximum suppression (NMS) is an important part of object detection process. First, it sorted based on the object's confidence, and selected the detection box M with the highest score. Then, if the overlap area between other detection boxes and M is greater than a certain proportion, the other detection boxes will be deleted. This process is recursively applied to the remaining check boxes. The formula of NMS algorithm is shown in Equation (2).

$$s_i = \begin{cases} s_i, iou(M, b_i) < N_t \\ 0, \mathrm{iou}(M, b_i) \geq N_t \end{cases} \quad (2)$$

Where, Si represents the confidence score of all detection boxes; Nt is an artificially set threshold; M is the detection box with the highest confidence score in each iteration; Bi is the detection box to be processed; Iou (M, BI) is the overlapping area between M and BI. According to formula (2): when the overlapping area Iou (M, BI) is greater than the artificially set threshold Nt, BI will be deleted. Otherwise, retain bi.



(a) Schematic diagram of the omission situation

(b) Schematic diagram of false detection

Figure 4. Is a schematic diagram of the shortcomings of The NMS algorithm

The disadvantage of the NMS algorithm is that the artificially set threshold Nt is difficult to determine. If the threshold is set too high, many objects with low confidence scores will be deleted directly; if the

threshold is set too low, many false objects will be detected, resulting in false detections. In addition, when bi is very close to the selected detection frame M and contains a object, the formula (2) still cannot achieve the effect of accurate detection. As shown in Figure 4(a), the detection results of the current image are marked with red boxes and blue boxes, respectively, and the scores of the two are 0.93 and 0.86, respectively. If the traditional non-maximum suppression method is adopted, the red box with the highest score will be preferentially selected, and the green box that is too large in the overlapping area will be deleted. This caused a missed inspection. Similarly, if the threshold is set too low, it will cause false detection as shown in Figure 4(b).

In order to improve the accuracy of object detection, we replaced the NMS algorithm with Soft-NMS algorithm. The Soft-NMS algorithm reduces the score of the adjacent detection frame through a function related to the overlap of the detection frame M, but it will not be completely deleted. Although the score has dropped, the adjacent detection frames are still in the sequence of the target detector. Therefore, the Soft-NMS algorithm effectively solves the problem of missed detection of objects. The core formula of its algorithm is the Gaussian penalty function, as shown in equation (3).

$$s_i = s_i \cdot e^{\frac{IoU^2(M,b_i)}{\sigma}} \tag{3}$$

Among them, si represents the confidence score of all detection frames; M represents the detection frame with the highest confidence score; bi is the unprocessed detection frame. If the IoU values of bi and M are larger, the value of confidence si decreases more. The formula will be applied to each subsequent iteration and update the confidence of all remaining detection frames.

III. Prepare Experimental and Analysis

The improved Faster R-CNN was used to conduct remote sensing image sea surface object detection experiment, and its detection process was shown in Figure 5. First, enter an image of any size and enter the ResNet101 convolutional network layer after certain adjustments. Secondly, the ResNet101 network layer outputs the feature map corresponding to the image. Again, RPN extracts candidate box feature blocks from the convolutional feature map. Then, the ROI pooling layer extracts the candidate block feature block and the feature map output by the convolution layer. Finally, the above results are sent to the fully connected layer, and the final object detection result is output after passing through the Soft-NMS.

C. Preparation of data sets

Experimental data were captured from Kaggle competition platform and Google Earth. Among them, the training set has a total of 192,556 images, including 150,000 images (77.9%) with no ship empty, and 42,556 images (22.1%) with ship images. Since there are many negative samples in the data set, it is necessary to select an effective graph from the training set.

In the data set, there are mostly cargo ships and passenger ships, and the number of container ships, warships, and aircraft carriers is relatively small, and most of the remote sensing data sets are overhead shots, and the orientation of the ships changes greatly. In order to make the trained model extensive, it is necessary to amplify the data set. There are four main amplification methods: flip horizontally or vertically, add random noise, rotate the image at different angles, and blur the image by the specified amount. The amplified data set is divided into three categories: training set, validation set and test set, and the division ratio is 65%, 15% and 20% respectively. After that, annotate the data set.

D. Experimental Environment

The experiment in this article is conducted under Linux system. In order to reduce the training time of the deep neural network model and improve the calculation speed, the Nvidia TITAN XP 12G graphics card is used, and CUDA9.0 and cuDNN7.0 are configured to call the GPU for acceleration. Use Tensorflow deep learning framework.

E. Training model

The main parameters of the detection network are shown in Table 2. Set up to save a model every 5000 iteration batches, and finally, use the obtained model to carry out the detection experiment.

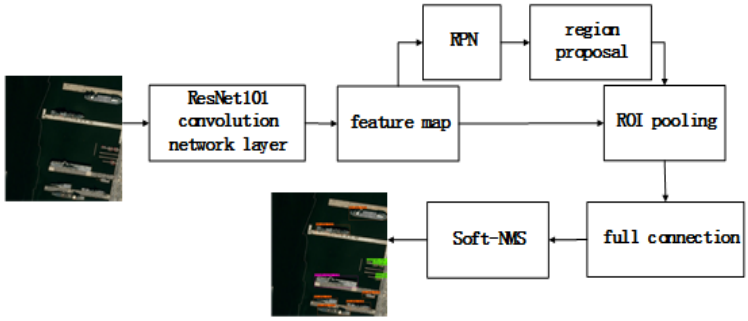


Figure 5. Improved object detection network

TABLE II. VALUES OF MAIN NETWORK PARAMETERS

Parameters	Values	Parameters	Values
LEARNING_RATE	1e-3	Batch_size	256
Anchor_Scales	[8,16,32]	Anchor_RATIOS	[0.7, 1.2, 1.3, 2.9]
ITERS	85000	num_classes	6
SOFT_NMS	1		

In Table 2, if the value of LEARNING_RATE is set to a small value, the rate of model convergence will become slower, and it is easy to fall into the local optimal value; if the setting is too large, the difficulty of model convergence will increase, the network parameters will change in a large range, and oscillate serious. After many experiments, set LEARNING_RATE to 1e-3.

In general, the larger the Batch_size, the better the model. However, in the actual training process, if the Batch_size is too large, the memory of the computer is insufficient, and if the Batch_size is too small, the trained network is difficult to converge. In this experiment, the Batch_size is set to 256. During the model training process, after each epoch training is completed, the network will output the total loss function value, classification loss function and loss function of the positioning frame of each epoch. If the total loss the function value and other loss function values are continuously decreasing, indicating that the network is converging and the Batch_size value is appropriate.

F. Experimental results and analysis

The experimental results include three parts: the detection results of the VGG-16 structure, the detection results of the ResNet101 structure, and the detection results of the algorithm in this paper. The experimental results are shown in Table 3.Table 3 Comparison of detection results of three network structures.

It can be seen from Table 3 that the detection results of the algorithm in this paper is best. Compared with the VGG-16 and ResNet101 structure methods, mAP has increased by 4.29% and 3.21%, respectively. The average accuracy of Passenger_ship has increased by 16% and 12.8%, respectively.

TABLE III. COMPARISON OF DETECTION RESULTS OF THREE NETWORK STRUCTURES

Detection method	Passenger_ ship	Cargo_ship	Container_ ship	Aircraft_ ship	War_ship	mAP
VGG-16 structure	50.9%	80.6%	92.4%	98.1%	92.8%	82.96%
ResNet101 structure	54.1%	81.0%	92.8%	99.3%	93.0%	84.04%
improved ResNet101 structure	66.9%	82.3%	93.65%	99.51%	93.9%	87.25%

It can be seen from Figure 6 that under the background of complex sea surface (such as: shoreline interference, cloud interference, the presence of small ships), the ships in the remote sensing image can be detected by the algorithm of this paper and marked in each picture. The box shows the predicted confidence score.

(a) result with shoreline interference (b) results With small ships (c) results with cloud interference

Figure 6. Test results

IV. CONCLUSION

In this paper, an improved Faster R-CNN algorithm for sea surface object detection is proposed, and the training and testing of the object detection network are carried out on the remote sensing image sea surface ship data set. The experimental results show that the improved Faster R-CNN algorithm improves the average accuracy of sea object, and solves the problems of low average accuracy and easy missing of small objects on the detection process of sea objects in complex sea environment.

REFERENCES

[1] Xu Fang. Research on key Technologies for Automatic Detection of Surface Objects in Visible light remote sensing images [D]. University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences), 2018.

[2] Zhang Zemiao, Huo Huan, Zhao Fengyu. Review of Object detection algorithms for deep convolutional neural networks [J]. Minicomputers, 2019, 40(09): 1825-1831.

[3] Yao Qunli, HU Xian, LEI Hong. Research Progress of Deep convolutional neural network in Object Detection [J]. Computer Engineering and Applications, 2016,54(17):1-9.

[4] Xu Junfeng, Zhang Baoming, Guo Haitao. A multi-feature fusion object-oriented multi-source remote sensing image change detection method [J]. Journal of surveying and mapping science and technology, 2015, 32(05):505-509.

[5] Sun Hao, Sun Xian, Wang Hongqi. Research on a high-resolution remote sensing image ship detection method [J]. Surveying and mapping science, 2013, 38(5):112-116.

[6] Wang Tengfei. Research on high-resolution Remote Sensing Image Deep Learning Ship Detection Technology [D]. Harbin Institute of Technology, 2017.

[7] Wang Jinchuan, TAN Xicheng, WANG Zhaohai, ZHONG Yanfei, Dong Huaping, Zhou Songtao, Cheng Boyi. Research on remote sensing image Object recognition Method based on Faster R-CNN Deep Network [J]. Journal of Earth Information Science, 2016, 20(10):1500-1508.

[8] Yao Hongge, Wang Cheng, Yu Jun, Bai Xiaojun, Li Wei. Small Object ship identification in complex satellite images [J]. Journal of remote sensing, 2020, 24(02): 116-125.

[9] Wu Jinliang, WANG Gang, Liang Shuo, Chen Jinyong, Gao Feng. Research on Ship Object Detection based on Mask R-CNN [J]. Radio Engineering, 2008, 48(11):947-952.

[10] He Yubo, LIU Kun. Surface Significance Object Detection based on convolutional neural network [J/OL]. Computer Engineering and Application:1-10[2020-06-09].

[11] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. Computer Science, 2015(99):1.

[12] Zhang Y, Sohn K, Villegas R,et al. Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction[C]. Computer Vision and Pattern Recognition. IEEE,2015:249-258

[13] Girshick R.Fast R-CNN[C].the IEEE International Conference on Computer Vision, December 7-13, 2015, Chile. New Jersey, IEEE Press, 2015:1440-1448.

[14] Engineering; Study Data from Northeast Electric Power University Provide New Insights into Engineering (Fingerprint location algorithm based on K-means for spatial farthest access point in Wi-Fi environment)[J]. Journal of Engineering, 2020.