



An International ICT R&D Journal Sponsored by ZTE Corporation

ISSN 1673-5188

CN 34-1294/TN

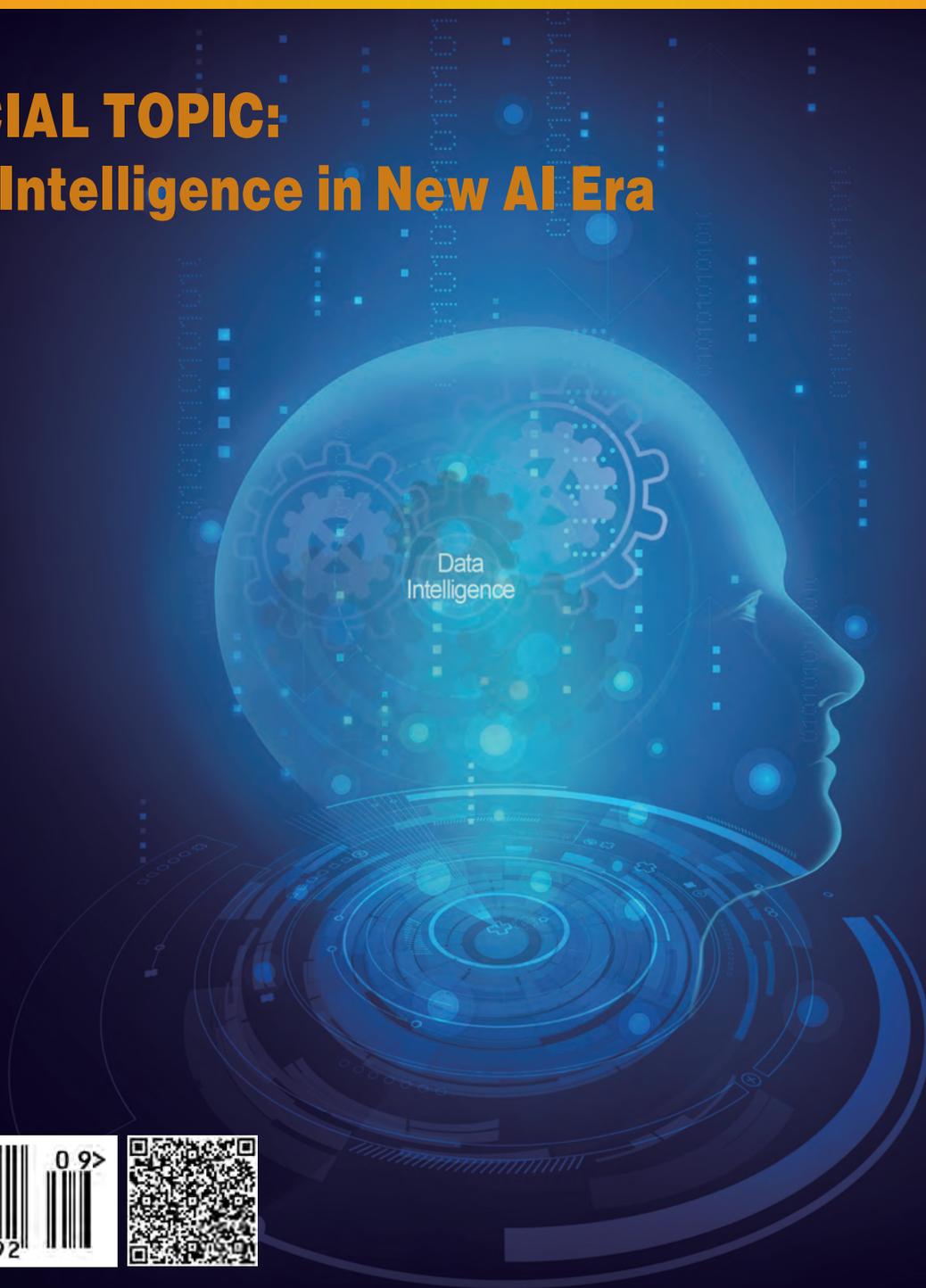
ZTE COMMUNICATIONS

中兴通讯技术(英文版)

<http://tech.zte.com.cn>

September 2019, Vol. 17 No. 3

SPECIAL TOPIC: Data Intelligence in New AI Era



ZTE COMMUNICATIONS
中兴通讯技术(英文版)

VOLUME 17 NUMBER 3 SEPTEMBER 2019

The 8th Editorial Board of ZTE Communications

Chairman

GAO Wen: Peking University (China)

Vice Chairmen

XU Ziyang: ZTE Corporation (China) | **XU Chengzhong:** University of Macau (China)

Members (Surname in Alphabetical Order):

| | |
|--------------------|--|
| AI Bo | Beijing Jiaotong University (China) |
| CAO Jiannong | Hong Kong Polytechnic University (China) |
| CHEN Chang Wen | The State University of New York at Buffalo (USA) |
| CHEN Yan | Northwestern University (USA) |
| CHI Nan | Fudan University (China) |
| CUI Shuguang | UC Davis (USA) and The Chinese University of Hong Kong, Shenzhen (China) |
| GAO Wen | Peking University (China) |
| GAO Yang | Nanjing University (China) |
| GE Xiaohu | Huazhong University of Science and Technology (China) |
| HWANG Jenq-Neng | University of Washington (USA) |
| Victor C. M. Leung | The University of British Columbia (Canada) |
| LI Guifang | University of Central Florida (USA) |
| LI Xiangyang | University of Science and Technology of China (China) |
| LIN Xiaodong | ZTE Corporation (China) |
| LIU Chi | Beijing Institute of Technology (China) |
| LIU Jian | ZTE Corporation (China) |
| LIU Ming | Institute of Microelectronics of the Chinese Academy of Sciences (China) |
| MA Jianhua | Hosei University (Japan) |
| MA Zheng | Southwest Jiaotong University (China) |
| NIU Zhisheng | Tsinghua University (China) |
| PAN Yi | Georgia State University (USA) |
| REN Fuji | Tokushima University (Japan) |
| REN Kui | Zhejiang University (China) |
| SHENG Min | Xidian University (China) |
| SONG Wenzhan | University of Georgia (USA) |
| SUN Huifang | Mitsubishi Electric Research Laboratories (USA) |
| SUN Zhili | University of Surrey (UK) |
| TAO Meixia | Shanghai Jiao Tong University (China) |
| WANG Haiming | Southeast University (China) |
| WANG Xiang | ZTE Corporation (China) |
| WANG Xiaodong | Columbia University (USA) |
| WANG Xiyu | ZTE Corporation (China) |
| WANG Yongjin | Nanjing University of Posts and Telecommunications (China) |
| WANG Zhengdao | Iowa State University (USA) |
| XU Chengzhong | University of Macau (China) |
| XU Ziyang | ZTE Corporation (China) |
| YANG Kun | University of Essex (UK) |
| YUAN Jinhong | University of New South Wales (Australia) |
| ZENG Wenjun | Microsoft Research Asia (China) |
| ZHANG Chengqi | University of Technology Sydney (Australia) |
| ZHANG Honggang | Zhejiang University (China) |
| ZHANG Jianhua | Beijing University of Posts and Telecommunications (China) |
| ZHANG Yueping | Nanyang Technological University (Singapore) |
| ZHOU Wanlei | University of Technology Sydney (Australia) |
| ZHUANG Weihua | University of Waterloo (Canada) |

CONTENTS

ZTE COMMUNICATIONS September 2019 Vol. 17 No. 3 (Issue 67)

Special Topic

Data Intelligence in New AI Era

Editorial 01

XU Chengzhong and QIAO Yu

A Lightweight Sentiment Analysis Method 02

As the most influential media for Chinese domestic movie ratings, Douban contains a huge amount of data, and one can understand users' perspectives towards these movies by analyzing these data. In this article, the authors study movie's critics from the Douban website, perform sentiment analysis on the data obtained by crawling, and visualize the results with a word cloud. They propose a lightweight sentiment analysis method which is free from heavy training and visualize the results in a more conceivable way.

YU Qingshuang, ZHOU Jie, and GONG Wenjuan

Big Data-Driven Residents' Travel Mode Choice: A Research Overview 09

Thanks to upgrades of urban infrastructures, many real-time location-tracking devices become available. These devices generate massive real-time data, which provides new opportunities to analyze and explain residents travel mode choice behavior more accurately and more comprehensively. This paper surveys the current research status of big data-driven residents' travel mode choice from three aspects: residents' travel mode identification, acquisition of travel mode influencing factors, and travel mode choice model construction. Finally, the limitations of current research and directions of future research are discussed.

ZHAO Juanjuan, XU Chengzhong, and MENG Tianhui

15 Face Detection, Alignment, Quality Assessment and Attribute Analysis with Multi-Task Hybrid Convolutional Neural Networks

This paper proposes a universal framework, termed as Multi-Task Hybrid Convolutional Neural Network (MHCNN), for joint face detection, facial landmark detection, facial quality, and facial attribute analysis. MHCNN consists of a high-accuracy single stage detector (SSD) and an efficient tiny convolutional neural network (T-CNN) for joint face detection refinement, alignment and attribute analysis. Since there is no public facial quality data and facial attribute data as we need, the authors also contribute two datasets, namely FaceQ and FaceA.

GUO Da, ZHENG Qingfang, PENG Xiaojiang, and LIU Ming

23 RAN Centric Data Collection for New Radio

Self-organizing network (SON) and minimization of driver tests (MDT) are functions designed for long term evolution (LTE) system. However, these functions do not support new features in new radio (NR) access technology, e.g., multiple radio access technology (RAT)-dual connectivity (MR-DC), central unit-distribute unit (CU-DU) split architecture, beam, etc. Therefore, how to support these features is a challenge for Industry. This paper provides analysis for these problems and provides the summary of SON/MDT functions progress in 3GPP. In addition, this paper also provides an initial thought on artificial intelligence (AI) algorithms applied to SON/MDT functions in NR, so called Smart Grid.

GAO Yin, LI Dapeng, HAN Jiren, LIU Zhuang, and LIU Yang

Submission of a manuscript implies that the submitted work has not been published before (except as part of a thesis or lecture note or report or in the form of an abstract); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors as well as by the authorities at the institute where the work has been carried out; that, if and when the manuscript is accepted for publication, the authors hand over the transferable copyrights of the accepted manuscript to *ZTE Communications*; and that the manuscript or parts thereof will not be published elsewhere in any language without the consent of the copyright holder. Copyrights include, without spatial or timely limitation, the mechanical, electronic and visual reproduction and distribution; electronic storage and retrieval; and all other forms of electronic publication or any other types of publication including all subsidiary rights.

Responsibility for content rests on authors of signed articles and not on the editorial board of *ZTE Communications* or its sponsors.
All rights reserved.

CONTENTS

ZTE COMMUNICATIONS September 2019 Vol. 17 No. 3 (Issue 67)

Review

Reinforcement Learning from Algorithm Model to Industry Innovation: A Foundation Stone of Future Artificial Intelligence 31

This paper first introduces the history and background of reinforcement learning, and then illustrates the industrial application and open source platforms. After that, the successful applications from AlphaGo to AlphaZero and future reinforcement learning technique are focused on. Finally, the artificial intelligence for complex interaction (e.g., stochastic environment, multiple players, selfish behavior, and distributed optimization) is considered and this paper concludes with the highlight and outlook of future general artificial intelligence.

DONG Shaokang, CHEN Jiarui, LIU Yong, BAO Tianyi, and GAO Yang

Research Paper

A Low-Cost Outdoor Fingerprinting Localization Scheme For Wireless Cellular Networks 42

This paper considers outdoor fingerprinting localization in LTE cellular Networks, which can localize non-cooperative user equipment (UE) that is unwilling to provide Global Positioning System (GPS) information. The authors propose a low-cost fingerprinting localization scheme that can improve the localization accuracy while reducing the computational complexity. Experiment results show that the performance is improved by the proposed localization scheme, and positioning errors corresponding to Cumulative Distribution Function (CDF) equaling to 67% and 95% are declined to 50 m and 150 m.

PEI Dengke, XU Xiaodong, QIN Xiaowei, LIU Dongliang, and ZHAO Chunhua

50 High Speed Polarization-Division Multiplexing Transmissions Based on the Nonlinear Fourier Transform

Polarization-division multiplexing (PDM) with modulation in the nonlinear frequency domain can compensate the optical fiber nonlinearity based on the nonlinear Fourier transform (NFT). In this paper, the authors combine PDM with the method of nonlinear frequency division multiplexing (NFDM) and demonstrate the achievable transmission rate by increasing the number of multiplexing nonlinear channels.

WANG Jia, ZHAO Yilong, HUANG Xin, and HE Guangqiang

56 A Service-Based Intelligent Time-Domain and Spectral-Domain Flow Aggregation in IP-over-EON Based on SDON

In this paper, the authors propose a service-based intelligent aggregation node selection and area division (ANS-AD) algorithm. Based on the ANS-AD algorithm, they introduce a time-domain and spectral-domain flow aggregation (TS-FA) algorithm. Moreover, they design a time-domain and spectral-domain flow aggregation module on software defined optical network (SDON) architecture.

NI Dong, LI Hui, JI Yuefeng, LI Hongbiao, and ZHU Yinan

63 Data-Driven Joint Estimation for Blind Signal Based on GA-PSO Algorithm

Without any prior information about related wireless transmitting nodes, joint estimation of the position and power of a blind signal combined with multiple co-frequency radio waves is a challenging task. In this paper, the authors propose a particle swarm optimization to estimate multiple co-frequency "blind" source nodes. To distract the mix signals precisely, a genetic algorithm is applied, and it further improves the estimation performance of the system.

LIU Shen, QIN Yuannian, LI Xiaofan, ZHAO Yubin, and XU Chengzhong

Serial parameters: CN 34-1294/TN*2003*Q*16*72*en*P* ¥20.00*5000*10*2019-09

Statement

This magazine is a free publication for you. If you do not want to receive it in the future, you can send the "TD unsubscribe" mail to magazine@zte.com.cn. We will not send you this magazine again after receiving your email. Thank you for your support.



Editorial: Special Topic on Data Intelligence in New AI Era



Guest Editor

XU Chengzhong, IEEE Fellow, is the Dean of the Faculty of Science and Technology, University of Macau (UM), Macao SAR, China and a Chair Professor of Computer Science of UM. He was a Chief Scientist of Shenzhen Institutes of Advanced Technology (SIAT) of Chinese Academy of Sciences, China and the Director of Institute of Advanced Computing and Digital Engineering of SIAT. He was also in the faculty of Wayne State University, USA for 18 years. Dr. XU's research interest is mainly in the areas of parallel and distributed systems, cloud and edge computing, and data-driven intelligence. He has published over 300 peer-reviewed papers on these topics with over 10k citations. Dr. XU served in the editorial

boards of leading journals, including *IEEE Transactions on Computers*, *IEEE Transactions on Cloud Computing*, *IEEE Transactions on Parallel and Distributed Systems*, and *Journal of Parallel and Distributed Computing*. He is the Associate Editor-in-Chief of *ZTE Communication*. He is the Chair of IEEE Technical Committee of Distributed Processing.



Guest Editor

QIAO Yu is a professor with the Shenzhen Institutes of Advanced Technology (SIAT), the Chinese Academy of Science, China and the director of Institute of Advanced Computing and Digital Engineering, China. His research interests include computer vision, deep learning, and robots. He has published more than 170 papers in international journals and conferences, including IEEE T-PAMI, IJCV, IEEE T-IP, IEEE T-SP, CVPR, ICCV, AAAI, and ECCV. He received LV Jiaxi Young Research Award from Chinese Academy of Sciences. He is a senior member of IEEE, and a youth science and technology innovation leader of Ministry of Science and Technology (MOST) of China. His group achieved the first runner-up at the ImageNet Large Scale Visual Recognition Challenge 2015 in scene recognition and the winner at the ActivityNet Large Scale Activity Recognition Challenge 2016 in video classification. His group also achieved top places in wide international vision challenges such as ChaLearn, LSun, and THUMOUS. He served as the program chair of IEEE IICIST 2014.

The new artificial intelligence (AI) era heavily depends on three converging forces: the advance of AI algorithms, the availability of big data, and the popularity of high performance computing platforms. Data-driven intelligence, or data intelligence, is a new form of AI technologies that leverages the power of big data and advanced learning algorithm. It is becoming an extremely active research area with broad area of applications such as computer vision, speech recognition, natural language processing, medial and healthy, intelligent transportation system, multimedia system, communication, and social network.

With the huge volume of data available in various domains, big data brings opportunities to boost the performance of AI system with advanced machine learning especially deep learning techniques. It has been widely verified that deep neural networks achieve significantly better performance than previous shallow models and even surpass human performance in certain specific tasks or datasets. One well-known example is ImageNet Large Scale Visual Recognition Challenge (ILSVRC) which aims to classify or detect objects in images from 1 000 categories. The state-of-the-art deep convolutional neural networks like squeeze - and - excitation networks (SE - Net) have achieved error rates lower than 3%, which is better than human performance (error rate 5.1%). These networks usually include a deep architecture with a huge number of parameters, which are optimized with one million training datasets. In nature language processing, recent language networks like BERT or XLNet leverage more than 100 GB text for training and achieve re-

markable performance on wide tasks like SQuAD, GLUE, and RACE. All these successes heavily rely on the large scale training data. In addition to the amount of data, the label or annotation of data is also important in supervised learning. Although unsupervised learning is desirable in many applications, supervised learning usually exhibits better performance. In the next, it is important to design effective learning algorithms in unsupervised, semi-supervised, or weakly-supervised setup. On the other hand, it also presents unprecedented challenges to manage and exploit big data for a variety of applications. Learning with big data is not easy, which always needs powerful models and efficient training algorithms. Take deep networks as an example. One may need to carefully design network architectures, training losses and strategies, and effectively exploit high performance computing platforms.

This special issue seeks original articles describing development, relevant trends, challenges, and current practices in the field of big data, artificial intelligence and their applications. After careful reviews, four papers have been selected for publication in this special issue.

The first paper is titled "A Lightweight Sentiment Analysis Method". It proposes a data driven approach to perform sentiment analysis of film's critics from the Douban website and visualize the results with a word cloud.

The second paper is a survey paper with the title of "Big Data-Driven Residents' Travel Mode Choice: A Research Overview". This paper surveys the studies of residents' travel mode identification, influencing factors acquisition and choice model construction using data driven approaches.

Face detection is a fundamental yet important problem in computer vision. The third paper "Face Detection, Alignment,

➔ To P. 08

A Lightweight Sentiment Analysis Method



YU Qingshuang¹, ZHOU Jie², and GONG Wenjuan¹

(1. China University of Petroleum (East China), Qingdao, Shandong 266000, China;

2. Operation Coordination Department of Tianjin Branch of CNOOC (China) Co., Ltd., Tianjin 300000, China)

Abstract: The emergence of big data leads to an increasing demand for data processing methods. As the most influential media for Chinese domestic movie ratings, Douban contains a huge amount of data and one can understand users' perspectives towards these movies by analyzing these data. In this article, we study movie's critics from the Douban website, perform sentiment analysis on the data obtained by crawling, and visualize the results with a word cloud. We propose a lightweight sentiment analysis method which is free from heavy training and visualize the results in a more conceivable way.

Keywords: web crawler; microblog; text sentiment analysis; word cloud

DOI: 10.12142/ZTECOM.201903002

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190920.2106.008.html>, published online September 20, 2019

Manuscript received: 2019-05-09

1 Introduction

Text, semantics and social analysis are means to mine users' opinions, the market trend, and other useful information. Currently, text, semantics and social analysis technologies have been widely used in many industries including finance, media and e-commerce. For example, TUDORAN used emotional analysis to show why users block ads [1]; LEE and KWON analyzed the psychology of the people who committed suicide by analyzing the Twitter data, and analyzed the causes so that certain medication is provided to decrease the number of people who committed suicide [2]. Also, in business and governments, people extract useful information that can improve the quality of decisions from massive amounts of data.

Therefore, how to efficiently and accurately analyze the information that reflect people's opinions has become a hot research topic. Recently most researchers use machine learning methods for sentiment analysis. For example, WANG et al. extended the text library by adding network terminology and Wikipedia to the corpus and used it to train convolutional neural networks for text-level sentiment analysis [3]. RASOOL et al. used the convolutional neural network to train the sentiment analysis model to compare the popularity of the two clothing

brands on Twitter [4].

However, the cost of sentiment analysis by training the neural network is high: 1) The data set used for training and testing is huge; 2) it takes a long time to train; 3) the development of training models has reached a bottleneck. This is cumbersome for adapting for unseen sentiment analysis and especially not applicable to small business users. HUSSEIN also raised the challenge of sentiment analysis—the accuracy of the analysis still needs to be improved [5]. To this end, we propose a lightweight sentiment analysis method based on SnowNLP that is a Python-written class library for processing Chinese content. SnowNLP performs sentiment analysis by segmenting sentences, part of speech tagging, and emotional judgments.

At the same time, the sentiment analysis methods trained by YANG et al. [3] and RASOOL [4] and others through neural networks cannot visualize the central words for movies and television, which is obviously inconvenient for audiences who want to understand the movie on a glimpse. To this end, we propose to explicitly display the central words of the movie through the combination of Jieba lexicon and word cloud.

We use the proposed method to solve the problem of using neural network to analyze small amount of data while taking time to train on a large dataset. Besides, it displays the central words from the movie review, and opens up new ideas and meth-

ods for the future analysis of small data in sentiment analysis.

2 Basics of Web Crawlers and Sentiment Analysis

2.1 Web Crawlers

The web crawler obtains the source code of the webpage (**Fig. 1**) by using the request library. After obtaining the source code of the webpage, we analyze the source code data. Common data processing methods are used including regular expressions and libraries that extract web page information based on web page node properties, CSS selectors or XPath, such as BeautifulSoup, pyquery, and lxml. These libraries can efficiently and quickly extract useful information from the source code in web pages, such as node attributes, text values, and others. Finally, the processed information is saved for subsequent use.

2.2 Text Sentiment Analysis

With the development of big data, the amount of information on the Internet is increasing and how to easily and accurately extract the information from massive online reviews has become a focus of research in the field of sentiment analysis. However, the text sentiment analysis mostly solves the English text, and is based on social platforms such as Twitter [6]–[8]. For this reason, we propose a Chinese-based sentiment analysis method based on the Chinese text sentiment analysis library.

2.2.1 SnowNLP Library

SnowNLP is a Python-written class library inspired by TextBlob. It can easily handle Chinese text. Unlike TextBlob, Natural language toolkit (NLTK) is not used in SnowNLP, and all algorithms are implemented under the framework. It includes some well-trained dictionaries and can reduce the amount of data set and the amount of training time required for training neural networks.

2.2.2 Jieba Word Segmentation

Now, open source tools for Chinese word segmentation include IK (Ik - analyzer), MMseg4j, THU Lexical Analyzer for

Chinese (THULAC), Ansj, Jieba, Han Language Processing (HanLP), etc. which are still being updated and maintained. Currently, Ansj, Jieba, and HanLP perform relatively better for segmentation.

Jieba support three word segmentation modes: the precise mode, which attempts to segment the sentence most accurately suitable for text analysis; full mode, which scans all the words in the sentence that can be performed very fast but cannot solve the ambiguity; and search engine mode, which further splits a long word, improves the recall rate, and is suitable for search engine segmentation. It supports traditional word segmentation and dictionaries. The core of the algorithm is to implement efficient word graph scan based on the prefix dictionary and generate a directed acyclic graph (DAG) composed of all possible formation. Dynamic programming is used to find the maximum probability path, and the maximum segmentation combination is found based on word frequencies. For unregistered words, an HMM model based on the formation ability is adopted, and it also uses the Viterbi algorithm.

The Jieba word segmentation method is based on the Trie tree structure to achieve word graph scanning, generates a directed acyclic graph of word formation. Jieba has a dictionary, which contains more than 20 000 words, including the number of occurrences and part of speech. And Trie tree is a well-known prefix tree. If the first few words of a phrase are the same, they have the same prefix, and we can use the Trie tree to store them. It has the advantage of fast search speed.

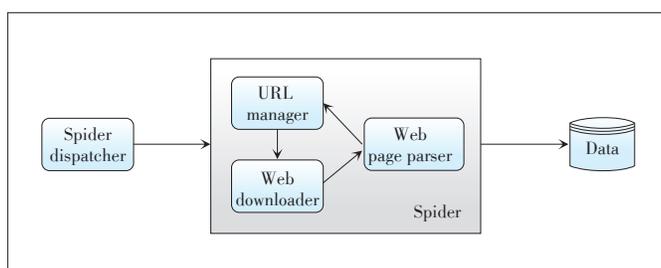
Data analysis results can be visualized through Jieba word segmentation and a word cloud which makes up for the shortcomings of deep neural networks.

3 Sentiment Analysis Based on SnowNLP

We will collect the data, process it through the SnowNLP class library, and use SnowNLP to perform word segmentation, part of speech tagging, abstract extraction, and text sentiment analysis. Sentiment analysis simply divides sentences into two categories: positive and negative. A histogram is drawn based on this probability, and the results of the sentiment analysis are explicitly displayed. The obtained data is presented in the central vocabulary of the data set through the Jieba word segmentation and the word cloud.

3.1 Collection of Comments

Comments are collected through the request library. The problem is that some websites will use anti-crawler strategies. The commonly used anti-crawler mechanism is mainly based on the following three types of information: user requests for Headers, user behavior, website directory and data loading methods. We discover that the Douban website uses the anti-crawler based on Headers. It checks and monitors the User-Agent of the Headers. We add Headers directly to the crawler. When crawling, the cookie file after login is saved in the cook-



▲ **Figure 1. Web crawler work pipeline.**

ie.txt file, and the browser's User-Agent is copied to the Headers. A user-agent is added to the headers and a cookie file is added to break through the anti-crawl mechanism and get the source code.

Firstly we should get the uniform resource locator (URL) of the website and analyze the URL to parse it on a specific page. For the homepage URL of the Douban Movie Critics¹, its prefix² is fixed, and the latter 27763742 is the ID of the movie, which is the number given by Douban. We can open the specified page to get the specified ID of the movie by analyzing the page URL of the second page³.

We discover that the value of the "start" is the number of pages; for example, the second page is 20, the third page is 30, and so on. The value of the "limit" indicates the number of movie reviews displayed on one page. This is the default value and cannot be modified. And the value of the "sort" indicates the latest hotspot comment sorting. And the value of the "status" indicates the comments of users who have seen this movie, but the comments written by users who have not seen this movie are not displayed.

3.2 Data Cleaning

The source code of the crawled webpage is processed, and the useless words are filtered out, leaving only the required information. This process mainly uses the library of bs4 for information processing.

Through the webpage check, we can know that the comments which we want are in the corresponding "p" field in the comment directory. So the latter code parsing is to extract the text contained in this part of code. Since each movie has a lot of reviews, for the timeliness of the information, only the first 200 pieces are obtained for analysis. The iterative processing is performed when the web page is processed, 20 pieces are fetched at a time, and 10 pages can be fetched. We use BeautifulSoup and regular expressions for data cleaning, leaving useful information for subsequent word frequency statistics. Finally, the cleaned string is input into the text file, which is convenient for word frequency analysis and data processing analysis.

3.3 Word Cloud Visualization

The obtained movie reviews are placed in a specific document to facilitate subsequent Jieba word segmentation and displayed in word cloud. After the content is crawled, the next step is to parse and display it. This part of the code applies to any text. For example, it can analyze the movie reviews, as well as the lyrics, news keywords, and so on. The commonly used Jieba function is its cut function, which can divide the entire sentence or the whole paragraph of the article according to the

words commonly used in Chinese, so as to facilitate the subsequent analysis of the keywords. In addition, WordCloud is also a great feature that displays the shape of the word cloud in a specific style and displays it in different text sizes depending on how often each word appears. The word cloud outlines include circle, cardioid, diamond, triangle-forward, triangle, pentagon, and star. After the word frequency of each word is obtained, there is a process of sorting the word frequency before the word cloud is displayed. Since the keywords and word frequency are composed of dictionaries, the sorted function is used in the sorting and the keywords are arranged according to the word frequency from large values to small values.

3.4 Part of Speech Analysis

SnowNLP is used to process text. It can be used for word segmentation, annotation, text sentiment analysis, etc. Sentiment analysis classifies text into two categories: the positive and negative. A predicted value of the method is the probability of the emotion.

4 Experimental Results

The movie review information is obtained by the crawler and the acquired information is cleaned by the BeautifulSoup class library. The useful information is retained and stored in the text. The information in the text is processed through the SnowNLP class library to analyze the emotional polarity, and the central subject of the text comment is explicitly displayed through the combination of Jieba word segmentation and word cloud. The experiments are carried out under the following environment: Windows10, Anaconda, WordCloud, Jieba word segmentation, and the SnowNLP class library.

4.1 Crawling Comments

The source code which is cleaned by BeautifulSoup and regular expressions is saved in the context.txt. An example is shown in Fig. 2.

4.2 Data Visualization Analysis

Five categories are picked. They are the animation, suspense, popular, unpopular, and newest. Ten movies are selected from each category. In total, 50 movies are analyzed and the results are visualized in Tables 1–5.

From the word cloud of different types of the movies, we found that there are some similarities. The same type of movie word clouds display mostly the same vocabulary, while the different types of movie word clouds mainly show different vocabulary. For example, the most displayed in the word cloud are the types of the movies, such as "animation" and "Cartoon" in the word cloud of the animation type, and "drama" and "suspense" in the word cloud of the suspense type. This is conducive to distinguishing different types of movies, which is of great significance in information recommendation. The word

¹ <https://movie.douban.com/subject/27763742/comments?status=P>

² <https://movie.douban.com/subject/>

³ https://movie.douban.com/subject/27763742/comments?start=20&limit=20&sort=new_score&status=P



▲ Figure 2. Comments after data cleaning.

clouds can also reflect the audiences’ emotional views on this movie, showing the words “like”, “exciting”, “classic”, “junk”, and so on. Moreover, there are many names of actors and protagonists in the word cloud of different types of the movies. This could help audiences choose the right movie to watch.

The difference is the emotional analysis of different types of movies. Since our experiment is to randomly select 10 movies on each type, all of them are random, that is, we can treat them as samples. Through the emotional analysis of these samples, we can know the reviews for the animation and popular movies are intensive and mostly positive comments. The suspense movies have many positive comments but mostly neutral words. The newest movies only have a few and mostly negative comments. Surprisingly, although there are not many movie reviews for the unpopular movies, most comments are positive. By analyzing the sentiment analysis graphs of different types of the movies, cultural-related workers can understand the audiences’ preferences for movie types so that they can make corresponding changes.

5 Discussion and Analysis

We learned from the experiment results that our method is applicable to different types of movies. It can analyze the expression emotions out of reviews on different types of movies. Combined with the network rating, we can see that the scores of high-score movies could be close to 1 in the sentiment analysis, which indicates that our method is highly credible. And at the same time through the word cloud we can know the central words of reviews on a movie, and it is helpful to conducive to information recommendation. The 50 pictures of the entire program ran down and took only 10 minutes (limited by the speed of the network and the anti-crawling of the website). In general, the proposed method is highly credible, easy to summarize and recommend, and takes less time.

Through this study, the method of sentiment analysis has been explored. From the large amount of data, the long-time machine learning, deep learning and training neural network,

▼ Table 1. Visualization analysis of experimental results from the animation movies

| Name of the movie | Visualized word cloud | Predicted sentiment probability |
|---------------------------|-----------------------|---------------------------------|
| The Monkey King | | |
| Up | | |
| Zootopia | | |
| Howl's Moving Castle | | |
| WALL·E | | |
| Spirited Away | | |
| The Lion King | | |
| Laputa: Castle in the Sky | | |
| The Legend of Sealed Book | | |
| Coco | | |

▼Table 2. Visualization analysis of experimental results from the suspense movies

| Name of the movie | Visualized word cloud | Predicted sentiment probability |
|-----------------------------|-----------------------|---------------------------------|
| Fight Club | | |
| Inception | | |
| The Butterfly Effect | | |
| Contratiempo | | |
| Witness for the Prosecution | | |
| Seven | | |
| The Lives of Others | | |
| Twelve Angry Men | | |
| Infernal Affairs | | |
| The Prestige | | |

▼Table 3. Visualization analysis of experimental results from the unpopular movies

| Name of the movie | Visualized word cloud | Predicted sentiment probability |
|------------------------------------|-----------------------|---------------------------------|
| The Grand Mansion Gate | | |
| British Museum presents: Hokusai | | |
| Dayo Wong Tze Wah Comedy Talk Show | | |
| Zur Person: Hannah Arendt | | |
| The Fantastic Mr Feynman | | |
| Billy Elliot | | |
| Curtain: Poirot's Last Case | | |
| The Little Prince | | |
| Daria in "Is it College Yet?" | | |
| Daria in "Is It Fall Yet?" | | |

▼Table 4. Visualization analysis of experimental results from the popular movies

| Name of the movie | Visualized word cloud | Predicted sentiment probability |
|-----------------------------------|-----------------------|---------------------------------|
| Bohemian Rhapsody | | |
| Avengers: Endgame | | |
| Gisaengchung | | |
| The Invisible Guest | | |
| Green Book | | |
| Hotel Mumbai | | |
| The Blind Melody | | |
| Pain&Glory | | |
| A Cool Fish | | |
| Spider-Man: Into the Spider-Verse | | |

▼Table 5. Visualization analysis of experimental results from the newest movies

| Name of the movie | Visualized word cloud | Predicted sentiment probability |
|-----------------------------------|-----------------------|---------------------------------|
| Bohemian Rhapsody | | |
| Avengers: Endgame | | |
| Capharnaïm | | |
| Gisaengchung | | |
| Green Book | | |
| Ready Player One | | |
| Dying to Survive | | |
| Shoplifters | | |
| Coco | | |
| Spider-Man: Into the Spider-Verse | | |

to the Python-based sentiment analysis method, the new ideas and methods of sentiment analysis are expanded. On the other hand, our experiments still have some shortcomings; for example, the accuracy is not very high, which has a possible solution of training the class library by adding data sets to SnowNLP and using short-term emotional polarity analysis only for small project data. We hope our research has a contribution to the research of textual sentiment analysis.

References

- [1] TUDORAN A A. Why do Internet Consumers Block Ads? New Evidence from Consumer Opinion Mining and Sentiment Analysis [J]. *Internet Research*, 2019, 29(1): 144–166. DOI: 10.1108/IntR-06-2017-0221
- [2] LEE S Y, KWON Y. Twitter as a Place Where People Meet to Make Suicide Pacts [J]. *Public Health*, 2018, 159: 21–26. DOI: 10.1016/j.puhe.2018.03.001
- [3] YANG X, XU S, WU H, BIE R. Sentiment Analysis of Weibo Comment Texts Based on Extended Vocabulary and Convolutional Neural Network [J]. *Procedia Computer Science*, 2019, 147: 361–368. DOI: 10.1016/j.procs.2019.01.239
- [4] RASOOL A, TAO R, MARJAN K, NAVEED T. Twitter Sentiment Analysis: A Case Study for Apparel Brands [J]. *Journal of Physics: Conference Series*, 2019, 1176(2): 022015. DOI: 10.1088/1742-6596/1176/2/022015
- [5] HUSSEIN D M E-D M. A Survey on Sentiment Analysis Challenges [J]. *Journal of King Saud University-Engineering Sciences*, 2016, 30(4): 330–338. DOI: 10.1016/j.jksues.2016.04.002
- [6] BAGHERI H, ISLAM M J. Sentiment Analysis of Twitter Data [DB/OL]. (2017-22-25). <https://arxiv.org/abs/1711.10377>
- [7] SAIF H, HE Y, FERNANDEZ M, et al. Contextual Semantics for Sentiment Analysis of Twitter [J]. *Information Processing & Management*, 2016, 52(1): 5–19. DOI: 10.1016/j.ipm.2015.01.005
- [8] THELWALL M, BUCKLEY K, PALTOGLOU G. Sentiment Strength Detection

for the Social Web [J]. *Journal of the Association for Information Science & Technology*, 2012, 63(1):163–173. DOI: 10.1002/asi.21662

Biographies

YU Qingshuang (yqs_18106301006@163.com) received the B.E. degree in Software Engineering from Qufu Normal University, China in 2019. He is currently pursuing a master's degree in computer science at China University of Petroleum (East China). His research interests include data mining and deep learning. He once participated in the National College Student Innovation Competition and won the second prize.

ZHOU Jie is currently working at the Tianjin branch of China National Offshore Oil (China) Co., Ltd. (CNOOC). After graduating in 2006, he joined the offshore oil industry and worked in China Petroleum Environmental Protection Services (Tianjin) Co., Ltd. and CNOOC Tianjin Branch. His main research interests include understanding and resolving the contradiction between the sensitive areas of the three provinces and one city, the environmental protection zone, the main functional zoning of the ocean, and the offshore oil and gas exploration and development.

GONG Wenjuan received the B.E. degree in software engineering from Shandong University, China in 2004, the M.S. degree in computer graphics from Shandong University, China in 2007, and the M.S. and Ph.D. degrees in information technology from Autonomous University of Barcelona, Spain in 2013. From 2013 to 2014, she was a postdoctoral researcher with the Oxford Brookes University, UK. She is currently with China University of Petroleum (East China). Her research interests include computer vision, audio processing, machine learning, and quantum machine learning. She has published 14 SCI-indexed papers.

←From P. 01

Quality Assessment and Attribute Analysis with Multi-Task Hybrid Convolutional Neural Networks” introduces multi-task hybrid convolutional neural networks for face detection, alignment, quality assessment and attribute estimation.

The last paper “RAN Centric Data Collection for New Radio” is from the communication area, which exploits self-orga-

nizing networks and minimization of driver tests to support deployment of new radio (NR) system and conduct performance optimization.

Finally, we would like to thank all the authors, the external reviewers for their contributions and efforts to organize this special issue in this esteemed journal.



Big Data–Driven Residents’ Travel Mode Choice: A Research Overview

ZHAO Juanjuan¹, XU Chengzhong², and MENG Tianhui¹

(1. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518000, China;

2. University of Macao, Macau SAR 999078, China)

Abstract: The research on residents’ travel mode choice mainly studies how traffic flows are shared by different traffic modes, which is the prerequisite for the government to establish transportation planning and policy. Traditional methods based on survey or small data sources are difficult to accurately describe, explain and verify residents’ travel mode choice behavior. Recently, thanks to upgrades of urban infrastructures, many real-time location-tracking devices become available. These devices generate massive real-time data, which provides new opportunities to analyze and explain resident travel mode choice behavior more accurately and more comprehensively. This paper surveys the current research status of big data-driven residents’ travel mode choice from three aspects: residents’ travel mode identification, acquisition of travel mode influencing factors, and travel mode choice model construction. Finally, the limitations of current research and directions of future research are discussed.

Keywords: intelligent transportation; travel modes choice; urban computing

1 Introduction

In recent years, with the rapid economic growth and the acceleration of urbanization process in China, modern urban transportation systems, especially metropolitan transportation systems are facing a series of problems such as inadequate bearing capacity, crowded traffic and air pollution. In order to improve the efficiency of urban transport systems and promote urban sustainable development, effective management strategies have to be taken. Moreover, accurately understanding urban resident’s travel mode behavior is the precondition for the government to make corresponding measures. The choice of residents’ travel mode determines the distribution of people and vehicles in the urban traffic network. Since travelers are autonomous, only by sufficiently understanding travelers’ choice behavior can traffic management measures be effectively formulated to lead the travelers to adjust their travel modes and further relieve traffic pressure.

Traditional studies on residents’ travel mode choice mainly rely on field experience and sampling survey to obtain the data. However, due to limited information, it is hard to accurately describe, explain and verify residents’ travel mode choice behavior. Recently, with the development of sensing technology and computing environment, long-term and continuous data can be collected, such as traffic flow, trajectory, traffic network, interest point and meteorological data. The multi-sources big data brings new perspectives to analyze and explain residents’ travel mode choice behavior.

Big data-driven residents’ travel mode modeling mainly contains three core issues: residents’ travel mode identification, residents’ travel mode influencing factors acquisition, and residents’ travel mode model construction (**Fig. 1**). Among them, acquisition of resident’s travel mode and influencing factors is the premise of constructing residents’ travel mode model. Residents’ travel mode identification identifies travelers’ transportation means based on the trajectory. The traveling means include walking, bicycle, bus, subway, car, and so on. Influencing factor acquisition mainly studies the algorithm of influencing factor extraction. There are many factors affecting residents’ travel mode choice behavior, such as the service level of traffic

DOI: 10.12142/ZTECOM.201903003

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190919.1452.002.html>, published online September 19, 2019

Manuscript received: 2019–05–09

This work was supported in part by National Natural Science Foundation of China (No. 61802387) and the Shenzhen Discipline Construction Project for Urban Computing and Data Intelligence.

facilities, urban design, transit user, and so on. Some factors need to be extracted by fusing multiple data sources. For example, we need to integrate bus Global Positioning System (GPS), bus operation time and smart card data to extract traffic congestion, which is important for evaluating service level of traffic facilities. The resident travel mode prediction model is used to discover the relationship between these factors and travel means, so that residents’ travel choice can be accurately predicted.

This paper is organized as follows. Sections 2–4 review the recent proposed methods from three aspects: residents’ travel mode identification, residents’ travel mode influencing factor acquisition, and residents’ travel mode choice model construction. The key limitations of current research and the challenges are discussed in Section 5. Finally, we have a concluding remark in Section 6.

2 Identification of Resident Travel Mode

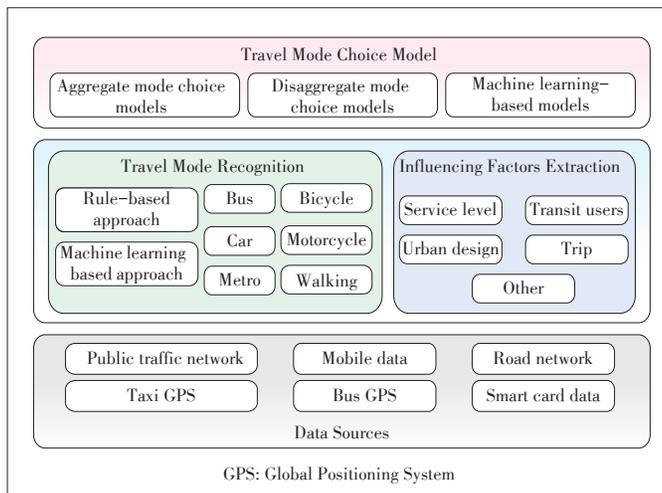
Residents’ travel modes can be divided into personal transport modes and public transport modes. Personal transport modes include walk, bicycle, electric vehicle, car, and other modes. Public transport modes mainly comprise ground bus and rail transit. Residents’ travel mode identification plays an important role in understanding users’ mobility and traffic situation. With the dramatic development of data processing and sensor technology (such as GPS, accelerometer, GSM and Bluetooth) over the last decade, we can collect huge amount of residents’ mobility information and environmental parameters. How to use the collected data to identify urban residents’ travel modes has become a hot but difficult research issue.

The process of identifying user’s travel mode can be summarized as three steps. First, user’s trip chain which reflects the relationship between trajectory and travel mode is extracted. Then the characteristics related to the travel mode, such as

speed, start, and stop time are extracted. Finally identification methods are used to identify the user’s travel mode. The identification methods can be divided into two categories: rule-based methods and machine learning-based methods.

Rule-based methods recognize different travel modes by setting different rules according to the logical characteristics of different travel modes. If the target satisfies the judgment conditions of a certain travel mode, it is classified into the travel mode. For example, STOPHER et al. [1], BOHTE et al. [2] and CHEN et al. [3] set different thresholds for average speed, maximum speed and travel time to distinguish different travel modes such as by walking, by bus and by car. However, due to its strong subjectivity (the rules and thresholds are set mainly based on experience), the accuracy of rule-based methods is not high, and the scalability is limited. So it is difficult to obtain high discriminating accuracy and scalability.

In recent years, the study of travel mode identification by machine learning, such as decision-tree, support vector machine, neural network, and stochastic forest, has become a hot research area. The accuracy and scalability have been greatly improved compared with the previous methods. The travel mode is identified using data from sensors e.g., GPS, accelerometer, Wi-Fi, and GSM. ZHENG et al. [4] collected the GPS location data of 45 objects for 6 months based on GPS device and compared the effectiveness of travel mode identification using different algorithms, including the decision-tree algorithm, Bayesian network and Support Vector Machine (SVM). The results indicated that the decision-tree algorithm is better than the other two [4]. Based on the assumption that using different travel modes make different vibrations, LAN et al. [5] used the output voltage of kinetic energy acquisition device as a signal source to detect the travel modes chosen by travelers in their daily travels. Experimental results show that this method is feasible [5]. ENDO et al. [6] put forward an automatic feature extraction method based on depth neural network method, to address the problem of low prediction accuracy caused by the artificial feature selection method and the noise of trajectory. In view of the fact that the existing machine learning methods were inadequate in explaining the results of residents’ travel modes, DAS et al. [7] put forward a hybrid knowledge driven method combining fuzzy logic with neural network. A Fuzzy expert system can give a reasoning scheme but lacks the adaptability and learning ability that the neural network has. Therefore, their combination can perfectly offset mutual weaknesses [7]. Due to the ambiguous situations, for instance, traffic lights, traffic jam, bus stops and weak signal reception, current techniques report high misclassification errors for inferring transportation modes. To overcome this problem, LOPEZ et al. [8] presented a method for detecting changes of transportation mode on a multimodal journey. They used a space transformation for extracting features that identify a transition between two transportation modes based on data collected from the Google Application Programming Interface (API) for hu-



▲ Figure 1. Framework of a big data–driven residents’ travel mode choice model.

man activity classification through a crowdsourcing-based application for smartphones. The results showed improvements on precision and accuracy in comparison to initial classification data outcomes [8]. Based on multiple data sources, there are some studies on travel mode identification. FENG et al. [9] used a Bayesian network model to identify the travel model based on acceleration and GPS mixed data. The results showed that the fusion of GPS data and acceleration data has better results than that of using one single data [9]. SHIN et al. [10] proposed a real-time travel mode identification algorithm based on the acceleration and location information collected from mobile phones. JAHANGIRI et al. [11] used multiple supervised learning methods, such as k-Nearest Neighbor (kNN), SVM, decision-tree algorithm, Bagging, and Random Forest, to identify users’ travel modes based on user mobile data collecting from mobile accelerometers and gyroscopes. The results revealed that Random Forest had the best prediction accuracy [11].

3 Acquisition of Influencing Factors

Residents’ travel mode choice behavior is influenced by many factors. These factors can be summed up in five categories [12]–[15]: service level of transportation facilities (fare, time, comfort, and reliability); urban design (functional zones, population density, surrounding environment, public transportation density, etc.); trip (start and end location, travel purpose, and travel distance); traveler (income, car ownership, sex and age, etc.); and else (weather, special events, etc.). So far, a lot of research advances have been achieved in acquiring influencing factors based on different sensor data. In the following, we introduce the latest research progresses from three aspects.

1) In terms of service level of transportation facilities.

BARABINO et al. estimated the punctuality of buses at each stop based on bus location data and passenger’s travel pattern [16], [17]. MA et al. identified and estimated the reliability of bus operation in different sections based on automatic vehicle positioning data and smart card data [18]. Based on bus location information, user’s request for bus and smart card data, HADJIMITRIOU et al. identified the bus stops where the waiting time is much more than the expected by considering bus service request volume and bus time reliability [19]. CHEPURRI et al. evaluated travel time variability as well as reliability using GPS based trajectory data [20]. ZHENG et al. analyzed road traffic conditions based on taxi GPS data [21]. Also based on smart card data, ZHANG et al. analyzed metro passengers’ fined-grained travel time, including the time of entering and exiting metro station, transferring time, and waiting time [22]–[24]. Overcrowding is one of the major causes of discomfort. CEAPA et al. extracted station crowding patterns based on historical automated fare collection systems data [25]. WANG et al. proposed a citywide and real-time model for estimating the travel time of any path [26], based on the GPS trajectories of vehicles received in current time slots and over a period of his-

tory as well as map data sources. ZHANG et al. estimated the congestion of buses based on passengers’ smart card data and bus GPS data [27]. ZHAO et al. proposed a probabilistic model to analyze how the passenger flows were dispatched to different routes and trains based on smart card and train operation timetables [28].

2) In terms of trip.

LI et al. proposed a method to predict and validate home and work locations of commuters by exploring underlying repeated travel patterns based on public transport smart card data [29]. Based on the internal spatiotemporal relationship within multi-day smart card transaction data, ZOU et al. proposed a center-point based algorithm to infer the home location for each cardholder, and a rule-based approach using the individual properties (home location and card type) of cardholders and the travel information (time and space) of each trip to identify trip purpose (work, school, shopping, and others) [30]. Furthermore, MA et al. proposed a mining method to identify transit commuters by leveraging spatial clustering and multi-criteria decision analysis approaches based on transit smart card data [31]. HUANG et al. presented an approach using spatial temporal attractiveness of Point of Interests (POIs) to identify activity-locations as well as durations from raw GPS trajectory [32]. FURLETTI et al. proposed a probability model to identify traveler’s purposes, such as going to work and studying, based on GPS trajectory and interest point data [33]. LIU et al. provided a practical framework for inferring the trip purposes of taxi passengers, where the probability of points of interest to be visited is modeled by Bayes’ rules, with both spatial and temporal constraints taken into consideration [34].

3) In terms of urban design.

LIU et al. proposed a scene classification framework to identify dominant urban land use type at the level of traffic analysis zone by integrating probabilistic topic models and support vector machine [35]. YUAN et al. identified urban functional zones based on the data of residents’ movement trajectory, POI, urban road network and so on [36]. YAO et al. established a novel framework to map urban population distributions at the building scale by integrating multisource geospatial big data [37].

4 Residents’ Travel Mode Choice Model

Constructing residents’ travel mode selection model is a classical research topic in traffic planning [38]. According to the analysis unit that is an individual or a group, there are two types of models, the aggregate model and disaggregate model. The aggregate model is proposed to model the behavior of a group of travelers. The disaggregate model is put forward later based on the utility theory and its analysis unit is an individual. Compared with the aggregate model, the disaggregate model has the characteristics of high prediction accuracy, which makes the disaggregate model the main research direction in

recent years.

The disaggregate model offers substantial advantage over the aggregate model as it models the behavior of individuals. The Logit model based on the theory of utility maximization is the first proposed disaggregate model. It is assumed that a traveler preference for a certain travel mode can be described by the “utility value” of the travel mode characteristics and socio-demographic attributes of the traveler. In the Logit model, the utility of an individual selecting a travel mode j is denoted as U_j and composed of two parts: the deterministic utility V_j of observable factors and the random utility ξ_j of unobservable factors. The random term ξ_j is assumed to be independent and subject to the same probability distribution (Gumbel distribution or extreme distribution). The Logit model is widely used to analyze the distribution of transport vehicles for inter-city and inner-city travelers [39], [40].

Because the Logit model assumes that the unobservable random utility ξ_j of choice branches are independent and subject to Gumbel distribution, there is no correlation among all the choices (Independence from Irrelevant Alternative). However, in some cases it is divorced from reality [41], [42]. In order to overcome the defects of the Logit model, many researchers have tried to improve the disaggregate model and proposed various and more advanced models. MINAL et al have given a survey of the research advances [43] about these models, such as Probit model and advanced Logit model. The Probit model assumes that ξ_j are not independent with each other, but this causes more parameters to calculate. When choice branches are greater than 2, it is very complex to solve the parameters. Although many approximation solutions, such as the simulation methods (Monte-Carlo simulation and McFadden), are proposed later, the qualitative analysis has uncertainty, inestimable error, and defects of complex parameter calibration process, which hinders the practicability of the model [44]–[46]. Therefore, some scholars have turned to study improved Logit models which have partial advantages of the Probit model. These improved Logit models can be generalized into two categories, the hierarchical Logit improvement model and the direct Logit improvement model. The nested Logit model is based on the multi-level partition of a travel mode, which is closer to reality than the multiple Logit model because of the correlation between the modes [43], [47]. The Dogit model is the representative of direct Logit models [48], which divides residents' choice of travel modes into two types, the forced choice behavior and free choice behavior. It assumes that the ratio of the forced choice behavior is fixed, which the free choice behavior is subject to the Logit model.

Along with the development of machine learning, some researchers use artificial neural networks (ANNs) to predict residents' travel mode choice [49], [50]. Compared with the Logit model, ANN models are easily applicable with their higher capability to identify nonlinear relationships between inputs and designated outputs to predict choice behaviors. OMRANI et al.

[51] presented four machine learning methods, namely artificial neural net-MLP, artificial neural net-RBF, multinomial logistic regression, and support vector machines, for predicting travel mode of individuals. The results reveal that the artificial neural networks perform better compared to other alternatives [51]. Moreover, LEE et al. [52] investigated the capability of four ANN models, including backpropagation neural networks (BPNNs), radial basis function networks (RBFNs), probabilistic neural networks (PNNs), and clustered probabilistic neural networks (CPNNs), and compared their prediction performance with a conventional multinomial logit model (MNL) for mode choice problems. The results show that ANN models outperform MNL, with prediction accuracies around 80% compared with 70% for MNL [52]. ANNs perform better than tradition models on modeling residents' travel mode choice, but they have the shortage that the parameters are hard to explain, which blocks its popularization in transportation applications.

5 Limitations of Current Research and Challenges

By now, a large number of related studies have been proposed to model residents' travel mode choice behavior. However, due to incomplete data, complex influence factors, unclear relationship between factors and residents' travel modes, there are still some limitations in current research stage, which we summarized as the following three points:

1) Residents' travel mode identification.

Current methods are able to achieve higher accuracy in identifying the travel mode based on relatively high-frequency sampling data, but they are not applicable to identify the travel modes of all residents in a whole city. First, in general, it is difficult to collect high-frequency location data of a large number of residents. Though we can obtain some time and space information of large number residents through various sensors, such as mobile phones and smart cards, for example, phone users do not use mobile phones all the time. The smart card can only capture a user's time and space information when he/she is using public transportation. Furthermore there is no unique identification between different sensors for the same user. That is, the method cannot be directly used when there is no user association between sensors and discontinuous sparse location data. Overall, current methods cannot be directly used under the condition of sparse and discontinuous location data, and no unique identification between different but related sensors.

2) Acquisition of influencing factors.

In terms of influencing factor acquisition, there are limitations in the following two aspects: a) Due to the sparsity and incompleteness of data, it is difficult to assess the influencing factors with great accuracy based on a single source or a small number of data sources. For example, the train capacity rate and passenger flow distribution in a metro network are the important indicators in comfort evaluation. The premise is to cap-

ture passengers’ route choice behavior in the subway network. However, we can only obtain each passenger’s time and station information when he/she enters and leaves metro stations. If such data as cell phone signaling of a passenger is integrated, more fine-grained location information of the passenger can be obtained, thus making a more accurate estimation of route choice behavior. ZHENG et al. have made a comprehensive overview of cross-domain data fusion technology, but relatively less in the analysis of residents’ travel modes [36]. b) The analysis results are dispersed and do not form a whole area. There are correlations between these factors, but current studies only focus on analyzing one or some of them, instead of paying more attention to the correlations between different factors.

3) Construction of residents’ travel mode choice model.

Many researchers are studying the construction of travel mode choice models, but there have been few accurate estimations of the residents’ travel modes due to limited data as well as the large gap between assumption and reality. For example, the most common definition of “utility value” is to consider the combination of fare/income, time in the car and walking time, while ignoring the influence of other factors.

6 Conclusions and Future Research

The emergence of big data provides us the possibility to deeply understand residents’ travel mode choice behavior. This paper summarizes the current research status from three aspects: residents’ travel mode identification, influencing factors acquisition and residents’ travel mode choice model construction. The future research directions include:

1) Knowledge fusion. The methods for residents’ travel mode identification and influencing factor acquisition based on a single data have been well explored. However, the methodology that can learn mutually reinforced knowledge from multiple data sources is still missing. Therefore, how to integrate heterogeneous cross-domain data sources to identify residents’ travel modes and extract potential influencing factors is a challenge that needs to be addressed in the future.

2) Residents’ travel mode prediction. Residents’ travel mode choice behavior is affected by multiple complex factors, including internal and external environments. The correlations between these factors, as well as between factors and resident travel modes, are still unclear. Therefore, how to discover these correlations and make an accurate prediction for residents’ travel mode choice is another challenge that needs to be studied in the future.

References

- [1] STOPHER P, FITZGERALD C, ZHANG J. Search for a Global Positioning System Device to Measure Person Travel [J]. *Transportation Research Part C Emerging Technologies*, 2008, 16(3): 350–369. DOI: 10.1016/j.trc.2007.10.002
- [2] BOHTE W, MAAT K. Deriving and Validating Trip Purposes and Travel Modes for Multi-Day GPS-Based Travel Surveys: A Large-Scale Application in the Netherlands [J]. *Transportation Research Part C*, 2009, 17(3): 285–297. DOI: 10.1016/j.trc.2008.11.004
- [3] CHEN C, GONG H, LAWSON C, et al. Evaluating the Feasibility of a Passive Travel Survey Collection in a Complex Urban Environment: Lessons Learned from the New York City Case Study [J]. *Transportation Research Part A Policy & Practice*, 2010, 44(10): 830–840. DOI: 10.1016/j.tra.2010.08.004
- [4] ZHENG Y, LIU L, WANG L, et al. Learning Transportation Mode from Raw GPS Data for Geographic Applications on the Web [C]//International Conference on World Wide Web, WWW 2008. Beijing, China, 2008:247–256. DOI: 10.1145/1367497.1367532
- [5] LAN G, XU W, KHALIFA S, et al. Transportation Mode Detection Using Kinetic Energy Harvesting Wearables [C]//IEEE International Conference on Pervasive Computing and Communication Workshops. Sydney, Australia, 2016: 1–4. DOI: 10.1109/PERCOMW.2016.7457048
- [6] ENDO Y, TODA H, NISHIDA K, et al. Deep Feature Extraction from Trajectories for Transportation Mode Estimation [M]. *Advances in Knowledge Discovery and Data Mining*. New York, USA: Springer International Publishing, 2016. DOI: 10.1007/978-3-319-31750-2_5
- [7] DAS R D, WINTER S. Detecting Urban Transport Modes Using a Hybrid Knowledge Driven Framework from GPS Trajectory [J]. *International Journal of Geo-Information*, 2016, 5(11): 207. DOI: 10.3390/ijgi5110207
- [8] LOPEZ A J, OCHOA D, GAUTAMA S. Detecting Changes of Transportation-Mode by Using Classification Data [C]//IEEE International Conference on Information Fusion. San Diego, USA, 2015: 2078–2083
- [9] FENG T, TIMMERMANS H J P. Transportation Mode Recognition Using GPS and Accelerometer Data [J]. *Transportation Research Part C Emerging Technologies*, 2013, 37(3): 118–130. DOI: 10.1016/j.trc.2013.09.014
- [10] SHIN D, ALIAGA D, TUNÇER B, et al. Urban Sensing: Using Smartphones for Transportation Mode Classification [J]. *Computers Environment & Urban Systems*, 2015, 53: 76–86
- [11] JAHANGIRI A, RAKHA H A. Applying Machine Learning Techniques to Transportation Mode Recognition Using Mobile Phone Sensor Data [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(5): 2406–2417. DOI: 10.1109/TITS.2015.2405759
- [12] ZHAO F, LI M T, CHOW L F, et al. FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access [R]. Tampa, USA: University of South Florida, 2002
- [13] RACCA D P, RATLEDGE E C. Factors that Affect and/or can Alter Mode Choice [R]. Newark, USA: University of Delaware, 2004.
- [14] YE X, PENDYALA R M, GOTTARDI G. An Exploration of the Relationship Between Mode Choice and Complexity of Trip Chaining Patterns [J]. *Transportation Research Part B Methodological*, 2007, 41(1): 96–113. DOI: 10.1016/j.trb.2006.03.004
- [15] MADHUWANTHI M, MARASINGHE A, RAJAPAKSE R P C J, et al. Factors Influencing to Travel Behavior on Transport Mode Choice—A Case of Colombo Metropolitan Area in Sri Lanka [J]. *International Journal of Affective Engineering*, 2015, 15(2): 63–72. DOI: 10.5057/ijae.IJAE-D-15-00044
- [16] BARABINO B, FRANCESCO M D, MOZZONI S. Rethinking Bus Punctuality by Integrating Automatic Vehicle Location Data and Passenger Patterns [J]. *Transportation Research Part A*, 2015, 75: 84–95. DOI: 10.1016/j.tra.2015.03.012
- [17] BARABINO B, FRANCESCO M D, MURRU R. An Offline Framework for Reliability Diagnosis by Automatic Vehicle Location Data [C]//Conference on Advanced Systems in Public Transport. Rotterdam, Netherlands, 2015. DOI: 10.1109/its.2016.2581024
- [18] MA Z, FERREIRA L, MESBAH M, et al. Modelling Bus Travel Time Reliability Using Supply and Demand Data from AVL and Smart Card Systems [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2015, 2533:17–27
- [19] HADJIDIMITRIOU S N, KAPARIAS I, DELL’AMICO M. Investigating Urban Bus Travel Time Reliability Patterns in London Using iBus Automatic Vehicle Locating and Live Bus Arrivals data [C]//6th Symposium of the European Association for Research in Transportation. Haifa, Israel, 2017
- [20] CHEPURI A, JAIRAM R, ARKATKAR S, et al. Travel Time Reliability-Based Performance Indicators Assessment for Bus Routes Using GPS-Based Bus Trajectory under Mixed Traffic Conditions [C]//3rd Conference of Transportation Systems Engineering and Management (CTSEM). Bangalore, India, 2016
- [21] ZHENG Y, LIU Y, YUAN J, et al. Urban Computing with Taxicabs [C]//13th ACM International Conference on Ubiquitous Computing. Beijing, China, 2011: 89–98. DOI: 10.1145/2030112.2030126
- [22] WAITING E, HOWEVER A S F. Splitting Travel Time Based on AFC Data: Estimating Walking, Waiting, Transfer, and In-Vehicle Travel Times in Metro System [J]. *Discrete Dynamics in Nature and Society*, 2015(1): 1–11. DOI:

- 10.1155/2015/539756
- [23] ZHANG F, ZHAO J, TIAN C, et al. Spatiotemporal Segmentation of Metro Trips Using Smart Card Data [J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(3): 1137–1149. DOI: 10.1109/TVT.2015.2409815
- [24] YUE Z, CHEN F, WANG Z, et al. Classifications of Metro Stations by Clustering Smart Card Data Using the Gaussian Mixture Model [J]. *Urban Rapid Rail Transit*, 2017, 30(2): 48–51
- [25] CEAPA I, SMITH C, CAPRA L. Avoiding the Crowds: Understanding Tube Station Congestion Patterns from Trip Data [C]//ACM SIGKDD International Workshop on Urban Computing, Beijing, China, 2012: 134–141. DOI: 10.1145/2346496.2346518
- [26] WANG Y L, ZHENG Y, XUE Y X. Travel Time Estimation of a Path Using Sparse Trajectories [C]//Knowledge Discovery and Data Mining (KDD'14). New York, USA, 2014: 25–34. DOI: /10.1145/2623330.2623656
- [27] ZHANG J, YU X, TIAN C, et al. Analyzing Passenger Density for Public Bus: Inference of Crowdedness and Evaluation of Scheduling Choices [C]//IEEE 17th International Conference on Intelligent Transportation Systems. Qingdao, China, 2014: 2015–2022. DOI: 10.1109/ITSC.2014.6958000
- [28] ZHAO J, ZHANG F, TU L, et al. Estimation of Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(4): 790–801. DOI: 10.1109/ITITS.2016.2587864
- [29] LI G, YU L, NG W S, et al. Predicting Home and Work Locations Using Public Transport Smart Card Data by Spectral Analysis [C]//IEEE 18th International Conference on Intelligent Transportation Systems. Gran Canaria, Spain, 2015: 2788–2793. DOI: 10.1109/ITSC.2015.445
- [30] ZOU Q, YAO X, ZHAO P, et al. Detecting Home Location and Trip Purposes for Cardholders by Mining Smart Card Transaction Data in Beijing Subway [J]. *Transportation*, 2016: 1–26. DOI: 10.1007/s11116-016-9756-9
- [31] MA X, LIU C, WEN H, et al. Understanding Commuting Patterns Using Transit Smart Card Data [J]. *Journal of Transport Geography*, 2017, 58: 135–145. DOI: 10.1016/j.jtrangeo.2016.12.001
- [32] HUANG L, LI Q, YUE Y. Activity Identification from GPS Trajectories Using Spatial Temporal POIs' Attractiveness [C]//2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks. San Jose, USA, 2010: 27–30. DOI: 10.1145/1867699.1867704
- [33] FURLETTI B, CINTIA P, RENSO C, et al. Inferring Human Activities from GPS Tracks [C]//ACM SIGKDD International Workshop on Urban Computing. Chicago, USA, 2013. DOI: 10.1145/2505821.2505830
- [34] LIU X. Inferring Trip Purposes and Uncovering Travel Patterns from Taxi Trajectory Data [J]. *Cartography & Geographic Information Science*, 2016, 43(2): 103–114. DOI: 10.1080/15230406.2015.1014424
- [35] LIU X, HE J, YAO Y, et al. Classifying Urban Land Use by Integrating Remote Sensing and Social Media Data [J]. *International Journal of Geographical Information Science*, 2017, 31(8): 1675–1696. DOI: 10.1080/13658816.2017.1324976
- [36] YUAN N J, ZHENG Y, XIE X, et al. Discovering Urban Functional Zones Using Latent Activity Trajectories [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2015, 27(3): 712–725. DOI: 10.1109/TKDE.2014.2345405
- [37] YAO Y, LIU X, LI X, et al. Mapping Fine-Scale Population Distributions at the Building Level by Integrating Multi-Source Geospatial Big Data [J]. *International Journal of Geographical Information Science*, 2017, 31(6): 1220–1244. DOI: 10.1080/13658816.2017.1290252
- [38] CASCETTA E. *Transportation Systems Engineering: Theory and Methods* [M]. Dordrecht, Netherlands: Kluwer Academic Publishers, 2001. DOI: 10.1007/978-1-4757-6873-2
- [39] ABUHAMOUD M A A, RAHMAT R A O K, ISMAIL A B. Modeling of Transport Mode in Libya: A Binary Logit Model for Government Transportation Encouragement [J]. *Australian Journal of Basic & Applied Sciences*, 2011, 5(5): 1291–1296
- [40] PRAVEEN KUMAR M, MALLIKARJUNA C. Mode Choice Modelling for Intercity Transportation in India: A Case of Guwahati to Five Metro Cities [J]. *International Journal of Earth Sciences and Engineering*, 2011, 4(6): 364–374
- [41] BHAT C R. A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice [J]. *Transportation Research Part B Methodological*, 1995, 29(6): 471–483. DOI: 10.1016/0191-2615(95)00015-6
- [42] HESS S. *Advanced Discrete Choice Models with Applications to Transport Demand* [D]. London, UK: Imperial College London, 2005
- [43] MINAL, SEKHAR C R. Mode Choice Analysis: The Data, the Models and Future Ahead [J]. *International Journal for Traffic and Transport Engineering*, 2014, 4(3): 269–285. DOI: 10.7708/ijte.2014.4(3).03
- [44] GHAREIB A H. Evaluation of Logit and Probit Models in Mode-Choice Situation [J]. *Journal of Transportation Engineering - Asce*, 1996, 122(4): 282–290. DOI: 10.1061/(ASCE)0733-947X(1996)122:4(282)
- [45] DOW J K, ENDERSBY J W. Multinomial Probit and Multinomial Logit: a Comparison of Choice Models for Voting Research [J]. *Electoral Studies*, 2004, 23(1): 107–122. DOI: 10.1016/S0261-3794(03)00040-4
- [46] BHAT C R, SARDESAI R. The Impact of Stop-Making and Travel Time Reliability on Commute Mode Choice [J]. *Transportation Research Part B-methodological*, 2006, 40(9): 709–730. DOI: 10.1016/j.trb.2005.09.008
- [47] ABDEL-ATY M, ABDELWAHAB H. Calibration of Nested Mode Choice Model for Florida [R]. Final Research Report, University of central Florida, 2001.
- [48] GAUNDRY M J, DAGENAIS M G. The Dogit Model [J]. *Transportation Research Part B-methodological*, 1979, 13(2): 105–111. DOI: 10.1016/0191-2615(79)90028-6
- [49] HENSHER D A, TU T T. A Comparison of the Predictive Potential of Artificial Neural Networks and Nested Logit Models for Commuter Mode Choice [J]. *Transportation Research Part E Logistics & Transportation Review*, 2000, 36(3): 155–172. DOI: 10.1016/S1366-5545(99)00030-7
- [50] CANTARELLA G E, DE LUCA S. Modeling Transportation Mode Choice Through Artificial Neural Networks [C]//Fourth International Symposium on Uncertainty Modeling and Analysis, 2003. College Park, USA, 2003: 84–90. DOI: 10.1109/ISUMA.2003.1236145
- [51] OMRANI H. Predicting Travel Mode of Individuals by Machine Learning [J]. *Transportation Research Procedia*, 2015, 10: 840–849. DOI: 10.1016/j.trpro.2015.09.037
- [52] LEE D, DERRIBLE S, PEREIRA F C. Comparison of Four Types of Artificial Neural Network and a Multinomial Logit Model for Travel Mode Choice Modeling [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2018. DOI:10.1177/0361198118796971

Biographies

ZHAO Juanjuan (jj.zhao@siat.ac.cn) received the Ph.D. degree from Chinese Academy of Sciences, China in 2017. She is an assistant professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. Her research interests include big data processing, data privacy, and urban computing.

XU Chengzhong received the Ph.D. degree from the University of Hong Kong, China in 1993. He is the Dean of the Faculty of Science and Technology, University of Macau, China and a Chair Professor of Computer Science of UM. He was a Chief Scientist of Shenzhen Institutes of Advanced Technology (SIAT) of Chinese Academy of Sciences and the Director of Institute of Advanced Computing and Digital Engineering of SIAT. He was also in the faculty of Wayne State University, USA for 18 years. Dr. Xu's research interest is mainly in the areas of parallel and distributed systems, cloud and edge computing, and data-driven intelligence. He has published over 300 peer-reviewed papers on these topics with over 10K citations. Dr. Xu served in the editorial boards of leading journals, including *IEEE Transactions on Computers*, *IEEE Transactions on Cloud Computing*, *IEEE Transactions on Parallel and Distributed Systems*, and *Journal of Parallel and Distributed Computing*. He is the Associate Editor-in-Chief of *ZTE Communication*. He is IEEE Fellow and the Chair of IEEE Technical Committee of Distributed Processing.

MENG Tianhui received the Ph.D. degree in computer science from Free University of Berlin, Germany in 2017. He is currently an assistant professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. His research interests include mobile edge computing, big data processing, blockchain and Internet of Things.



Face Detection, Alignment, Quality Assessment and Attribute Analysis with Multi-Task Hybrid Convolutional Neural Networks

GUO Da^{1,2*}, ZHENG Qingfang^{3,4*}, PENG Xiaojiang^{1,2}, and LIU Ming^{3,4}

(1. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. ZTE Corporation, Shenzhen, Guangdong 518057, China;
4. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen, Guangdong 518057, China)

Abstract: This paper proposes a universal framework, termed as Multi-Task Hybrid Convolutional Neural Network (MHCNN), for joint face detection, facial landmark detection, facial quality, and facial attribute analysis. MHCNN consists of a high-accuracy single stage detector (SSD) and an efficient tiny convolutional neural network (T-CNN) for joint face detection refinement, alignment and attribute analysis. Though the SSD face detectors achieve promising results, we find that applying a tiny CNN on detections further boosts the detected face scores and bounding boxes. By multi-task training, our T-CNN aims to provide five facial landmarks, facial quality scores, and facial attributes like wearing sunglasses and wearing masks. Since there is no public facial quality data and facial attribute data as we need, we contribute two datasets, namely FaceQ and FaceA, which are collected from the Internet. Experiments show that our MHCNN achieves face detection performance comparable to the state of the art in face detection data set and benchmark (FDDB), and gets reasonable results on AFLW, FaceQ and FaceA.

Keywords: face detection; face alignment; facial attribute; CNN; multi-task training

DOI: 10.12142/ZTECOM.201903004

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190920.2104.004.html>, published online September 20, 2019

Manuscript received: 2019-06-11

1 Introduction

Face analysis has been widely-used in many applications such as face beautification system, face based access system, and video anti-terrorism system. Although great progresses have been made in recent, detecting and aligning abnormal faces such as occlusion faces as well as analyzing their attributes in surveillance are very challenging due to low resolution, lack of abnormal training data, etc.

Generally, face detection, face alignment, facial quality assessment, and facial attribute recognition are considered as separate face analysis tasks which may have their own task-de-

pendent models. For face detection, from traditional Viola-Jones (VJ) face detector [1] and deformable part model (DPM) based face detector [2] to recent convolutional neural networks (CNN) based face detectors [3]–[9], the performance of face detection has been improved significantly. Among all the CNN based face detectors, those detectors evolved from anchor-based object detectors (e.g. single shot multibox detector (SSD) [10], Faster Region CNN (R-CNN) [11]), such as single shot scale-invariant face detector (S³FD) [12] and Face Region CNN (R-CNN) [4], are superior to pure CNN face detectors [7], because anchor-based detectors can naturally leverage the context information. For face alignment, CNN based methods have also achieved promising results [12]–[16]. However, most of the alignment methods must be initialized by the provided face bounding box in advance, which presents a great demand of joint face and landmark detection [7], [17]. For facial quality assessment, traditional methods [18] mainly apply local binary patterns (LBP) [19] or histograms of oriented gradients (HOG)

This work was supported by ZTE Corporation and State Key Laboratory of Mobile Network and Mobile Multimedia Technology.
* Both authors contributed equally to this work.

[20] features with a support vector machine (SVM) classifier, while a few works with CNN obtain state-of-the-art performance [21], [22]. For facial attribute recognition, [23] introduces the CelebA dataset with 40 facial attributes ranging from smiling to gender, and subsequently many deep learning based methods are developed for facial attribute analysis [23]–[25]. Unfortunately, CelebA does not contain the attribute of wearing face mask which we are interested in.

In this paper, we address several face analysis tasks including face detection, face alignment, facial quality assessment, and facial attribute recognition in the wild. Specifically, we propose a Multi-Task Hybrid Convolutional Neural Network (MHCNN) which unifies all the tasks in a framework. MHCNN is comprised of two parts, namely a single stage detector (SSD) and an efficient tiny CNN (T-CNN). Compared to pure CNN face detectors, the SSD based face detector ensures high baseline accuracy on challenging face images in the wild. Instead of performing multi-task learning with SSD like [17], we apply a tiny CNN which is more feasible for multi-task face analysis. We argue that a CNN operated on cropped faces brings complementary information to SSD. Given an image, the MHCNN first detects all the faces with a SSD based face detector and then refines both the scores and bounding boxes with T-CNN. Since the T-CNN is applied in individual faces, it is straightforward to add multiple tasks upon it. We here add face alignment, facial quality assessment, and facial attribute recognition. In addition, we introduce a facial attribute dataset, i. e. FaceA, which contains two highly-concerned attributes in surveillance, namely wearing sunglasses and wearing masks. We also introduce a human-based facial quality assessment dataset, i. e. FaceQ, where low-quality cases include occlusion, low-resolution, large pose, etc. We evaluate our face detection performance on the well-known face detection data set and benchmark (FDDB) dataset, and demonstrate our T-CNN on our FaceA and FaceQ.

The remained of this paper is organized as follows. In Section 2, we review related work on face detection and multi-task learning. We introduce our MHCNN and its training strategy in Section 3. Our collected datasets are introduced in Section 4. We present experimental results in Section 5 and conclude the paper in Section 6.

2 Related work

We mainly review the face detection and multi-task learning for face analysis in this section. One can refer to [23]–[26] and [27] for face image quality assessment and facial attribute recognition, respectively.

2.1 Face Detection

Face detection has been a well-studied field of computer vision. According to the used features, face detection methods can be roughly divided into two categories, namely hand-craft

feature based methods and CNN feature based methods.

1) Hand-craft feature based methods. The cascaded face detector proposed by VIOLA et al. [1] (VJ detector) obtains good performance in simple scenarios with real-time efficiency. Due to the relatively weakness of Haar-like features, the VJ detector degrades significantly in real-world applications with larger visual variations of human faces. Some works improved the VJ detectors by replacing the Haar-like features with more advanced hand-crafted ones [28]–[30], which need more computational cost. Another popular pipeline of face detection is based on DPM [2], [31], [32]. It performs relatively better than VJ detector in the wild but it is more computationally expensive and usually requires expensive annotation in the training stage.

2) CNN feature based methods. Since the remarkable success of CNN in object classification [33], many progresses have been made for face detection [3]–[9]. These CNN-based methods can be mainly concluded as three categories, namely cascaded CNN based face detection, two-stage region-based face detection, and single-stage face detection. The cascaded CNN based face detection pipeline, which inherits the advantage of the VJ detector, utilizes several small networks from simple to complex to detect faces and regress face boxes in a coarse-to-fine manner [5]–[7]. Two-stage region-based face detection pipeline is mainly transferred from region-based object detectors, like R-CNN [34], Fast R-CNN [35], and Faster R-CNN [11]. This method mainly includes two stages, namely proposal generation and classification. The single-stage face detection pipeline directly generates face boxes and scores from dense anchor boxes [8], [9]. The face detection model for finding tiny faces [7] trains separate detectors for different scales. S³FD [12] presents multiple strategies to improve the performance of small faces. Single stage headless (SSH) [9] models the context information by large filters on each prediction module. PyramidBox [36] utilizes contextual information with improved SSD network structure. The advantage of single-stage face detectors is that it can use the context semantic information to assist in detecting faces, which is difficult for cascade face detectors. So, we introduce single-stage detection architecture into cascade face detector to get higher performance.

2.2 Multi-Task Learning

There are some existing works attempting to jointly solve the problem of face detection, alignment and facial attribute in a single model. YANG et al. [6] train deep convolution neural networks for facial attribute recognition to obtain high response in face regions which further yield candidate windows of faces. ZHANG et al. [37] proposed to use facial attribute recognition as an auxiliary task to enhance face alignment performance using deep convolutional neural network. CHEN [38] et al. apply random forest based on the features of pixel value difference to jointly conduct alignment and detection, but these handcraft features are low-level features and greatly limit its perfor-

mance. Multitask cascaded convolutional network (MTCNN) [7] leverages a cascaded architecture with three stages of shallow to deep convolution networks to jointly predict face and landmark locations in a coarse-to-fine manner, but the performance of MTCNN is limited by cascade architecture. So, we propose MHCNN with single-stage architecture in cascade face detector and joint multi-task learning to improve the performance of face detector.

3 Multi-Task Hybrid Convolutional Neural Networks

In this section, we first overview the proposed MHCNN, and then detail the two parts of MHCNN, namely the SSD-based face detector and the tiny multi-task CNN. We finally describe the training process of MHCNN.

3.1 Overview

Fig. 1 illustrates the pipeline of our MHCNN. The MHCNN consists of an SSD) and an efficient T-CNN. Given an image, the MHCNN first detects all the faces with the SSD based face detector and then refines both the scores and bounding boxes with T-CNN. T-CNN is also responded on facial landmark regression, facial quality assessment, and facial attribute recognition. We merge these facial tasks by considering that 1) it is more efficient than training multiple networks for each task, 2) some tasks, such as face classification and face attribute recognition, could be complementary with each other, and 3) T-CNN is performed on individual faces which could be complementary with an SSD-based face detector.

3.2 The SSD-Based Face Detector

We resort to the recent S³FD [12] as the first part of

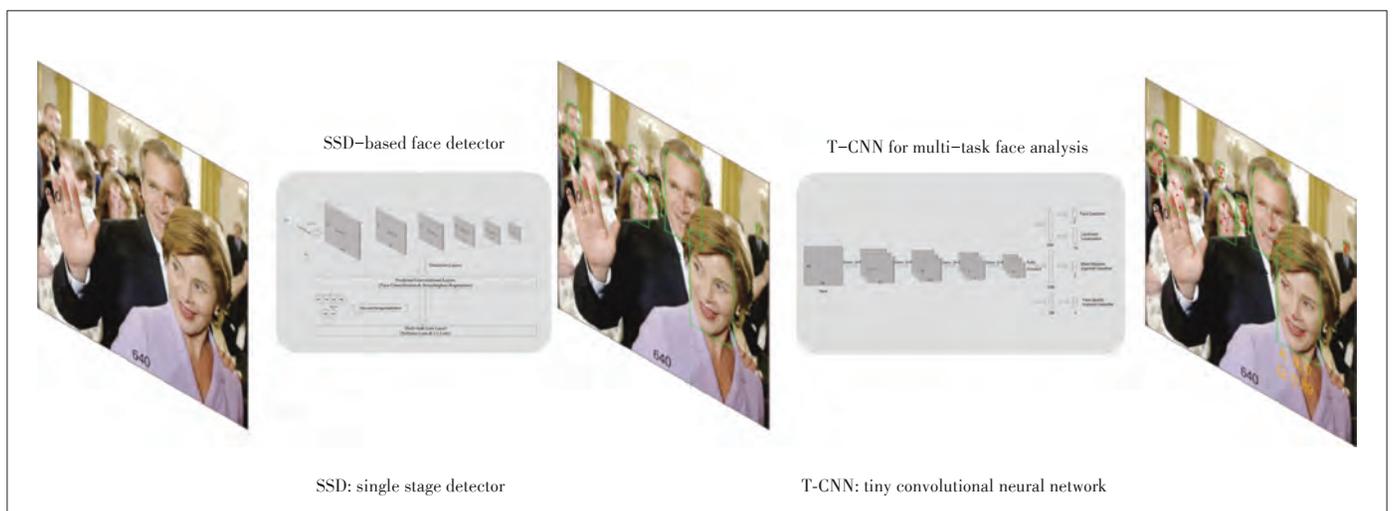
MHCNN since it is a state-of-the-art SSD-based face detector which is robust to face scale variation. **Fig. 2** shows the architecture of S³FD. Both S³FD and original SSD use VGG16 [39] as their backbone and pretrained on ImageNet. Compared to original SSD, S³FD has several adjustments. We briefly review these adjustments as follows.

1) Anchor design. Instead of generating anchors with different scale and ratio for each detection layer in the original SSD, S³FD generates one anchor scale for each detection layer since faces can be simply approximated as squares. Specifically, anchor scales range from 16 to 512 pixels with equal-proportion interval principle with strides increasing from 4 to 128 pixels. This strategy aims to guarantee that there are adequate features in different layers for detection.

2) Detection layers. One of the main challenges for face detection is to detect faces from tiny size to very large size. Unlike object detection of SSD in ImageNet, S³FD moves forward the detection layers in order to detect tiny faces. Specifically, the detection layers of S³FD are conv3_3, conv4_3, conv5_3, conv_fc7, conv6_2, and conv7_2. The norms of conv3_3, conv4_3, and conv5_3 are fixed to 10, 8, and 5 by L2 normalization for better training.

3) Max-out of background. Since negative anchors dominate the shallow layers with massive types, S³FD adds a max-out layer for conv3_3 to relax the imbalance of positive and negative anchors. The max-out layer views the background label as several different labels and only takes the most active one for classification.

4) Scale compensation anchor matching. To match more tiny faces for anchors, S³FD decreases the jaccard overlap threshold from 0.5 to 0.35 in order to increase the average number of matched anchors, and further sorts these anchors with jaccard overlap higher than 0.1 and selects top-N as matched anchors.



▲ **Figure 1.** The pipeline of the proposed Multi-Task Hybrid Convolutional Neural Network (MHCNN). It consists of an SSD-based face detector for high-accuracy detection performance and a T-CNN for detection refinement and multi-task face analysis.

5) Training. We use the training set of the WIDER FACE [40] to train the SSD-based detector, and use the same data augmentation strategies as S³FD, including color distort, random crop, and horizontal flip. The input size of network is fixed to 640×640. We use smooth L1 loss for face bounding box regression and softmax loss for face/non-face classification. We apply non-maximum suppression (NMS) to remove the highly overlapped results with a threshold of 0.7.

3.3 The Multi-Task Tiny CNN

We design an efficient T-CNN for the second part of MHCNN. The T-CNN aims to further refine the candidates from the SSD-based face detector, detect facial landmarks, assess the face quality, and recognize two importance facial attributes.

Fig. 3 presents the architecture of T-CNN. This architecture is inspired by the MTCNN [7]. Specifically, we use off-the-shelf O-Net of MTCNN as the architecture while add more tasks. T-CNN takes as input the 48×48 face regions, and output results for four face tasks as follows.

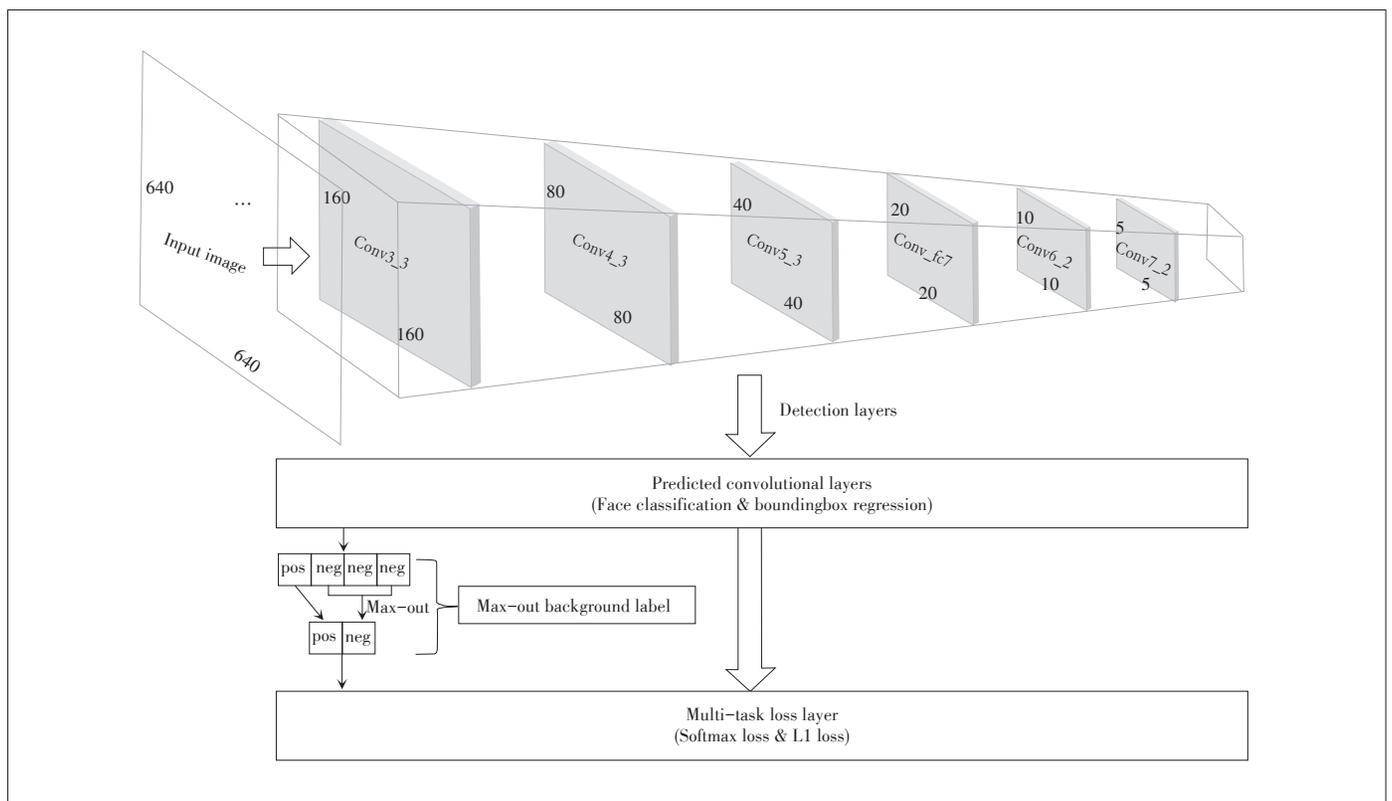
1) Face classification. We find that there are a number of non-face cases in the detections of the first part of MHCNN, which are mainly caused by hard negative contexts and low qualities, such as a person with back face. These cases can be relaxed by directly classifying the face regions. We believe

that adding a refinement T-CNN could be complementary with the SSD-based face detector. The face scores from both S³FD and T-CNN will be averaged to provide the final detections.

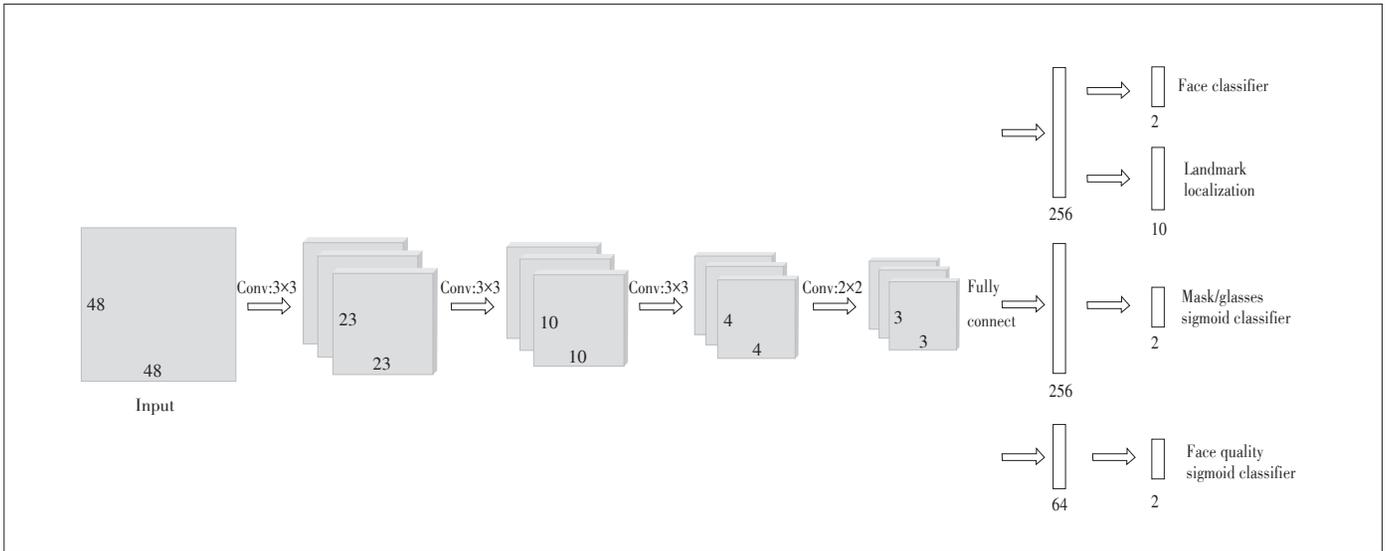
2) Landmark localization. We also predict five landmarks at eyes, nose, and mouth as in MTCNN. As shown in MTCNN, adding landmark localization is helpful for face recognition. We explain that landmarks can be viewed as a post validation of faces.

3) Attribute classification. Facial attribute is naturally a multi-label task. In this paper, we only concern about two important attributes for surveillance applications, i. e. wearing face masks and wearing sunglasses. A sigmoid layer is responded to each attribute.

4) Face quality classification. Face quality can impact the face/non-face classification scores in practice. We add face quality task as a two-class (i. e. high quality and low quality) classification problem since it is hard to annotate accurate quality scores for human. We consider two issues for face quality classification: a) face quality assessment served as a filter for subsequent face recognition system since too many low-quality and unrecognizable faces can impact the communication of front devices and cloud devices; b) As shown in Fig. 3, we use a separate fully-connected (FC) layer for face quality assessment because this task mainly depends on non-semantic information and it brings negative influence to other tasks if



▲ Figure 2. The architecture of single stage detector (SSD)-based face detector.



▲ Figure 3. The architecture of tiny convolutional neural network (T-CNN).

sharing the same FC layer in practice.

5) Training. Since T-CNN is partially served as a face/non-face refinement of the SSD-based detector in our MHCNN framework, we need to collect training data according to the results of SSD-based detector. To this end, we first calculate the Intersection-over-Union (IoU) ratio between the detections from the SSD-based detector and ground-truth faces on the training set of WIDER FACE, and then select these detections with IoU above 0.4 as positives and those less than 0.35 as negatives. The number of total face/non-face training data for T-CNN is 60000 with a ratio of 1:3. For facial landmark localization, we use the CelebA dataset which annotated with five facial landmarks, and apply random crop and gaussian blur as two data augmentation strategies. Euclidean loss is used for training facial landmark regression. We use our FaceA and FaceQ to train facial attribute recognition and face quality assessment. As for the training of facial attributes, the sigmoid layer with binary cross-entropy loss is used.

4 FaceA and FaceQ

This section details the collection and annotation of our FaceA and FaceQ datasets. To our knowledge, there is no public dataset for face attribute recognition with both wearing face masks and wearing sunglasses, and there is also no public dataset for face quality assessment in the wild. To meet our research, we collect the FaceA and the FaceQ datasets, and will make it publicly available to promote this area.

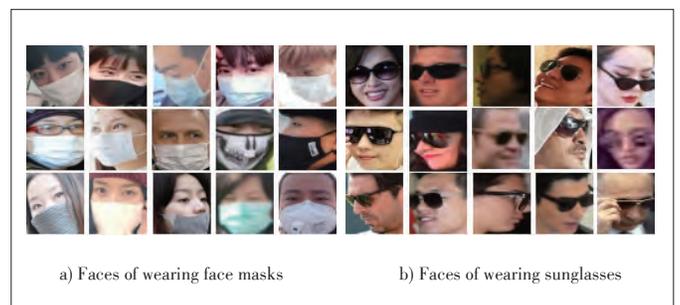
1) Collection. We make a python script and start from crawling “wearing sunglasses” and “wearing face mask” in image searching engine such as baidu (www.baidu.com) and google (www.google.com). We find it is hard to collect a large scale of data for wearing face mask since this case usually happens in

surveillance. Totally, we crawled 1 409 and 1 335 images for “wearing sunglasses” and “wearing face mask”, respectively. After crawling, we then feed these images into the first part of our MHCNN and crop the detected faces for further annotation.

2) FaceA. With the detected faces, we find there are many noises which are not human faces (e.g. cartoon and animation) or without the expected attributes. We manually remove these noises, and finally the FaceA dataset consists of 1 072 faces with sunglasses and 663 faces with face mask. The FaceA also includes a background category which contains 630 faces without wearing things. We randomly split both classes with 8:2 as training set and test set. **Fig. 4** shows some examples of FaceA. We note that these faces are mostly with large head poses which could be challenging for recognition.

3) FaceQ. With the crawled images, we find there are a lot of faces that neither belong to “wearing sunglasses” nor “wearing face mask”, and that there are a number of faces with either high or low resolution. Thus, we collect FaceQ from the same source with FaceA but with three rules:

- Face resolution: Blurred and tiny faces are divided into low-quality class.



▲ Figure 4. Examples of our FaceA dataset.

- Head pose: We collect faces as high-quality class only if their eyes can be seen clearly and their resolution are larger than 80×80 . In other word, profile faces are not selected as high-quality class.
- Occlusion: Occluded faces are selected as the low-quality class except for the faces that are only occluded by wearing sunglasses and masks.

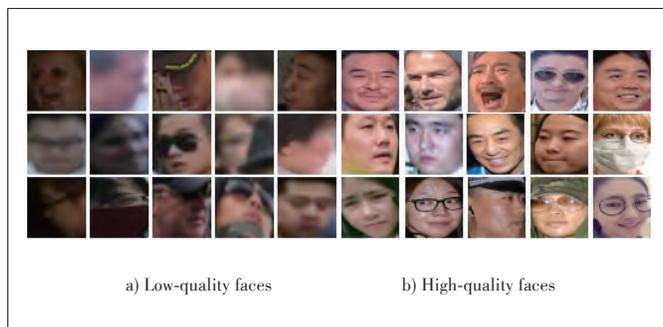
After manually selection, we totally obtain 1001 high-quality faces and 1 097 low-quality faces. We also randomly split both classes with 8:2 as training set and test set. Some examples of FaceQ are shown in Fig. 5.

5 Experiments

In this section we first present the implementation details, and then analyze the effectiveness of our joint multi-task training for the face detection, and further evaluate the final model on Fddb face detection benchmark and our own benchmarks. Finally, we evaluate the time of inference of our MHCNN.

5.1 Implementation Details

We use Caffe toolbox for implementation of our MHCNN. For the SSD-based face detector, we follow the training setting of S³FD, using the pretrained VGG16 to initialize, and the other layers are randomly initialized with the “Xavier” method [13]. We fine-tune the pretrained model using stochastic gradient descent (SGD) with 0.9 momentum, 0.0005 weight decay and batch size 32. We train 80k iterations by using 10^{-3} learning rate, then continue training for 20k iterations with 10^{-4} and 10^{-5} learning rates. For T-CNN, we convert different datasets to hdf5 format for joint multi-task training, and train the model using SGD with 0.9 momentum and 0.0005 weight decay. The batch size of each task is 512, and we concatenate all data for joint training. Due to the fact that face quality assessment task depends on different information (i. e. low-level information) compared to the other tasks, we train T-CNN in two stages. First, we train the face classification, facial landmark localization, and face attribute recognition tasks with 10^{-1} learning rate for 200 iterations. Then we freeze the weights and only train the face quality recognition part of T-CNN with 10^{-1} learning



▲ Figure 5. Examples of our FaceQ dataset.

rate for another 500 iterations.

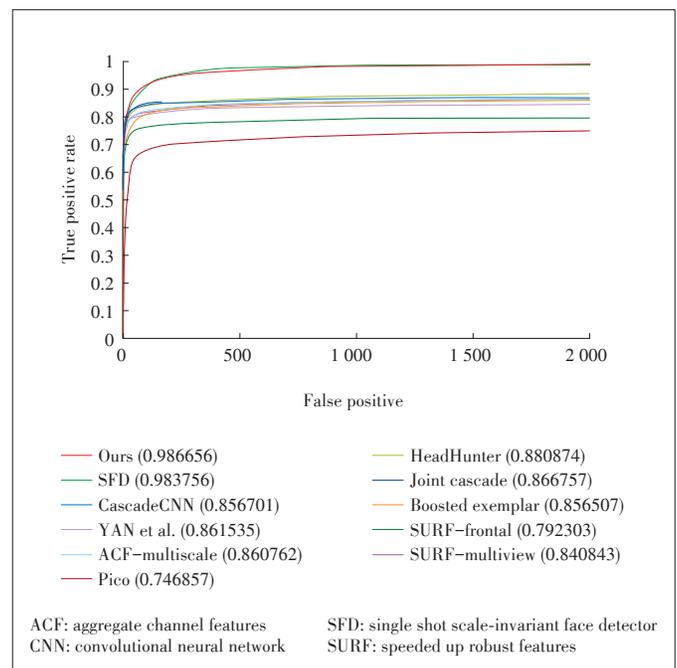
5.2 Evaluation on Face Detection Task

We evaluate and compare the face detection performance of our MHCNN on the Fddb dataset. Fddb contains 5 171 face annotations in 2 845 images. We compare our face detector to the state-of-the-art methods [2], [3], [12], [38], [41]–[45]. Table 1 shows the recall ratio at 2 000 false positive and Fig. 6 compares the receiver operating characteristic (ROC) curves to several state-of-the-art methods. Although our T-CNN aims for multi-task facial analysis, compared to the original S³FD, our extra T-CNN provides complementary information for face/non-face classification which boosts S³FD by around 0.3%. It is worth noting that a tiny improvement is difficult on the nearly-saturated Fddb. From Fig. 6, we observe that our MHCNN

▼ Table 1. Comparison of our MHCNN on Fddb

| Methods | Recall |
|------------------------|---------------|
| Cascade CNN [3] | 85.67% |
| ACF-multiscale [41] | 86.08% |
| YAN et al. [42] | 86.15% |
| Faster R-CNN [11] | 96.10% |
| S ³ FD [12] | 98.37% |
| MHCNN | 98.66% |

ACF: Aggregate Channel Features R-CNN: Region CNN
 CNN: Convolutional Neural Network S³FD: Single Shot Scale-invariant Face Detector
 MHCNN: Multi-task Hybrid CNN



▲ Figure 6. Discontinuous ROC curves on the Fddb dataset.

mainly improve the true positive rate at low false positive rate, which means the MHCNN has higher scores for true faces than S³FD. This character is practical in real applications.

5.3 Ablation Study of T-CNN

We make an ablation study of our T-CNN on the facial attribute task. We perform experiments on the collected FaceA dataset, and use the sigmoid scores for each attribute with the best threshold searched on the training set.

Table 2 presents the results of ablation study on FaceA. We find several observations from Table 2. First, adding the facial landmark localization task improves both attribute tasks, where the gains for sunglasses and face mask are 0.43% and 2.6%, respectively. Second, training with all attribute tasks and landmark localization achieves the best results on FaceA. Third, the overall results with multiple tasks are relatively high though we only use a low-resolution input, which demonstrates the efficiency of our T-CNN.

1) Visualization on FaceA. **Fig. 7** visualizes some false positives on FaceA. We find that “wearing mask” is easily confused by large-pose faces with heavy hair, partial occlusion, and sunglasses; “wearing sunglasses” is confused by wearing common glasses, partial occlusion, and wearing mask.

2) Face quality assessment with T-CNN. For face quality assessment task, we evaluation our T-CNN on the FaceQ dataset. We compare a well-known and popular method in real applications, i.e. LBP feature with SVM. In this method we use the circular LBP operator with 8 sampling points in a circular area with radius 2, and divide face image into 7×5 to calculate LBP histogram, getting a 2 065 dim feature vector for each face image. Using the LBP face image features, we train a SVM model with radial basis kernel function (RBF) to predict either the normalized comparison scores. The cost of SVM is set at 1.5 and the gamma for RBF at 6.82. **Table 3** presents the comparison between MHCNN and LBP+SVM. We find that T-CNN outperforms LBP+SVM by 3.34%, which demonstrates its effectiveness. As a traditional method, LBP+SVM is also promising on this task which may be explained by that the face quality assessment task mainly depend on low-level texture information.

3) Visualization on FaceQ. **Fig. 8** shows some false positives on FaceQ. We find that most of the faces with serious occlusion by sunglasses and masks are recognized as low-quality faces, which may make sense since they are not suitable for recognition by human; several smooth profile faces also have low quality scores since they have little texture information; these low-quality faces with small occlusion by sunglasses are easily categorized into high-quality faces, which may be explained by the fact that these faces can provide relatively rich texture information.

5.4 Inference Time

During inference, we set 0.5 as the face confidence threshold in both parts of MHCNN. We perform the inference in 10

real-world surveillance images with 1 080×1 920 scales and report the average time. We first downscale the images to 320×568 and then use our MHCNN. The inference time of the SSD-

▼ **Table 2. Ablation study of T-CNN on the FaceA dataset**

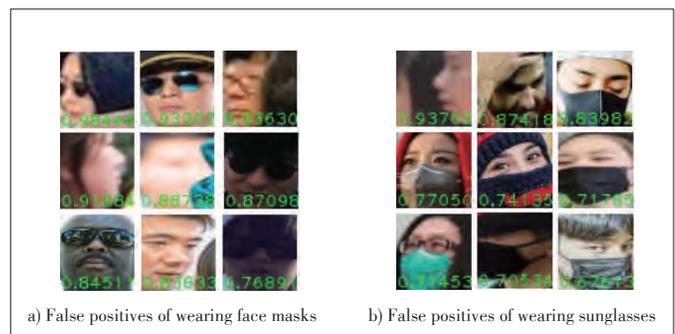
| Methods (task setting) | Accuracy of Sunglasses (Threshold = 0.5) | Accuracy of Mask (Threshold = 0.5) |
|---|--|------------------------------------|
| T-CNN (sunglasses) | 76.14% | ---- |
| T-CNN (sunglasses + landmarks) | 76.57% | ---- |
| T-CNN (masks) | ---- | 83.30% |
| T-CNN (masks + landmarks) | ---- | 85.90% |
| T-CNN (sunglasses + masks + landmarks) | 98.70% | 99.35% |

T-CNN: tiny CNN

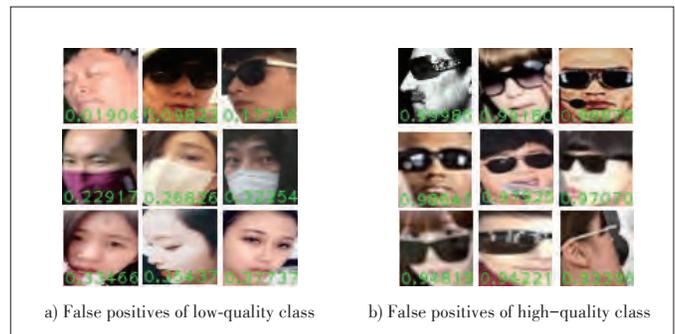
▼ **Table 3. Evaluation on the FaceQ dataset**

| Methods | Accuracy of Face Quality (Best threshold) |
|--------------|---|
| LBP+SVM | 78.52% |
| T-CNN | 81.86% |

LBP: local binary patterns SVM: support vector machine T-CNN: tiny CNN



▲ **Figure 7. False positives on FaceA test set. The score of images are the probability to predict wearing-mask and wearing-sunglasses.**



▲ **Figure 8. False positives on FaceQ test set. The faces with higher scores are predicted to the high-quality class and those with lower scores are predicted to the low-quality class.**

based detector and T-CNN are around 22 ms/frame and 2 ms/frame in NVIDIA TITAN Xp, respectively. Overall, our MHCNN can run 40 FPS in NVIDIA TITAN Xp for four face tasks including resizing computational time.

6 Conclusions

In this paper, we propose MHCNN for face detection, facial landmark detection, facial quality, and facial attribute analysis. We combine the single stage detector and CNN-based detector to boost the performance of face detection and implement multi-task learning. Our MHCNN achieves real-time performance in NVIDIA GPU for four face tasks. Additionally, we contribute two datasets on face attribute and face quality assessment. Experiments show that our MHCNN achieves the state-of-the-art on Fddb benchmark and gets reasonable results on FaceQ and FaceA.

References

- [1] VIOLA P, JONES M J. Robust Real-Time Face Detection [J]. *International Journal of Computer Vision*, 2004, 57(2): 137–154. DOI: 10.1023 / B: VISI.0000013087.49260.fb
- [2] MATHIAS M, BENENSON R, PEDERSOLI M, et al. Face Detection Without Bells and Whistles [C]//*European Conference on Computer Vision*. Zurich, Switzerland, 2014: 720–735. DOI: 10.1007/978-3-319-10593-2_47
- [3] ZHU C, ZHENG Y, LUU K, et al. CMS-RCNN: Contextual Multi-Scale Region-Based CNN for Unconstrained Face Detection [M]. *Deep Learning for Biometrics*. Cham, Switzerland: Springer, 2017: 57–79
- [4] JIANG H, LEARNED-MILLER E. Face Detection with the Faster R-CNN [C]//*12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. Washington DC, USA, 2017: 650–657. DOI: 10.1109/FG.2017.82
- [5] LI H, LIN Z, SHEN X, et al. A Convolutional Neural Network Cascade for Face Detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA, 2015: 5325–5334. DOI: 10.1109/CVPR.2015.7299170
- [6] YANG S, LUO P, LOY C C, et al. From Facial Parts Responses to Face Detection: A Deep Learning Approach [C]//*IEEE International Conference on Computer Vision*. Santiago, Chile, 2015: 3676–3684. DOI: 10.1109/ICCV.2015.419
- [7] ZHANG K, ZHANG Z, LI Z, et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks [J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1499–1503. DOI: 10.1109/LSP.2016.2603342
- [8] ZHANG S, ZHU X, LEI Z, et al. Faceboxes: A CPU Real-Time Face Detector with High Accuracy [C]//*2017 IEEE International Joint Conference on Biometrics (IJCB)*. Denver, Colorado, USA, 2017: 1–9. DOI: 10.1109 / BTAS.2017.8272675
- [9] NAJIBI M, SAMANGOUEI P, Chellappa R, et al. SSH: Single Stage Headless Face Detector [C]//*IEEE International Conference on Computer Vision*. Venice, Italy, 2017: 4875–4884. DOI: 10.1109/ICCV.2017.522
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector [C]//*European Conference on Computer Vision*. Amsterdam, The Netherlands, 2016: 21–37. DOI: 10.1007/978-3-319-46448-0_2
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [C]//*Advances in Neural Information Processing Systems*. Montreal, Canada, 2015: 91–99. DOI: 10.1109 / TPAMI.2016.2577031
- [12] ZHANG S, ZHU X, LEI Z, et al. S³FD: Single Shot Scale-Invariant Face Detector [C]//*IEEE International Conference on Computer Vision*. Venice, Italy, 2017: 192–201. DOI: 10.1109/ICCV.2017.30
- [13] GLOROT X, BENGIO Y. Understanding the Difficulty of Training Deep Feed-forward Neural Networks [C]//*13th International Conference on Artificial Intelligence and Statistics*. Sardinia, Italy, 2010: 249–256.
- [14] SUN Y, WANG X, TANG X. Deep Convolutional Network Cascade for Facial Point Detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Portland, USA, 2013: 3476–3483. DOI: 10.1109/CVPR.2013.446
- [15] ZHU X, LEI Z, LIU X, et al. Face Alignment Across Large Poses: A 3D Solution [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 2016: 146–155. DOI: 10.1109/CVPR.2016.23
- [16] FENG Z H, KITTLER J, AWAIS M, et al. Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018: 2235–2245. DOI: 10.1109/CVPR.2018.00238
- [17] ZHUANG C, ZHANG S, LEI Z, et al. FLDet: A CPU Real-Time Joint Face and Landmark Detector [C]// *IAPR International Conference on Biometrics (ICB)*. Crete, Greece, 2019
- [18] BHARADWAJ S, VATSA M, SINGH R. Can Holistic Representations be Used for Face Biometric Quality Assessment? [C]//*IEEE International Conference on Image Processing*. Melbourne, Australia, 2013: 2792–2796. DOI: 10.1109 / ICIP.2013.6738575
- [19] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002 (7): 971–987. DOI: 10.1109/TPAMI.2002.1017623
- [20] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection [C]//*International Conference on Computer Vision & Pattern Recognition (CVPR'05)*. San Diego, USA, 2005, 1: 886–893. DOI: 10.1109/CVPR.2005.177
- [21] HERNANDEZ-ORTEGA J, GALBALLY J, FIERREZ J, et al. FaceQnet: Quality Assessment for Face Recognition Based on Deep Learning [DB/OL]. (2019-04-03). <https://arxiv.org/abs/1904.01740>
- [22] NASROLLAHI K, MOESLUND T B. Face Quality Assessment System in Video Sequences [C]//*European Workshop on Biometrics and Identity Management*. Roskilde, Denmark, 2008: 10–18. DOI: 10.1007/978-3-540-89991-4_2
- [23] LIU Z, LUO P, WANG X, et al. Deep Learning Face Attributes in the Wild [C]//*IEEE International Conference on Computer Vision*. Santiago, Chile, 2015: 3730–3738. DOI: 10.1109/ICCV.2015.425
- [24] HAN H, JAIN A K, WANG F, et al. Heterogeneous Face Attribute Estimation: A Deep Multi-Task Learning Approach [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(11): 2597–2609. DOI: 10.1109/TPAMI.2017.2738004
- [25] RANJAN R, SANKARANARAYANAN S, CASTILLO C D, et al. An All-in-One Convolutional Neural Network for Face Analysis [C]//*12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. Washington DC, USA, 2017: 17–24. DOI: 10.1109/FG.2017.137
- [26] ZHANG Z, LUO P, LOY C C, et al. Learning Deep Representation for Face Alignment with Auxiliary Attributes [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(5): 918–930. DOI: 10.1109 / TPAMI.2015.2469286
- [27] BEST-ROWDEN L, JAIN A K. Learning Face Image Quality from Human Assessments [J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(12): 3064–3077. DOI: 10.1109/TIFS.2018.2799585
- [28] ZHANG L, CHU R, XIANG S, et al. Face Detection Based on Multi-Block LBP Representation [C]//*International Conference on Biometrics*. Seoul, South Korea, 2007: 11–18. DOI: 10.1007/978-3-540-74549-5_2
- [29] ZHU Q, YE H M C, CHENG K T, et al. Fast Human Detection Using a Cascade of Histograms of Oriented Gradients [C]//*IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. New York, USA, 2006, 2: 1491–1498. DOI: 10.1109/CVPR.2006.119
- [30] PHAM M T, GAO Y, HOANG V D D, et al. Fast Polygonal Integration and its Application in Extending Haar-Like Features to Improve Object Detection [C]//*IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, USA, 2010: 942–949. DOI: 10.1109/CVPR.2010.5540117
- [31] YAN J, LEI Z, WEN L, et al. The Fastest Deformable Part Model for Object Detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, USA, 2014: 2497–2504. DOI: 10.1109/CVPR.2014.320
- [32] RAMANAN D, ZHU X. Face Detection, Pose Estimation, and Landmark Localization in the Wild [C]//*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Rhode Island, USA, 2012: 2879–2886. DOI: 10.1109 / cvpr.2012.6248014
- [33] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet Classification with Deep Convolutional Neural Networks [C]//*Advances in Neural Information Processing Systems*. Lake Tahoe, USA, 2012: 1097–1105. DOI: 10.1145/3065386
- [34] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for

➔ To P. 49



RAN Centric Data Collection for New Radio

GAO Yin, LI Dapeng, HAN Jiren, LIU Zhuang, and LIU Yang

(Algorithm Department, ZTE Corporation, Shanghai 201203, China)

Abstract: Self-organizing network (SON) and minimization of driver tests (MDT) are functions designed for Long Term Evolution (LTE) system. SON is designed for network deployment by automatic configuration. MDT is designed for network performance evaluation by automatic signalling procedure. However, these functions do not support new features in new radio (NR) access technology, e.g., multiple radio access technology (RAT)-dual connectivity (MR-DC), central unit-distribute unit (CU-DU) split architecture, beam, etc. Therefore, how to support these features is a challenge for the industry. This paper provides analysis for these problems and provides the summary of SON/MDT functions progress in 3GPP. The analysis includes sub functions such as inter/intra system mobility robustness enhancement, inter/intra system mobility load balance, measurement qualities and mechanism of MDT, energy saving mechanism and procedure, RACH procedure optimization, PCI selection optimization, coverage and capacity optimization, and quality of service (QoS) monitoring mechanism. In addition, this paper also provides an initial thought on artificial intelligence (AI) algorithms applied to SON/MDT functions in NR, so called Smart Grid.

Keywords: NR; SON; MDT; data collection

DOI: 10.12142/ZTECOM.201903005

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190924.1129.004.html>, published online September 24, 2019

Manuscript received: 2019-07-31

1 Introduction

Self-organizing networks (SON) and minimization of driver tests (MDT) have been introduced in Long-Term Evolution (LTE) system [1] to support deployment of the system and performance optimization. From the field application, benefits have been identified when performing network planning and self-optimization by applying SON and MDT features. It's worthy to investigate whether to introduce these mature technologies to achieve better performance in new radio (NR) system [2], [3]. In May 2019, 3GPP completed a study project for this purpose with the name "Study on RAN - Centric Data Collection and Utilization for LTE and NR" [4], [5]. In addition, a new working project "New WID on support of SON and MDT for NR" [6] has been started from middle of 2019. In this paper, we analyze SON and MDT features and introduce the corresponding standard progress in 3GPP. In addition, this paper also provides the initial thought

on artificial intelligence (AI) algorithms applied to SON/MDT functions in NR, so called Smart Grid. At the end of the paper, conclusions and future work are proposed.

2 Overview

The first SON features, physical cell identifier (PCI) allocation and automatic neighbor relations (ANR) [1], were introduced in early version of LTE. Success of these two features encouraged further study on the topic and resulted in a Rel-9 work item that eventually enabled three SON features: mobility robustness optimization (MRO) [1], mobility load balancing (MLB) [1] and random access channel (RACH) optimization [1]. MRO and MLB have been turned out to be the key enablers of LTE and further enhanced in the following releases to match the increasing LTE complexities. Besides ANR, MRO, MLB and RACH optimization, other features enabling particular aspects of network self-optimization have been discussed

and enabled in separate studies or working projects, including MDT [7], energy saving (ES) [1], interference cancellation (ICIC, eICIC) [1], etc. These SON features are performed based on the statistic of massive data from network and user equipment (UE), which can be regarded as the pioneers of data usage in Radio access network (RAN). How to support these features have also been discussed in the related study projects. For example, introducing SON and MDT features in NR should take into account new architecture such as multiple radio access technology (RAT)-dual connectivity (MR-DC) architecture, central unit-distribute unit (CU-DU) split architecture, beam level resource management, RRC-INACTIVE state, etc [8], [9].

3 SON and MDT Features and Evolution in 3GPP

3.1 Mobility Robustness Optimization

MRO aims at detecting and enabling correction of mobility related problems which will deteriorate user experience and waste network resources. In LTE, there are three kinds of problems, i.e., connection failure due to intra-LTE or inter-RAT (radio access technology) mobility, unnecessary handover to another RAT, and inter-RAT ping-pong.

With the deployment of NR, as shown in **Fig. 1**, an evolved Node B (eNB) could be connected with Evolved Packet Core (EPC) and 5G Core (5GC) at the same time. The handover between two eNBs may be either intra-system or inter-system. Similarly, for the handover between eNB and gNodeB (gNB), it could be also either intra-system or inter-system [10].

According to Fig. 1, the handover procedures in NR include the following four types:

- Intra-NR mobility, i.e., handover between gNBs via Xn interface

- Intra-system/inter-RAT mobility, i.e. handover from gNB to eNB 2 via Xn interface
- Inter-system/inter-RAT mobility, i.e., handover from gNB to eNB 1 via S1 and NG interfaces
- Inter-system/intra-RAT mobility, i.e., handover from eNB 1 to eNB 2 via S1 and NG interfaces.

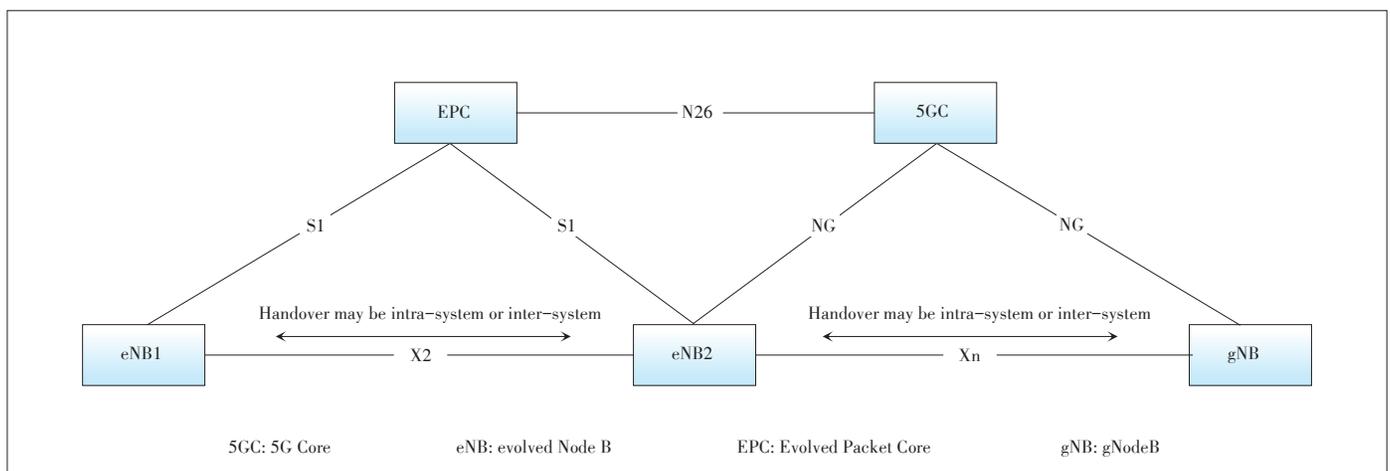
In NR, the above mobility related problems still exist. Just as shown in Fig. 1, the radio link failure report and handover report could be transferred over Xn, S1 and NG interfaces.

There are three kinds of connection failure to intra-system or inter-system mobility, i.e., too late handover, too early handover, and handover to a wrong cell. To detect these events, the RRC procedures as radio link failure (RLF) indication and handover report can be used. With the report from UE, the network side could adjust the handover report related parameters to solve the above mobility failure cases.

For detecting inter-system unnecessary handover, UE continues to evaluate the received measurement report with the coverage and quality condition during the inter-system handover preparation phase. With the measurement report, the RAN node in the source system could sent the inter-system unnecessary handover report to the RAN node in the target system.

For Inter-system ping-pong, the UE history information can be used to detect this event. With the history information, the target new radio-radio access network (NG-RAN) node indicates the occurrence of the potential ping-pong cases to the source NG-RAN node.

In addition to the legacy problems, there are some new issues that need to be studied with the new features and architecture introduced in NR. For example, the new information or measurements should be further defined and collected, such as the beam and slice level measurements. In case of MR-DC, the secondary node (SN) can be configured with Signalling Radio Bearer 3 (SRB3), and can trigger the SN change autonomously. And the SN change related failure can be considered, such as the failure due to too late SN change, failure due to too early



▲ Figure 1. Different handover scenarios in new radio.

SN change, and failure due to change to wrong SN [11]. The RLF report and handover report can be introduced between the main node (MN)/SN and SN. Furthermore, another possible enhancement on MRO is the successful handover report comprising a set of measurement results during the handover. Upon reception of a successful handover report, the receiving node is able to analyze whether its mobility configuration needs adjustment or not in order to optimize the handover procedure [12].

3.2 Mobility Load Balancing

Load balance is another key feature in SON. With this function, the network can distribute the cell load evenly among cells, transfer part of the traffic from congested cell and offload users from one cell or carrier or RAT. This kind of optimization aims to increase system capacity and improve user experience. In addition, this function is able to minimize the human intervention in network management and optimization tasks.

In NR Rel-15, the new architecture and new features are introduced, i.e., CU-DU split and CP-UP separation, the network slicing, MR-DC. All of these features should be taken into account for load sharing and load balancing optimization. Moreover, the features have impact on different interfaces, such as load management over F1 for CU-DU split, load management over E1 for CP-UP separation and load management over X2 and Xn for MR-DC. However, the slice level metrics, such as slice level radio resource utilization, should be further studied to confirm whether they are useful to adjust the cell selection/reselection and handover configuration. As for a wide-band carrier, cell capacity is defined as using the whole capacity of the wide-band carrier for each individual cell belonging to this wide-band carrier [13].

The load related information can be summarized as different level: cell level, beam level and hardware level, as shown in **Table 1** [14].

During the Rel-16 normative phase, the following metrics need to be further specified and discussed.

- Composite available capacity (CAC) per cell (downlink (DL)/uplink (UL)), reported via X2, Xn and F1 interfaces
- Cell Level Load (DL/UL/ supplementary uplink (SUL)) reported via X2, Xn, and F1 interfaces
- Transport network layer (TNL) Load reported via X2, Xn, F1, and E1 interfaces
- Beam level load indication
- Hardware load related information
- Per slice/band level reporting
- For EN-DC case, the EN-DC X2 interface shall be enhanced to support load information report from SgNB to MeNB
- Both periodic and event-triggered procedures shall be supported for load information report.

3.3 Minimization of Driver Tests

Driver test technology can check network coverage and improve the quality of service (QoS) of the mobile network. It is a

commonly used tool for operators. However, the driver test technology has many limitations, such as high cost and being difficult for many regional operators' equipment to reach. The purpose of MDT is to use mobile phones as road test equipment. Multiple use cases for NR MDT have been identified in the study stage in Rel-16. Logged MDT and immediate MDT [7] are recommended to be supported for NR MDT. In NR, these modes also support RRC_INACTIVE state. For tracing [15], the management based mechanism and signaling based mechanism in LTE [16] can be reused in NR MDT, both of which can be either logged MDT or immediate MDT. The MDT data reported from UE and the RAN may be used to monitor coverage problems and to verify QoS in the network. In Rel-16, the WLAN/Bluetooth measurement can be collected to improve positioning accuracy for indoor MDT. In addition to location and time information collection, UE orientation in a global coordinate system can be collected in MDT if the information is available. The mentioned use cases, including coverage optimization, QoS verification and WLAN/Bluetooth measurement of LTE, are taken as the baseline of NR. Take the QoS verification function as an example. When UE is in RRC_CONNECTED state, the network retrieves the UE measurements of on-going traffic and uses such information to decide which area is the hot spot and lack of capacity.

▼ **Table 1. Load related information**

| Load category | Load related information | X2 | Xn | F1 | E1 |
|-----------------|-----------------------------------|----|----|----|----|
| Cell level | CAC | ✓ | ✓ | ✓ | |
| Cell level | Cell level load (DL/UL/SUL) | ✓ | ✓ | ✓ | |
| Cell level | Cell level load per slice | | ✓ | ✓ | |
| Cell level | Cell level load per band | ✓ | ✓ | ✓ | ✓ |
| Cell level | PRB usage for SUL | ✓ | ✓ | ✓ | |
| Cell level | Number of RRC connections | ✓ | ✓ | | |
| Beam level | PRB usage for SSBs | ✓ | ✓ | ✓ | |
| Beam level | CAC per SSB beam | ✓ | ✓ | ✓ | |
| Hardware level | HW load indicator | ✓ | ✓ | ✓ | ✓ |
| Hardware level | CU-CP hardware load | ✓ | ✓ | | |
| Hardware level | CU-UP hardware load | ✓ | ✓ | | ✓ |
| Hardware level | CU hardware load (CU-CP+CU-UP) | ✓ | ✓ | | |
| Hardware level | CU-UP (hardware) load per slice | | | | ✓ |
| Hardware level | CU-UP max number of DRB supported | | | | ✓ |
| Hardware level | DU hardware load | | | ✓ | |
| Hardware level | DU DRB capacity | | | ✓ | |
| Transport level | TNL load indicator | ✓ | ✓ | ✓ | ✓ |

CAC: composite available capacity
 CU-CP: central unit control plane
 CU-UP: central unit user plane
 DL: downlink
 DRB: data radio bearer
 DU: distributed unit
 HW: hardware

PRB: Physical Resource Block
 RRC: radio resource control
 SSB: synchronization signal block
 SUL: supplementary uplink
 TNL: transport network layer
 UL: uplink

The following measurements shall be supported for MDT performance in Rel-16:

- DL signal quantity measurement results for the serving cell and for intra-frequency/inter-frequency/inter-RAT neighbor cells (logged MDT and immediate MDT)
- Power HeadRoom (PHR in case of immediate MDT)
- Received interference power measurement (In case of immediate MDT)
- Data volume measurement separately for DL and UL (In case of immediate MDT)
- Scheduled IP throughput for MDT measurement separately for DL and UL (In case of immediate MDT)
- Packet delay measurement separately for DL and UL (In case of immediate MDT)
- Packet loss rate measurement separately for DL and UL (In case of immediate MDT)
- Received Signal Strength Indicator (RSSI) measurement by UE (logged MDT and immediate MDT, for WLAN/Bluetooth measurement)
- Round Trip Time (RTT) measurement by UE (logged MDT and immediate MDT, for WLAN measurement).

In addition, MDT procedure takes LTE as baseline. In addition, further evolution on the procedure needs to be consider in Rel-16:

- Immediate MDT configurations are supported for the DC scenario.
- Logged MDT configurations can come from SN in the DC scenario.
- The existing MDT framework is the baseline for the secondary cell group (SCG) cells related MDT configuration
- The triggers for MDT measurements associated to the main cell group (MCG) and SCG are separate
- MN-SN coordination is required for MDT measurements' configuration and reporting in the DC framework
- If SRB3 is not configured, SN related measurements are transmitted to MN via SRB1/2 and then forwarded to SN
- If SRB3 is configured, MN related measurements are transmitted to MN via SRB1/2, SN related measurements are transmitted to SN via SRB3

The mechanisms of LTE RLF and Connection Establishment Failure (CEF) reporting are taken as baseline for 5G NR. To address the use case of coverage optimization in NR, RLF report and accessibility measurements are enhanced by including synchronization signal block (SSB) related information, which consists of SSB index and preambles sent for each tried SSB in the RACH information report. Additionally, the SSB index of the downlink beams of both serving cell and neighbor cells and SUL carrier information can be included in the NR CEF report.

3.4 RAN Energy Saving

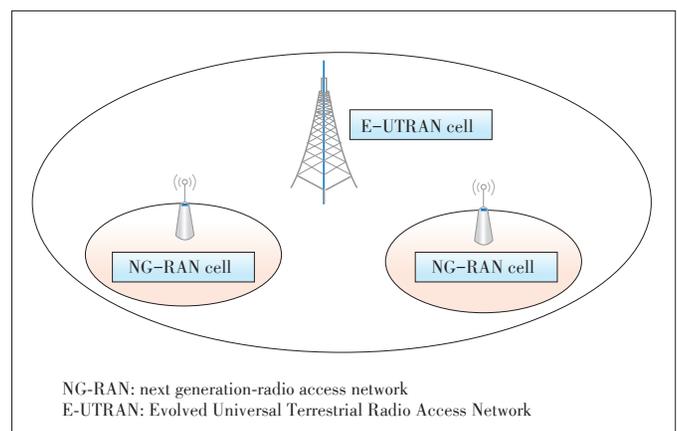
In LTE, RAN energy saving can be achieved by cell switch on/off according to cell load or by operate and maintenance

(OAM). The capacity booster cell can be switched off when its capacity is no longer needed and to be re-activated on a need basis. The eNB owning a capacity booster cell autonomously decides to switch such a cell off to lower energy consumption; the peer eNBs are informed by the eNB owning the concerned cell about the switch - off actions over the X2 interface by means of the eNB configuration update procedure; the eNBs owning non-capacity boosting cells may request a re-activation over the X2 interface by means of the cell activation procedure.

Similar as in LTE, the RAN energy saving function in NR allows, for example in a deployment where capacity boosters can be distinguished from the cells providing basic coverage, energy consumption optimization to enable a NR cell to provide additional capacity. In addition, new architectural features in NR, including CU/DU split and dual connectivity, should also be considered. The Rel-15 3GPP specifications support the following energy saving scenarios: switch - off of NR cells and their reactivation by X2AP signalling from eNB in case of EN-DC; deactivation of capacity cells and their reactivation of neighboring coverage cells by XnAP signalling in NG-RAN; activation of NR cells and their deactivation by gNB-CU via F1AP signalling in the CU-DU separation architecture.

In Rel-16, inter-system inter-RAT energy saving solution is proposed to support efficient 4G and 5G energy saving. **Fig. 2** shows one of the prominent scenarios in real deployment for the inter-system inter-RAT energy saving case from operators' point of view [17], where Evolved Universal Terrestrial Radio Access Network (E-UTRAN) cell associated eNB and new radio-radio access network (NR-RAN) cell associated gNB are connected to EPC and NGC, respectively, i.e., both eNB and gNB are deployed standalone.

As depicted in Fig. 2, the eNB cells are deployed to provide basic coverage, while the gNB cells boost the capacity. The gNB cell providing capacity booster can autonomously decide to switch off. This decision is typically based on cell load information or commanded by OAM. The eNB providing basic coverage can decide to re-activate the gNB cell on a per-demand



▲ **Figure 2.** A typical inter-system inter-RAT deployment scenario.

basis. Since there is no direct interface between eNB and gNB, energy saving for this case requires S1/NG signalling enhancements. Furthermore, it is up to operator's deployment whether inter-system inter-RAT energy saving solution needs to be applied [18], and energy saving solution should not affect user coverage and service.

Furthermore, to evaluate the energy efficiency of RAN node, European Telecommunications Standards Institute (ETSI) Energy Saving (ES) 203 228 (Environmental Engineering (EE); assessment of mobile network energy efficiency) defines the high-level energy efficiency KPI as $EE_{MN,DV} = DV_{MN}/EC_{MN}$ [5], in which the data energy efficiency of mobile network is the ratio of the performance indicator (i.e., Data Volume DV_{MN}) and the energy consumption (EC_{MN}). The data volume per site could be measured by using the Packet Data Convergence Protocol (PDCP) packet data unit Packet Data Unit (PDU). These measurements calculate the UL and DL PDCP data volumes for an aggregated base station or a gNB central unit user plane (CU-UP) in the CP-UP separation architecture. When looking at future cases, it could be plausible to assume that one CU-UP may serve gNB-DUs at different sites. In this case, the measurements could be taken as reference to define new measurements, taken at gNB-DU and measuring the PDCP data throughput incoming to and outgoing from a gNB-DU.

3.5 RACH Optimization

RACH optimization allows network to optimize and select optimized parameters for random access, which mitigates a number of access attempts, mitigates interference among RACH attempts, and/or mitigates uplink interference due to RACH. In LTE system, RACH parameter optimization such as RACH configuration and RACH preamble grouping is performed by UE reported information and by PRACH parameter exchange between eNBs.

In NR, the RACH optimization in LTE could be taken as baseline. RACH optimization is supported by UE reported information and by PRACH parameters of normal UL carrier and SUL carrier exchange over Xn between gNBs and F1 between gNB-CU and gNB-DU.

In addition, In NR system, the NG-RAN node configures the number of SSBs mapped to each PRACH occasion and the number of preambles linked to each SSB. Therefore, the content of the RACH information report possibility comprises of the following [5]:

- SSB indexes and RACH preambles sent on each tried SSB
- Frequency and beam quality of each tried SSB
- Elapsed time from the last measurement prior to the beam selection time
 - RACH preambles sent on SUL and NUL
 - Fallbacks between contention-based RACH access (CBRA) and contention-free RACH access (CFRA) contention detection indication.

The above RACH information report should also be applied

to the SN node for the MR-DC case.

Upon receiving the polling message from the NG-RAN node, UE reports RACH information within a UE information response message. The NG-RAN node takes the RACH report into account, as well as other node related information, in order to optimize the RACH parameters. The setting of RACH parameters that can be optimized includes RACH configuration (resource unit allocation), RACH preamble split, RACH back-off parameter setting, and RACH transmission power control parameter setting.

3.6 Selection of Physical Cell Identifier

PCI is distributed to each cell as a basic feature in SON. With the help of PCI, UE can distinguish different wireless signals from different cells. However, the PCI of different cells could be same because of the limited number of PCI. To minimize the effect of PCI collision, the PCI selection is introduced. In LTE, there are two kinds of PCI assignment algorithms for eNB's PCI selection, i.e., the centralized and distributed algorithms. With the centralized algorithm, the eNB selects the specific PCI value from OAM as its PCI. As for the distributed algorithm, the eNB randomly selects its PCI value from the restricted list of PCIs [19].

In NR, the mechanism of PCI selection in LTE is taken as baseline. For further discussion, the definition of information and the procedures to support PCI conflict discovery and PCI selection should be studied in Rel-16.

As the split gNB architecture, e.g. CU/DU split, is introduced in NR, the PCI assignment algorithm should be evolved based on the new network architecture. The centralized PCI selection can be divided into Option 1a and Option 1b, while the distributed PCI selection can be divided into Option 2a and Option 2b.

- Option 1a: CU detects PCI conflict and indicates to OAM via DU. OAM assigns a new PCI.
- Option 1b: CU detects PCI conflict and indicates to OAM directly. OAM assigns a new PCI.
- Option 2a: CU detects PCI conflict and indicates to DU. DU reassigns a new PCI.
- Option 2b: CU detects PCI conflict and reassigns a new PCI.

The solution to down-selection for the split gNB case will be decided in the normative phase [20].

3.7 Coverage and Capacity Optimization

In RAN, there is a trade-off between coverage and capacity optimization. The increase of the capacity will lead to the degradation of the service coverage, and vice versa. The objective of coverage and capacity optimization (CCO) is to provide the required capacity in the targeted coverage areas and to minimize the interference and maintain an acceptable quality of service in an autonomous way. CCO allows the system to periodically adapt to the changes in traffic (i.e., load and location) and the

radio environment by automatically adjusting coverage for the cells that serve a certain area for a particular traffic situation.

In NR, the beam based antenna structures with multi-dimensional RF parameters make it complex to find the mapping between network configurations with target coverage and capacity performance. By using the data collected in the RAN network, e.g. UE measurements, performance measurements, events and other monitoring information, and by taking into account beamforming and massive multiple input multiple output (MIMO)-related information, the operator could identify the coverage and capacity problems, such as coverage holes, weak pilot pollution, overshoot coverage, and DL and UL channel coverage mismatch and further perform coverage and capacity optimization. A coverage hole is an area where the signal level Signal Noise Rate (SNR) of both serving and allowed neighbor cells is below the level needed to maintain basic service, i.e., coverage of Physical Downlink Control Channel (PDCCH). Coverage holes are usually caused by physical obstructions such as new buildings and hills, by unsuitable antenna parameters, or just by inadequate RF planning. UE in a coverage hole will suffer from call drop and radio link failure. In order to identify a coverage hole, for example, RAN network needs to collect necessary measurements with location (e.g. GPS attitude with latitude). Based on these measurements, RAN node can identify the area where the downlink signal cannot fulfill the need for UE. As another example, overshoot occurs when coverage of a cell reaches far beyond what is planned. It can occur as an “island” of coverage in the interior of another cell, which may not be a direct neighbor. Reasons for overshoot may be reflections in buildings or across open water, lakes, etc. UE in this area may suffer call drops or high interference. In order to identify the overshoot issue, RAN network collects measurements with location and with cell information.

For Rel-16 discussion, CCO should focus on the coverage planning. There are two types of network coverage adjustment: long term cell RF parameter tuning and short term cell coverage switching among pre-configurations. The long term cell RF parameter tuning is usually hosted by OAM and relies on UE radio measurements (e.g., Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), call drops statistics, etc.) that are collected by MDT or performance management function. The short term cell coverage switching is implemented by the NG-RAN node. The dynamic cell coverage configuration change function will be supported by exchanging the cell coverage change information between two neighboring NG-RAN nodes, including the cell coverage state, cell deployment status indicator, and cell replacing information. In the CU-DU split architecture, the gNB-CU should provide relevant information to the gNB-DU, leaving the freedom to the gNB-DU to address such problems [21].

3.8 QoS Monitoring

The QoS requirements specified for particular services (such

as Ultra-Reliable and Low Latency Communications (URLLC) services, vertical automation communication services, and Vehicle-to-Everything (V2E) mandate QoS guarantees from the network. However, the network may not be able to always guarantee the required QoS. One reason for this shortcoming is that the latency and/or packet error rates increase due to interference in a radio cell. In this case, it is critical that the application and its server are notified in a timely manner. Hence, the 5G system should be able to support QoS monitoring/assurance for URLLC services, V2X and vertical automation. QoS monitoring function [22]–[30] also needs to consider the impact of CU-DU split architecture. There are two candidate solutions [22], [30] to achieving the requirements of URLLC services:

1) Solution 1: Using newly defined packets as monitoring packets.

As shown **Fig. 3**, a new payload type of QoS monitoring packet (QMP) in the tunnelling protocol for user plane (GTP-U) header (between the user plane function (UPF) and RAN) is introduced. Based on this new QMP, the network can calculate the round trip latency. However, this solution may introduce extra normative work on the specification.

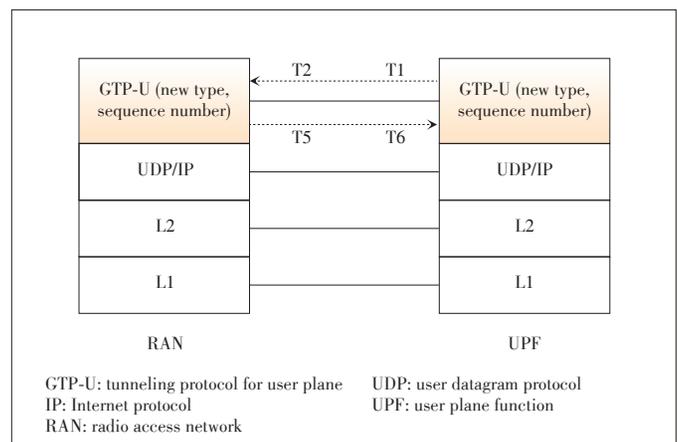
2) Solution 2: Using actual service packets as monitoring packets.

In this solution, the monitoring packet is sampled from the URLLC service packets of UE and UPF based on the measurement period, following the QoS monitoring policy. It is assumed that one-way UL packet delay plus one-way DL packet delay could be regarded equal to the round trip packet delay. Then the network calculates the path delay and provides the whole latency.

The solution to down-selection will be decided in the normative phase.

3.9 Application of Artificial Intelligence

The future 5G corresponds to a highly heterogeneous network, including multiple RATs, multiple cell layers, multiple spectrum bands, different devices and services, etc. Conse-



▲ Figure 3. Packet delay measurement.

quently, the overall RAN planning and optimization processes that constitute a key point for the success of the 5G concept will exhibit tremendous complexity. The legacy systems such as 2G, 3G, and 4G already started the path towards a higher degree of automation in the planning and optimization processes through the introduction of SON/MDT functionalities.

Nowadays, AI is the science and engineering of making machines as intelligent as humans, which can be applied to optimize communication networks. The inclusion of AI-based tools enables to shift the evolution of the SON paradigm in 5G towards a more proactive approach of exploiting the huge amount of data available and of incorporating additional dimensions coming from the characterization of end-user experience and end-user behavior [31]–[35].

The smart grid method [36] is an example to show how to achieve AI-based 5G SON. The smart grid is obtained by measuring UE’s intra-frequency measurement results and then divides the serving cell with multiple grids according to the signalling measurement difference. This is a kind of different space partition method, which is different from the traditional space partition method based on the geographical location. Some AI algorithms can be used to construct the smart grids, e.g., the clustering algorithm.

The NG-RAN node can use these AI algorithms to directly predict the RSRP values of cells on neighboring frequencies for each UE.

We use the load balance use case as an example on how to use smart grids (Fig. 4). Firstly, each grid of the serving cell will be established based on intra-frequency measurement of UE under this serving cell and the grid can be marked with RSRP of the serving cell and other 3 RSRPs of intra-frequency neighbor cells.

After the grids have been established, all the UE measurement reports will be used to setup the footprint information of

those grids, e.g., the UE list under this Grid, the inter-frequency neighbor cell list for each carrier (which can be ordered by RSRP). While other possible useful footprint information are not excluded, which depends on the RRM usage.

For load balance case, when the serving cell is overloaded, the RAN node can select some of the UEs under this serving cell, then check the corresponding selected UE via UE ID, and look up the the Footprint information of the Grid which this UE belongs to, there is no need to trigger the inter-frequency measurement towards UE, the RAN node can decide the proper inter-frequency neighbor cell to offload this UE according to the load information of the neighboring cells and the inter-frequency neighbor cell list information of each grid.

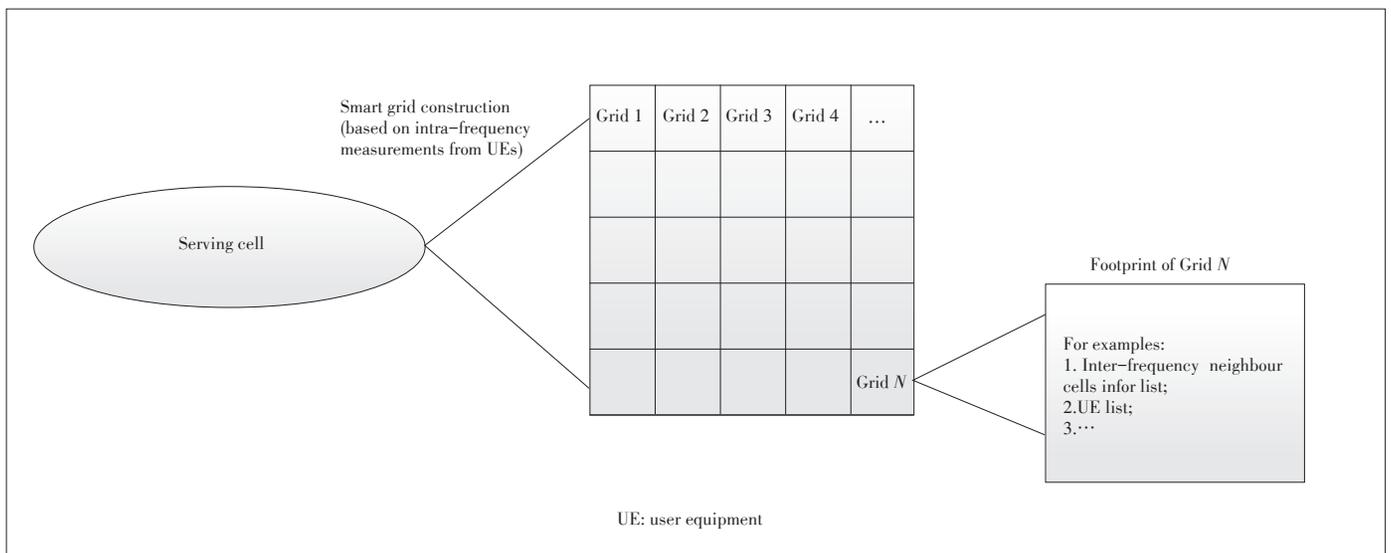
The smart grid method can bring lots of benefits to improve the network performance and user experience, e.g., accelerating load balance, improving carrier selection, improving inter-frequency handovers and increasing the user experienced data rate.

4 Conclusions

In this paper, the progress of SON and MDT evolution of NR in 3GPP is analyzed, which include MRO, load balancing and load sharing optimization, MDT, RAN energy saving, RACH optimization, PCI selection, coverage and capacity optimization, Qos monitoring, and application of AI. The function evolution has taken into account NR new features such as MR-DC architecture, CU-DU split architecture, beam level measurements, and measurement sensors. Based on the evolution of these SON and MDT features, better performance can be achieved in NR system.

References

[1] 3GPP. EUTRAN Overall Description: TS 36.300 f.5.0 [S]. 2019



▲ Figure 4. Smart grid method.

- [2] 3GPP. NR and NG-RAN Overall Description: TS 38.300 f.5.0 [S]. 2019
- [3] 3GPP. Revised WID on New Radio Access Technology: RP-172109 [S]. 2017
- [4] 3GPP. Study on RAN-Centric Data Collection and Utilization for LTE and NR: RP-182105 [S]. 2018
- [5] 3GPP. Study on RAN-Centric Data Collection and Utilization for LTE and NR: TR37.816 [S]. 2019
- [6] 3GPP. New WID on Support of SON and MDT for NR: RP-191594 [S]. 2019
- [7] 3GPP. Minimization of Drive Tests (MDT): TS 37.320 [S]. 2019
- [8] GAO Y, HAN J R, LIU Z, et al. General Architecture of Centralized Unit and Distributed Unit for New Radio [J]. ZTE Communications, 2018, 16(2): 23–31. DOI: 10.3969/j.issn.1673-5188.2018.02.005
- [9] HUANG H, LIU Y, LIU Z, et al. Mechanism of Fast Data Retransmission in CU-DU Split Architecture of 5G NR [J]. ZTE Communications, 2018, 16(3): 40–44. DOI: 10.19729/j.enki.1673-5188.2018.03.007
- [10] 3GPP. Discussion on Mobility Robustness Optimization in 5G System: R3-190289 [S]. 2019
- [11] 3GPP. SN Change Failure in Case of MR-DC: R3-192255 [S]. 2019
- [12] 3GPP. Discussion on Mobility Robustness Optimization in 5G system: R3-193186 [S]. 2019
- [13] 3GPP. TP for Load Balancing for SON in NR (Solution Description): R3-192259 [S]. 2019
- [14] 3GPP. TP for Load Management: R3-193188 [S]. 2019
- [15] 3GPP. Trace Control and Configuration Management: TS 32.422 b.7.0 [S]. 2019
- [16] 3GPP. Trace Concepts and Requirement: TS 32.421 b.7.0 [S]. 2019
- [17] 3GPP. Energy Saving in NR: R3-191829 [S]. 2019
- [18] 3GPP. Consideration on Inter-RAT Energy Saving R3-191453: [S]. 2019
- [19] 3GPP. PCI Selection Solution for NR: R3-192321 [S]. 2019
- [20] 3GPP. TP to 37.816 for PCI Selection for NR: R3-193247 [S]. 2019
- [21] 3GPP. TP to TR38.716 on Coverage and Capacity Optimization Solution for NR: R3-193246 [S]. 2019
- [22] 3GPP. Study on Enhancement of Ultra-Reliable Low-Latency Communication (URLLC) Support in the 5G Core Network: TR 23.725 [S]. 2019
- [23] 3GPP. 5G Performance Measurement: TR 28.552 [S]. 2019
- [24] 3GPP. LS on QoS Monitoring SA2: R3-192499 [S]. 2019
- [25] 3GPP. Initial Consideration on RAN-Centric Data Collection and Utilization for LTE and NR: R3-185578 [S]. 2018
- [26] 3GPP. RAN3 Impact to Support URLLC QoS Monitoring: R3-191476 [S]. 2019
- [27] 3GPP. Reply LS on Support URLLC QoS Monitoring: R3-191477 [S]. 2019
- [28] 3GPP. Discussion on QoS Monitoring for URLLC: R3-191770 [S]. 2019
- [29] 3GPP. TP for NG-U Delay Measurement for QoS Monitoring: R3-192592 [S]. 2019
- [30] 3GPP. Discussion on QoS Monitoring for URLLC: R3-193051 [S]. 2019
- [31] O'SHEA T J, HOYDIS J. An Introduction to Machine Learning Communications Systems [DB/OL]. (2017-02-02). <https://arxiv.org/abs/1702.00832v1>
- [32] FU Y, WANG S, WANG C-X, et al. Artificial Intelligence to Manage Network Traffic of 5G Wireless Networks [J]. IEEE Network, 2018, 32(6): 58–64. DOI: 10.1109/MNET.2018.1800115
- [33] LI R P, ZHAO Z F, ZHOU X, et al. Intelligent 5G: When Cellular Networks Meet Artificial Intelligence [J]. IEEE Wireless Communications, 2017, 24(5): 175–183. DOI:10.1109/MWC.2017.1600304WC
- [34] PÉREZ-ROMERO J, SALLENTO O, FERRÚS R, et al. Knowledge-Based 5G Radio Access Network Planning and Optimization [C]//International Symposium on Wireless Communication Systems (ISWCS). Poznan, Poland, 2016. DOI: 10.1109/ISWCS.2016.7600929
- [35] ZHANG Y, SHENG M, LI J D. Big Data-Driven "Artificial Intelligence" Wireless Network [J]. ZTE Technology Journal, 2018, 24(2): 2–5. DOI: 10.3969/j.issn.1009-6868.2018.02.001
- [36] 3GPP. Smart Grid Method for SON: R3-191455 [S]. 2019

Biographies

GAO Yin (gao.yin1@zte.com.cn) received the master's degree in circuit and system from Xidian University, China in 2005. Since 2005 she has been with the Research Center of ZTE Corporation and engaged in the study of 4G/5G technology. She has authored/co-authored about hundreds of proposals for 3GPP meetings and journal papers in wireless communications and has filed more than 100 patents. She has been elected as the 3GPP RAN3 Vice Chairman from August 2017.

LI Dapeng received the M.S. degree in computer science from University of Electronic Science and Technology of China in 2003. He is currently a senior researcher with ZTE Corporation and mainly focuses on the research and implementation of wireless access network system.

HAN Jiren received the master's degree in wireless communication systems from University of Sheffield, UK in 2016. He is an advanced research engineer at the Algorithm Department, ZTE Corporation. His research focuses on next generation radio access network.

LIU Zhuang received the master's degree in computer science from Xidian University, China in 2003. He is currently a senior 5G research engineer at the R&D center, ZTE Corporation. His research interests include 5G wireless communications and signal processing.

LIU Yang received the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications, China in 2016. He was a visiting scholar at Department of Electrical and Computer Engineering of North Carolina State University, USA from 2013 to 2015. He is currently an advanced 5G research engineer at the R&D center, ZTE Corporation. His research interests include statistical signal processing, information theory and performance optimization for wireless communication networks.



Reinforcement Learning from Algorithm Model to Industry Innovation: A Foundation Stone of Future Artificial Intelligence

DONG Shaokang, CHEN Jiarui, LIU Yong, BAO Tianyi, and GAO Yang

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210008, China)

DOI: 10.12142/ZTECOM.201903006

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190920.2105.006.html>, published online September 20, 2019

Manuscript received: 2019-07-10

Abstract: Reinforcement learning (RL) algorithm has been introduced for several decades, which becomes a paradigm in sequential decision-making and control. The development of reinforcement learning, especially in recent years, has enabled this algorithm to be applied in many industry fields, such as robotics, medical intelligence, and games. This paper first introduces the history and background of reinforcement learning, and then illustrates the industrial application and open source platforms. After that, the successful applications from AlphaGo to AlphaZero and future reinforcement learning technique are focused on. Finally, the artificial intelligence for complex interaction (e.g., stochastic environment, multiple players, selfish behavior, and distributes optimization) is considered and this paper concludes with the highlight and outlook of future general artificial intelligence.

Keywords: artificial intelligence; decision-making and control problems; reinforcement learning

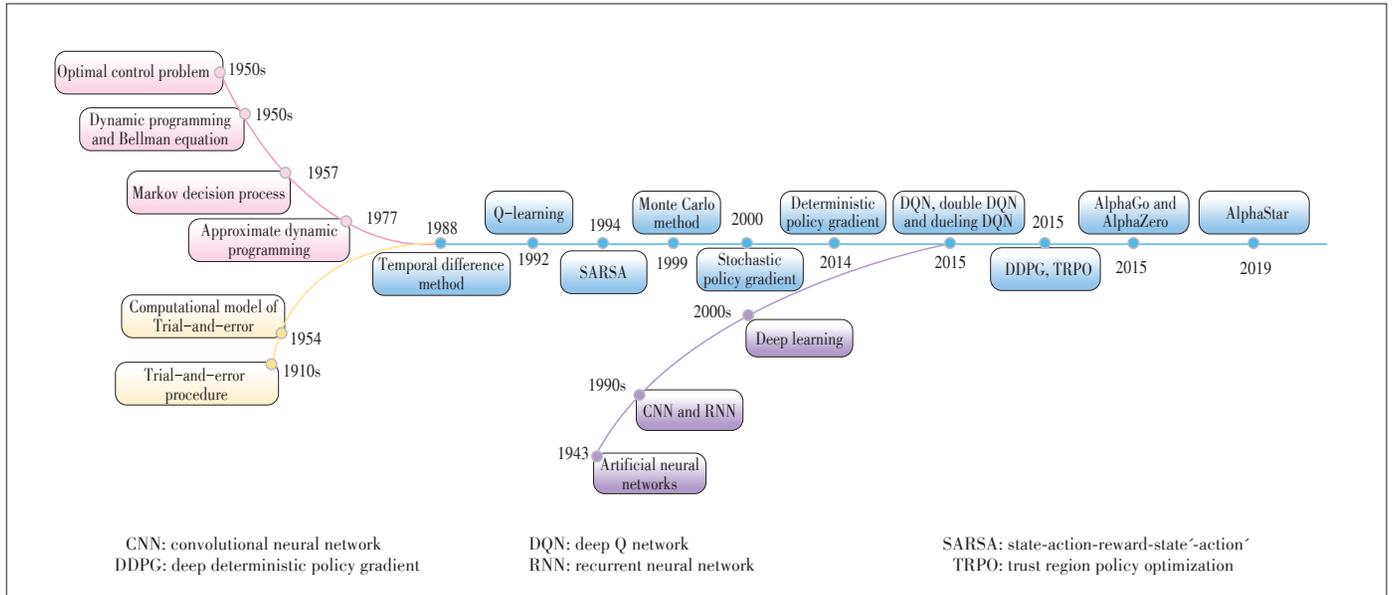
1 Introduction

Reinforcement learning (RL) originates from the trial-and-error (TE) procedure, which was first conducted in the animal learning by Thorndike in 1998 [1]. In 1954, the first concept of the computational model of TE was introduced [2]. Another term “optimal control” was first used to control a dynamical system to reach the goal of a controller. In the 1950s, Bellman introduced the dynamic programming (DP) method to solve the optimal control problems by the Bellman equation [3] and expressed this kind of control problems as Markov decision processes (MDPs) [4]. Later on, Werbos [5] introduced an approximate dynamic programming method as adaptive critic designs (ACDs) in 1977. After a decade, Sutton introduced the temporal difference (TD) method [6] in 1988, marking the point at which the TE procedure and the control problem became inter-related and then produced the field of traditional reinforcement learning. In 1992, Watkins combined the TD learning and MDPs together to solve the optimal policy, which is the well-known algorithm Q-learning [7]. In 1994, Rummery introduced the on-policy version of TD learning as the state-action-reward-state-action (SARSA) algorithm [8]. As for a general condition in MDPs, Thrun introduced a partially observable Markov decision process (POMDPs) and designed the Monte Carlo method [9] to solve it in 1999. At the same time, Sutton introduced another policy-based perspective to solve the reinforcement learning problems and proved the stochastic policy gradient

(SPG) [10] computational formula. In 2014, David Silver introduced the deterministic policy gradient (DPG) [11], the essence of which is to maximize the Q value function.

With the rapid development of deep learning, Google DeepMind utilized the deep network to approximate Q value function to address the problem of continuous state space, which was the origin of modern deep reinforcement learning and called deep Q network (DQN) [12]. After that, several improvements related to DQN emerged, such as double Q-network [13], prioritized experience replay [14] and dueling network [15]. In 2015, Google combined DQN and DPG to introduce the deep deterministic policy gradient (DDPG) [16], which extended the deep reinforcement learning (DRL) method to continuous action space control problem. Besides that, Google Gorilla introduced the asynchronous distributed reinforcement learning framework [17] in 2015. Further, John Schulman introduced Trust Region Policy Optimization (TRPO) [18] that is effective for optimizing large nonlinear policies. The development trajectory is shown in Fig. 1.

In the industry filed, Google’s DeepMind designed the AlphaGo and the AlphaZero to beat the best professional Go players KE Jie and LEE Sadol [19]. Then, more and more Atari games and MuJuCo physic problems [20] and 3D maze games [21] got great scores through deep reinforcement learning. In 2017, OpenAI announced that the reinforcement learning agent could beat the best game player on the online game Dota 2. In 2018, Tencent introduced a hierarchical macro strategy model [22] for a popular 5v5 multiplayer online battle (MOBA)



▲ Figure 1. The development trajectory of reinforcement learning technique.

game “Honor of Kings” to achieve a 48% winning rate against human player teams which are ranked top 1%. In 2019, Google’s DeepMind designed AlphaStar to beat two professional players in Warcraft II.

2 Background

2.1 Preliminary

Reinforcement learning is a field of machine learning inspired by psychology, in which the optimal control problem focuses on how to make the agent take different actions in an environment to maximize the cumulative long-term returns. In most situations, it is not possible for the agents to know the optimal action directly. In other words, the agent needs to interact with the environment to learn the optimal policy with the immediate reward feedback given and its reliability to the environment makes the study challenging and interesting. The interactions between the agent and environment are described by three essential elements: state s , action a and reward r , the relationship of which is illustrated in Fig. 2. The state of the environment in the time step t is S_t . At the time the agent takes an action A_t . Then the environment feedbacks a reward R_t and transits to another state S_{t+1} .

The reinforcement learning task as Fig. 2 can be formulated as a MDP $M = \langle S, A, T, R \rangle$, where S is the state set of environment, A is the action set, $T: S \times A \times S \rightarrow [0, 1]$ is the state transition probability and $R: S \times A \rightarrow \mathbb{R}$ is the immediate reward. A policy $\pi: S \times A \rightarrow [0, 1]$ is a mapping from states to probabilities of selecting each possible action. If the agent follows a policy π , the probability of taking action a in state s is $\pi(a|s)$. The goal of an agent is to learn the optimal policy π^* in order to get

more long-term returns $G = \sum_{t=0}^{\infty} \gamma^t r_t$, where $\gamma \in [0, 1]$ is the discounted factor. The agent becomes farsighted when γ approaches 1 while the agent becomes shortsighted when γ is close to 0.

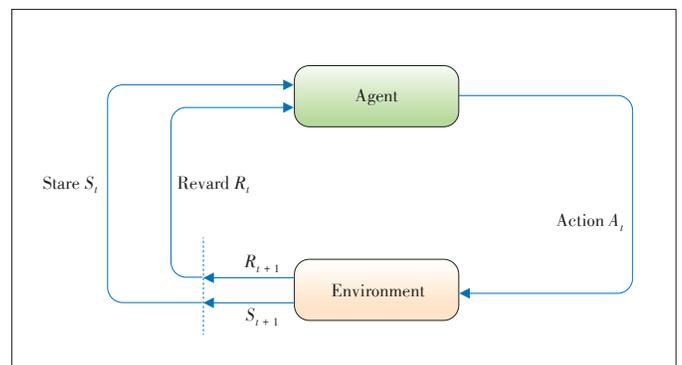
The state value function is defined to measure the utility of a state. For the MDP above, the state value function under the policy π is defined as

$$V_{\pi}(s) = \mathbb{E}_{\pi} [G_t | S_t = s] = \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t r_t | S_t = s \right], \quad (1)$$

where $\mathbb{E}_{\pi}[\cdot]$ denotes the expected value of the random variable return when the agent is under the policy π . Similarly, the state-action value function is defined as

$$Q_{\pi}(s, a) = \mathbb{E}_{\pi} [G_t | S_t = s, A_t = a] = \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t r_t | S_t = s, A_t = a \right]. \quad (2)$$

The agent always tries to find the optimal policy. A policy is better than or equal to other policies if and only if the state val-



▲ Figure 2. The interaction between the agent and environment.

ue function under this policy is greater than or equal to that under other policies. In other words, the mathematical expression can be formulated as

$$\pi \geq \pi' \Leftrightarrow V_{\pi}(s) \geq V_{\pi'}(s), \forall s. \quad (3)$$

Therefore, there always exists at least one policy that is better than or equal to all the others, which is defined as the optimal policy π^* . The corresponding state value function and state-action value function is denoted as

$$V^*(s) = \max_{\pi} V_{\pi}(s) \forall s, \quad (4)$$

$$Q^*(s, a) = \max_{\pi} Q_{\pi}(s, a) \forall (s, a). \quad (5)$$

Therefore, all kinds of reinforcement learning algorithms tend to obtain the optimal policy. Based on the prior knowledge of reward R and dynamic T , we can classify the methods into model-based and model-free. The model-based methods can have access to all the information of MDP, in which the well-known algorithm is dynamic programming (illustrated in the next section). The model-free methods contain the Monte Carlo and the temporal difference algorithms. The temporal difference algorithms can further be classified as the value-based and the policy-based. The value-based methods find the optimal policy through value function, in which the famous algorithm is SARSA and Q-learning depending on whether it is on-policy or off-policy. The other kind of policy-based algorithms converges to the optimal policy through policy gradient. Besides that, the policy-based and value-based methods can be combined into the actor-critic methods. Through the deep learning network, Q-learning and actor-critic methods can be extended to DQN and DDPG in order to address large-scale problems. The classification of different reinforcement learning

methods is shown in Fig. 3.

2.2 Reinforcement Learning Methods

In this section, we will illustrate the three basic reinforcement learning methods: dynamic programming, Monte Carlo, and temporal difference. The dynamic programming method needs information about the reward and dynamic of the environment, while the other two are model-free methods.

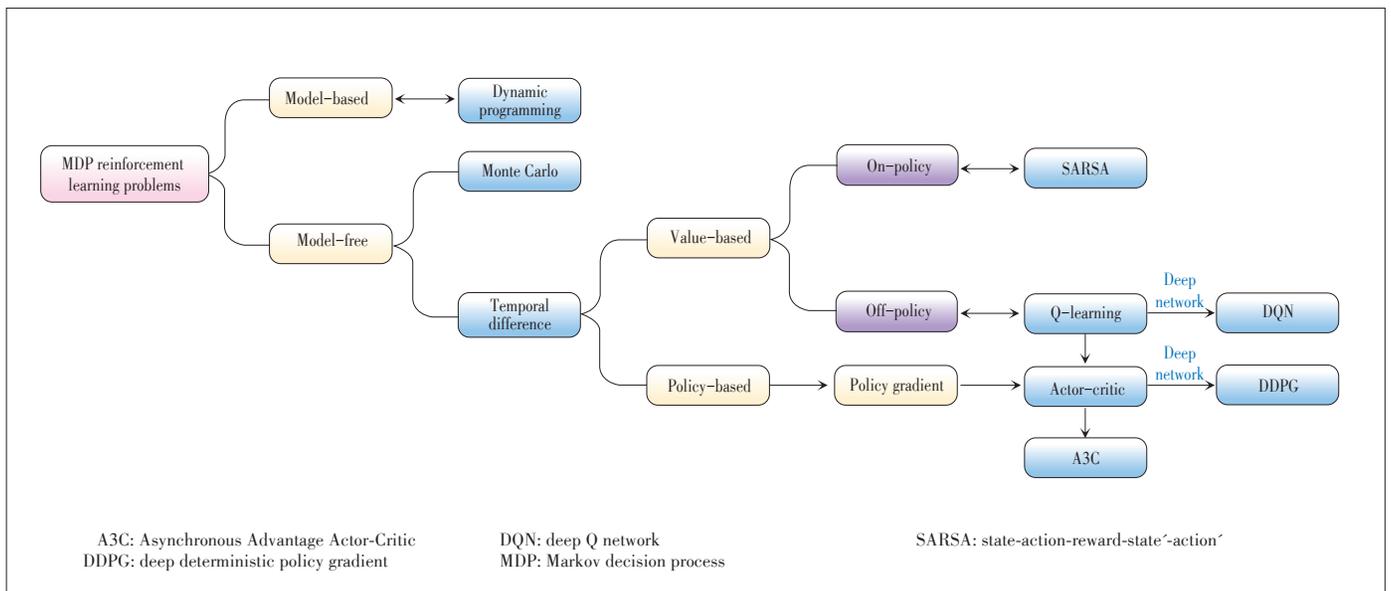
2.2.1 Dynamic Programming Method

The key idea of DP is utilizing the Bellman equation. We can utilize DP to compute the value function illustrated above. Based on the definition of the state value function, we can obtain the Bellman optimality equation as

$$\begin{aligned} V^*(s) &= \max_{a \in A(s)} Q_{\pi^*}(s, a) = \\ &= \max_{a \in A(s)} \mathbb{E}_{\pi^*} [G_t | S_t = s, A_t = a] = \\ &= \max_{a \in A(s)} \mathbb{E}_{\pi^*} [R_{t+1} + \gamma G_{t+1} | S_t = s, A_t = a] = \\ &= \max_{a \in A(s)} \mathbb{E}_{\pi^*} [R_{t+1} + \gamma V^*(S_{t+1}) | S_t = s, A_t = a] = \\ &= \max_{a \in A(s)} \sum_{s', r} p(s', r | s, a) [r + \gamma V^*(s')], \end{aligned} \quad (6)$$

where $p(s', r | s, a)$ is the state transition probability under the reward r , and $\gamma \in [0, 1]$ is the discounted factor. Obviously, we can obtain that the state transition probability $p(s' | s, a) = \sum_{r \in R} p(s', r | s, a)$. Then we will illustrate how to use DP to converge to the optimal policy.

In the first step, we should compute the state value function V_{π} for any arbitrary policy π , which is called the policy evaluation. Similar to Eq. (6), we can obtain the iterative equation as



▲ Figure 3. The classification of different reinforcement learning methods.

the updated rule

$$V_{k+1}(s) = \mathbb{E}_\pi [R_{t+1} + \gamma V_k(S_{t+1}) | S_t = s] = \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a) [r + \gamma V_k(s')]. \quad (7)$$

In this case, we can rigorously prove that $V_k = V_\pi$ is a fixed point for this iterative equation. Therefore, the updated equation can converge to the state value function under any arbitrary policy π .

Then, we need to compute the optimal policy instead of an arbitrary policy, called the policy improvement. Similarly, we can take the action as a greedy way to obtain a new policy π' that is better than or equal to the current policy π as

$$\begin{aligned} \pi'(s) &= \arg \max_a Q_\pi(s, a) = \\ & \arg \max_a \mathbb{E}_\pi [R_{t+1} + \gamma V_\pi(S_{t+1}) | S_t = s, A_t = a] = \\ & \arg \max_a \sum_{s',r} p(s',r|s,a) [r + \gamma V_\pi(s')]. \end{aligned} \quad (8)$$

It can be proved that the greedy policy π' is better than the current policy π as $V_{\pi'}(s) \leq V_\pi(s), \forall s$. So far, we can utilize the policy evaluation to get the stable value function and then policy improvement to obtain a better policy until the state value function and policy are both stable, which is called the policy iteration. Then, the current policy is obviously the optimal policy and the value function is the optimal value function.

2.2.2 Monte Carlo Method

Monte Carlo method is a kind of model-free method, so we cannot utilize the reward and state transition probability to compute the state value function. Instead, the state value function is estimated by repeatedly generating episodes and averaging the returns under policy π as

$$V_\pi^{MC}(s) = \mathbb{E}_\pi [G_t | S_t = s]. \quad (9)$$

Although we do not utilize the information related to the model, it must satisfy the assumption that for any states and actions, the sampling times must approach infinity. Only under that condition, we can ensure that the Monte Carlo value function V_π^{MC} converges to the true state value function. Then, we need to modify the greedy method as ϵ -greedy to take the policy improvement in order to guarantee the assumption:

$$\pi'(als) = \begin{cases} 1 - \epsilon + \epsilon / |A(s)|, & \text{if } a = \arg \max_k Q_\pi(s, a_k) \\ \epsilon / |A(s)|, & \text{if } a \neq \arg \max_k Q_\pi(s, a_k) \end{cases}, \quad (10)$$

where $|A(s)|$ denotes the number of actions taken in the state s . In this case, we can trade off the exploration-exploitation dilemma, but the policy of evaluation and exploration is different, called the off-policy. The fact is that the off-policy method is sometimes unstable in continuous state - space problems. Therefore, we usually design the on-policy method through im-

portance sampling.

2.2.3 Temporal Difference Value-Based Method

Similar to the Monte Carlo method, the temporal difference method is also a model-free method. However, what makes it different from the Monte Carlo method is that it needs bootstrapping, which means the estimated value function V_{k+1} is updated based on the last estimated value function V_k . The simplest temporal difference update is that

$$V(s_t) \leftarrow V(s_t) + \alpha [R_{t+1} + \gamma V(s_{t+1}) - V(s_t)], \quad (11)$$

where $\alpha \in (0, 1)$ is the learning rate. As the category in Monte Carlo, the temporal difference method is also divided into on-policy and off-policy, which are the well-known algorithms SARSA and Q-learning.

In SARSA, we update the state-action value function for the current policy and take the action based on the state-action value function. The algorithm can converge to an optimal policy with probability 1 as long as all states and actions are visited an infinite number of times. The update equation is

$$Q(S, A) \leftarrow Q(S, A) + \alpha [R + \gamma Q(S', A') - Q(S, A)], \quad (12)$$

where S' and A' denote the next state and next action. It is obvious that the policy of evaluation and exploration is the same, so SARSA belongs to the on-policy methods.

Q-learning is an off-policy method because the evaluation of state-action value function is maximizing over all those actions possible in the next state instead of the current Q function. The update equation is

$$Q(S, A) \leftarrow Q(S, A) + \alpha [R + \gamma \max_a Q(S', a) - Q(S, A)], \quad (13)$$

where S' denotes the next state and next action. The convergence condition of Q-learning is the same as SARSA.

2.2.4 Temporal Difference Policy-Based Method

Although the value-based methods have made great achievement in many fields, such as robotics and video games, they cannot be applied in the situation of continuous state and action space. In this situation, there exists a new policy-based method, in which the policy is explicitly represented by the function approximator and updated based on the gradient of expected reward [10]. Let $J(\pi_0)$ denote the performance objective as an expectation as

$$\begin{aligned} J(\pi_0) &= \int_S \rho^\pi(s) \int_A \pi_0(als) r(s, a) dads = \\ & \mathbb{E}_{s \sim \rho^\pi, a \sim \pi} [r(s, a)], \end{aligned} \quad (14)$$

where $\rho^\pi(s)$ is the discounted state distribution as $\rho^\pi(s) = \int_s \sum_{t=1}^{\infty} \gamma^{t-1} p_1(s) p(s \rightarrow s', t, \pi) ds$. Policy gradient algorithms search for a local maximum in $J(\pi_0)$ by ascending the gradient of the policy w.r.t. parameter θ as $\Delta \theta = \alpha \nabla_\theta J(\pi_\theta)$.

Sutton has proved the update rule of stochastic policy gradient as

$$\begin{aligned} \nabla_{\theta} J(\pi_{\theta}) &= \int_S \rho^{\pi}(s) \int_A \nabla_{\theta} \pi_{\theta}(a|s) Q^{\pi}(s, a) da ds = \\ &= \mathbb{E}_{s \sim \rho^{\pi}, a \sim \pi_{\theta}} [\nabla_{\theta} \log \pi_{\theta}(a|s) Q^{\pi}(s, a)]. \end{aligned} \quad (15)$$

The stochastic policy gradient is surprisingly simple and does not depend on the gradient of the state distribution [10]. In 2014, David Silver [11] introduced DPG as

$$\begin{aligned} \nabla_{\theta} J(\pi_{\theta}) &= \int_S \rho^{\pi}(s) \nabla_{\theta} \pi_{\theta}(s) \nabla_a Q^{\pi}(s, a) |_{a = \pi_{\theta}(s)} ds = \\ &= \mathbb{E}_{s \sim \rho^{\pi}} [\nabla_{\theta} \pi_{\theta}(s) \nabla_a Q^{\pi}(s, a) |_{a = \pi_{\theta}(s)}]. \end{aligned} \quad (16)$$

The difference of DPG and SPG is that the essence of deterministic policy gradient is to maximize the Q value function, which is similar to the value-based method.

2.2.5 Temporal Difference Actor-Critic Method

Actor-critic method combines the value-based method and policy-based method, in which the actor network (policy network) outputs the action selection and critic network (value network) evaluates the performance of action. The well-known DDPG algorithm utilizes the actor-critic idea to address the continuous control problems [16]. In the part of actor, the parameter θ^{ρ} is updated to maximize Q value and parameter θ^{μ} is updated to make actor more likely to select this action. Therefore, the updated rule is as

$$\nabla_{\theta^{\mu}} J \approx \frac{1}{N} \sum_i \nabla_a Q^{\pi}(s, a | \theta^{\rho}) \Big|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s, \theta^{\mu}) \Big|_{s_i}. \quad (17)$$

As for the part of critic, it is similar to the DQN (illustrated in the next section). The parameter is updated to minimize the TD loss

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta^{\rho}))^2, \quad (18)$$

where $y_i = r_i + \gamma Q(s_{i+1}, \mu(s_{i+1} | \theta^{\rho}))$.

Another benchmarking work Asynchronous Advantage Actor-Critic (A3C) [17] is desired to solve the convergence problem of actor-critic method. It creates multiple parallel environments, in which there exist many thread-specific agents updating the global parameters together. The agents in parallel do not interfere with each other, and the parameter update of the main structure is disturbed by the discontinuity of the update submitted by the thread-specific agents, so the correlation of the update is reduced to accelerate the convergence rate. Different from DDPG, A3C updates the policy gradient through advantage as

$$\nabla_{\theta} J = \nabla_{\theta} \log \pi(a_i | s_i; \theta) A(s_i, a_i; \theta_v), \quad (19)$$

where the advantage is computed by N -steps TD error as $A(s_i, a_i; \theta_v) = \sum_{k=0}^{N-1} \gamma^k r_{t+k} + \gamma^N V(s_{t+k}; \theta_v) - V(s_i; \theta_v)$ and the parameter θ_v represents for each local thread-specific agent.

Because the learning rate of policy gradient is dependent on the specific environment and not easily determined, PPO [23],

[24] utilizes the ratio of new policy and old policy in order to restrict the update. The critic network is also similar to the AC3 method to minimize TD error, while the actor network updates the parameter by KL penalty as

$$\nabla_{\theta} J = \sum_{t=1}^N \frac{\pi_{\theta}(a_t | s_t)}{\pi_{old}(a_t | s_t)} \hat{A}_t - \lambda KL[\pi_{old} | \pi_{\theta}], \quad (20)$$

where \hat{A}_t is also the estimated advantages as A3C.

After illustrating these reinforcement learning algorithms, we summarize the characteristics of the algorithms and differentiate them based on the policy, bootstrapping and model property in **Table 1**.

2.3 Deep Reinforcement Learning

Deep learning can help reinforcement learning address the high-dimension problem when the states and actions are continuous variables in the environment. In 2015, Mnih [12] trains a deep neural network to develop a novel artificial agent to interact with the environment, as shown in **Fig. 4**. In other words, DQN utilizes the neural networks to interpret the complex state representation instead of the traditional MDP. Of course, the goal of network to optimize also satisfies the Bellman equation, of which the loss function is defined as

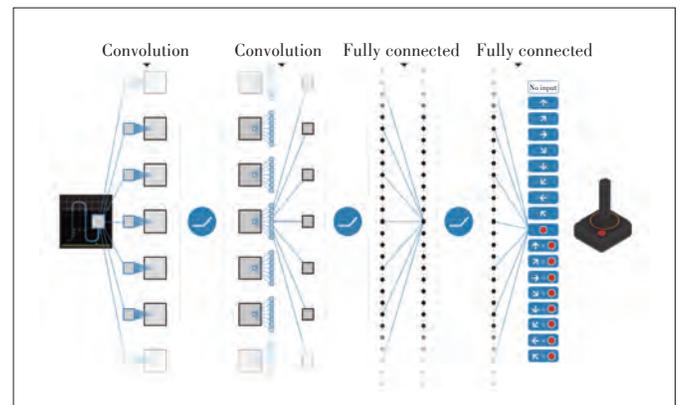
$$L(\theta) = \left(r + \gamma \max_{a'} \hat{Q}(s', a'; \theta^-) - Q(s, a; \theta) \right)^2, \quad (21)$$

where \hat{Q} is the target value function and Q is the evaluation val-

▼ **Table 1. Comparisons between reinforcement learning (RL) methods**

| Category | DP | On-policy MC | Off-policy MC | SARSA | Q-learning |
|---------------|----|--------------|---------------|-------|------------|
| Model-based | ✓ | | | | |
| Model-free | | ✓ | ✓ | ✓ | ✓ |
| On-policy | | ✓ | | ✓ | |
| Off-policy | | | ✓ | | ✓ |
| Bootstrapping | ✓ | | | ✓ | ✓ |

DP: dynamic programming
SARSA: state-action-reward-state-action'
MC: Monte Carlo



▲ **Figure 4. Neural network of the deep Q-learning [12].**

ue function. The purpose of designing the two value function is to eliminate the correlation between the samples due to the unstable property of neural networks. Besides that, generated samples from the environment are stored in the experience replay memory, the data of which are randomly fed to train the network. Note that, only the evaluation network with the parameter β is trained, the target network with the parameter β' is just a duplicate of β but with the delayed update as $\beta' \leftarrow \beta$ for every N steps.

With the development of different scenario requirements, DQN also has some variants. The first algorithm is called double deep Q-network (DDQN) [13], the idea of which is to separate the optimal action with the evaluation step in order to avoid the overestimation of Q-value function. Therefore, the loss function is modified as

$$L(\theta) = \left(r + \gamma \hat{Q} \left(s', \max_a \hat{Q}(s', a'; \theta^-); \theta^- \right) - Q(s, a; \theta) \right)^2. \quad (22)$$

Besides that, Tom Schaul [14] proposes a prioritized experience replay method to reinforce the goal-related samples with greater TD error. The corresponding sampling probability is the normalization of the absolute TD error as $p_i = |\delta_i| = |r + \gamma \max_a \hat{Q}(s', a'; \theta^-) - Q(s, a; \theta)|$. In addition, Ziyu Wang [15] introduces a dueling network to represent the state value function and state-dependent action advantage function. Then the Q-value function is computed as

$$Q(s, a; \theta, \alpha, \beta) = V(s; \theta, \beta) + \left(A(s, a; \theta, \alpha) - \frac{1}{|A|} \sum_a A(s, a'; \theta, \alpha) \right), \quad (23)$$

where θ denotes the parameter of the convolutional layers, while α and β are the parameters of the two streams of fully-connected layers.

Therefore, it is obvious that the development of reinforcement learning is tracking the increasing demands when applied to the real industry field. In the following sections, we will illustrate the application and open source platforms about reinforcement learning.

3 Reinforcement Learning Industry and Application

3.1 Famous Pioneering Companies

On top of all the research companies, DeepMind is the first that should be mentioned, which was founded in London in 2010 and had been acquired by Google in 2014. The core research filed is about health artificial intelligence (AI), the energy efficiency of Google's data centers and ethics & society. The significant milestones for reinforcement learning research are introducing the AlphaGo, AlphaZero, AlphaStar and Reinforcement Learning with TensorFlow (TFRL) library. The library collects key algorithmic components that we have used internally for a large number of our most successful agents

such as DQN, DDPG, and the Importance of Weighted Actor Learner Architecture. Another company OSARO was founded in 2015 in San Francisco, which is devoted to developing products based on proprietary deep reinforcement learning technology in order to process large amounts of unstructured data and efficiently learn complex control tasks. The company VocallIQ was formed in March 2011 and had been acquired by Apple to exploit technology developed by the Spoken Dialogue Systems Group at University of Cambridge. VocallIQ builds a platform for voice interfaces, making it easy for everybody to voice-enable their devices and apps. AI is extremely difficult to be applied in business because the systems cannot start learning on their own. The company CogitAI is solving this problem by self-learning AI platform that learns through its interaction with the real world, which will make the perception and behavior of daily routine become more intelligent and experienced. The Japanese company Preferred Networks was founded in 2014 and then invested by Toyota, with the main products related to automatic driving, medical health and AI in the manufacturing industry. At last, the World's First Deep Reinforcement Learning Platform for the Enterprise Bonsai was acquired by Microsoft in 2018, which focuses on optimizing the Azure public cloud service platform.

3.2 Representative Applications

Nowadays, the reinforcement learning techniques have been applied in many industries including robotics, automatic driving, natural language processing, computer vision, finance, healthcare, smart grid, intelligent transportation systems, and so on. In robotics, deep reinforcement learning can help to address the high-dimensional control problems, which has become not only a major research topic but also an efficient method to build products for industrial robots. Usually, the RL can train to control robotic arms and legs to accomplish the carrying or motion tasks. A famous experiment about the autonomous helicopter is designed by imitation learning, a branch of reinforcement learning. With a higher level goal, reinforcement learning can also be applied in automatic driving. Only relying on clever models with high-quality training data and carefully thinking out objective functions, the agent can learn a more comprehensive set of skills for driving. Up to now, the Tesla automatic driving cars can independently achieve the cruise, steering, line-changing and parking tasks, which have been produced and put into application in some areas.

As in natural language processing, the researchers face a sequential prediction problem where future observations (visited states) depend on previous actions. This is challenging because it violates the common independently and identically distributed (i.i.d.) assumptions made in statistical learning. For example, naively training the agent on the gold labels alone would unrealistically teach the agent to make decisions under the assumption that all previous decisions were correct, potentially causing it to over-rely on information from past actions [25].

Especially for the multiple rounds of question & answer and semantic analysis system, reinforcement learning methods can achieve great performance because the context semantic information can be considered. Therefore, it is clearly seen that the RL model is giving more human-like answers than tradition systems. Yangfeng Ji [26] proposes an adversarial learning procedure to where they train a generative model to produce response sequences and a discriminator to distinguish between the human-generated dialogues and the machine-generated ones. The outputs from the discriminator are then used as rewards for the generator, pushing the system to generate human-like dialogues.

There is a huge potential for reinforcement learning in finance. Apart from just playing games, it seems reasonable for such a framework to have meaningful applications in finance and trading due to the following reasons: the size of states in the finance environment may be large and continuous. Actions may have long-term consequences, which means the agent should have further insight. Your trader actions can affect the current market environments. Financial management is the continuous investment of new funds into different financial products with the goal of achieving maximum returns. Therefore, when facing these challenges, the reinforcement learning framework is an effective tool to provide smarter solutions for fund management. Another application is in the search and recommender systems, which usually utilize the multi-arm bandit (MAB, a special reinforcement learning method) algorithm to address the exploration-exploitation dilemma. Then Alibaba models the browsing and purchasing behavior of users as MDP in the search scene [27]. What takes Taobao search to the next level is the 20% improvement of the performance in the real-time learning and decision computing system using reinforcement learning. In the recommended scenario, Alibaba applies deep reinforcement learning and adaptive online learning to analyze the users' behaviors and the commodity characteristics, which helps users find devoted goods quickly and improves the effect index by 10%–20%.

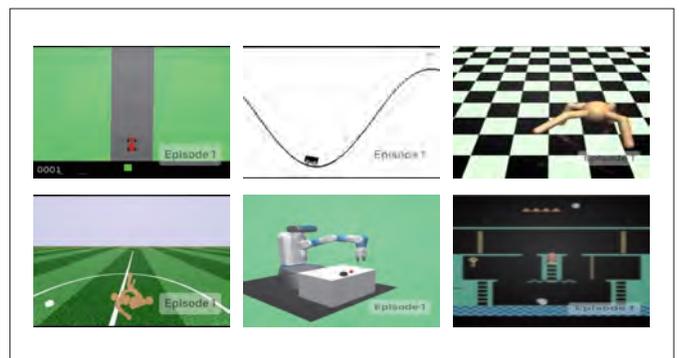
Finally, a closely related application to everyday life is healthcare. Nearly all large companies in the healthcare field have already begun to use deep reinforcement learning technology in practice. Researchers have also found the potential of reinforcement learning in the medical image screening for diagnosis detection, medical chatbots, and clinical decision-making simulation. For example, in robot-assisted surgery, doctors usually guide the robot to operate instruments by remote control. With the applications of computer vision models to observe the surgical environment and reinforcement learning methods to learn the surgeon's movements, the robustness and adaptability of robot-assisted surgery are effectively improved. Additionally, reinforcement learning can also enhance the efficiency of ICU rescue. When rescuing and caring for intensive care patients, doctors are often caught in a dilemma: blood test indicators can provide critical information for patient rescue,

but too frequent tests may aggravate the risk of illness and increase the cost of treatment. A research group from Princeton University finds that the reinforcement learning algorithm mentioned above can improve the ability of machine learning to select clinical test items and get the most rewards by optimizing the order of clinical test. Therefore, reinforcement learning is a practical tool that can help companies optimize their service provision, improve the standard of care, generate more revenue, and decrease risk.

4 Open Source Platform

Reinforcement learning has become a hot topic all over the world, but where can we learn and practice the RL algorithms? The good news is that many companies have opened up resources and platforms. The famous OpenAI Gym is a toolkit for developing and comparing reinforcement learning algorithms, which includes a series of continuously growing and improving environments (e.g. simulated robots and Atari). It helps teach agents to do everything, such as walking and playing Ping Pong. In addition, Gym is compatible with other numerical calculation libraries, such as tensorflow and theano library. The environments include different scenes, such as Atari, Box2D, Classic control, MuJoCo, Roboschool, Robotics, and Toy text. **Fig. 5** shows several representative games in Gym.

Arcade Learning Environment (ALE) is an object-oriented open source platform, which provides an interface to hundreds of Atari 2600 game environments, each one different, interesting, and designed to be a challenge for human players [28]. ALE is easy to add any game agent based on the object-oriented framework. In addition, the simulation core is decoupled from the game screen rendering and sound module to improve the simulation efficiency. It supports multiple operating systems (OS X, Linux) and can be cross-language developed. Finally, the platform also has some rules to evaluate different algorithms. Another platform VizDoom is an artificial intelligence research platform based on the 3D shooting game Doom, which is used for reinforcement learning from original visual



▲ Figure 5. The Gym games (from left to right and top to bottom: CarRacing, Mountainair, Ant, RoboschoolHumanoidFlagrunHarder, FetchPickAndPlace and MontezumaRevenge).

information and mainly used to study machine vision learning, especially deep reinforcement learning. And the Google DeepMind Lab is also a first-person perspective 3D gaming platform designed specifically for research into general-purpose artificial intelligence and machine learning systems. It can be used to study how agents learn to perform complex tasks in a large, partially visual and visually diverse environment. The other platform The Open Racing Car Simulator (TORCS) is an open-source 3D car racing simulator available for many systems. TORCS is designed to enable pre-programmed AI drivers to race against one another while allowing the user to control a vehicle using either a keyboard, mouse or wheel input.

The other game platform StarCraft II Learning Environment (SC2LE) is a StarCraft II artificial intelligence research environment jointly released by DeepMind and Blizzard Entertainment, aiming to further promote the development of AI by developing learning systems capable of solving complex problems. SC2LE can also choose the mini-game maps, which can first evaluate the algorithm in a simpler environment. The configurations can set player and time limits, whether to use the game outcome or curriculum score and a handful of other things. It allows Python codes to communicate with the API and this client will be the main focus of this guide, from which we can get spatial features from the map and call actions that mimic how humans interact with the game.

5 An Insight: From AlphaGo to AlphaZero

Several years ago, the milestone in the field of artificial intelligence was the appearance of AlphaGo [19] that beat the professional player Lee Sedol. The challenge of Go for classic methods is its enormous search space and the difficulty in evaluating board positions and moves. The Go game contains approximately b^d ($b \approx 250, d \approx 150$) possible move sequences, which results in the failure of brute search. Therefore, the AlphaGo combines the supervised learning and reinforcement learning with Monte Carlo tree search (MCTS) to reduce the

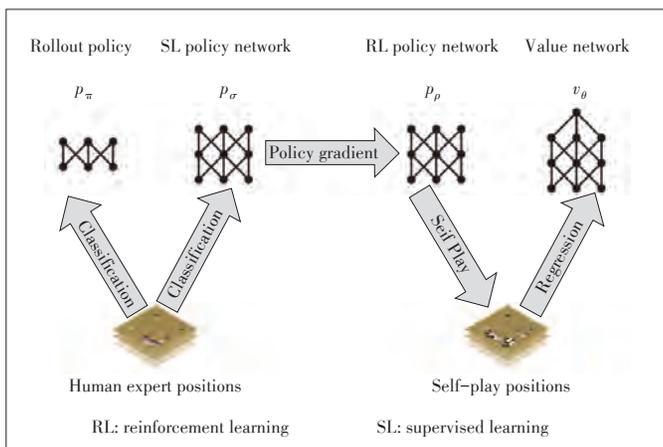
space. There are four networks in total: rollout policy network, supervised learning (SL) policy network, RL policy network, and value network. The relationship between these networks is shown in Fig. 6, in which the rollout and SL policy network are trained by the expert data. The target of them is to predict the most likely moves of human by learning the conditional probability $p(a|s)$, where s denotes the state of the chessboard and a denotes the move position. The difference of them is that the SL policy network is then used to train the RL policy network, while the rollout policy network sacrifices the accuracy for quick simulation in order to provide the possibility of MCTS. The training of SL policy network is based on CNN, where the inputs are the chessboard features ($19 \times 19 \times 48$) and outputs are the probability of different positions. The training process cost about 3 weeks by 50 Graphics Processing Units (GPUs) through the expert game data in Kiseido Go Server (KGS) platform.

Another important network is the RL policy network, the basic idea of which is using self-play to formulate the new experience that human cannot image. The detailed steps are as follows:

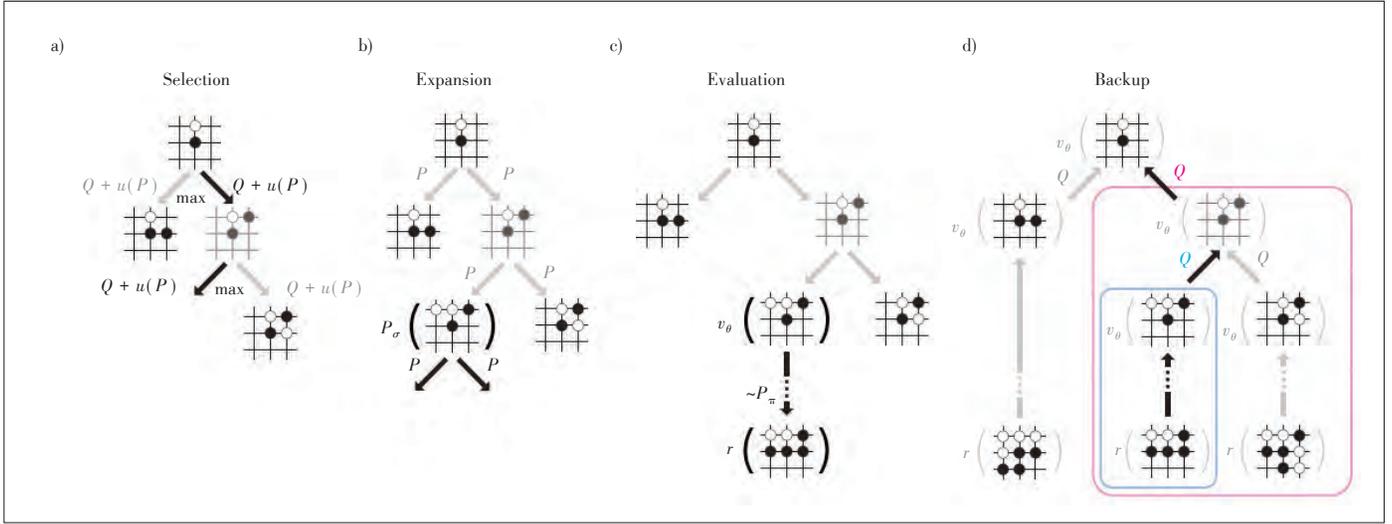
- 1) Set the structure of RL policy network identical to the SL policy network and its weights are initialized to the same values.
- 2) Put this network to the opponent pool.
- 3) Make the current RL policy network to play against a random version in the opponent pool and then utilize the reinforcement algorithm to maximize the expected outcome by updating the parameters as $\Delta \rho \propto \partial \log P_{\rho}(a_t|s_t) z_t / \partial \rho$, where ρ denotes the parameters in the RL policy network and $z_t = \pm r(s_t)$ is the terminal reward at the end of the game as +1 for winner else -1 for the loser.
- 4) Put the current network into the opponent pool after every 500 iterations and return to step 2 recursively.

Through the RL training process, the winning rate can achieve 80% against the SL policy network.

The last value network is used to compute the optimal value functions and predict the winner of the games. This neural network has a similar architecture to the policy network and outputs a certain action. The value function is regressed by the parameter θ , which minimizes the mean squared error (MSE) between the predicted values. Finally, the AlphaGo takes actions by the MCTS algorithm (Fig. 7). When MCTS is faced with the current state, AlphaGo will simulate some kinds of strategy to play a few steps or all the way to the end by itself. More and more simulation can make the AlphaGo deduce more accurately. Because of the policy network, the search space has been narrowed down to the high-probability actions. The edge of tree stores the basic information such as value function $Q(s, a)$, visited times $N(s, a)$, and probability $P(s, a)$. The first step of MCTS is to select the appropriate actions as $a_t = \arg \max_a (Q(s_t, a) + u(s_t, a))$. What is different from the tradi-



▲ Figure 6. The different networks of AlphaGo [19].



▲ Figure 7. The process of MCTS [19].

tional RL methods is the additive term $u(s, a) \propto P(s, a) / (1 + N(s, a))$ that is proportional to the prior probability $P(s, a)$ and inversely proportional to the visited times $1 + N(s, a)$. When the leaf node in L depth is visited at time step L , the expansion step should be conducted and the new leaf node keeps the same probability of its parent node. The third step evaluation is to compute the reward of the leaf node; therefore, we are able to use the value network to compute as well as the rollout network to run to the end and obtain a winning or losing the reward. Finally, we can back up the Q value through the sub-tree as

$$N(s, a) = \sum_{i=1}^n 1(s, a, i), \quad (24)$$

$$Q(s, a) = \frac{1}{N(s, a)} \sum_{i=1}^n 1(s, a, i) V(s_i^i), \quad (25)$$

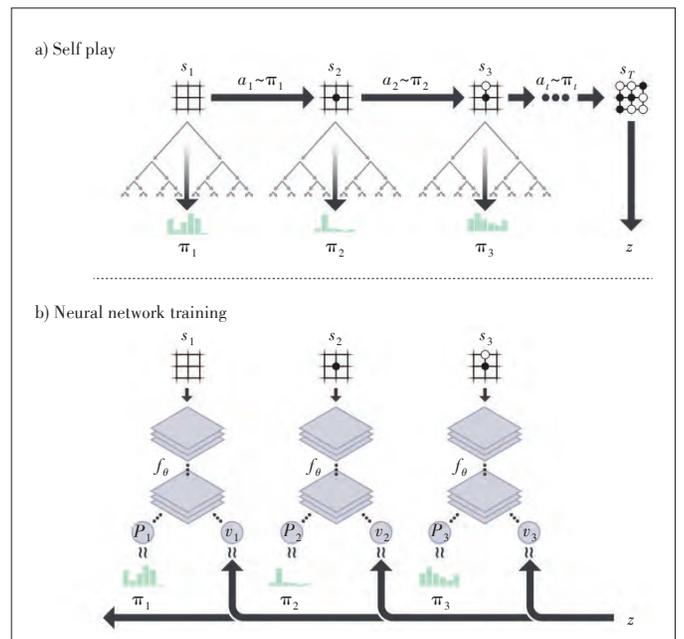
where s_i^i is the leaf node in the i -th simulation and $1(s, a, i)$ indicates whether an edge (s, a) was traversed. Through the backup process, we can recursively take the appropriate action in the selection step.

In conclusion, AlphaGo needs the human experience to predict the position in SL policy network. Combined with the RL policy network, AlphaGo can generate new experience through self-play, which can achieve a better win rate. Besides that, the MCTS and value work can be used to get a more accurate result through the search process. However, the drawback of AlphaGo is that it needs the supervision of human game data, which is expensive or even unavailable. Moreover, human data may also cause an inaccurate bias of the deep network. Another shortcoming is that AlphaGo uses four networks, which is very complex and costs more times to converge. Therefore, AlphaZero [28] is introduced to address these problems.

AlphaZero is learned from a random game, using self-play

reinforcing learning, without human data or supervision. The policy and value networks are combined with only a single network, the input of which is the direct position of black and white instead of the manually designed features. The detailed process is shown in Fig. 8. A self-play game has many different states $S_1, S_2, S_3, \dots, S_T$. In each state, the neural network is training to output the action probability vector P_i and the predicted value v_i . Its goal is to minimize the difference between policy vector P_i and the MCTS probabilities π_i and also that of the predicted value v_i and the end reward z . The loss function is

$$L(\theta) = (z - v)^2 - \pi^T \log P + c \|\theta\|^2. \quad (26)$$



▲ Figure 8. The network of AlphaZero [28].

This neural network improves the efficiency of Monte Carlo tree search, which enhances the accuracy of move selection and stronger self-play in the next iteration. The final performance can achieve 100–0 against AlphaGo.

6 Future Techniques

Reinforcement learning has become more and more important, but there are also several problems which need to be addressed in the future. As we all know, deep reinforcement learning has continuous large state and action spaces, resulting in the long endless iteration times to learn a stable converged policy. Moreover, the real-world environment is more complex and nonstationary, because there are many other agents that also need to learn an optimal policy in the environment. Therefore, future reinforcement learning research must contain transfer learning and multi-agent reinforcement learning.

6.1 Transfer in Reinforcement Learning

At present, in many sub-fields of machine learning (such as neural networks and reinforcement learning), transfer learning has made some remarkable progress. In real-world environments, many learning tasks may share some similar features, resulting in similar policies. Therefore, the agent is not necessary to learn from scratch but with a prior bias. To enhance the performance, the research of transfer in reinforcement learning should answer the following questions: whether to transfer, what to transfer, how to transfer and how to evaluate the transfer performance. Whether to transfer depends on the similarity of source and target tasks, which contains the state space, action space, dynamics, and rewards. The transfer methods are usually classified into behavior transfer and knowledge transfer. The behavior transfer means to utilize the previously learned policy and common sub-tasks, including the policy transfer [29], option transfer [30], and hierarchical reinforcement learning [31]. While the knowledge transfer contains the value function transfer [32], relational reinforcement learning [33], and heuristic information transfer [34]. Different from behavior transfer, the knowledge transfer focuses on the essence of tasks and tries to learn a general model to address the problem, which is more similar to our minds but also more difficult. Finally, the common emulation of transfer in reinforcement learning is the jump-start (the initial enhancement), learning speed, and asymptotic performance.

6.2 Multi-Agent Reinforcement Learning and Game

Multi-agent system (MAS) also attracts many researchers, because many complex tasks cannot be modeled as the single-agent environment. For example, the famous prisoner's dilemma is based on game theory, and complex environments such as *Warcraft II* are the fields of multi-agent reinforcement learning. In MAS, agents communicate and interact with other agents, of which the MDP can be generalized to a stochastic

game, or a Markov game. As usual, the multi-agent reinforcement learning algorithms can be classified into cooperative, competitive and game types, which results in the famous algorithms combining the reinforcement learning with game theory, including minimax-Q [35], Correlated Q-Learning [36], Nash Q-Learning [37], Value - Decomposition Networks (VDN) [38], Mixing Q-network(QMIX) [39], Multi-agent Deep Deterministic Policy Gradient (MADDPG) [40], and so on. However, VDN and QMIX can only work in fully cooperative multi-agent environments. Although MADDPG has become a benchmarking work in multi-agent environments, there are also many challenges in real-world environments. For example, MADDPG cannot deal with the game environment and the opponent model prediction is impractical in partial observability environment. Up to now, the main challenges of multi-agent reinforcement learning are about the non-stationary environment, partial observability MDP (POMDP), training schemes, continuous space, and an equilibrium solution.

7 Conclusions

This paper presents an overview of the background and development of reinforcement learning, industry applications, and open source platforms. Since the remarkable milestone in AI beating professional Go players, we illustrate the famous AlphaGo and AlphaZero in details. Finally, we introduce the future technologies such as transfer reinforcement learning, multi-agent reinforcement learning, and game in complex interaction environment including stochastic environment, multiple players, selfish behavior and distributes optimization. In conclusion, we can look forward that future general artificial intelligence can be achieved through deep learning and reinforcement learning.

References

- [1] THORNDIKE E L. Animal Intelligence: An Experimental Study of the Associate Processes in Animals [J]. *American Psychologist*, 1998, 53(10): 1125. DOI: 10.1037/0003-066X.53.10.1125
- [2] MINSKY M L. *Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain Model Problem* [M]. New Jersey, America: Princeton University Press, 1954
- [3] BELLMAN R. On the Theory of Dynamic Programming [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1952, 38(8): 716. DOI: 10.1073/pnas.38.8.716
- [4] BELLMAN R. A Markovian Decision Process [J]. *Journal of Mathematics and Mechanics*, 1957: 679–684. DOI: 10.2307/2343663
- [5] WERBOS P. Advanced Forecasting Methods for Global Crisis Warning and Models of Intelligence [J]. *General System Yearbook*, 1977, 22: 25–38
- [6] SUTTON R. Learning to Predict by the Methods of Temporal Differences [J]. *Machine Learning*, 1988, 3(1): 9–44. DOI: 10.1007/BF00115009
- [7] WATKINS C J C H, DAYAN P. Q-Learning [J]. *Machine Learning*, 1992, 8: 279–292
- [8] RUMMERY G A, NIRANJAN, M. *On-Line Q-Learning Using Connectionist Systems* [M]. Cambridge, England: the University of Cambridge, 1994
- [9] THRUN S. Monte Carlo POMDPs [C]//*Advances in Neural Information Process-*

- ing Systems. Cambridge, USA, 2000
- [10] SUTTON R S, MCALLESTER D A, SINGH S P, et al. Policy Gradient Methods for Reinforcement Learning with Function Approximation [C]//Advances in neural information processing systems. Cambridge, USA, 2000
- [11] SILVER D, LEVER G, HEES N, et al. Deterministic Policy Gradient Algorithms [C]//31th International Conference on Machine Learning. Beijing, China, 2014
- [12] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human - Level Control Through Deep Reinforcement Learning [J]. Nature, 2015, 518(7540): 529. DOI: 10.1038/nature14236
- [13] VAN HASSELT H, GUEZ A, SILVER D. Deep Reinforcement Learning with Double Q-Learning [C]//13th AAAI Conference on Artificial Intelligence, Phoenix, USA, 2016
- [14] SCHAUL T, QUAN J, ANTONOGLU I, et al. Prioritized Experience Replay [DB/OL]. (2015-11-18). <https://arxiv.org/abs/1511.05952>
- [15] WANG Z, SCHAUL T, HESSEL M, et al. Dueling Network Architectures for Deep Reinforcement Learning [C]//International Conference on Machine Learning. New York, USA, 2016. DOI: 10.1155/2018/2129393
- [16] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous Control with Deep Reinforcement Learning [DB / OL]. (2015 - 09 - 09). <https://arxiv.org/abs/1509.02971>
- [17] NAIR A, SRINIVASAN P, BLACKWELL S, et al. Massively Parallel Methods for Deep Reinforcement Learning [DB/OL]. (2015-07-15). <https://arxiv.org/abs/1507.04296>
- [18] SCHULMAN J, LEVINE S, ABBEEL P, et al. Trust Region Policy Optimization [C]//31th International Conference on Machine Learning. Lille, France, 2015
- [19] SILVER D, HUANG A, MADDISON C J, et al. Mastering the Game of Go with Deep Neural Networks and Tree Search [J]. Nature, 2016, 529(7587): 484. DOI: 10.13140/RG.2.2.18893.74727
- [20] DUAN Y, CHEN X, HOUTHOOFT R, et al. Benchmarking Deep Reinforcement Learning for Continuous Control [C]//33rd International Conference on Machine Learning, New York, USA, 2016
- [21] BEATTIE C, LEIBO J Z, TEPLYASHIN D, et al. Deepmind Lab [DB/OL]. (2016-12-12). <https://arxiv.org/abs/1612.03801>
- [22] WU B. Hierarchical Macro Strategy Model for Moba Game AI [C]//AAAI Conference on Artificial Intelligence. Honolulu, USA, 2019. DOI: 10.1609/aaai.v33i01.33011206
- [23] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal Policy Optimization Algorithms [DB/OL]. (2017-07-20). <https://arxiv.org/abs/1707.06347>
- [24] HEES N, SRIRAM S, LEMMON J, et al. Emergence of Locomotion Behaviours in Rich Environments [DB/OL]. (2017 - 07 - 07). <https://arxiv.org/abs/1707.02286>
- [25] CLARK K. Neural Coreference Resolution [EB/OL]. <https://cs224d.stanford.edu/reports/ClarkKevin.pdf>
- [26] JI Y, TAN C, MARTSCHAT S, et al. Dynamic Entity Representations in Neural Language Models [C]//2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017. DOI: 10.18653/v1/D17-1195
- [27] BELLEMARE M G, NADDAF Y, VENESS J, et al. The Arcade Learning Environment: An Evaluation Platform for General Agents [J]. Journal of Artificial Intelligence Research, 2013, 47: 253-279. DOI: 10.1613/jair.3912
- [28] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the Game of Go Without Human Knowledge [J]. Nature, 2017, 550(7676): 354. DOI: 10.1038/nature24270
- [29] FERNÁNDEZ F, VELOSO M. Probabilistic Policy Reuse in a Reinforcement Learning Agent [C]//5th International Joint Conference on Autonomous Agents and Multiagent Systems. Hakodate, Japan, 2006. DOI: 10.1145 / 1160633.1160762
- [30] MCGOVERN A, BARTO A G. Automatic Discovery of Subgoals in Reinforcement Learning using Diverse Density [C]//8th International Conference on Machine Learning. Williams College, USA, 2001.
- [31] DIETTERICH T G. Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition [J]. Journal of Artificial Intelligence Research, 2000, 13: 227-303.
- [32] MAHADEVAN S. Proto-Value Functions: Developmental Reinforcement Learning [C]//22nd International Conference on Machine Learning. Bonn, Germany, 2005: 553-560. DOI: 10.1145/1102351.1102421
- [33] MADDEN M G, HOWLEY T. Transfer of Experience Between Reinforcement Learning Environments with Progressive Difficulty [J]. Artificial Intelligence Review, 2004, 21(3-4): 375-398. DOI: 10.1023/B:AIRE.0000036264.95672.64
- [34] DRIESSENS K, RAMON J, CROONENBORGH S T. Transfer Learning for Reinforcement Learning Through Goal and Policy Parametrization [C]//ICML Workshop on Structural Knowledge Transfer for Machine Learning. Pittsburgh, USA, 2006
- [35] LITTMAN M L. Markov Games as a Framework for Multi-Agent Reinforcement Learning [M]. Machine Learning Proceedings 1994. Burlington, USA: Morgan Kaufmann, 1994: 157-163. DOI: 10.1016/B978-1-55860-335-6.50027-1
- [36] GREENWALD A, HALL K, SERRANO R. Correlated Q-Learning [C]//20th International Conference on Machine Learning. Washington D. C., USA, 2003
- [37] HU J and WELLMAN M P. Multiagent Reinforcement Learning: Theoretical Framework and an Algorithm [C]//15th International Conference on Machine Learning, Madison, USA, 1998
- [38] SUNEHAG P, LEVER G, GRUSLYS A, et al. Value-Decomposition Networks for Cooperative Multi-Agent Learning [DB/OL]. (2017-06-16). <https://arxiv.org/abs/1706.05296>
- [39] RASHID T, SAMVELYAN M, WITT C S, et al. QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning [C]//35th International Conference on Machine Learning. Stockholm, Sweden, 2018
- [40] LOWE R, WU Y, TAMAR A, et al. Multi-Agent Actor-Critic for Mixed Cooperative - Competitive Environments [C]//Conference on Neural Information Processing Systems (NIPS). Long Beach, USA, 2017: 6379-6390.

Biographies

DONG Shaokang (shaokangdong@gmail.com) obtained his B.S. degree from the Advanced Class of Huazhong University of Science and Technology, China in 2018. He is currently a Ph.D. student in the Department of Computer Science and Technology, Nanjing University, China. His research interests include machine learning, reinforcement learning, and multi-armed bandits.

CHEN Jiarui obtained his B.S. degree from Dongbei University of Finance and Economics, China in 2018. He is currently a master student in the Department of Computer Science and Technology, Nanjing University, China. His research interests include machine learning, multi-agent reinforcement learning, and game.

LIU Yong received a B.S degree in communication engineering from China Agricultural University, China in 2017. He is currently a master student in the Department of Computer Science and Technology, Nanjing University, China. His current research interests include reinforcement learning, multi-agent learning, and transfer learning.

BAO Tianyi is an undergraduate student currently studying in the University of Michigan, USA. She studies computer science and psychology and will receive her B.S. degree in 2020. Her current research interests include the machine learning and human-computer interaction.

GAO Yang received the Ph.D. degree in computer software and theory from the Department of Computer Science and Technology, Nanjing University, China in 2000. He is a professor with the Department of Computer Science and Technology, Nanjing University. His current research interests include artificial intelligence and machine learning. He has published over 100 papers in top international conferences and journals.



A Low-Cost Outdoor Fingerprinting Localization Scheme For Wireless Cellular Networks

Abstract: This paper considers outdoor fingerprinting localization in LTE cellular Networks, which can localize non-cooperative user equipment (UE) that is unwilling to provide Global Positioning System (GPS) information. We propose a low-cost fingerprinting localization scheme that can improve the localization accuracy while reducing the computational complexity. Firstly, a data filtering strategy is employed to filter the fingerprints which are far from the target UE by using the Cell-ID, Timing Advance (TA) and eNodeB environment information, and the distribution of TA difference is analyzed to guide how to use TA rationally in the filtering strategy. Then, improved Weighted K Nearest Neighbors (WKNN) are implemented on the filtered fingerprints to give the final location prediction, and the WKNN is improved by removing the fingerprints that are still far away from the most of the K neighbors. Experiment results show that the performance is improved by the proposed localization scheme, and positioning errors corresponding to Cumulative Distribution Function (CDF) equaling to 67% and 95% are declined to 50 m and 150 m.

Keywords: fingerprinting localization; TA; filtering strategy; improved WKNN

PEI Dengke, XU Xiaodong,
QIN Xiaowei, LIU Dongliang,
and ZHAO Chunhua

(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, Anhui 230027, China)

DOI: 10.12142/ZTECOM.201903007

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190920.2103.002.html>, published online September 20, 2019

Manuscript received: 2019-05-18

1 Introduction

Benefiting from the fast growing of wireless techniques, location-based services (LBS) have made great development in various fields, such as emergency rescue, reconnaissance survey, and intelligent transportation systems. Several localization approaches have been analytically presented in literature for in-building and outdoor environments. Global Positioning System (GPS) is considered as one of the most well-known outdoor localization techniques [1], since it can provide acceptable localization accuracy as long as the target user equipment (UE) can receive navigation messages from at least four visible satellites. With the ubiquitous terrestrial wireless cellular network, LBS can also be provided by several eNodeBs through Time of Arrival (TOA), Time Difference of Arrival (TDOA) and Angle of Arriv-

al (AOA) measurements. It appears apparently an important way to localize non-cooperative UE which is unwilling to provide GPS information, but the network has to be equipped with multiple antennas and additional location measurement units (LMUs) in general.

As a cost-saving counterpart, the pioneer framework of fingerprinting localization is proposed in [2], where no additional hardware modification is needed. The process of fingerprinting localization can be constituted by two phases, namely off-line construction of fingerprinting database and on-line localization. During the off-line phase, the received signal strengths from adjacent access points are tabulated as location-related fingerprints, and the fingerprints with the corresponding physical coordinates are stored in a database. While during the on-line localization phase, target user reports the observed signal strengths so that the user's location can be estimated by matching in the database. As for wireless cellular networks, UE at a specific place can receive information of multiple Reference Signal Received Powers (RSRPs) from multiple eNodeBs, and the tuple of these RSRP measurements is usually

This work was supported by the ZTE Industry-Academia-Research Cooperation Funds under Grant No. 20160722-01.

employed as the fingerprint. In literature, [3] presents a probabilistic Received Signal Strength Indicator(RSSI) based fingerprinting localization system for GSM wireless cellular networks, which needs a large amount of fingerprinting information to build the probabilistic model for a particular location. In this regard, the database construction becomes laborious and time costly, and the localization accuracy is not satisfactory since RSSI is the only information in use. For the sake of decreasing computational complexity as wells as increasing localization accuracy, [4], [5] and [6] use the Cell-ID as an additional information to form the fingerprint, and deterministic localization algorithms, such as Weighted K Nearest Neighbors (WKNN) are used, where fewer fingerprinting records are measured as compared with the probabilistic localization one. However, this kind of fingerprinting localization scheme may eliminate those physically much closer fingerprints so as to induce attenuated localization accuracy.

In this paper, we put forward a low-cost outdoor fingerprinting localization scheme for the fourth generation (4G) wireless systems, which is shown to have decreased positioning error and computational load in contrast to the counterparts. During the off-line phase, database construction can be simply implemented through the UE measurement report (MR) that is gathered at the network side. Different from the most existing works in which fingerprints just contain RSRP, the fingerprint used in this paper is comprised of RSRP, Cell-ID, and Timing Advance (TA). What's more, we also derive a theoretical analysis to demonstrate the usage of TA. Therefore, we would like to shed more light on the improvement during on-line localization phase, and the localization result is shown to be better than the related work [4] with the same dataset. The main technical contributions are listed as follows:

1) A database filtering strategy is proposed to exclude the fingerprints which are far from the target UE before matching, where we preserve only the fingerprints which have the same serving Cell-ID, approximate TA and approximate eNodeB environment for the target UE. Furthermore, the potential range of TA at a particular location is theoretically analyzed to help us construct a proper filtering threshold. In this way, the positioning error and computational load can be decreased to some extent since during the matching process, fingerprints with geographic positions closer to target UE have a larger probability to participate in the computation.

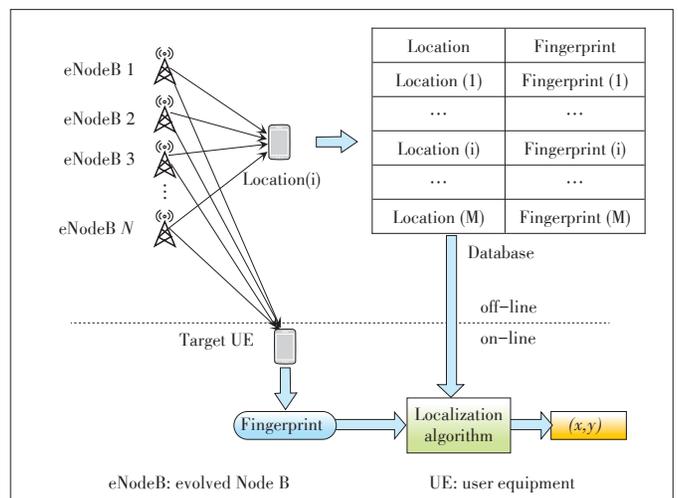
2) WKNN is employed to compute the geographic position of the target UE according to the fingerprinting database. K nearest neighbors are selected from the database according to the distance between target UE and their RSRP measurements. However, due to the intrinsic intense fluctuation of RSRP measurements, two adjacent fingerprints with the minimum Euclidean distance do not always mean that they are close to each other in geographic position. Therefore, we further improve the WKNN matching algorithm by removing those anomaly fingerprints that are still possibly far away from the others among K

neighbors.

2 Architecture of Fingerprinting Localization

Obeying the two- phase strategy of off-line construction of fingerprinting database and on-line localization, **Fig. 1** depicts an architecture of fingerprinting localization for 4G wireless cellular networks. In this paper, the RSRPs and cellular network information from eNodeBs received by the measured UE are denoted as fingerprint, and every fingerprint and the corresponding geographic coordinates provided by GPS (AGPS) modules are stored in the database (Fig. 1). For clarity, the stored samples with sufficient information about geographic coordinate in the database are called reference points (RPs). In the fingerprinting database, detailed information about a RP can be expressed as **Table 1**, where x, y represent the longitude and latitude of the corresponding RP respectively, and the $Cell-ID_n$ stands for the Cell-ID of the severing eNodeB, and the value of the TA for the severing eNodeB is m . N is the number of total eNodeBs measured by all of the RPs in the database. $RSRP_i$ represents the RSRP of the i th eNodeB. Note that, for one RP, the number of detectable eNodeB is l , so for $N - l$ unreachable eNodeBs, the corresponding RSRPs will be set as a minimum level of detectable signal strength empirically [4], which is -140 dBm in this paper.

As shown in Table 1, almost all location-related parameters have been considered in this paper. But unfortunately, none of



▲ **Figure 1. An architecture of fingerprinting localization for 4G system.**

▼ **Table 1. Information of reference points**

| Longitude | Latitude | Severing Cell-ID | TA | Detectable eNodeBs | Cell-ID ₁ | ... | Cell-ID _i | ... | Cell-ID _N |
|-----------|----------|----------------------|-----|--------------------|----------------------|-----|----------------------|-----|----------------------|
| x | y | Cell-ID _n | m | l | RSRP ₁ | ... | RSRP _i | ... | RSRP _N |

RSRP: Reference Signal Received Power

TA: Timing Advance

them has an explicit relationship with the location of target UE in reality due to the complicated characteristics of the propagation channel. For instance, the instantaneous values of RSRP and TA from a single eNodeB are often sensitive to the environment so that they vary distinctively even for multiple observations at the same place. It is reported in [7] that RSRP, Cell-ID and TA information is used to classify UE, while in this paper, we would like to further perform a new application, in which more comprehensive information is used to calculate the geographical coordinates of UE.

3 Proposed Fingerprinting Localization Scheme

In this section, the procedure of online fingerprinting localization will be introduced. Fig. 2 illustrates the proposed scheme, which consists of a data filtering module followed by an improved WKNN algorithm.

3.1 A Data Filtering Strategy

As mentioned above, thousands of RPs have been stored in the fingerprinting database according to their MRs. Given a fingerprint from the target UE, it is computationally wasteful to use WKNN to figure out the most approximate RP in the whole dataset directly with a brute-force searching. On the other hand, RSRP information is commonly used by WKNN during the matching process, but in general RSRP fluctuates too sharply to obtain expected results. In this regard, a data filtering method should be introduced beforehand, and the main purpose of data filtering is to perform the matching process in a constrained region so as to decrease the localization error as well as the computational load. In the sequel, additional steady information is used to construct the constrained region, including Cell-ID, TA and eNodeB environment.

The filtering strategy is divided into three steps. Firstly, we

use Cell-ID to eliminate candidate RPs which have different serving eNodeB as the target UE, since UEs belonging to distinctive eNodeBs are generally far from each other to a large extent. However, there is at least one specific case that the Cell-ID based data filtering strategy is improper and should be skipped. That is, when the target UE is at the edge of an eNodeB, where the knowledge can be read by RSRP no more than a specified threshold in this paper, because the serving RSRP is below the specified thresholds mainly due to the high propagation loss caused by excessive distance between the eNodeB and the target UE. And the thresholds is set to -110 dBm according to [8].

Secondly, TA is employed to further filtrate improper RPs within the candidates obtained through the first filtering step. Due to the immunity of TA against the variation of local propagation environment, we keep RPs in the candidate set which have approximate TAs with the target UE and remove the offenders according to the following rule

$$\left| \hat{t}_{\frac{UE}{TA}} - \hat{t}_{\frac{RP}{TA}} \right| \leq \Delta T, \tag{1}$$

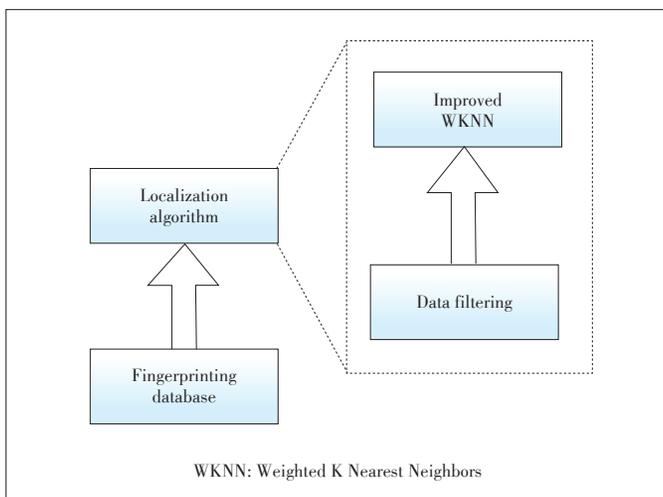
where $\hat{t}_{\frac{UE}{TA}}$ and $\hat{t}_{\frac{RP}{TA}}$ stand for measured TA of the target UE and certain RP respectively. And $\hat{t}_{\frac{UE}{TA}} - \hat{t}_{\frac{RP}{TA}}$ is called TA difference in this paper, which is an integer value and can be designed following the analytical result of TA difference in the next section.

The third and the last filtering steps are established based on the fact that the target UE and its neighbor RPs should be in similar eNodeB environment, and the similarity is reflected both in the number of total detected eNodeBs and the number of eNodeBs with the same Cell-IDs. In this work, the strong correlation between the target UE and RPs in the candidate set holds if the difference about the number of total detected eNodeBs is less than four, where at least three detected eNodeBs have the same Cell-IDs as the target UE [4]. After all, RPs violating this rule will be eliminated from the candidate set. Now the candidate set selected by the three filtering steps is defined as \mathfrak{R} .

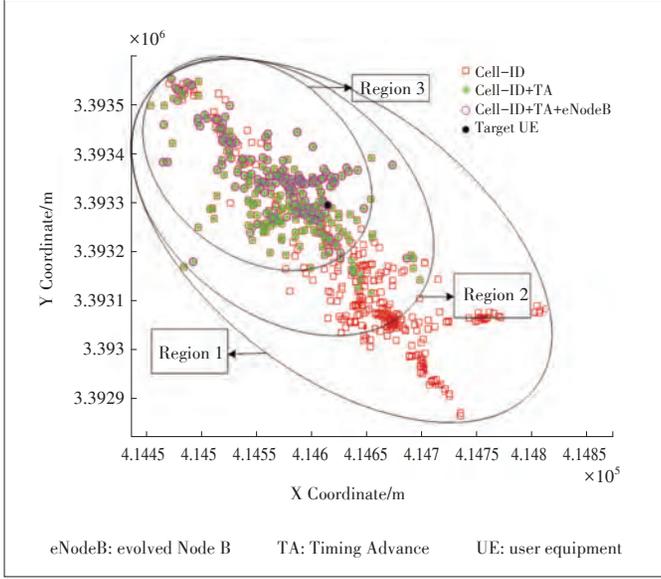
To end this discussion, a numerical example is presented to show the effectiveness of the data filtering strategy in Fig. 3. We highlight three regions in this figure to clarify such a filtering process. It is clearly that the regions become more precise with further rigorous filtering conditions. The following matching process can now be performed on the candidate set \mathfrak{R} , therefore the localization error can be reduced and the number of candidates is decreased from ten thousands to less than one hundred for about 1 000 times.

3.2 Improved WKNN Algorithm

Conventional WKNN algorithm is widely used in fingerprint-



▲ Figure 2. The proposed fingerprinting localization scheme.



▲ Figure 3. Filtering Results.

ing localization and has shown the superior performance due to many advantages [9], such as being easy to understand, being robust to noise, and high accuracy in localization. However, it is inefficient in terms of computational complexity if one performs WKNN directly on a large training dataset. Fortunately, this disadvantage can be alleviated by the data filtering strategy presented above.

Because of the drastic fluctuation of instantaneous RSRPs, RPs which have high fingerprinting similarity with the target UE may be far from the UE geographically. In this regard, among K neighbors chosen by WKNN for the target UE, some should be looked as erasable anomaly RPs to further improve the performance of localization accuracy. In the sequel, we present a way to remove these anomaly RPs from K neighbors and name the modified WKNN algorithm as an improved one.

When K nearest neighbors have been selected, the improved WKNN algorithm continues to compute the geographic distance for every pair of them. We assume that RPs which are far away from the others are regarded as anomalous RPs, and they should not participate in location calculation. The procedure of the improved WKNN algorithm is summarized as follows:

1) Calculate the distance between the target UE and each filtered RP in the candidate set according to the following equation

$$e(k) = \left(\sum_{i=1}^l (UE_i - RP_i(k))^2 \right)^{1/2}, k = 1, 2, 3, \dots, M, \quad (2)$$

where $e(k)$ is the Euclidean distance between target UE and the k th RP in the sense of RSRP. M is the total number of RPs in the candidate set \mathfrak{R} , and UE_i represents for the measured RSRP of the detectable eNodeB by the target UE. l is the number of the detected eNodeBs by the target UE.

2) Sort $\{e(k)\}$ to choose K RPs as the nearest neighbors with

K lowest distances; in this paper, the value of K equals to 7. Then we calculate the geographic distance between any two nearest neighbors and eliminate T anomaly RPs which are far away from the others for more than 300 m in general.

3) Compute the weight coefficients through

$$w_{\tilde{k}} = \frac{1/(e(\tilde{k}) + 0.001)}{\sum_{\tilde{k}=1}^{K-T} 1/(e(\tilde{k}) + 0.001)}, \quad (3)$$

where \tilde{k} represents the k th RP of the rest $K-T$ nearest neighbors after erasing anomaly. The resulting weight coefficients are used to proceed to calculate the geographic coordinate of the target UE. The relationship between the target UE and the rest $K-T$ nearest neighbors can be formulated as

$$\begin{cases} \tilde{x}_{UE} = \sum_{\tilde{k}=1}^{K-T} w_{\tilde{k}} x_{\tilde{k}} \\ \tilde{y}_{UE} = \sum_{\tilde{k}=1}^{K-T} w_{\tilde{k}} y_{\tilde{k}} \end{cases}, \quad (4)$$

where $(\tilde{x}_{UE}, \tilde{y}_{UE})$ denotes the geographic coordinate of the target UE, and $(x_{\tilde{k}}, y_{\tilde{k}})$ stands for the geographic coordinate of the \tilde{k} th neighbor.

4 Analysis of TA Difference

The threshold of TA difference plays an important role on the performance of the proposed data filtering strategy. To shed more light on how to design this threshold, we will give an analytical result in the following.

TA can be used as an approximation for TOA [10], but TA represents the round trip time. Therefore, $|t_{TA}^{UE} - t_{TA}^{RP}|/2$ equals to the discrete value of $|t_{TOA}^{UE} - t_{TOA}^{RP}|$, where t_{TA}^{UE} and t_{TA}^{RP} are the TA of the target UE and RP respectively. Similarly, t_{TOA}^{UE} and t_{TOA}^{RP} represent the TOA of the target UE and RP. Based on this perspective, if we can infer the fluctuation range of TOA, we can get the fluctuation range of the TA.

TA in [11] can be modeled by

$$\hat{t}_{TA} = t_{TA} + \xi, \quad (5)$$

where \hat{t}_{TA} and t_{TA} are the measured TA and actual value of TA respectively, and t_{TA} can be seen as the mean value of \hat{t}_{TA} . ξ is the error corrupting the true value and can be modeled as a random variable obeying normal distribution. As for TOA, it is reasonable to construct the observation model as

$$\hat{t}_{TOA} = t_{TOA} + \omega, \quad (6)$$

where \hat{t}_{TOA} and t_{TOA} are the measured TOA and actual value of TOA respectively and ω can also be modeled as a Gaussian variable.

$$\omega \sim N(0, \sigma^2). \quad (7)$$

Given two independent identically distributed random variables \hat{t}_{TOA1} and \hat{t}_{TOA2} obeying (6), we can easily get the Gaussian distribution of the TOA difference:

$$\hat{l}_{TOA1} - \hat{l}_{TOA2} \sim N(0, 2\sigma^2). \quad (8)$$

In the same way, given two independent identically distributed random variables \hat{l}_{TA1} and \hat{l}_{TA2} , TA difference shall obey the Gaussian distribution with standard deviation $\hat{\sigma}$ and $\hat{\sigma}^2 = 8\sigma^2$ can be gotten according to the relationship between TOA and TA. It appears that if the standard deviation of the measured TOA can be calculated, the distribution of TA can thus be acquired. Hence in the next step, we attempt to explore an approximate solution about the standard deviation of the measured TOA.

In the multipath channel environments, TOA means the signal arrival time about first path, so the fluctuation range of the TOA cannot be greater than τ , the delay spread of the propagation channel. In nature, the delay spread τ can be interpreted as the difference arriving time between the earliest multipath component (typically the line-of-sight component) and the latest multipath component. On the other hand, according to the three-sigma rule, more than 99.73% samples lie within a band around the mean in a normal distribution $N(\mu, \sigma^2)$ with a width six standard deviations. In short, random variable distributes in range $[\mu - 3\sigma, \mu + 3\sigma]$ with an asymptotic probability to 1. Therefore, the relationship between τ and the fluctuation range of the TOA can be described as

$$\tau \geq \hat{l}_{TOA_{\max}} - \hat{l}_{TOA_{\min}} = 6\sigma, \quad (9)$$

where $\hat{l}_{TOA_{\max}}$ and $\hat{l}_{TOA_{\min}}$ mean the maximal TOA the minimum TOA respectively. In this paper, τ can be calculated according to [12]

$$\tau = \exp(A \times L + B), \quad (10)$$

where $A = 0.038$, $B = 2.3$, and L is path loss (dBm). In this way, we can obtain the upper bound of the standard deviation of the measured TOA, which is $\tau/6$, and the standard deviation of TA difference can then be described as $\hat{\sigma} \leq \sqrt{2}/3 \exp(A \times L + B)$. Taking the upper bound as the variance, the distribution of TA difference can thus be formulated by

$$\hat{l}_{TA1} - \hat{l}_{TA2} \sim N(0, (\frac{\sqrt{2}}{3} \exp(A \times L + B))^2). \quad (11)$$

Defining $\Delta T = \delta\hat{\sigma}$ as a threshold of TA difference for data filtering strategy, and the corresponding confidence intervals as γ , we have:

$$\begin{aligned} p(|\hat{l}_{TA}^{UE} - \hat{l}_{TA}^{RP}| \leq \Delta T) &= p(|l_{TA}^{UE} - l_{TA}^{RP}| \leq \delta\hat{\sigma}) = \\ &= \int_{-\delta\hat{\sigma}}^{\delta\hat{\sigma}} \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp(-\frac{x^2}{2\hat{\sigma}^2}) dx = \\ &= \int_{-\infty}^{\delta\hat{\sigma}} \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp(-\frac{x^2}{2\hat{\sigma}^2}) dx - \int_{-\infty}^{-\delta\hat{\sigma}} \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp(-\frac{x^2}{2\hat{\sigma}^2}) dx = \\ &= \text{erf}(\frac{\delta}{\sqrt{2}}) = \gamma, \end{aligned} \quad (12)$$

where δ and γ are positive, and γ should be no greater than 1. Apparently, δ can be given by:

$$\delta = \sqrt{2} \text{erf}^{-1}(\gamma). \quad (13)$$

In this paper, we set $\gamma = 0.9$, which means that 90% RPs that are at the same location as target UE will be selected through this TA difference threshold and that $\delta = 1.65$ according to (13), thus the filtering rule can be described as

$$|\hat{l}_{TA}^{UE} - \hat{l}_{TA}^{RP}| \leq 1.65 \times \frac{\sqrt{2}}{3} \exp(A \times L + B). \quad (14)$$

From (14), the path loss of the actual environment has a great influence on the threshold ΔT . To compute the path loss, Standard Propagation Model (SPM) [13], [14] is considered in this paper, which can be formulated by

$$L_{SPM} = K_1 + K_2 \times \log(d) + K_3 \times \log(h_{te}) + K_4 \times (\text{Diffract}) + K_5 \times \log(d) \times \log(h_{re}) + K_6 \times h_{re} + K_7 \times f(\text{clutter}), \quad (15)$$

where K_1 is the constant offset related to frequency; d is the distance between the receiver and the transmitter (m); K_2 is the multiplying factor for $\log(d)$; h_{te} is the effective height of the transmitter antenna (m); K_3 is the multiplying factor for $\log(h_{te})$; K_4 is the multiplying factor for diffraction calculation; K_5 is the multiplying factor for $\log(d) \times \log(h_{re})$; h_{re} is the effective height of the receiver mobile antenna; K_6 is the multiplying factor for $\log(h_{re})$; $f(\text{clutter})$ is the average of weighted losses due to clutter.

For users within a particular cell, the maximum pass loss often happens when they are at the cell edge. In an LTE system, h_{te} is about 30 m usually, and K_3, K_5, K_6 are often set separately with default values 5.83, -6.55, and 0. K_4 and K_7 can be zero each as in [13] and [15]. K_1 and K_2 are adjusted in our situation as $K_1 = 25, K_2 = 37$. Under the consumption of $d=400$ m for the effective cell radius in an urban area, the maximum path loss can eventually be calculated as 104.7 dBm according to (15). Therefore, the resulting threshold. $\Delta T = 1.65 \times \hat{\sigma} = 414$ ns. In LTE system, TA is practically quantized in bit period T_b with $T_b = 520$ ns. Therefore, the quantized threshold used in this paper should be set as 1 TA. By this filtering rule we can select most of the fingerprints who are close to the target UE, and exclude the fingerprints who are far from the UE geographically.

5 Experimental Results

To validate the performance of the proposed fingerprinting localization scheme, some experiments will be presented in this section by using the experiment data collected at some districts in Sichuan Province, China. There are 42 588 MR messages with exact geographic coordinates. 20% of them are selected randomly as testing data, and 80% of them are used to construct database. The experiments include three parts. The first part demonstrates the contributions of filtering on localization accuracy and computational burden. The second part com-

compares different matching algorithms and demonstrates the superiority of the improved WKNN. The third part tests the localization scheme by doing experiments in different districts with different experimental data.

5.1 Effect of the Data Filtering Strategy

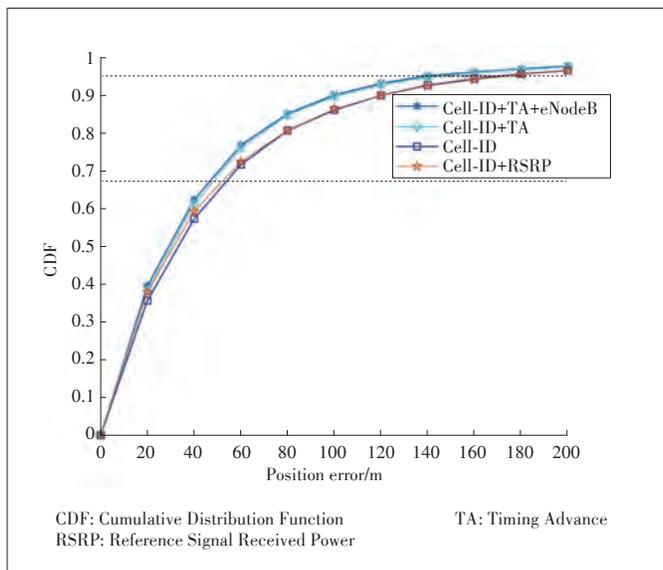
We examine the effect of the proposed filtering strategy on improvement of localization accuracy by using different filtering strategies in the experimental area. The experimental area is 35 000×35 000 m², containing 34 070 RPs and 8 518 test data (target UEs). Different filtering strategies are as follows:

- 1) Select all the RPs which have the same Cell-ID with the target UE in the filtering process
- 2) Select all the RPs which have the same Cell-ID and similar TA with the target UE in the filtering process
- 3) Select all the RPs which have the same Cell-ID, similar TA and similar eNodeB environment with the target UE in the filtering process
- 4) Use the filtering strategy in [4], where TA information is not considered.

▼ **Table 2. Positioning error with different filtering strategies**

| Filter criteria | Location error | | | Time/s |
|-------------------|----------------|-------------|--------------|--------|
| | 67% error/m | 95% error/m | Mean error/m | |
| Cell-ID | 52.6 | 170.66 | 58.7 | 562 |
| Cell-ID+TA | 46.8 | 146.8 | 52.4 | 401 |
| Cell-ID+TA+eNodeB | 44.4 | 139.2 | 48.7 | 104 |
| Cell-ID+RSRP | 50.3 | 169.45 | 57.2 | 204 |

RSRP: Reference Signal Received Power TA: Timing Advance



CDF: Cumulative Distribution Function TA: Timing Advance
RSRP: Reference Signal Received Power

▲ **Figure 4. Positioning error with different filtering strategies.**

Table 2 shows the positioning errors and runtime of the different filtering strategies, where 67% error and 95% error mean the positioning errors corresponding to CDF = 67% and CDF = 95%, the mean error is the mean positioning error, and the time is the total runtime for the 8 518 test data. **Fig. 4** is the CDF of the positioning error, where the legend Cell-ID, TA and eNodeB represent the remaining RPs after the corresponding step of data filtering. We use Cell-ID+RSRP to stand for the filtering strategy presented in [4]. From Fig. 4 and Table 2, it is clear that the positioning error of the proposed scheme has been reduced obviously by adding the TA filtering criterion, and the time spent for positioning has been reduced obviously by adding the eNodeB environment information. By this way, the mean positioning error declines to 50 m, and the positioning error corresponding to CDF = 67% and CDF = 95% are less than 50 m and 150 m, which have achieved the localization requirements of Federal Communication Commission (FCC), the United States, whereby mobile operators should be able to determine the location of a UE within 50–100 m in 67% cases and 150–300 m in 95% cases. And the filtering strategy presented in [4] has a similar positioning errors as the proposed one where only the first data filtering step is performed.

5.2 Performance of the Improved WKNN

Here we compare different matching algorithms to examine the performance of improved WKNN. Firstly, we compare different matching algorithms containing WKNN, Support Vector Machine (SVM), and Back Propagation Neural Networks (BPNN). The areas of experimental regions are chosen as 100×100 m², 400×400 m² and 1 000×1 000 m², respectively. The localization results of different matching algorithms are shown in **Table 3**, where we can see that WKNN shows superior performance in all of the experiments. The performance of SVM and BPNN can be acceptable when the experimental area is small, but with the increase of the experimental area, the performance of BPNN and SVM becomes worse than WKNN. More clearly, the mean localization error for WKNN is 33 m, such as the experimental results with the region of 1 000×1 000 m², but the

▼ **Table 3. Comparison of different matching algorithms**

| | | 67% error/m | 95% error/m | Mean error/m |
|------------------------------|------|-------------|-------------|--------------|
| 100 × 100 m ² | WKNN | 23 | 43 | 18 |
| | SVM | 23 | 45 | 19 |
| | BPNN | 25 | 56 | 22 |
| 400 × 400 m ² | WKNN | 27 | 79 | 27 |
| | SVM | 29 | 73 | 27 |
| | BPNN | 39 | 77 | 35 |
| 1 000 × 1 000 m ² | WKNN | 30 | 100 | 33 |
| | SVM | 46 | 155 | 49 |
| | BPNN | 70 | 191 | 70 |

BPNN: Back Propagation Neural Network WKNN: Weighted K Nearest Neighbors
SVM: Support Vector Machine

mean localization errors for SVM and BPNN are 49 m and 70 m respectively.

Then the contrastive experiment between WKNN and improved WKNN has been operated. The area of experimental region is 35 000 × 35 000 m², containing 34 070 RPs and 8 518 test data (target UEs). Localization results of WKNN and improved WKNN are showed in **Table 4**. Just as expected, the localization performance is improved by using improved WKNN, especially with the CDF=95%, which testifies that the improved WKNN is effective for removing the anomalous RPs who are harmful to localization accuracy.

5.3 Application to Other Districts

In this section, we conduct experiments in different districts to test the robustness of the proposed localization scheme. The experimental results in **Table 5** show that the positioning errors corresponding to CDF = 67% are almost the same between District A, District B, and District C, and the positioning errors corresponding to CDF = 95% have a little difference but also in an acceptable range, and the mean errors also keep jarless between different districts. Generally speaking, there is no significant difference in positioning errors between different districts, which declares that the localization scheme proposed in this paper is robust, and the parameter setting is reasonable. Furthermore, the experimental results can meet the FCC requirements in all of the experimental districts.

6 Conclusions

In this paper, we present an outdoor fingerprinting localization scheme for wireless cellular networks, which is low-cost and high-efficiency with localization accuracy meeting the FCC requirements for the case of urban environments. In the off-line database construction phase, we use the MRs obtained from the network side instead of laborious war driving. In the online localization phase, firstly, a data filtering strategy is pro-

posed to improve the performance of localization, which selects the RPs that own the same Cell-ID, similar TA and eNodeB environment with the target UE, and the distribution of TA difference has been analyzed to guide how to use TA rationally in the filtering strategy. Secondly, WKNN has been improved by computing the geographic distance for every pair of K nearest RPs, and the RPs who are far away from the others should not participate in location calculation. Some experiments have been conducted and experimental results demonstrated that the localization scheme proposed could decrease the localization error as well as the computational load, and the mean positioning error declines to 50 m, and positioning errors corresponding to CDF= 67% and CDF = 95% are less than 50 m and 150 m.

In the future, we will investigate how to further improve the outdoor localization accuracy and analyze the bound of outdoor localization accuracy for wireless cellular networks.

References

- [1] IBRAHIM M, YOUSSEF M. CellSense: A Probabilistic RSSI-Based GSM Positioning System [C]// Global Telecommunications Conference. Miami, USA, 2010. DOI:10.1109/GLOCOM.2010.5683779
- [2] YOUSSEF M A, AGRAWALA A, SHANKAR A U. WLAN Location Determination via Clustering and Probability Distributions [C]//IEEE International Conference on Pervasive Computing & Communications. Fort Worth, USA, 2003. DOI: 10.1109/PERCOM.2003.1192736
- [3] BSHARA M, ORGUNER U, GUSTAFSSON F, et al. Fingerprinting Localization in Wireless Networks Based on Received-Signal-Strength Measurements: A Case Study on WiMAX Networks [J]. IEEE Transactions on Vehicular Technology, 2010, 59(1): 283–294. DOI: 10.1109/tvt.2009.2030504
- [4] FAYAZ S, SARRAFIAN S. Location Service for Wireless Network Using Improved RSS-Based Cellular Localization [J]. International Journal of Electronics, 2014, 101(6): 763–778. DOI: 10.1080/00207217.2013.794492
- [5] SCHROTH G, HUITL R, CHEN D, et al. Mobile Visual Location Recognition [J]. IEEE Signal Processing Magazine, 2011, 28(4): 77–89. DOI: 10.1109/MSP.2011.940882
- [6] ZHANG J, HALLQUIST A, LIANG E, et al. Location-Based Image Retrieval for Urban Environments [C]//18th IEEE International Conference on Image Processing. Brussels, Belgium, 2011. DOI: 10.1109/ICIP.2011.6116517
- [7] SHI L, WIGREN T. AECID Fingerprinting Positioning Performance [C]//IEEE Global Telecommunications Conference. Honolulu, USA, 2009. DOI: 10.1109/GLOCOM.2009.5425928
- [8] GOMEZ-ANDRADES A, BARCO R, SERRANO I, et al. Automatic Root Cause Analysis Based on Traces for LTE Self-Organizing Networks [J]. IEEE Wireless Communications, 2016, 23(3): 20–28. DOI: 10.1109/MWC.2016.7498071
- [9] MOGHATAIEE V, DEMPSTER A G. Indoor Location Fingerprinting Using FM Radio Signals [J]. IEEE Transactions on Broadcasting, 2014, 60(2): 336–346. DOI: 10.1109/TBC.2014.2322771
- [10] YOST G P, PANCHAPAKESAN S. Improvement in Estimation of Time of Arrival (TOA) from Timing Advance (TA) [C]// IEEE International Conference on Universal Personal Communications. Florence, Italy, 1998. DOI: 10.1109/ICUPC.1998.733714
- [11] ROTH J D, TUMMALA M, MCEACHEN J C, et al. Maximum Likelihood Geolocation in LTE Cellular Networks Using the Timing Advance Parameter [C]// IEEE International Conference on Signal Processing & Communication Systems. Montreal, Canada, 2017. DOI: 10.1109/ICSPCS.2016.7843379
- [12] ITU. Propagation Data and Prediction Methods for the Planning of Short-Range Outdoor Radio Communication Systems and Radio Local Area Networks in the

▼Table 4. Comparison between WKNN and the improved WKNN

| Experimental areas | Matching algorithm | 67% error/m | 95% error/m |
|--------------------------------|--------------------|-------------|-------------|
| 35 000 × 35 000 m ² | WKNN | 46 | 147 |
| | Improved WKNN | 44 | 139 |

WKNN: Weighted K Nearest Neighbors

▼Table 5. Localization in different districts

| | 67% error/m | 95% error/m | Mean error/m |
|------------|-------------|-------------|--------------|
| District A | 44 | 140 | 50 |
| District B | 44 | 132 | 48 |
| District C | 43 | 127 | 50 |

Frequency Range 300 MHz to 100 GHz [S], 2013.

- [13] De Freitas P R, FILHO H T. Parameters Fitting to Standard Propagation Model (SPM) for Long Term Evolution (LTE) Using Nonlinear Regression Method [C]// IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA). Anecy, France, 2017: 84–88. DOI: 10.1109/CIVEMSA.2017.7995306
- [14] RANI M S, BEHARA S, SURESH K. Comparison of Standard Propagation Model (SPM) and Stanford University Interim (SUI) Radio Propagation Models for Long Term Evolution (LTE), International Journal of Advanced and Innovative Research (IJAIR), 2012, 1(6): 221–228
- [15] XU H, SHI C, ZHANG W, et al. Field Testing, Modeling and Comparison of Multi Frequency Band Propagation Characteristics for Cellular Networks [C]// IEEE International Conference on Communications. Kuala Lumpur, Malaysia, 2016. DOI: 10.1109/ICC.2016.7510961

Biographies

PEI Dengke (pdke1@mail.ustc.edu.cn) received the B.S. degree from China University of Geosciences, China in 2016. She is currently pursuing the B.E. degree at University of Science and Technology of China. Her research interest is localization and deep learning.

XU Xiaodong received his B.Eng. degree and Ph.D. degree in electronic and information engineering from University of Science and Technology of China (USTC) in 2000 and 2007, respectively. Since 2007, he has been a faculty mem-

ber with the Department of Electronic Engineering and Information Science, USTC, where he is currently an associate professor. His research interests include array signal processing, wireless communications, and information-theoretic security.

QIN Xiaowei received the B.S. and Ph.D. degrees from the Department of Electrical Engineering and Information Science, University of Science and Technology of China (USTC) in 2000 and 2008, respectively. Since 2014, he has been a member of staff in Key Laboratory of Wireless-Optical Communications of Chinese Academy of Sciences at USTC. His research interests include optimization theory, service modeling in future heterogeneous networks, and big data in mobile communication networks.

LIU Dongliang received the B.S. degrees from the Telecommunications Engineering College, Beijing University of Posts and Telecommunications (BUPT), China, in 2004. Since 2006, he has been an Algorithms Researcher of the Algorithms Department in ZTE Corporation. His research interests include positioning theory, network optimization theory and big data in mobile communications.

ZHAO Chunhua received the M.S. degrees from the College of Electronic and Communication Engineering, Harbin Institute of Technology (HIT), China in 2002. Since then, she has been working on RRM algorithms design and network optimization in mobile communications. She has been an algorithms researcher at the Algorithms Department of ZTE Corporation since 2009. Her research interest is big data analytics for network optimization areas of experimental regions are chosen as 100×100 in mobile communications.

← From P. 22

Accurate Object Detection and Semantic Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 580–587. DOI: 10.1109/CVPR.2014.81

- [35] GIRSHICK R. Fast R-CNN [C]//IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440–1448. DOI: 10.1109/ICCV.2015.169
- [36] TANG X, DU D K, HE Z, et al. Pyramidbox: A Context-Assisted Single Shot Face Detector [C]//European Conference on Computer Vision (ECCV). Munich, Germany, 2018: 797–813. DOI: 10.1007/978-3-030-01240-3_49
- [37] ZHANG Z, LUO P, LOY C C, et al. Facial Landmark Detection by Deep Multi-Task Learning [C]//European Conference on Computer Vision. Zurich, Switzerland, 2014: 94–108. DOI: 10.1007/978-3-319-10599-4_7
- [38] CHEN D, REN S, WEI Y, et al. Joint Cascade Face Detection and Alignment [C]//European Conference on Computer Vision. Zurich, Switzerland, 2014: 109–122. DOI: 10.1007/978-3-319-10599-4_8
- [39] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [DB/OL]. (2014 - 09 - 04). <https://arxiv.org/abs/1409.1556>
- [40] YANG S, LUO P, LOY C C, et al. Wider Face: A Face Detection Benchmark [C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vega, USA, 2016: 5525–5533. DOI: 10.1109/CVPR.2016.596
- [41] YANG B, YAN J, LEI Z, et al. Aggregate Channel Features for Multi-Vision Face Detection [C]//IEEE International Joint Conference on Biometrics. Clearwater, USA, 2014: 1–8. DOI: 10.1109/BTAS.2014.6996284
- [42] YAN J, ZHANG X, LEI Z, et al. Face Detection by Structural Models [J]. Image and Vision Computing, 2014, 32(10): 790–799. DOI: 10.1016/j.imavis.2013.12.004
- [43] MARKUS N, FRLJAK M, PANDZIC I S, et al. Object Detection with Pixel Intensity Comparisons Organized in Decision Trees [DB/OL]. (2013-05-20). <https://arxiv.org/abs/1305.4537>
- [44] LI H, LIN Z, BRANDT J, et al. Efficient Boosted Exemplar-Based Face Detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1843–1850. DOI: DOI: 10.1109/CVPR.2014.238
- [45] LI J, ZHANG Y. Learning Surf Cascade for Fast and Accurate Object Detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 3468–3475. DOI: 10.1109/CVPR.2013.445

Biographies

GUO Da received the B.Eng. from the Computer Engineering College, JiMei University, China in 2018. He is currently a master student at the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. His research direction is face detection and recognition based on deep learning.

ZHENG Qingfang received the B.S. degree in civil engineering from Shanghai Jiao Tong University, China in 2002 and Ph.D. degree in computer science from Institute of Computing Technology, Chinese Academy of Science, China in 2008. He is currently the chief scientist of video technology with ZTE Corporation. His research interests include computer vision, multimedia retrieval, image/video processing, with a special focus on low power embedded application and large-scale cloud application.

PENG Xiaojiang (xj.peng@siat.ac.cn) received his Ph.D. from School of Information Science and Technology from Southwest Jiaotong University, China in 2014. He currently is an associate professor at the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. He was a postdoctoral researcher at Idiap Institute, Switzerland from 2016 to 2017, and was a postdoctoral researcher in LEAR Team, INRIA, France, working with Prof. Cordelia Schmid from 2015 to 2016. He serves as a reviewer for IJCV, TMM, TIP, CVPR, ICCV, AAAI, IJCAI, FG, Image and Vision Computing, IEEE Signal Processing Letter, Neurocomputing, etc. His research focus is in the areas of action recognition and detection, face recognition, facial emotion analysis, and deep learning.

LIU Ming received the M.Sc. degree from Harbin Engineering University, China in 2011. He is currently a senior engineer with ZTE Corporation. His research interests include object detection, tracking and recognition.



High Speed Polarization-Division Multiplexing Transmissions Based on the Nonlinear Fourier Transform

Abstract: Polarization-division multiplexing (PDM) with modulation in the nonlinear frequency domain consisting of the discrete and/or continuous spectrum has been recently regarded as a useful method to be utilized in optical fiber communication system. It can compensate the optical fiber nonlinearity based on the nonlinear Fourier transform (NFT). In this paper, we combine PDM with the method of nonlinear frequency division multiplexing (NFDM) and demonstrate the achievable transmission rate by increasing the number of multiplexing nonlinear channels. For the selected subcarriers (i.e. 32, 64, and 128), the transmission rates are 64 Gbit/s, 76.8 Gbit/s, and 109.7 Gbit/s respectively by applying 64-quadrature amplitude modulation (64-QAM) on the nonlinear continuous spectrum. For the transmission distance shorter than 1 200 km, the transmission rate of 128-NFDM PDM system can even reach up to 153.6 Gbit/s.

Keywords: fiber optics communications; multiplexing; nonlinear optical signal processing

WANG Jia, ZHAO Yilong, HUANG Xin, and HE Guangqiang

(State Key Laboratory of Advanced Optical Communication Systems and Networks, Electronic Engineering Department, Shanghai Jiao Tong University, Shanghai 200240, China)

DOI: 10.12142/ZTECOM.201903008

<http://kns.cnki.net/kcms/detail/34.1294.TN.20190628.1108.002.html>, published online June 28, 2019

Manuscript received: 2018-08-31

1 Introduction

With the development of modern information technology, optical fiber communication plays an increasingly important role in information transmission. However, the presence of nonlinearity in optical fiber greatly hinders the improvement of transmission capacity in optical fiber communication systems [1]. When a single-channel transmission rate reaches 50 Gbit/s, nonlinearity has become a major factor affecting system performance [2].

In 1993, HASEGAWA and NYU [3] first used the nonlinear Fourier transform (NFT) to transform the signal from the time domain to the nonlinear frequency domain consisting of nonlinear discrete spectrum and/or nonlinear continuous spectrum for coding and multiplexing, which can overcome the communication constraints caused by optical fiber nonlinearity. The method is now called eigenvalue communication because of the invariance of eigenvalues. With the development of coherent detection and digital signal processing (DSP) technology recently, any complex signals can be generated and processed in communication systems. Thus, the optical fiber nonlinearity

compensation coherent optical communication system based on the NFT algorithm has attracted more and more attention [4], [5].

The nonlinearity compensation algorithm based on NFT has the advantages of stable transmission and independent algorithm complexity of transmission distance. It does not affect the mixed use of various linear modulation methods [6], [7]. Taking into consideration overcoming the nonlinear effect, the transmission speed of the systems can be further improved. In 2014, YOUSEFI and KSCHISCHANG carried out theoretical modeling and numerical derivation of the NFT algorithm, and put forward the feasibility of nonlinear frequency division multiplexing (NFDM) [8]–[10]. In 2017, GOOSSENS and YOUSEFI did the similar work in polarization-division multiplexing (PDM) system and then applied it to PDM-NFDM simulation system, achieving the data rate of 44.8 Gbit/s by 112 nonlinear subcarriers and 64-quadrature amplitude modulation (64-QAM) format [11].

In this paper, the transmitted data are modulated only on the nonlinear continuous spectrum through 64-QAM, i.e. there are 6 bits per symbol. We also combine the PDM with NFDM (PDM-NFDM) and demonstrate the achievable transmission

rate by increasing the number of multiplexing nonlinear channels [2] through simulation, followed by the analysis of the transmission performance of different schemes. The simulation results show the data rate is obviously increased and is more than 100 Gbit/s by utilizing 128 subcarriers, with a Q-factor about 8.19 dB at 960 km away.

2 Theory

2.1 Channel Model

The propagation of light in two polarizations in a standard single mode fiber (SSMF) is described by the coupled nonlinear Schrödinger equation (CNLSE) [11], [12]. Without considering the noise and polarization mode dispersion (PMD), the model can be expressed as following [13]–[15]:

$$\frac{\partial \vec{Q}(\tau, l)}{\partial l} = -\frac{\alpha}{2} \vec{Q}(\tau, l) - j \frac{\beta_2}{2} \frac{\partial^2 \vec{Q}(\tau, l)}{\partial \tau^2} + j \frac{8}{9} \gamma \left\| \vec{Q}(\tau, l) \right\|^2 \vec{Q}(\tau, l), \quad (1)$$

where $\vec{Q}(\tau, l)$ is the 2×1 Jones vector containing the complex envelopes $Q_1(\tau, l)$ and $Q_2(\tau, l)$ of the signal components in two polarizations, τ represents time, l represents the distance along the optical fiber, α denotes the loss coefficient of the optical fiber, and β_2 and γ are constants denote fiber dispersion and Kerr nonlinearity respectively.

The system (1) is obviously not integrable because of the fiber loss term. Since the NFT is based on the integrability of NLSE [8]–[10], measures should be taken to avoid the effect of fiber attenuation. Here we use the Erbium Doped Fiber Amplifier (EDFA) to compensate the fiber loss in transmission system and adopt the idea of the lossless path-average (LPA) in [16] to derive the following model without loss term [17], [18]:

$$\frac{\partial \bar{\vec{Q}}(\tau, l)}{\partial l} = -j \frac{\beta_2}{2} \frac{\partial^2 \bar{\vec{Q}}(\tau, l)}{\partial \tau^2} + j \frac{8}{9} \gamma_1 \left\| \bar{\vec{Q}}(\tau, l) \right\|^2 \bar{\vec{Q}}(\tau, l), \quad (2)$$

where the $\bar{\vec{Q}}$ and γ_1 is defined as:

$$\bar{\vec{Q}} = \vec{Q} \sqrt{G(l)}, \quad (3)$$

$$\gamma_1 = \gamma \left(\frac{1}{L_{span}} \int_0^{L_{span}} G(x) dx \right), \quad (4)$$

where $G(x) = e^{2 \int_0^x g(l) dl}$ and $g(l) = -\frac{\alpha}{2}$.

For simplicity, it is useful to normalize (2). Let:

$$P = \frac{|\beta_2|}{\frac{8}{9} \gamma_1 T_0^2}, T_0 = \frac{T_{FWHM}}{2 \ln(1 + \sqrt{2})}, Z_0 = \frac{T_0^2}{|\beta_2|}, \quad (5)$$

where T_{FWHM} which represents the full width at half maximum of the pulse is a free parameter relating to the bitrate [19], [20]. Introducing the normalized variables:

$$\bar{q} = \frac{\bar{\vec{Q}}}{\sqrt{P}}, t = \frac{\tau}{T_0}, Z = \frac{l}{Z_0}, \quad (6)$$

the normalized CNLSE is derived as follows [8]–[10]:

$$j \frac{\partial \bar{q}(t, z)}{\partial z} = -\frac{1}{2} \frac{\partial^2 \bar{q}(t, z)}{\partial t^2} - \left\| \bar{q}(t, z) \right\|^2 \bar{q}(t, z). \quad (7)$$

Through the nonlinear Fourier transform, the difficulty of analysis can be simplified and the nonlinear effect can be compensated.

2.2 PDM-NFDM Scheme

According to the theory of two-dimensional electrical signals in [21], $\bar{q}(t, z)$ in (7) can be represented by a series of scattering coefficients or data (also called nonlinear spectrum). In general, the nonlinear spectrum consists of nonlinear discrete spectrum and nonlinear continuous spectrum. The process of obtaining the corresponding nonlinear spectrum from a time domain signal is called NFT, and the inverse process is known as inverse nonlinear Fourier transform (INFT). For a two-dimensional signal, its nonlinear continuous spectrum is defined as [21]–[23]:

$$q_i(\lambda) = \frac{b_i(\lambda)}{a(\lambda)}, \lambda \in \mathbb{R}, i = 1, 2, \quad (8)$$

where $a(\lambda)$, $b_1(\lambda)$, and $b_2(\lambda)$ are called the nonlinear Fourier coefficients [24]. If necessary, the reader can refer to [11] for calculation of these coefficients in detail. The evolution of the nonlinear continuous spectrum along the optical fiber can be expressed as [2], [17], [25]–[27]:

$$q_i(\lambda, z) = q_i(\lambda, 0) \cdot e^{2i\lambda^2 z}, i = 1, 2. \quad (9)$$

It is observed that the effect of fiber dispersion and Kerr nonlinearity on the signal's nonlinear spectrum can be regarded as a simple phase rotation. The principle of eigenvalue communication is actually to use this property to compensate the fiber nonlinearity and dispersion in the nonlinear frequency domain, so is the PDM-NFDM transmission system. At the transmitter of the transmission scheme, the bit information to be transmitted is encoded on the nonlinear spectrum which can be discrete spectrum and/or continuous spectrum [28]–[33]. Subsequently, a specific time domain signal can be calculated through INFT and then be sent into the optical fiber. At the receiver, the received time domain signal is then processed by NFT to obtain the corresponding nonlinear spectrum, $q(\lambda, z)$, when the transmission distance is z . According to (9), the initial nonlinear continuous spectrum $q(\lambda, 0)$ can be recovered through a phase compensation and thus we can obtain the bit information by decoding the recovered nonlinear spectrum. It can be found that the complexity of the nonlinear compensation algorithm based on NFT in the coherent optical communication system is independent of the transmission distance and does not affect the mixed use of various linear modulation

schemes, so it can be utilized in NFDM, PDM and other transmission systems.

3 Simulation

Fig. 1 shows the scheme of simulation system and data processing process based on the theories above. Details about the specific data processing can be found in the following sections.

3.1 Transmitter

This part mainly introduces the process of data processing on the transmitter. For the generated transmitted bits, we first start with a string-to-parallel transformation for data mapping of 64-QAM format. In this way, each 6 bits of data will correspond to a special amplitude and phase in total 64 cases. Next, we modulate the transmitted information on the nonlinear continuous spectrum. For each polarization component, the spectrum is shown in the following formula [17], [25], [29].

$$U_i(\lambda, m) = A \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} C_{i,m,k} \frac{\sin(\lambda T_p + k\pi)}{\lambda T_p + k\pi} e^{-2jm\lambda T_1}, i = 1, 2. \quad (10)$$

In the upper formula, λ is the nonlinear frequency, m is the data block index, $U_i(\lambda, m)$ denotes the synthetic continuous spectrum, A is the power control parameter, N represents the number of multiplexed nonlinear channels, and $C_{i,m,k}$ is a complex number drawn from the 64-QAM constellation diagram. In addition, $T_p = 2ns/T_0$ represents the useful data duration, where T_0 is the parameter given in (5) and T_1 denotes the total

data duration taking into consideration the delay in the system transmission. For the selected subcarriers (i. e. 32, 64, and 128), the corresponding T_1 are $6ns/T_0$, $10ns/T_0$, and $14ns/T_0$ respectively. Thus, when the baud rate is set to 0.5 Gbaud/s, the transmission rates of the PDM-NFDM systems are 64 Gbit/s, 76.8 Gbit/s and 109.7 Gbit/s, respectively. According to the theory in [11] and [34], $U_i(\lambda, m)$ obtained above is actually within the area called U-domain, so we should subsequently convert the data from U-domain to nonlinear Fourier domain. The relationship between them is [34]:

$$q_i(\lambda) = \sqrt{e^{|U_i(\lambda)|^2} - 1} \cdot e^{j\angle(U_i(\lambda))}, i = 1, 2. \quad (11)$$

After that, we can calculate the two-dimensional time domain signal through INFT, inversely normalize the signal, carry out the electro-optic conversion, and then send the corresponding optical signal together with the protection interval to the optical fiber.

3.2 Optical Fiber

The standard single mode fiber is used here for simulation. The dispersion and nonlinear coefficients of the fiber are calculated by the following formulas [19]–[20]:

$$\beta_2 = -\frac{\lambda^2}{2\pi c} D_\lambda, \quad (12)$$

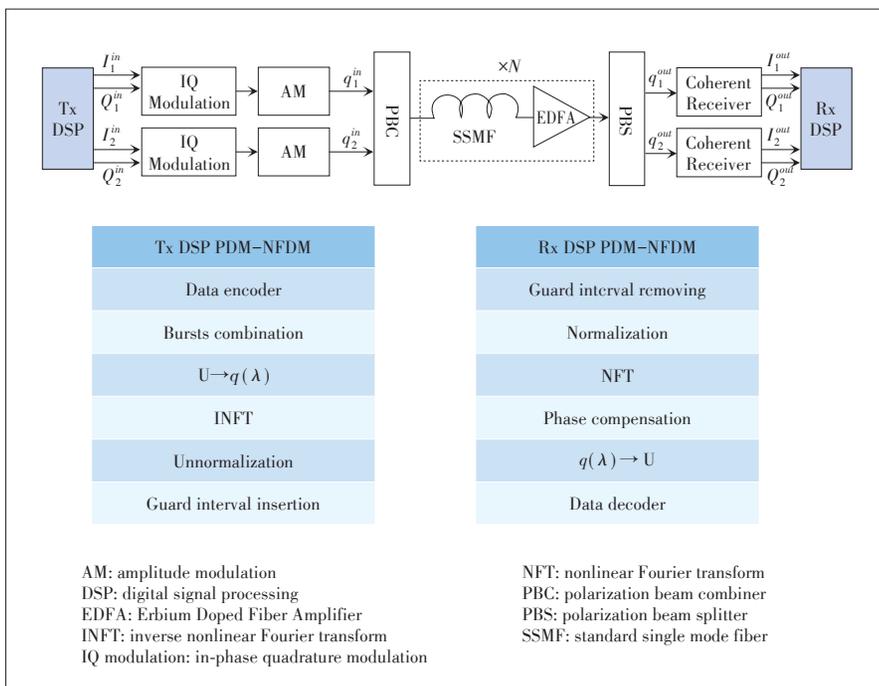
$$\gamma = \frac{2\pi n_2 f_{ref}}{c A_{eff}}, \quad (13)$$

where λ is the wavelength of the reference light, f_{ref} is the reference frequency, D_λ denotes the dispersion coefficient in terms of wavelength, c is the velocity of light in vacuum, n_2 represents the nonlinear refractive index, and A_{eff} refers to the effective core area of the fiber for nonlinearity calculations. In our simulation, the parameters related to β_2 and γ are set as: $f_{ref} = 193.1$ THz, $D_\lambda = 16e^{-6}$ s/m², $n^2 = 2.6e^{-20}$ m²/W, and $A_{eff} = 80.0e^{-12}$ m² and then useful parameters can be calculated through (12) and (13).

As mentioned in Section 2.1, the system (1) is obviously not integrable in the presence of fiber loss and thus NFT does not apply. For the sake of lossless transmission model, we utilize EDFAs to equalize the fiber attenuation each 80 km and the fiber loss is set as 0.2 dB/km. The length of total optical fiber could be increased by loop links.

3.3 Receiver

At the receiver, we first demodulate the received optical signal by a coherent demodulator and subsequently carry on the



▲ **Figure 1.** PDM-NFDM transmission system with data processing process in the transmitter and receiver.

optic-electro conversion. By this way, we can get the time domain electrical signal of two polarization components. Before we carry on the NFT to the two-dimensional time domain signal, we first normalize it according to (5). For the method that the evolution of nonlinear spectrum is just a simple phase rotation as expressed in (9), we then compensate the effect of fiber dispersion and Kerr nonlinearity by this way to obtain $q(\lambda, 0)$. Later, we refer to (11) to recovery the data from nonlinear Fourier domain to U-domain, i.e. the formula in (11) will be rewritten as:

$$U_i(\lambda) = \sqrt{\log(1 + |q_i(\lambda)|^2)} \cdot e^{i\angle(q_i(\lambda))}, i = 1, 2. \quad (14)$$

Due to the presence of phase noise and frequency offset, the carrier recovery algorithm [35] is necessary. Then the transmission performance of the system can be analyzed by drawing the 64-QAM constellation diagram, decoding the symbols and calculating the symbol error rate (SER) or Q-factor.

4 Results

In this paper, we first demonstrate the achievable transmission rate by increasing the number of multiplexing nonlinear channels through simulation. For the subcarriers utilized (i.e., 32, 64, and 128), the transmission rates are 64 Gbit/s, 76.8 Gbit/s and 109.7 Gbit/s respectively by 64-QAM on the nonlinear continuous spectrum. We then show it is necessary to eliminate the frequency offset and phase noise in the NFDM transmission systems, especially when the number of multiplexing channels is large. Finally, we mainly analyze the high-speed transmission performance of system with 128 subcarriers, in which case the transmission rate of the PDM-NFDM system reaches 109.7 Gbit/s.

4.1 Reachable Transmission Rate

Table 1 shows some parameters we used in different PDM-NFDM transmission systems. The achieved spectral efficiency of each system and the corresponding SER with the launched power around -3 dBm (-3.22 dBm for the case of 32-subcarriers, -2.75 dBm for 64-subcarriers and -3.07 dBm for 128-subcarriers) are also given in the table. It is worth noticing that the launched power here is not the optimal power, which means that each transmission system may be able to achieve a more reliable transmission performance in other cases. In this sec-

tion, just phase noise and additive white Gaussian noise (AWGN) are taken into consideration and thus the optical signal to noise ratio (OSNR) used here are all over 18 dB.

It can be inferred from Table 1 that the transmission rate of optical fiber communication system increases significantly with the increase of multiplexed nonlinear channels, but a higher spectral efficiency is achieved in the PDM-NFDM system with 32 subcarriers than other cases, which is largely due to the effect of optical fiber dispersion. It is well known that different frequency components of the optical signal transmitted in the optical fiber will travel at different speeds, which inevitably cause signal distortion after reaching a certain distance. The signal distortion is shown as pulse broadening in the time domain. The phenomenon is certain to get serious with the increase of the number of multiplexing channels or signal bandwidth and then a longer zero guard interval during the long distance transmission is required, thus decreasing the effective transmission rate correspondingly. Therefore, the guard interval used to reducing the inter-symbol interference is an important influencing factor to the balance of transmission rate and transmission distance.

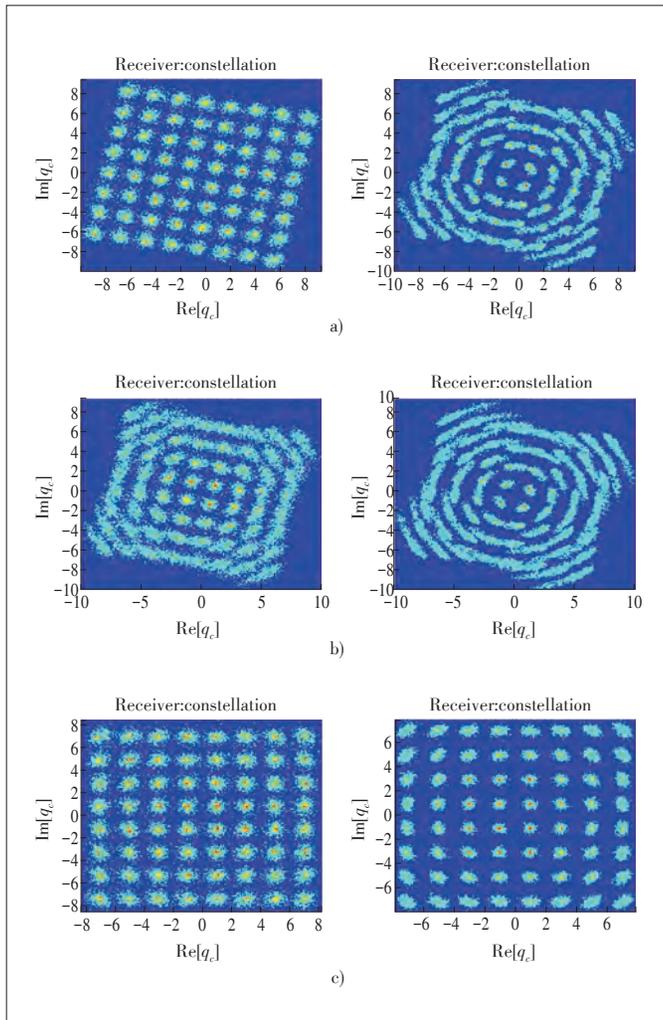
4.2 The Effect of Frequency Offset

Carrier synchronization is a very important problem in traditional orthogonal frequency division multiplexing (OFDM) systems, as is the case with NFDM systems. The carrier frequency offset caused by devices or nonlinear channels or other factors can damage the orthogonality between the subcarriers, thus introducing the subcarriers interference, as well as the rotation of the signal constellation diagram. Unsurprisingly, these effects will cause the decision error of a demodulation signal to a large extent and thus the carrier frequency offset should be compensated in the coherent NFDM systems. Here, the effect of frequency offset as well as the influence of laser linewidth is given through the simulation results in both cases where the number of subcarriers is 32 and 128, just as shown in **Fig. 2**. It should be noted that the launched power is approximately -3.07 dBm and the transmission distance is 960 km for all the simulation systems in this section. At the same time, in order to clearly observe the effect of frequency offset and phase noise on transmission performance, the additive noise of the channel is not considered in this section.

We simply set the laser linewidth in the simulation system

▼ **Table 1. The parameters used in different PDM-NFDM transmission systems**

| Channels | Baud rate /Gbaud | Guard interval /ns | Modulation format | Bandwidth /GHz | Transmission rate /(Gbit/s) | Spectral efficiency /(bit/s/Hz) | Distance /km | Q-factor |
|----------|------------------|--------------------|-------------------|----------------|-----------------------------|---------------------------------|--------------|----------|
| 32 | 0.5 | 4 | 64-QAM | 16 | 64 | 4 | 1 200 | 6.234 |
| 64 | 0.5 | 8 | 64-QAM | 32 | 76.8 | 2.4 | 1 200 | 7.405 |
| 128 | 0.5 | 12 | 64-QAM | 64 | 109.7 | 1.71 | 1 200 | 6.751 |
| 128 | 0.5 | 8 | 64-QAM | 64 | 153.6 | 2.4 | 960 | 8.195 |



▲ Figure 2. The constellation diagram at the distance of 960 km for the case of 32 subcarriers (left column) and 128 subcarriers (right column); a) linewidth=0, w/o carrier recovery; b) linewidth=1 kHz, w/o carrier recovery; c) linewidth=1 kHz, with carrier recovery.

to 0 when the effect of frequency offset is studied only. At the receiver, we obtain the constellation diagram as Fig. 2a after utilizing the theory of NFT for compensation of fiber dispersion and Kerr nonlinearity. It can be inferred that the rotation of signal constellation diagrams here is due to the frequency offset that is estimated to be about 0.25 GHz for the case of 32 subcarriers and 0.36 GHz for 128 subcarriers, which severely affect the data decision. To be exact, the symbol error rate is calculated as 0.5254 in 32-NFDM PDM system and 0.3721 in 128-NFDM PDM system, neither of which can be served as an effective transmission of optical fiber communication system. Furthermore, the influence of laser linewidth is increased, i.e. the corresponding parameters in the simulation system are set to 1 kHz, and the receiver results are finally obtained as shown in Fig. 2b. As we can see from the diagrams, the additional phase noise is introduced to the transmission systems because

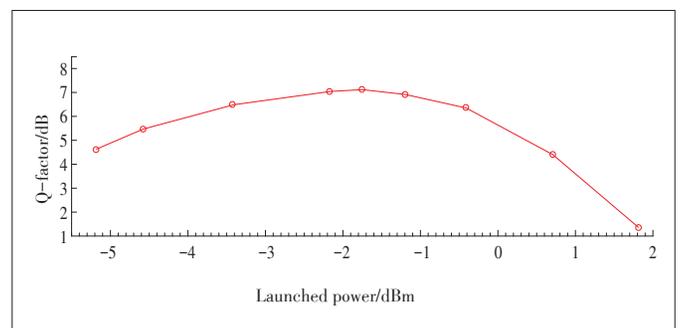
of the laser linewidth, which increases the SER to 0.5681 in the case of 32 subcarriers and 0.5945 for 128 subcarriers, further degrading the transmission performance of the system. For the sake of effective transmission, we then adopt the carrier recovery algorithm [35] at the receiver after the compensation for fiber nonlinearity based on NFT and successfully improve the transmission performance (Fig. 2c). Under this situation, the SERs are calculated as $1.01e-2$ and $1.53e-5$, respectively.

We can conclude from the figures that the PDM - NFDM scheme based on NFT well compensates the negative effects of fiber nonlinearity on the optical fiber communication by simultaneously solving the problem of unsynchronization at both transmitters and receivers of the transmission systems. At the same time, it is obvious that the 128-NFDM PDM transmission system outperforms the 32-NFDM PDM case in this section, which is largely due to the launched power. In fact, different optimum launched powers will be applied to different transmission systems. However, we carry on the above simulation cases at the same launched power approximately -3.07 dBm for simplicity, which means that this value is not be in the best performance range for 32-NFDM PDM system while near the best performance range for 128-NFDM PDM system. As an example, the optimal power for 128-NFDM PDM system is given in the next section.

4.3 The Optimum Launched Power

In this part, we conduct the simulation of the relationship between Q-factor and the launched power by the PDM - NFDM transmission system with 128 multiplexed channels. Fig. 3 shows the corresponding results with all the transmission distances at 1 200 km and OSNR over 18 dB.

As shown in Fig. 3, the Q-factor of this specific fiber nonlinear compensation communication system based on NFT reaches the maximum close to 7 dB when the launched optical power is around -1.75 dBm. That is to say, there is a choice of optimum launched power for this 128-NFDM PDM transmission system to achieve its best compensation performance, which can be utilized as a notice in engineering applications. In addition, we also give the comparison of the two received constellation diagrams with the launched power of -1.7563 dBm and



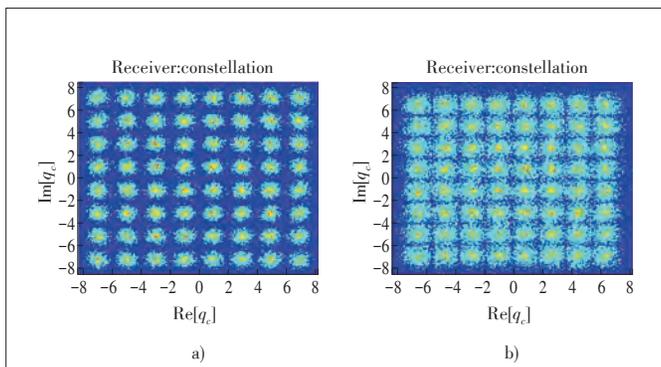
▲ Figure 3. Q - factor versus launched power with the case of 128 subcarriers.

1.8092 dBm respectively for a more obvious display (**Fig. 4**).

5 Conclusions

We present the process of modulating the necessary transmitted data on the nonlinear continuous spectrum through 64-QAM format, transmitting them through the optical fiber communication system and then compensating the effect of Kerr nonlinearity at the receiver. We also combine the PDM with NFDM technology to improve the transmission capacity and demonstrate the achievable transmission rate by increasing the number of multiplexing nonlinear channels through simulation. Moreover, we analyze the transmission performance of the different schemes. As a result, for the selected subcarriers (i.e., 32, 64, and 128), the transmission rates are 64 Gbit/s, 76.8 Gbit/s and 109.7 Gbit/s respectively. For the transmission distance shorter than 1 200 km, the transmission rate of 128-NFDM PDM system can even reach up to 153.6 Gbit/s. In addition to the effects of fiber dispersion and nonlinearity which can be compensated by the theory of NFT, carrier frequency offset and phase noise can degrade the system performance severely, especially when the number of multiplied nonlinear channels is large. It is demonstrated that a better transmission performance could be obtained for the PDM-NFDM transmission system by taking measures to compensate the frequency offset and phase noise simultaneously.

Aside from the launched power, many other factors will also have effects on the performance of this nonlinearity compensation system unavoidably, various noise included. In this paper, we just carry on a research related to the AWGN and phase noise, thus we still have a lot to do in the future to improve the transmission performance of high-speed optical fiber communication with various noise taken into account.



▲ **Figure 4.** Constellation diagram at the receiver a) with the launched power at -1.7563 dBm; b) with the launched power at 1.8092 dBm.

References

- [1] ESSIAMBRE R J, KRAMER G, WINZER P J, et al. Capacity Limits of Optical Fiber Networks [J]. *Journal of Lightwave Technology*, 2010, 28(4): 662–701. DOI: 10.1109/JLT.2009.2039464
- [2] SON T Le, VAHID A, HENNING B. High Speed Precompensated Nonlinear Fre-

- quency-Division Multiplexed Transmissions [J]. *Journal of Lightwave Technology*, 2018, 36(6):1296–1303. DOI: 10.1109/JLT.2017.2787185
- [3] HASEGAWA A, NYU T. Eigenvalue Communication [J]. *Journal of Lightwave Technology*, 1993, 11(3): 395–399. DOI: 10.1109/50.219570
- [4] PRILEPSKY J E, DEREYANKO S A. Nonlinear Inverse Synthesis and Eigenvalue Division Multiplexing in Optical Fiber Channels [J]. *Physical Review Letters*, 2014, 113(1): 013901. DOI: 10.1103/PhysRevLett.113.013901
- [5] SON T L, PRILEPSKY J E, ROSA P, et al. Nonlinear Inverse Synthesis for Optical Links with Distributed Raman Amplification [J]. *Journal of Lightwave Technology*, 2016, 34(8): 1778–1786. DOI: 10.1109/JLT.2015.2511084
- [6] TURITSYNA E G, TURITSYN S K. Digital Signal Processing Based on Inverse Scattering Transform [J]. *Optics Letters*, 2013, 38(20): 4186–4188. DOI: 10.1364/OL.38.004186.
- [7] SON T Le, PRILEPSKY J E, TURITSYN S K. Nonlinear Inverse Synthesis for High Spectral Efficiency Transmission in Optical Fibers [J]. *Optics Express*, 2014, 22(22): 26720–26741. DOI: 10.1364/OE.22.026720
- [8] YOUSEFI M I, KSCHISCHANG F R. Information Transmission Using the Nonlinear Fourier Transform, Part I: Mathematical Tools [J]. *IEEE Transactions on Information Theory*, 2014, 60(7): 4312–4328. DOI: 10.1109/TIT.2014.2321143
- [9] YOUSEFI M I, KSCHISCHANG F R. Information Transmission Using the Nonlinear Fourier Transform, Part II: Numerical Methods [J]. *IEEE Transactions on Information Theory*, 2014, 60(7): 4329–4345. DOI: 10.1109/TIT.2014.2321151
- [10] YOUSEFI M I, KSCHISCHANG F R. Information Transmission Using the Nonlinear Fourier Transform, Part III: Spectrum Modulation [J]. *IEEE Transactions on Information Theory*, 2014, 60(7): 4346–4369. DOI: 10.1109/TIT.2014.2321155
- [11] GOOSSENS J W, YOUSEFI M I, JAOUEN Y, et al. Polarization-Division Multiplexing Based on the Nonlinear Fourier Transform [J]. *Optics Express*, 2017, 25(22): 26437–26452. DOI: 10.1364/OE.25.026437
- [12] MENYUK C R. Pulse Propagation in An Elliptically Birefringent Kerr Medium [J]. *IEEE Journal of Quantum Electronics*, 1989, 25(12): 2674–2682. DOI: 10.1109/3.40656
- [13] WAI P K A, MENYUK C R. Polarization Mode Dispersion, Decorrelation, and Diffusion in Optical Fibers with Randomly Varying Birefringence [J]. *Journal of Lightwave Technology*, 1996, 14(2): 148–157. DOI: 10.1109/50.482256
- [14] MARCUSE D, MENYUK C R, WAI P K A. Application of the Manakov-PMD Equation to Studies of Signal Propagation in Optical Fibers with Randomly Varying Birefringence [J]. *Journal of Lightwave Technology*, 1997, 15(9): 1735–1746. DOI: 10.1109/50.622902
- [15] MENYUK C R, MARKS B S. Interaction of Polarization Mode Dispersion and Nonlinearity in Optical Fiber Transmission Systems [J]. *Journal of Lightwave Technology*, 2006, 24(7): 2806–2826. DOI: 10.1109/JLT.2006.875953
- [16] SON T Le, PRILEPSKY J E, TURITSYN S K. Nonlinear Inverse Synthesis Technique for Optical Links with Lumped Amplification [J]. *Optics Express*, 2015, 23(7): 26720–26741. DOI: 10.1364/OE.23.008317
- [17] SON T Le, HENNING B. 640.5 Gbaud Nonlinear Frequency Division Multiplexed Transmissions with High Order Modulation Formats [J]. *Journal of Lightwave Technology*, 2017, 35(17): 3692–3698. DOI: 10.1109/JLT.2017.2718105
- [18] HASEGAWA A, KODAMA Y. *Solitons in Optical Communications* [M]. Oxford, U.K.: Oxford Univ. Press, 1995
- [19] VPIsystems. *Photonic Modules Reference Manual* [Z]. Holmudd, USA, 2002: 337–350
- [20] MYNBAEV D K, SCHEINER L L. *Fiber-Optic Communications Technology* [M]. Cambridge, USA: Academic press, 2008
- [21] MANAKOV S V. On the Theory of Two-Dimensional Stationary Self-focusing of Electromagnetic Waves [J]. *Soviet Journal of Experimental and Theoretical Physics*, 1974, 38(2): 248–253.
- [22] DEGASPERIS A, LOMBARDO S. *Integrability in Action: Solitons, Instability and Rogue Waves* [M]. New York City, USA: Springer International Publishing, 2016
- [23] GAIARIN S, PEREGO A M, SILVA E P D, et al. Dual-Polarization Nonlinear Fourier Transform-Based Optical Communication System [J]. *Optica*, 2018, 5(3): 263–270. DOI: 10.1364/OPTICA.5.000263
- [24] GUI T, CHAN T H, LU C, et al. Alternative Decoding Methods for Optical Communications Based on Nonlinear Fourier Transform [J]. *Journal of Lightwave Technology*, 2017, 35(9): 1542–1550. DOI: 10.1109/JLT.2017.2654493
- [25] SON T L, VAHID A, HENNING B. Nonlinear Signal Multiplexing for Communication Beyond the Kerr Nonlinearity Limit [J]. *Nature Photonics*, 2017, 11: 570–577. DOI: 10.1038/NPHOTON.2017.118
- [26] TURITSYN S K, PRILEPSKY J E, SON T Le, et al. Nonlinear Fourier Transform for Optical Data Processing and Transmission: Advances and Perspectives

➔ **To P. 62**



A Service-Based Intelligent Time-Domain and Spectral-Domain Flow Aggregation in IP-over-EON Based on SDON

Abstract: The rapid growth of IP traffic has contributed to wide deployment of optical devices in elastic optical network. However, the passband shape of wavelength selective switches (WSSs) that are used in reconfigurable optical add-drop multiplexer (ROADM)/optical cross connect (OXC) is not ideal, causing the narrowing of spectrum. Spectral narrowing will lead to signal impairment. Therefore, guard-bands need to be inserted between adjacent paths which will cause the waste of resources. In this paper, we propose a service-based intelligent aggregation node selection and area division (ANS-AD) algorithm. For the rationality of the aggregation node selection, the ANS-AD algorithm chooses the aggregation nodes according to historical traffic information based on big data analysis. Then the ANS-AD algorithm divides the topology into areas according to the result of the aggregation node selection. Based on the ANS-AD algorithm, we propose a time-domain and spectral-domain flow aggregation (TS-FA) algorithm. For the purpose of reducing resources' waste, the TS-FA algorithm attempts to reduce the insertion of guard-bands by time-domain and spectral-domain flow aggregation. Moreover, we design a time-domain and spectral-domain flow aggregation module on software defined optical network (SDON) architecture. Finally, a simulation is designed to evaluate the performance of the proposed algorithms and the results show that our proposed algorithms can effectively reduce the resource waste.

Keywords: IP-over-EON; time-domain; spectral-domain; flow aggregation; big data analysis; SDON

NI Dong¹, LI Hui¹, JI Yuefeng¹,
LI Hongbiao², and ZHU Yinan²

(1. Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Wireline Product Operation Division, ZTE Corporation, Shenzhen, Guangdong 518057, China)

DOI: 10.12142/ZTECOM.201903009

[http://kns.cnki.net/kcms/detail/34.1294.](http://kns.cnki.net/kcms/detail/34.1294.TN.20190923.1804.002.html)

TN.20190923.1804.002.html, published online September 23, 2019

Manuscript received: 2019-07-10

1 Introduction

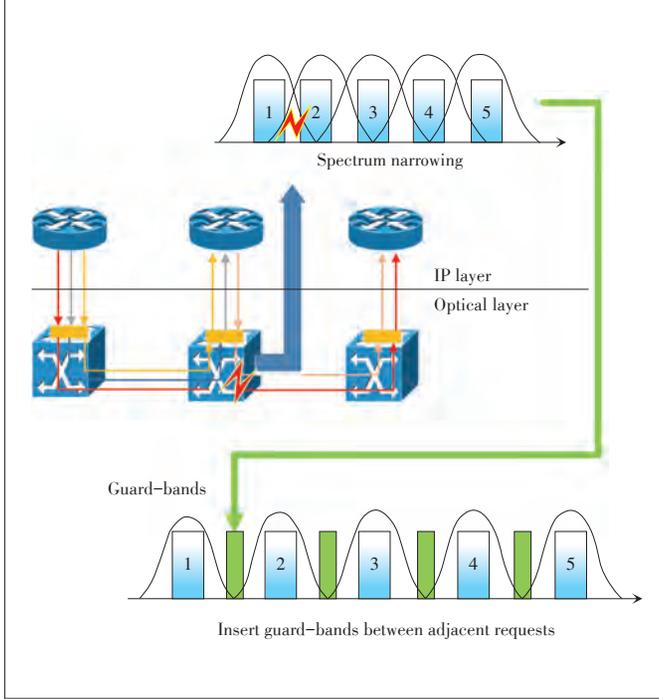
With the emerging requests for Internet applications, cloud services, big data, and more, the 5G-oriented transport network faces many technical challenges such as high capacity, low latency, network fragmentation, and intelligence [1]. IP over elastic optical networks (IP-over-EON) is a promising multilayer network architecture that can achieve efficient IP traffic accommodation at the optical layer [2].

Software defined optical network (SDON) decouples the application plane, data plane and control plane, and offers a centralized control of all networking elements via software pro-

gramming through a controller [3]. The implementation of new allocation method based on SDON just needs to add the corresponding modules in the SDON controller, not affecting the current network and allocation method, which is an important solution to meet the intelligent request of the next-generation optical transport network and has a wide range of application prospects in 5G.

Now, the rapid growth of IP traffic has contributed to wide deployment of optical devices in EON [4]. However, the passband shape of wavelength selective switches (WSSs) which are used in reconfigurable optical add-drop multiplexer (ROADM)/optical cross connect (OXC) is not ideal, causing the narrowing of spectrum [5]. Spectral narrowing will lead to signal impairment. Therefore, guard-bands need to be inserted between adjacent paths. As shown in **Fig. 1**, there are five requests. Each request occupies three spectral slots and the guard-bands between adjacent requests occupies two spectral slots. The ratio

This work has been funded by ZTE Industry-Academia-Research Cooperation Funds under Grant No. 2017110031005226.



▲ Figure 1. Spectrum narrowing may occur when adding/dropping optical paths.

of the guard-bands bandwidth to the total bandwidth is $2 \times 4 / (3 \times 5 + 2 \times 4) = 34.8\%$. When a large number of requests emerge in the network, it is inefficient to construct a light-path for each request [6], which needs to insert a lot of guard-bands and causes the waste of resources.

In order to meet the high-capacity requirements in 5G, it is necessary to save spectrum resources. An effective method is to reduce the insertion of guard-bands. In [7], the authors propose a traffic dispatching algorithm which chooses two aggregation nodes according to the topology with the largest degree. Then they make sure the requests which go through the aggregation nodes have the same routing and no guard-bands are inserted in the candidate path. When the services of a network are unevenly distributed, this method of aggregation node selection is unreasonable.

In this paper, we focus on a service-based intelligent time-domain and spectral-domain flow aggregation in IP-over-EON based on SDON. First, we propose a service-based intelligent aggregation node selection and area division (ANS-AD) algorithm based on big data analysis idea. Then, according to the result of the aggregation node selection and area division, we propose a time-domain and spectral-domain flow aggregation (TS-FA) algorithm based on SDON architecture. In order to obtain the most suitable number of aggregation nodes, we simulate the resource usage under different aggregation ratios (the number of aggregation nodes/the total number of nodes in the topology) and obtain the aggregation nodes under the minimum

resource usage.

2 Service-Based Intelligent Aggregation Node Selection and Area Division Algorithm

2.1 Service-Based Intelligent Aggregation Node Selection Algorithm

With the development of big data, it has generated enormous publicity at home and abroad. Big Data Analytics (BDA) is at the heart of big data ideas and methods [8]. It refers to the analysis of a large variety of data with real content and the process of finding hidden patterns, unknown correlations and other useful information that can help decision-making.

Based on the idea of BDA, we propose a service-based intelligent aggregation node selection algorithm. The approach is to collect a huge amount of historical service information from the network and extract the path information of each service.

Suppose the set of historical services is

$$R = \{ R_1, R_2, R_3, \dots, R_i, \dots, R_M \}, \quad (1)$$

where M presents the number of historical services, while R_i is the i th service.

We represent the path node set of R_i as

$$Path_{R_i} = \{ s_i, v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{im}, t_i \}, \quad (2)$$

where s_i is the source node of R_i , v_{ij} is the intermediate node of R_i , and t_i is the destination node of R_i .

Then we calculate the historical traffic of each node according to (3):

$$his_tra_k = 2 \times \sum_{j=1}^M num_k_intra + 1 \times \sum_{j=1}^M num_k_sord, \quad (3)$$

where if node k is the intermediate node of the j th path, $num_k_intra = 1$, otherwise $num_k_intra = 0$. If node k is the source or destination node of the j th path, $num_k_sord = 1$, otherwise $num_k_sord = 0$.

From (3) we can get each node's his_tra and sort the nodes in descending order of his_tra . We select the top ranked nodes as aggregation nodes. Suppose the aggregation node set is

$$V_{agg} = \{ v_{1agg}, v_{2agg}, v_{3agg}, \dots, v_{iagg}, \dots, v_{kagg} \}, \quad (4)$$

where k indicates the number of aggregation nodes and v_{iagg} represents the aggregation node.

2.2 Area Division Algorithm Based on Aggregation Node Selection

After the aggregation node selection, we divide the topology into areas. The number of areas is equal to the number of aggregation nodes. The initial set of nodes for each area is represented as

$$\begin{aligned} area_1 &= \{ v_{1agg} \}, area_2 = \{ v_{2agg} \}, \dots, area_i = \{ v_{iagg} \}, \dots, \\ area_k &= \{ v_{kagg} \}, \end{aligned} \quad (5)$$

where $area_i$ represents the area.

Assume that the non-aggregation node set is

$$V_{non_agg} = \{v_{1non}, v_{2non}, v_{3non}, \dots, v_{jnon}, \dots, v_{N-Knon}\}, \quad (6)$$

where N is the number of nodes in the topology, K indicates the number of aggregation nodes, and v_{jnon} represents the j th non-aggregation node.

The distance from v_{jnon} to v_{iagg} is recorded as

$$len < path_{v_{jnon}, v_{iagg}} >, \quad (7)$$

where $path_{v_{jnon}, v_{iagg}}$ is the shortest path from v_{jnon} to v_{iagg} .

The minimum distance from v_{jnon} to each aggregation node is calculated as

$$\min \left\{ len < path_{v_{jnon}, v_{1agg}} >, len < path_{v_{jnon}, v_{2agg}} >, \dots, len < path_{v_{jnon}, v_{Kagg}} > \right\}. \quad (8)$$

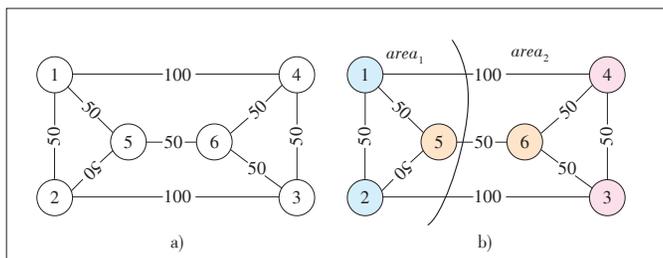
Then v_{jnon} is assigned to the same area as the aggregation node which corresponds to the minimum distance.

Until all nodes are partitioned, the area division ends. Then virtual direct links between the aggregation nodes are established. We use the shortest path algorithm to calculate the candidate path for any two aggregation nodes. For the optical links on the aggregation candidate paths, we allocate the spectral resources to the non-aggregation requests and the aggregation requests in a ratio of 2:1.

2.3 Illustrative Example

A topology is shown in **Fig. 2a**. For the convenience of explanation, we use five historical paths as shown in **Table 1** to calculate the historical traffic of each node. From (3), we calculate the his_tra of each node and get $his_tra_1 = 2, his_tra_2 = 3, his_tra_3 = 1, his_tra_4 = 1, his_tra_5 = 5, his_tra_6 = 4$, e.g., node 5 is the intermediate node of $Path_1$ and $Path_2$, and the source node of $Path_3$, so $his_tra_5 = 2 \times (1 + 1) + 1 = 5$. Sort the nodes in descending order of his_tra and we get the node sorting set as $node_sorting = \{5, 6, 2, 1, 3, 4\}$. We choose two nodes from $node_sorting$ as aggregation nodes, then the aggregation node set is $V_{agg} = \{5, 6\}$ and the non-aggregation node set is $V_{non_agg} = \{1, 2, 3, 4\}$.

As shown in **Fig. 2b**, we divide the topology into areas and



▲ **Figure 2.** a) Six-node topology; b) aggregation node selection and area division result.

▼ **Table 1.** Historical path information

| Path | Path node set |
|----------|---------------|
| $Path_1$ | {1,5,6} |
| $Path_2$ | {2,5,6} |
| $Path_3$ | {5,6} |
| $Path_4$ | {1,2,3} |
| $Path_5$ | {6,4} |

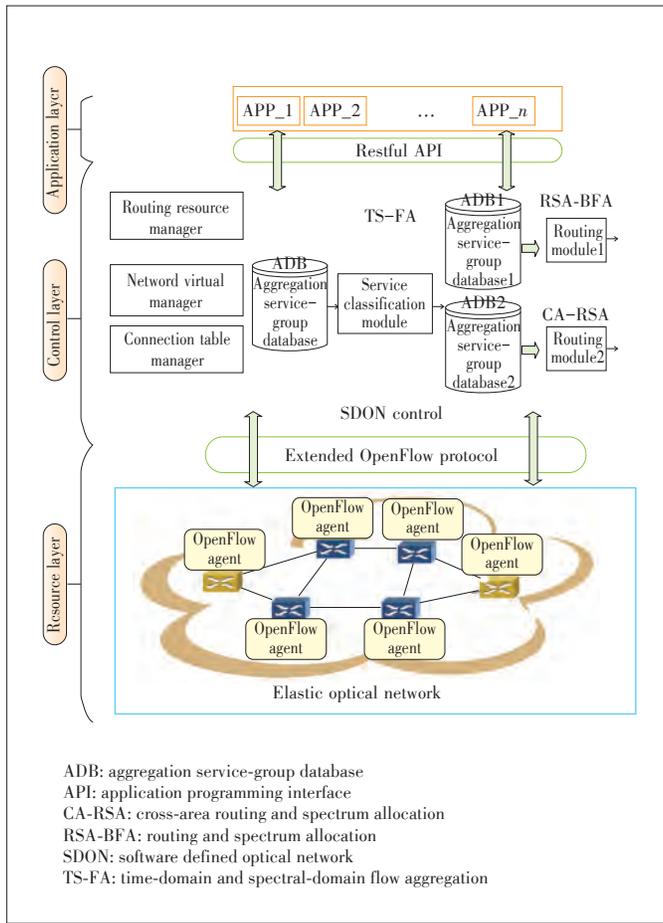
get $area_1 = \{5, 1, 2\}, area_2 = \{6, 3, 4\}$, e.g., $len < path_{1,5} > = 50, len < path_{1,6} > = 50 + 50 = 100, \min \{ len < path_{1,5} >, len < path_{1,6} > \} = 50$ so node 1 and node 5 are divided into the same area. Suppose the number of the spectral slots of each optical link is 300, the first 200 spectral slots between node 5 and node 6 are used for non-aggregation requests, and the last 100 spectral slots are used for aggregation requests.

3 Time-Domain and Spectral-Domain Flow Aggregation Module and Algorithm Based on SDON

3.1 Time-Domain and Spectral-Domain Flow Aggregation Module

Software defined network (SDN) is a new type of network architecture with forwarding control separation and software programming. SDON, the extension of SDN in optical network, offers a global view of network resources, enabling more optimized configuration strategy for the network [8]. SDON architecture mainly consists of resource layer, control layer, and application layer. The resource layer is composed of optical network resources, such as OXC, ROADM, and bandwidth-variable transponders (BVTs). Each optical device is logically connected with its own OpenFlow agent. The control layer is the core of the SDON architecture, responsible for programming the physical hardware at the resource layer based on the requests from the application layer. SDON controller is the logical entity that implements this function and it communicates with optical devices by the OpenFlow agent using the extended OpenFlow protocol. The application layer is composed of applications. An application can submit the network behavior requiring a request to the controller in a programmable manner via the restful application programming interface (API).

Based on the proposed ANS-AD algorithm, we design a TS-FA module on SDON architecture (**Fig. 3**). There are several submodules in the TS-FA module, including service classification module, routing and spectrum allocation (RSA-BFA) module and cross-area routing and spectrum allocation (CA-RSA) module. The implementation of these modules depends on the controller's unified management and control of traffic and the



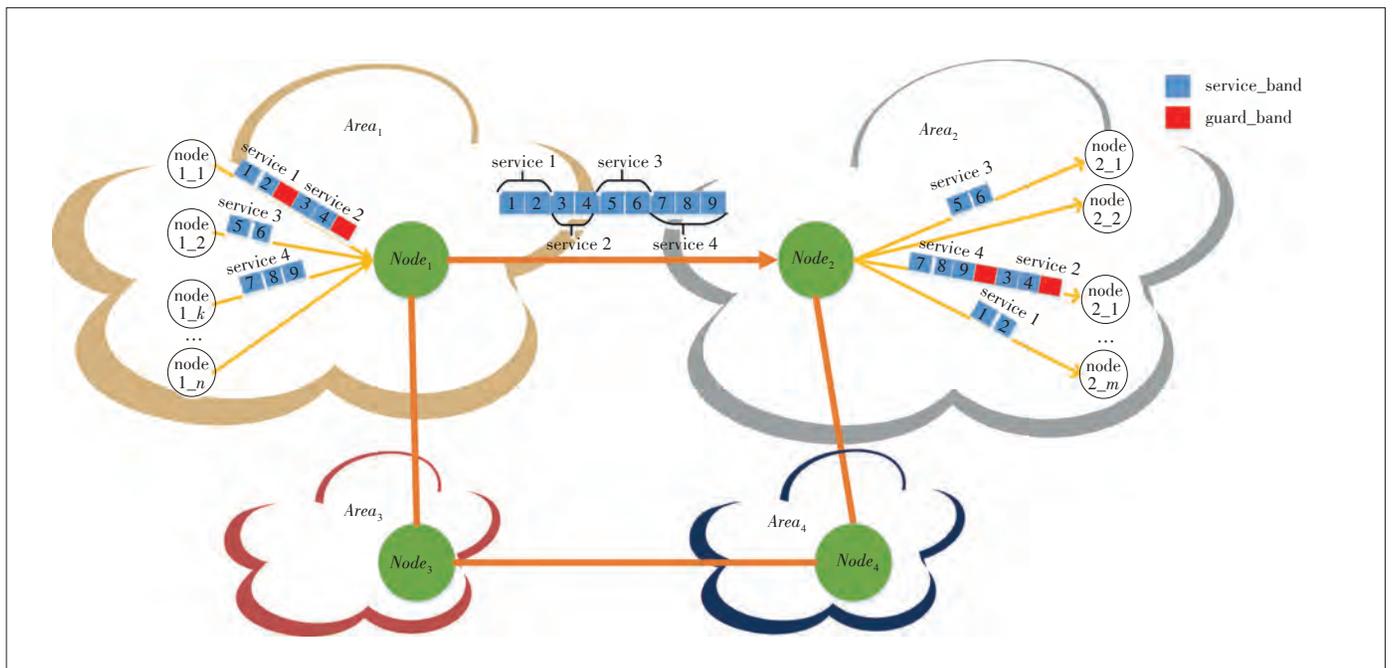
▲ Figure 3. TS-F module on SDON architecture.

flexibility of the underlying hardware [8]. The TS-F module uses three databases to describe and store arrival requests, including aggregation service-group database (ADB), aggregation service - group database1 (ADB1), and aggregation service - group database2 (ADB2).

We set an aggregation time slot for time domain aggregation. ADB stores the arrival requests in a time slot. The service classification module classifies the requests in ADB, then stores them in the ADB1 and ADB2. The requests of the source node and destination node in the same area are stored in ADB1, the requests of the source node and destination node in different areas are stored in ADB2. The RSA-BFA module is responsible for the routing and spectrum allocation of the requests in ADB1. The CA-RSA module is responsible for the routing and spectrum allocation of the requests in ADB2.

3.2 Time-Domain and Spectral-Domain Flow Aggregation Algorithm

Based on the ANS-AD algorithm and the TS-F module proposed above, we propose a TS-F algorithm. The routing and spectrum allocation strategy used before flow aggregation is recorded as RSA-BFA algorithm. When the requests in a time slot arrive, classifying each request according to the request is the cross-area or same-area. Suppose a request in the time slot is $R_i(s, d, B_i)$, where s_i indicates the source node, d_i is the destination node and B_i represents the bandwidth of the request. If s_i and d_i are in the same area, $R_i(s_i, d_i, B_i)$ is a same-area request, otherwise $R_i(s_i, d_i, B_i)$ is a cross - area request. If $R_i(s_i, d_i, B_i)$ is a same-area request, it uses the RSA-BFA algorithm for routing and spectrum allocation. If it is a cross-area



▲ Figure 4. An illustrative example of the cross-area routing and spectrum allocation (CA-RSA) strategy.

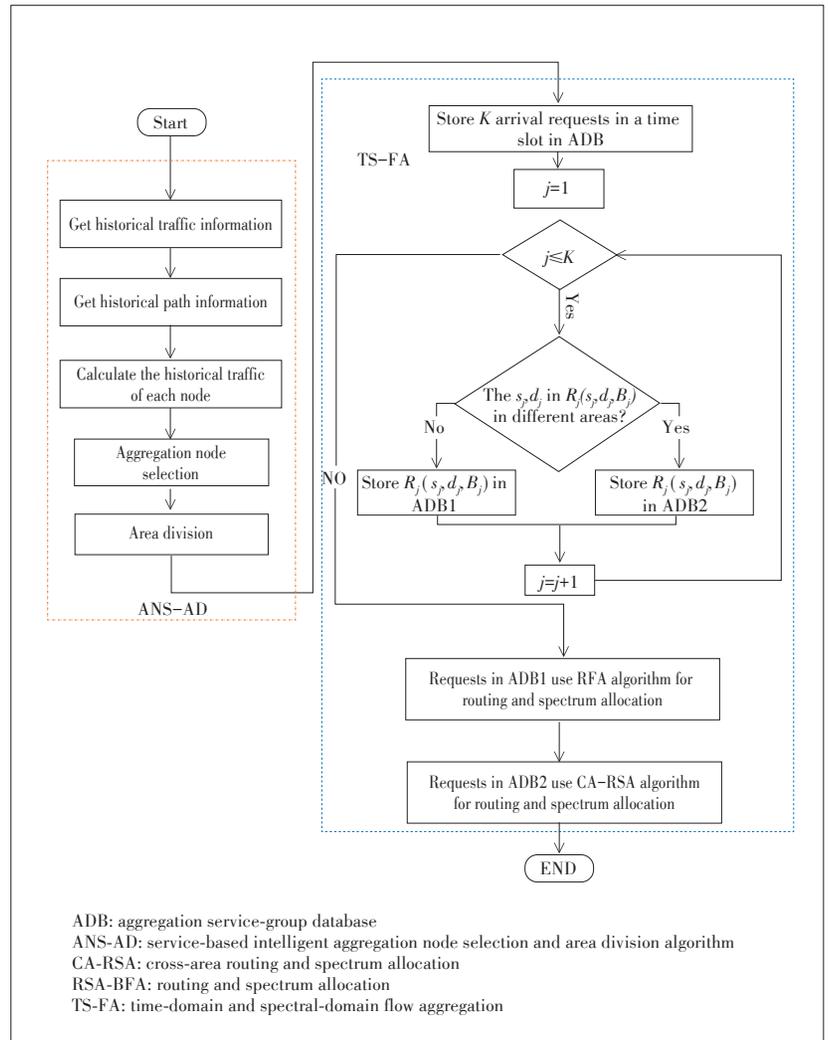
request, it performs routing and spectrum allocation according to the CA-RSA strategy. We use the following example to describe the CA-RSA strategy.

There are four requests from $area_1$ to $area_2$ in a time slot, n non-aggregation nodes in $area_1$ and m non-aggregation nodes in $area_2$ (Fig. 4). The aggregation node in $area_1$ is $Node_1$ and the aggregation node in $area_2$ is $Node_2$. $Node_1$ and $Node_2$ have the functions of aggregation and de-aggregation. All requests from $area_1$ to $area_2$ firstly go to $Node_1$ according to the RSA-BFA algorithm and aggregate together in $Node_1$. Then the aggregation flow is transmitted along the virtual direct link between $Node_1$ and $Node_2$. After that, the aggregation flow de-aggregates on $Node_2$ and the requests independently transmit to respective destination nodes according to the RSA-BFA algorithm with $Node_2$ as the source node. The guard-bands need to be inserted between adjacent paths in non-aggregation links. There is no need to insert guard-bands between adjacent paths in the virtual direct link. The collection of the requests in a time slot realizes the time-domain aggregation and the transmission of requests in the virtual direct link realizes the spectral-domain aggregation. Fig. 5 shows the ANS-AD and TS-FA algorithms.

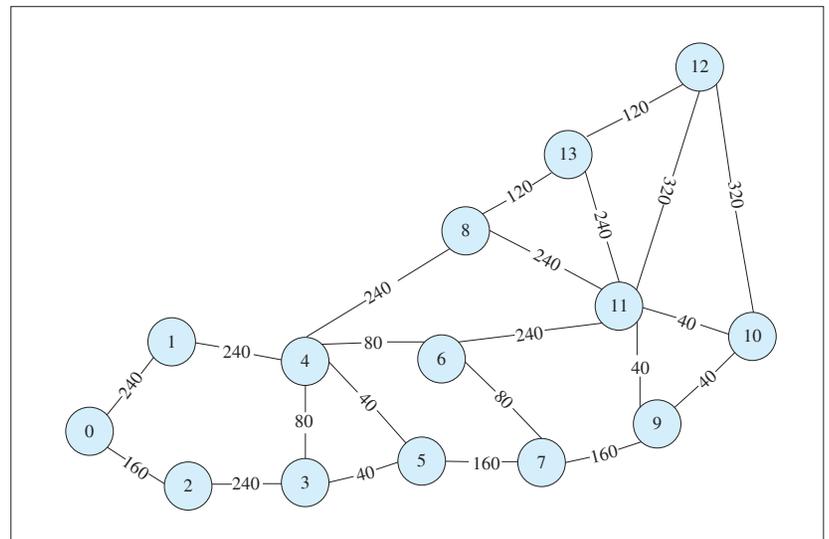
The introduction of the TS-FA solution requires adding the aggregation combiners/separators in aggregation nodes. The core of the network with ROADMs is kept unchanged thus offering a time-domain and spectral-domain flow aggregation of the requests. Moreover, the combination of time domain and spectral domain offers the finest possible granularity and thus flexibility in resource management. This ideally matches the requirement of high capacity [9].

4 Performance Evaluation

In this section, we design a simulation to evaluate the performance of the ANS-AD algorithm and the TS-FA algorithm on the Japan topology (14 nodes and 22 links) as shown in Fig. 6. We set 750 spectral slots on each optical link. The bandwidth of each spectral slot is 12.5 G. The routing algorithm before flow aggregation is the shortest path algorithm and the spectrum allocation algorithm is First-Fit. The routing and spectrum allocation strategy used before flow aggregation is recorded as the RSA-BFA algorithm. Each request has a uniform distribution of spec-



▲ Figure 5. The ANS-AD and TS-FA algorithms.



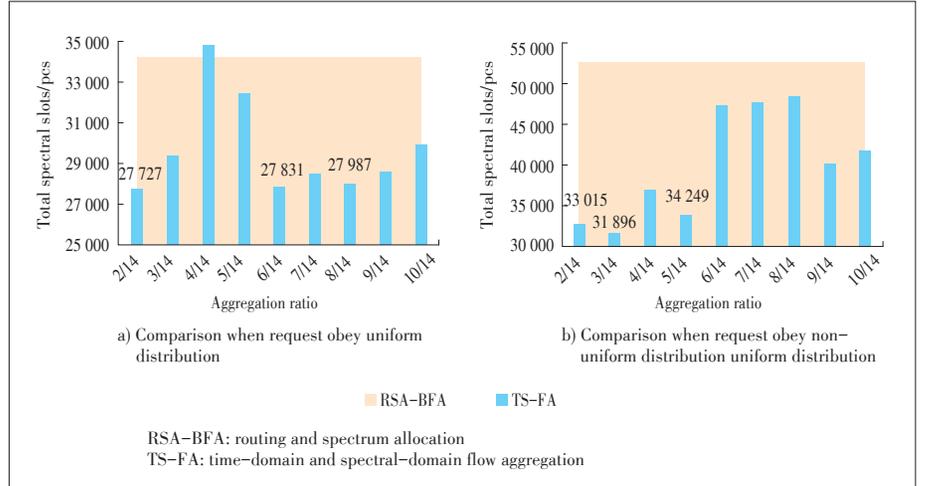
▲ Figure 6. 14-node Japan topology.

tral slots from 1 to 10. We generate 100 000 historical services for aggregation node selection. Then we generate arrival requests of twenty time slots for time-domain and spectral-domain flow aggregation and the number of requests in each time slot is uniformly distributed from 40 to 60. The service distribution for time-domain and spectral-domain flow aggregation is the same as the aggregation node selection and area division.

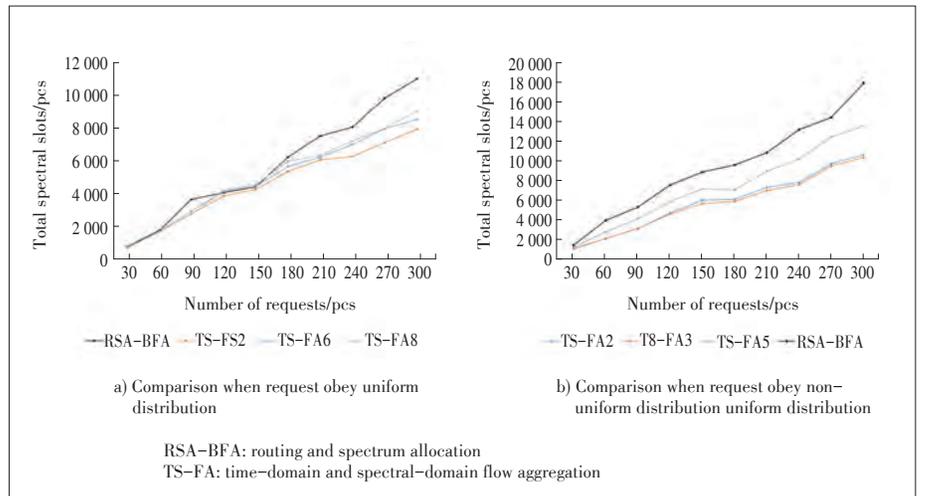
We generate two service distributions to simulate the ANS-AD algorithm and the TS-FA algorithm. Then we calculate the total resource usage of the requests in 20 time slots under different aggregation ratio. Moreover, we compare the resource usage by the TS-FA algorithm with the RSA-BFA algorithm. The results are shown in **Fig. 7**. When the source node and the destination node of each request obey uniform distribution from 0 to $node_num$, where $node_num$ refers to the number of nodes in the topology (**Fig. 7a**). It can be seen from Fig. 7a that the TS-FA algorithm uses less resources than the RSA-BFA algorithm except the aggregation ratio 4/14. We compare the TS-FA algorithm when the aggregation ratio is 2/14, 6/14, and 8/14 with the RSA-BFA algorithm by resource usage under different number of requests (**Fig. 8a**). The results show that the TS-FA algorithm obtains lower resource usage than the RSA-BFA algorithm. When the aggregation ratio is 2/14, the TS-FA algorithm uses the least resources.

When the source node of each request obeys uniform distribution from 0 to $node_num/2$ and the destination node of each request obeys uniform distribution from $(node_num)/2 + 1$ to $node_num$, the results are shown in **Fig. 7b**. It can be seen that the TS-FA algorithm uses less resources than the RSA-BFA algorithm under any aggregation ratios. We compare the TS-FA algorithm when the aggregation ratio is 2/14, 3/14, 5/14 with the RSA-BFA algorithm by resource usage under different number of requests as shown in **Fig. 8b**. The results show that the TS-FA algorithm obtains lower resource usage than the RSA-BFA algorithm. When the aggregation ratio is 3/14, the TS-FA algorithm uses the least resources.

From Fig. 7, we can also see that in the first service distribution when the aggregation ratio is 4/14, the TS-FA algorithm



▲ **Figure 7. Comparison of resource usage between TS-FA and RSA-BFA under different aggregation ratio.**



▲ **Figure 8. Comparison of resource usage under different number of requests.**

uses more resources than the RSA-BFA algorithm. The reason for this result is that some cross-area requests have to take more hops in the case of aggregation. At this service distribution and aggregation ratio, the bandwidth saved by aggregation is lower than the wasted bandwidth by more hops.

5 Conclusions

In this paper, we propose a service-based intelligent ANS-AD algorithm. Then, a TS-FA module and algorithm based on SDON are detailed. The algorithms attempt to reduce bandwidth waste and increase network capacity. The simulation results show that under different service distribution, the ANS-AD algorithm and TS-FA algorithm can effectively reduce the resource waste. Our proposed algorithms provide some ideas for service-based network deployment.

References

[1] WONG E, GRIGOREVA E, WOSINSKA L, et al. Enhancing the Survivability and Power Savings of 5G Transport Networks Based on DWDM Rings [J]. *IEEE/OSA Journal of Optical Communications & Networking*, 2017, 9(9): D74–D85. DOI: 10.1364/JOCN.9.000D74

[2] TANAKA T, INUI T, KADOHATA A, et al. Adaptive and Efficient Multilayer Elastic Optical Network Planning [C]//42nd European Conference on Optical Communication. Düsseldorf, Germany, 2016.

[3] THYAGATURU A S, MERCIAN A, MCGARRY M P, et al. Software Defined Optical Networks (SDONs): A Comprehensive Survey [J]. *IEEE Communications Surveys & Tutorials*, 2016, 18(4): 2738–2786. DOI: 10.1109/COMST.2016.2586999

[4] THOUÉNON G, BETOULE C, KHODASHENAS P S, et al. Electrical v/s optical aggregation in multi-layer optical transport networks [C]//IEEE International Conference on Photonics in Switching. Florence, Italy, 2015: 28–30. DOI: 10.1109/PS.2015.7328942

[5] YANG H, WANG R, WRIGHT P, et al. Impact of WSS Passband Narrowing Effect on the Capacity of the Flexible-Spectrum Networks [C]//IEEE Optical Fiber Communications Conference and Exhibition. Los Angeles, USA, 2017: W11.5. DOI: 10.1364/OFC.2017.W11.5

[6] LI X, TANG Y, GUO J F, et al. A Novel BV-WSS Based Flow Aggregation (BV-WSS-FA) Scheme for Finer-Grained Services in Elastic Optical Networks [C]//16th International Conference on Optical Communications and Networks (ICOCN), Wuzhen, China, 2017: 1–3. DOI:10.1109/ICOCN.2017.8121320

[7] DUAN Z W, LI H, LIU Y Z, et al. Effective Distance Adaptation Traffic Dispatching in Software Defined IP over Optical Network [C]//Proc. SPIE 10464, AOPC 2017: Fiber Optic Sensing and Optical Communications. Bellingham, USA: SPIE, 2017: 1046420. DOI: 10.1117/12.2285596

[8] LYU C J, LI H, JI Y F, et al. Holding-Time-Aware Asymmetric Spectrum Allocation in Virtual Optical Networks [C]// Proc. SPIE 10464, AOPC 2017: Fiber Optic Sensing and Optical Communications. Bellingham, USA: SPIE, 2017: 104641V. DOI: 10.1117/12.2285545

[9] GAVIGNET P, LE ROUZIC E, PINCEMIN E, et al. Time and Spectral Optical Aggregation for Seamless Flexible Networks [C]//IEEE International Conference on Photonics in Switching. Florence, Italy, 2015: 43–45. DOI: 10.1109/PS.2015.7328947

← **From P. 55**

[J]. *Optica*, 2017, 4(3): 307–322. DOI: 10.1364/OPTICA.4.000307

[27] ZAKHAROV V E, SHABAT A B. Exact Theory of 2-Dimensional Self-Focusing and One-Dimensional Self-Modulation of Waves in Nonlinear Media [J]. *Soviet Journal of Experimental and Theoretical Physics*, 1972, 34: 62–69.

[28] SON T L, HENNING B, VAHID A. Demonstration of 640.5 Gbaud Nonlinear Frequency Division Multiplexed Transmission with 32QAM Formats [C]//Optical Fiber Communications Conference and Exhibition (OFC). Los Angeles, USA, 2017: W3J.1.

[29] VAHID A, HENNING B. Experimental Demonstration of Nonlinear Frequency Division Multiplexed Transmission [C]//European Conference on Optical Communication (ECOC). Valencia, Spain, 2015. DOI: 10.1109/ECOC.2015.7341903

[30] VAHID A, SON T L, HENNING B. Demonstration of Fully Nonlinear Spectrum Modulated System in the Highly Nonlinear Optical Transmission Regime [C]//European Conference on Optical Communication (ECOC). Duesseldorf, Germany, 2016: 18–22.

[31] HE G-Q, WANG L-N, LI C-Y, et al. Spectral Function Modulation Based on Nonlinear Frequency Division Multiplexing [J]. *Scientific Reports*, 2017, 7: 6058. DOI:10.1038/s41598-017-06427-1

[32] WANG L-N, LIU S-Y, LI C-Y, et al. A Combination of Eigenvalue and Spectral Function Modulation in Nonlinear Frequency Division Multiplexing [C]//OSA Nonlinear Optics Topical Meeting (NLO). Waikoloa, Hawaii, USA, 2017: NW4A.21.

[33] LIU S-Y, WANG L-N, LI C-Y, et al. Spectral Function Modulation Based on Nonlinear Fourier Transform [C]//OSA Nonlinear Optics Topical Meeting (NLO). Waikoloa, Hawaii, USA, 2017: NW4A.22

[34] YOUSEFI M I, YANGZHANG X. Linear and Nonlinear Frequency-Division Multiplexing [C]//2016 42nd European Conference on Optical Communication (ECOC). Duesseldorf, Germany, 2016: 342–344

[35] PFAU T, HOFFMANN S, NOÉ R. Hardware-Efficient Coherent Digital Receiver Concept with Feedforward Carrier Recovery for M-QAM Constellations [J]. *Journal of Lightwave Technology*, 2009, 27(8): 989–999. DOI: 10.1109/JLT.2008.2010511

Biographies

NI Dong received the bachelor’s degree in communication engineering from Jilin University, China in 2017. She is currently a graduate student in Information and Communication Engineering at Beijing University of Posts and Telecommunications, China. In the period of graduate study, her main research field is the theory and technology of IP over elastic optical networks.

LI Hui (lihui@bupt.edu.cn) is an associate professor in the Broadband Network Innovation Research Center Lab, School of Information and Telecommunication, Beijing University of Posts and Telecommunications, China. Her main research field is optical network planning and optimization.

JI Yuefeng is a Ph.D. supervisor and professor of Beijing University of Posts and Telecommunications, China. His research interests include optical communication and broadband networks.

LI Hongbiao has been working in the Wireline Product Operation Department of ZTE Corporation since 1999. He is currently a member of ZTE’s strategic expert group and the chief planner of the transport network, mainly engaging in communication transport network technology research and planning work.

ZHU Yinan graduated from Nanjing University of Posts and Telecommunications, China in 2015. At present, he is engaged in the planning of transport network solutions in ZTE Corporation and is mainly responsible for SDN and IP+optical coordination. He has also been a CCIE since 2014.

Biographies

WANG Jia (jwang_wj@sjtu.edu.cn) received her B.S. degree in communication engineering from Nanjing Tech University, China in 2017. She is currently a postgraduate in the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China. Her main research interest is nonlinear optics.

ZHAO Yilong is pursuing his bachelor degree in electronic science and technology at Shanghai Jiao Tong University, China. His research interests include fiber nonlinearity mitigation algorithm and artificial intelligence.

HUANG Xin is pursuing his bachelor degree in information engineering at Shanghai Jiao Tong University, China. His research interests include optical communication, nonlinear photonics, and embedded system.

HE Guangqiang received his Ph.D. degree in communication and information system from Shanghai Jiao Tong University, China in 2006. He joined the Department of Electronic Engineering, Shanghai Jiao Tong University as a lecturer in 2006. From 2009 to 2010, he was a visiting scientist in Department of Physics and Astronomy, University of Rochester, New York, USA. Since December 2011, he has been an associate professor in the Department of Electronic Engineering, Shanghai Jiao Tong University. He has published over 50 SCI papers and held 5 patents. His research interests include quantum information processing, quantum entanglement, quantum cryptography, and nonlinear optics.



Data-Driven Joint Estimation for Blind Signal Based on GA-PSO Algorithm

Abstract: Without any prior information about related wireless transmitting nodes, joint estimation of the position and power of a blind signal combined with multiple co-frequency radio waves is a challenging task. Measuring the signal related data based on a group distributed sensor is an efficient way to infer the various characteristics of the signal sources. In this paper, we propose a particle swarm optimization to estimate multiple co-frequency “blind” source nodes, which is based on the received power data measured by the sensors. To distract the mix signals precisely, a genetic algorithm is applied, and it further improves the estimation performance of the system. The simulation results show the efficiency of the proposed algorithm.

Keywords: Particle Swarm Optimization (PSO); Genetic Algorithm (GA); spatially distributed sensor; blind signal detection

LIU Shen^{1,2,3}, QIN Yuannian¹,
LI Xiaofan², ZHAO Yubin³,
and XU Chengzhong⁴

(1. Guilin University of Electronic Technology, Guilin, Guangxi 541004, China;

2. Shenzhen Institute of Radio Testing & Tech., Shenzhen, Guangdong 518000, China;

3. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518000, China;

4. University of Macau, Macau SAR 999078, China)

DOI: 10.12142/ZTECOM.201903010

[http://kns.cnki.net/kcms/detail/34.1294.](http://kns.cnki.net/kcms/detail/34.1294.TN.201900910.0953.002.html)

TN.201900910.0953.002.html, published online September 10, 2019

Manuscript received: 2019-01-21

1 Introduction

As mobile broadband traffic and end-user demands continue to grow, the radio frequency spectrum has become an expensive and limited resource for wireless communications. Therefore, how to reasonably allocate spectrum resources, avoid conflicts, and detect interference have become an extremely worthwhile research issue. Spatially distributed sensor networks can effectively detect a certain frequency band in a certain area [1]. The power information can be used in a variety of applications, such as indoor positioning, interference detection, signal recognition, and cognitive radio systems. As a major threat to wireless communications, interference creates significant usage and financial impact on users and operators. Interference is generally processed by several related techniques, such as interference monitoring, interference detection and isolation, interference classification, interference localization, and interference mitigation. Interference localization is what we are concerned about.

General interference signals are blind source signals, with unknown transmission power, which also brings difficulties to interference location [2]. For some communication systems, in-

terference sources can cause very serious consequences if they cannot be located and eliminated in real time [3]. For instance, the signals in the communication based train control system (CBTC) communicate with the base station in the 2.4 GHz public band and are easily interfered by other devices in the same band. Especially when there are non-protocol signals in the environment, the system signals may be interfered by them. Once the same frequency signals are mixed, the signal sources are difficult to locate. Therefore, it is necessary to use appropriate methods to detect blind signals. Currently, common spectrum sensing techniques include matched filter detection, cyclostationary detection, and energy detection, which are mainly based on physical layer signal processing. Several methods were proposed to solve the problem of blind signal detection, by the use of a path loss model. The most used is the theoretical analysis method of using received power measurements at spatially distributed monitors [4], [5], [6]. Some scholars have proposed a method to detect the co-frequency blind signal by using the topology of the receiver and the mathematical models [6]. Some methods use k-means clustering algorithm combined with Particle Swarm Optimization (PSO) to estimate the position of nodes [2]. Other methods use SDR equipment for inter-

ference detection [7]. From the simulation results, the PSO has better adaptability and performance when dealing with such interference problems [8], [9]. The PSO has caught great attention and various improved algorithms are continuously proposed, such as the local PSO through changing the topology of the particle population. There are some methods combined with other algorithms such as Genetic Algorithm (GA) and clustering algorithms to improve performance [10], [11]. The PSO is then extended to use for the multi-objective estimation problem [10]. Genetic algorithms are also widely used in smart applications [12]–[14]. Similarly, improvements to genetic algorithms have never stopped, such as the elite genetic algorithm, adaptive genetic algorithm, multi-objective genetic algorithm, and hybrid genetic algorithm [15]–[18].

However, there are few works concerning the data-driven method using sensors or detecting equipment.

According to the received signal power measurement and free-space path loss model, the sensor network can be utilized to estimate the relative position of the target node based on the mutual measurements from other nodes in the network. A natural question that then arises is what if only the general path loss model is utilized, this makes it infeasible to locate one or multiple co-frequency nodes when the transmit power is unknown [19], [20].

In this paper, we utilize the sensor measurement data which contains the spatial diversity of the source signals to estimate the unknown node's coordinates and power based on a proposed particle swarm optimization algorithm. We formulate the sum of squared error of the mathematical model of the sensor system as the objective function of the PSO algorithm and find the minimum value. According to the path loss model and the limited range of the spatially distributed sensor receiving area, the unknown nodes' transmit power range is estimated. The major contributions are listed as follows:

1) We use spatially distributed sensor networks and PSO to study the joint estimation of the position and power of multiple unknown nodes. In the absence of transmitter power information, a unique estimation method is proposed based solely on the received power information on the sensor. For locating multiple co-frequency unknown nodes, the received power is the superposition of the co-frequency unknown nodes power. The mathematical model of this sensor system is non-convex, and the function has multiple minimum points. We reduce the range of the objective function solution by pre-estimating the power and coordinates of unknown nodes to reduce the number of local optimal solutions and improve the accuracy of joint estimation.

2) We find that the ratios of power between unknown nodes will affect the final estimate of the algorithm. To avoid the wrong estimation of the PSO, we designed a unique GA and combined it with the PSO, which greatly reduced the probability of the error estimation of the PSO and improved the estimation performance of the system. When a signal has higher power,

it is easy to obtain a smaller estimation error. However, the signal with lower power cannot be estimated accurately. The greater the ratio of the power of the signals, the more pronounced this phenomenon is. Moreover, when multiple signal power values are similar, the PSO may be confused. In this case, the PSO algorithm cannot distinguish the phenomenon of each signal because the PSO falls into the local optimal solution. Therefore, we combine the GA with the PSO algorithm to avoid the pre-maturity of the algorithm and enable it to jump out of the trap of the local optimal solution.

Through simulation, we found that the combined performance of PSO has been significantly improved after combining GA. The estimated error is reduced by a maximum of 0.6 m, and the estimation accuracy is improved by nearly 30%. Compared with the PSO with the exhaustive method, the performance gap is about 30%.

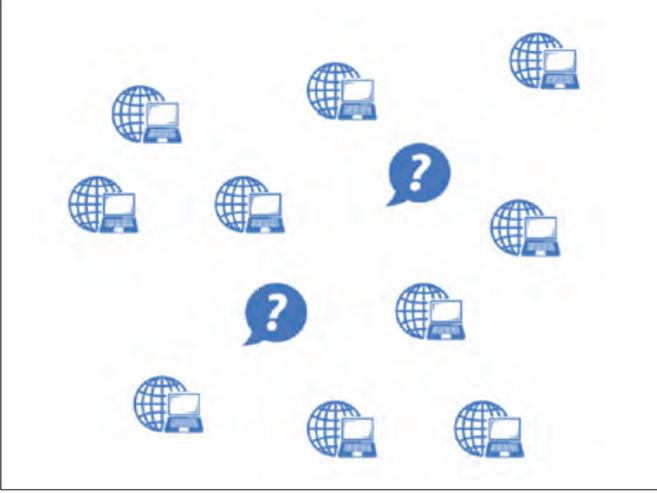
The rest of this paper proceeds as follows. Section 2 introduces the system model. The PSO algorithm improved according to system characteristics is proposed in Section 3. Section 4 mainly deals with unique GA and the method combined with PSO. In Section 5, the algorithm combined with the proposed PSO and GA is simulated and verified. Section 6 concludes this paper.

2 System Model

We assume a scenario where there are M transmitters and N sensor nodes within a certain square area. There is no reflective medium in the area, and there will be no multipath effect. In general, spatially distributed sensor networks are placed according to a certain spatial topology. This usually enables the system to have better estimation performance and to preliminarily position the target node. However, in order to make the study without loss of generality, we randomly placed the sensor in the hypothetical area and recorded its coordinates. The locations of the transmitters and sensors are represented in Cartesian coordinates. The uppercase notation denotes the coordination and power of transmitters such as (X_1, Y_1) and (P_{T1}) , while the lowercase denotes the coordinates and power of sensors such as (x_1, y_1) and (p_{r1}) . For example, a diagram of the square region containing M transmitters and N receivers is shown in **Fig. 1**. Here, we only consider the line-of-sight channels from each transmitter to each sensor.

$$p_r = \frac{P_T}{\beta d^\alpha}, \quad (1)$$

where d denotes the Euclidean distance between the transmitter and the sensor, p_r denotes the measured power of sensor, P_T denotes the transmitted power of the blind signal, and β is a known constant selected based on the carrier frequency and the antenna structure. In practice, $\alpha \in [2, 6]$ depending on the environment. Here, we make the value of $\alpha = 2$. Since the estimation system only utilizes the received power level of the sen-



▲ Figure 1. M transmitters (question marks) and N sensors placed within a certain square area.

sor, we do not consider short-term fading.

2.1 Multiple Transmitter System

Since we chose PSO as the primary tool to solve joint estimation problems, we transform the system model into an objective function that serves as the basis for particle search. We define that the objective function is the sum of the squared difference between the actual measured power of each sensor and the predicted received power based on the estimated transmitter power and coordinates and causes the population of particles to search for its minimum value.

$$F = \sum_{i=1}^N \left(p_{r,i} - \sum_{j=1}^M \frac{P_{T,j}}{\beta d_{i,j}^\alpha} \right)^2, \quad (2)$$

where $p_{r,i}$ denotes the measured power of sensor i . $P_{T,j}$ denotes the transmitted power of the transmitter j . $d_{i,j}^\alpha$ denotes the Euclidean distance between the transmitter j and the sensor i . Our objective function is not convex concerning the estimated transmitter locations. Note that there may be many local optimal solutions when the number of sensor nodes is small, and the estimation error does not approach zero endlessly as the number of sensor nodes increases.

3 Particle Swarm Optimization

PSO is a population-based evolutionary algorithm. The classical PSO is used for balancing the weights of a neural network. The basic unit of the PSO population is the particle. The algorithm forms the search behavior through the interaction between each particle. Each particle represents a solution to the objective function and corresponds to a fitness function value. The merits of the fitness function value directly determine the quality of the solution. The target position is considered a global optimal solution.

PSO has two important concepts:

- 1) Exploration means that particles leave the original search trajectory to a certain extent and search in a new direction, which reflects the ability to develop into unknown regions.
- 2) Exploitation refers to the fact that particles continue to perform more detailed searches on the original search trajectory to some extent.

Population searches for optimal solutions in exploration and development. To control these two kinds of search better, the inertia weight is introduced. Its value can adjust the global and local search ability of PSO. When α is large, global optimization ability tends to be strong. On the contrary, the local search ability is enhanced. In this paper, we utilize the equation with inertia weights. The topology of the population directly determines the way in which particles interact. The classical PSO is a global topology and the learning samples for each particle are all other particles in the population. Subsequently, scholars proposed a variety of local PSO algorithms with excellent performance, and found that local PSO algorithm has better performance in local search. But what our system needs most is the global search ability of the algorithm. Therefore, it is more appropriate for us to choose a global topology.

We generate a certain number of particle populations, and then initialize the particles and randomly generate their speeds and positions. All particles are initialized, giving them random values in each dimension, as shown in Equ. (3), where k is the number of the particle and j is the number of the transmitter. Then we calculate the fitness function value of each particle according to Equ. (4), where F^k denotes the fitness function of particle k . i is the number of the sensor, and update the particles' speeds and positions according to Eqs. (5) and (6) in each iteration until the end of the algorithm. The equations and pseudo code of the standard PSO (Algorithm 1) are as follows.

$$\begin{cases} x^k = [X_1^k, Y_1^k, P_{T,1}^k, \dots, X_M^k, Y_M^k, P_{T,M}^k] \\ v^k = [\dot{X}_1^k, \dot{Y}_1^k, \dot{P}_{T,1}^k, \dots, \dot{X}_M^k, \dot{Y}_M^k, \dot{P}_{T,M}^k] \end{cases}, \quad (3)$$

$$F^k = \sum_{i=1}^N \left(p_{r,i} - \sum_{j=1}^M \frac{P_{T,j}^k}{\beta [(X_j^k - x_i)^2 + (Y_j^k - y_i)^2]} \right)^2, \quad (4)$$

$$v^k(t+1) = \omega \times v^k(t) + c_1 \times r_1 \times (pbest^k(t) - x^k(t)) + c_2 \times r_2 \times (gbest(t) - x^k(t)), \quad (5)$$

$$x^k(t+1) = x^k(t) + v^k(t+1), \quad (6)$$

where $v^k(t)$ denotes the speed of particle k at iteration time t . x^k denotes the position of the particle k . The $pbest^k$ is the optimal position experienced by the particle k . The $gbest$ denotes the position of the particle that operates optimally in the population. c_1 and c_2 denote the acceleration constants of the particles. In order to make the particles have stronger self-search-

ing ability in the initial stage, they can converge to the global optimum at the later stage. We linearize c_1 and c_2 so that c_1 has a larger initial weight and gradually decreases. And c_2 is the opposite.

$$\begin{cases} c_1 = \frac{-t}{T} + 2 \\ c_2 = \frac{t}{T} + 1 \end{cases} \quad (7)$$

r_1 and r_2 are two random numbers distributed uniformly in $[0, 1]$. ω is generally a constant, and the linearly decreasing weight is currently used more frequently as the value of the inertia weight.

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min}) \times t}{T}, \quad (8)$$

where T denotes the total number of iterations and t denotes the number of iterations that have taken place.

The population size of particles affects the performance of the algorithm. Different estimation problems apply to different population sizes, and it is generally proportional to the dimension of the objective function. Some results suggest that using $(2m + 1)^2$ particles work better, where m is the number of dimensions the objective function has. For our simulation, each unknown transmitter needs to estimate three dimensions (x, y, P_T) . Therefore, for each additional transmitter, there are three dimensions added. We need to change the population of particles according to the different dimensions of the objective function. Otherwise, too few particles may miss the global optimal solution, while too many particles may produce many repeated results, reducing the efficiency of the algorithm. So we use $(2m + 1)^2$ particles, which is enough for most of the estimation problems. Using excess particles within the algorithm can indeed make the estimation more accurate. However, it will cause the algorithm calculation volume to increase, affecting the overall performance of the system. In Section 4, with the help of the exhaustive method, we use a large number of particle population to approximate the estimation error limit of the objective function in this paper.

Algorithm 1. Particle Swarm Optimization

- 1: Establish the objective: minimize $\{F^k\}$.
- 2: Initialize, randomly generated particle position and velocity.
- 3: **While** $t \leq T$ // T is the maximum number of iterations
- 4: **If** $F^k < pbest^k$ **then**
- 5: $pbest^k = x^k$
- 6: **If** $pbest^k < gbest$ **then**
- 7: $gbest = pbest^k$
- 8: **End if**
- 9: **End if**

10: Update particle. Update particle velocity and position according to Eqs. (5) and (6).

11: **End While**

To evaluate the effectiveness of the proposed joint estimation technique, we attempted to simulate the performance of three to twenty sensors and one to four transmitters based on different systems. As mentioned earlier, each additional transmitter will add 3-dimensions that need to be estimation to the system. For the estimation problem, the number of equations should be more than or equal to the estimated dimensions. In other words, every time an unknown transmitter is added, at least three receivers need to be added. Increasing the number of transmitters and receivers does increase the computational burden of the simulation, but the proposed method has no inherent limits to the dimensionality it can handle. However, as the dimension increases, the performance of the PSO deteriorates. In the standard PSO, the learning object of each particle is a particle in the population that has the best fitness function value. When the dimensions are large, it is difficult to ensure that the global-best particle can search for the optimal solution in every dimension.

After a large number of simulation experiments using only the particle swarm algorithm, we found that the ratio between the powers of multiple transmitters had different effects on the results. The impact can be divided into two situations. In each case, the algorithm falls into a special local minimum. In the first case, when there are two or more transmitters with similar actual power, the particle swarm algorithm may easily match the estimated power of one transmitter with the estimated position of another transmitter with similar power. Because the objective function is a multi-dimensional estimate, when the estimated value of some of the dimensions is close to the actual value, the overall fitness function value is too small. The algorithm tends to converge and falls into a local minimum.

In the second case, the high power transmitter estimates better than the low power transmitter. This situation is caused by the objective function. High-power transmitters have a greater impact on fitness function values than low- power transmitters. Although the dimensions are synchronous estimates, the convergence of various dimensions of the high power transmitter is significantly faster than the dimensions of the low power transmitter. When the various dimensions of the high-power transmitter obtain relatively accurate estimates, the fitness function value is already small and the algorithm tends to converge.

4 Genetic Algorithm

The genetic algorithm is a computational model that simulates the natural evolution of Darwin's biological evolution theory and the biological evolution process of the genetic mechanism. It is a method to search for optimal solutions by simulat-

ing natural evolutionary processes. The GA begins with a population that represents a potential solution set of problems, while a population consists of a certain number of individuals genetically encoded. Each is an entity with a characteristic chromosome. Chromosomes are the main carrier of genetic material, that is, a collection of multiple genes whose internal manifestation is a combination of genes that determine the external representation of the shape of the individual. Therefore, mapping from phenotype to genotype, that is, coding work, needs to be implemented at the beginning. Since the work of gene encoding is complex, we tend to simplify it. Here, we combine GA and PSO, which share a population. Next, we genetically encode each particle according to the category of information estimated on each dimension of the particle. The specific gene coding method is shown in Fig. 2.

The figure shows the situation where there are two unknown transmitters in the detection area. The coordinate (X_1, Y_1) of the transmitter 1 is used as a gene, and the coordinate (X_2, Y_2) of the transmitter 2 is used as another gene, and the transmission power of each transmitter is independently used as a gene.

In this paper, the PSO and the GA share an objective function. After one iteration of the PSO, the particle population is sent to the GA for a round of selection. We calculate the survival probability based on the fitness function value of each particle. Since the algorithm solves the problem of finding the minimum value of the objective function, it can be considered that the smaller the fitness function value is, the better the estimation result of the particle is. We give a larger probability of survival for particles with smaller fitness function values, while particles with larger fitness function values are given a lower probability of survival.

Then, the roulette strategy and the elitism are joined to select the paternal particle population. The previously calculated survival probability, at this time serving the roulette strategy, is used to select the paternal particle population.

The genes in the GA do not necessarily reflect the nature of the problem to be solved. Therefore, the genes are not necessarily

independent of each other. Simply hybridizing is likely to destroy a better combination. In this way, the purpose of accumulating better genes is not achieved, but the original good genes are destroyed. Rudolph used the finite Markov chain theory to prove that the Canonical Genetic Algorithm (CGA), which uses only three genetic operators of crossover, mutation, and selection (proportional selection), cannot converge to the global optimal value.

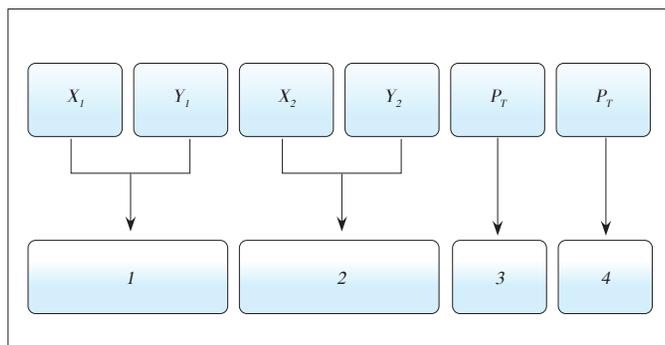
Therefore, we added elitism to the algorithm, so that the particles with the best value of fitness function in the particle population output from the PSO are directly crossed into the progeny particles without crossover, mutation, and selection.

It is assumed that the number of particle output from the PSO to the GA is L , with $L = (2m + 1)^2$. We select $2(L - 1)$ paternal particles by roulette strategy. The two pairs are combined to perform gene cross-interchange, and each pair of paternal particles produces only one progeny particle. The number of progeny particle population reaches $L - 1$, and finally, the optimal particles retained by the elitism. The number of particles guaranteed to be re-entered into the PSO is still L . The specific roulette and elitism is shown in Algorithm 2.

Algorithm 2. Roulette and Elitism

- 1: Calculate F^k of each particle.
 - 2: Select the particle k with the $\min(F^k)$ and output it to the progeny population.
 - 3: Calculate the survival probability:

$$P_{\text{survival}} = ((\max(F^k) + \min(F^k)) - F^k (1 : L, 1)) / \text{sum}((\max(F^k) + \min(F^k)) - F^k (1 : L, 1))$$
 - 4: Create a roulette based on P_{survival} .
 - 5: **While** $ii \leq 2(L - 1) // 2(L - 1)$ is the number of paternal particles. And $ii = 1$.
 - 6: Select a paternal particle by roulette
 - 7: $ii = ii + 1$
 - 8: **End While**
-



▲ Figure 2. The coordinate estimates of unknown node 1 are combined into one gene. The coordinate estimates of unknown node 2 are combined into one gene. Their respective power estimates are separated into a single gene.

Because the PSO is difficult to distinguish several blind source signals with similar power, it is extremely easy to fall into the local minimum. Therefore, when the genes are crossed, we allow the coordinate loci to be interchanged and the power loci can be interchanged, although this may make the progeny particle gene obtained after the crossover completely different from the paternal particle gene. Because of the existence of elite strategies, we do not need to worry that the PSO will lose the original possible correct estimation results.

We hope that when the PSO particles tend to converge, some particles can still jump out of the existing estimation results. The particles may quickly and accurately estimate the position and power of the high-power transmitter, especially when the power difference between the two transmitters is very large.

Therefore, the entire PSO algorithm tends to converge when the low power transmitter does not obtain an accurate estimate. And the estimated result falls into a local minimum. So the design of genetic mutation becomes extremely critical. The mutation may occur on any one or more genes of the particle, and the mutation is completely random in the range of values of the gene. We do not make any human intervention in the values and direction of mutation. The mutation probability is set to increase linearly with the number of iterations, as shown in Equ. (9), where P_{muta} denotes the mutation probability.

$$P_{muta} = \frac{t}{T} \times 0.1 + 0.05. \tag{9}$$

The specific operation of the particle group gene crossover and mutation is shown in **Fig. 3**.

Referring to Fig. 3, the first two genes are the coordinate estimate genes, and the last two are the power estimate genes. Each of the two paternal particles that exchange genes takes out a coordinate estimate gene and a power estimate gene. These four random genes are randomly combined to form progeny particles with new gene sequences. Taking two unknown nodes as an example, the genetic composition of the progeny particles is 64 cases. And the genes of the progeny particles

are still arranged in the order of the genes of the paternal particles. Genetic mutation may also occur in progeny particles obtained after crossing.

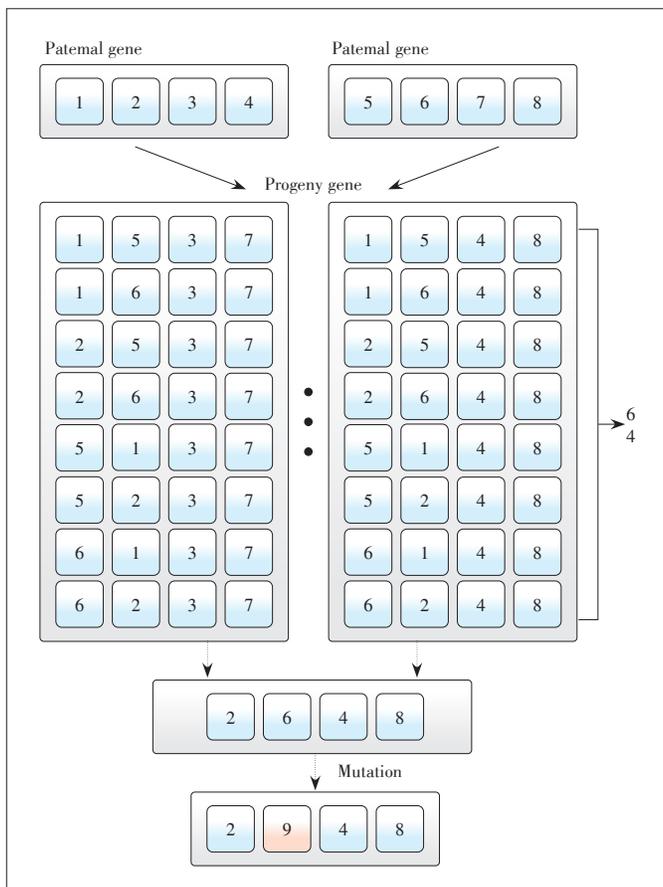
The computational complexity of the GA-PSO algorithm is indeed higher than the PSO algorithm. For the traditional PSO algorithm, the time complexity is $O(N \times m \times (2m + 1)^2 \times T)$. That is the product of the number of receivers, the number of objective function dimensions, the total number of particles, and the number of iterations. In the GA-PSO algorithm, the main computational complexity still comes from the PSO part. And iterations do not run the GA algorithm every time. Instead, every other iteration of h runs the GA algorithm. Through simulation experiments, we found that when the value of h is less than 5, the algorithm works better. In the experimental part that follows, our default h has a value of 5. Therefore, the time complexity of the GA-PSO algorithm in this paper is $O(N \times m \times (2m + 1)^2 \times T + (1 + 4((2m + 1)^2 - 1)) \times T/5)$.

5 Simulation Results and Discussion

In this work, we set up a Monte-Carlo based simulation. We jointly estimate the power and position of 2 unknown transmitters. We fix the number of spatially distributed receivers to 20 and ensured that each receiver has the same weight and assume that the size of the test area is 10 m×10 m.

Before performing the simulation, we must first make a basic estimate of the power and location of the blind source nodes within the system. Some scholars have proposed to use the received power of spatially distributed receivers for clustering, and use the k-means algorithm to quickly narrow the search range of PSO. However, this proposed method is for known nodes with the same power. The problems we face are obviously much more complicated. When the power difference between two unknown transmitters is too large, we may no longer be able to use clustering to determine the approximate location of the blind source node. Because the blind source node with low power is likely to disappear in the signal strength of the high power blind source node. In this paper, we do not pre-estimate the position of the particle in order to more directly show the improvement of the performance of the PSO in dealing with this kind of problem. But we still make a preliminary estimate of the power of the unknown node. First, we set the lower limit of the estimated power to zero. Then we select the transmitter with the lowest received power, assuming that it is at the maximum distance from the transmitter (according to the estimated area, $d_{max} = 10\sqrt{2}$). Substituting the minimum received power and the maximum distance into Equ. (1), what is obtained is the upper limit of the transmission power. Note that this upper limit is the upper limit of the sum of the powers of multiple transmitters.

In the simulation, we assume that the transmit power of the unknown transmitter 1 is always 100 mw, while the transmit power of the unknown transmitter 2 is from 10 mw to 90 mw.



▲ Figure 3. Particle gene exchange and genetic mutation process of two unknown nodes.

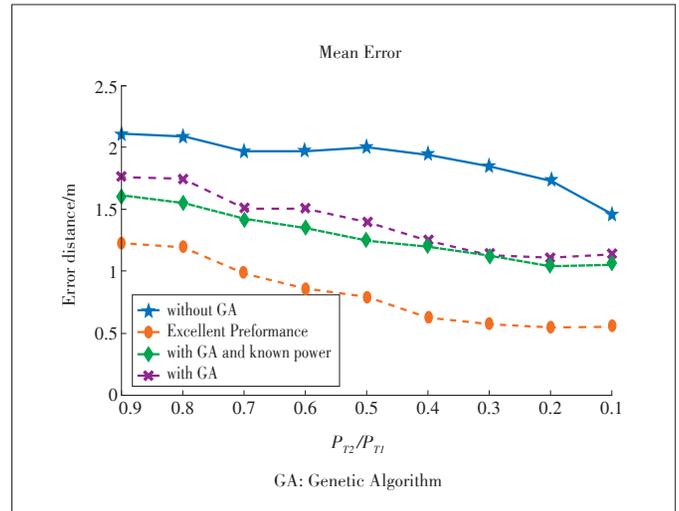
The power of each transmitter is still unknown to the system. We use P_{T2}/P_{T1} as the independent variable and the average estimation error as the dependent variable. In order to form a comparison, we also add the position estimation error of the GA-PSO algorithm under known power. At the same time, we use the idea of the exhaustive method to maximize the particle population of the PSO (excellent performance), and test the lower error limit of the objective function.

From Fig. 4, the estimation performance of the PSO algorithm is greatly improved after adding the GA. Especially when the value of P_{T2}/P_{T1} is in the range of [0.3, 0.7]. The average error is reduced by up to about 0.6 m, and the overall performance is increased by about 30%. However, when the power of the unknown nodes is similar or the difference is large, the system estimates that the performance improvement is slightly reduced. Compared with the results obtained by the exhaustive method, the algorithm proposed in this paper still has a certain gap and there is still room for improvement in the future. However, the difference between the joint estimate and the estimate with known power is not very large.

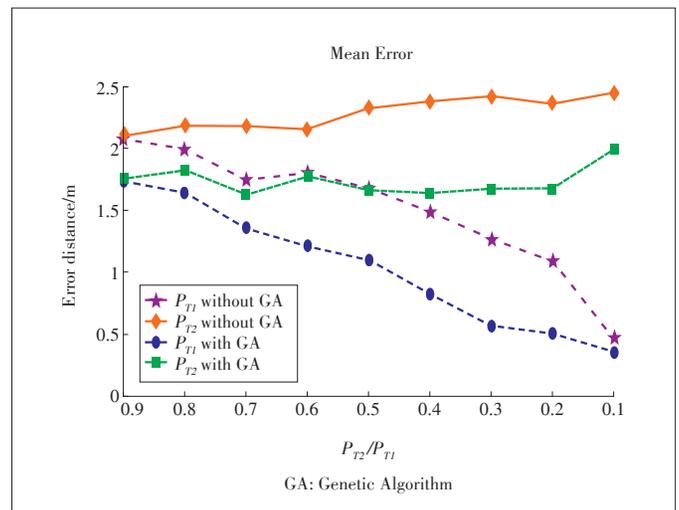
As shown in Fig. 5, as P_{T2}/P_{T1} gradually decreases, the difference between the estimation errors of the unknown node 2 and the unknown node 1 gradually increases. Although the genetic PSO algorithm has reduced both, it cannot suppress this trend. Fig. 6 shows the case where the power estimation error varies with the ratio of P_{T2} and P_{T1} . When the power of two unknown nodes is similar, the power estimation error is only about 10%. And when the power difference between the two is large, the power estimation error is also steady.

The main reason for this trend is the design of the objective function. High power transmitters have a greater impact on the objective function than small power transmitters. Both PSO and GA are algorithms for finding suboptimal solutions. They only care if the fitness function value is small enough. As long as the power and position estimates of a high power transmitter are accurate, the fitness function value can be reduced to a small enough. Therefore, even if the power and position of the small power transmitter are not accurate enough, the algorithm will tend to converge.

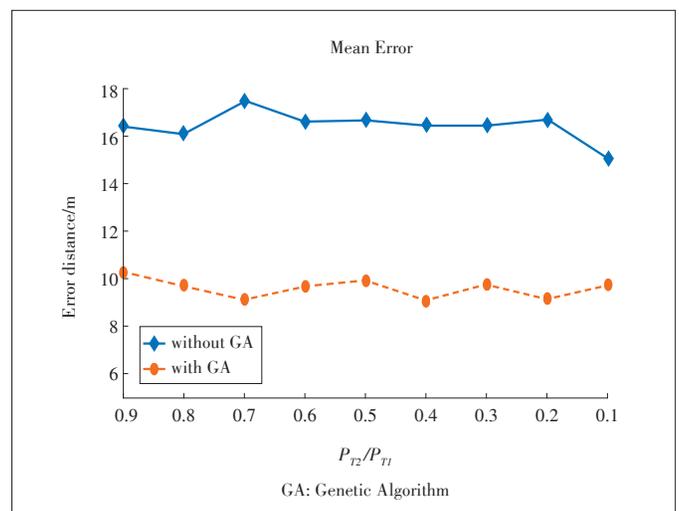
We also conduct an artificial observation analysis of the error results and find that there are still cases where the two unknown nodes estimate the coordinate interchange because they fall into the local minimum. Comparing the values of the fitness function of the local minimum and the global optimal solution, it is found that their fitness function values are indeed very close. The introduction of the GA algorithm can make the algorithm jump out of the local minimum and can effectively guide the estimated value to the global optimal solution. However, there is still no guarantee that it will not fall into other local optimal solutions after jumping out. This shows that the objective function as a non-convex function has many non-inferior solutions. In future research, we can extend this problem from single-objective PSO to multi-objective PSO, and new de-



▲ Figure 4. Mean estimation error distance of all transmitters.



▲ Figure 5. Mean estimation error distance of each transmitter.



▲ Figure 6. Mean estimation error power of each transmitter.

velopment may be made.

6 Conclusions

Joint estimation of power and positions for blind signals is essential to detect the co-frequency interference signal. We propose a data-driven based PSO algorithm combined with a special GA to estimate multiple co-frequency wireless blind source nodes with different powers. To avoid the local minimum problem, we use GA to help particle population jump out of local minima. The simulation results show that the proposed GA-PSO algorithm has better estimation performance than the classical PSO algorithm and can reduce the error distance by up to 30%.

References

- [1] MOKIN V B, SKORYNA L M, YASCHOLT A R, et al. Method for Selecting the Ranking Criteria for Monitoring Stations of the Status of Spatially Distributed Systems and for Defining the Priority of Their Location [C]. IEEE First Ukraine Conference on Electrical and Computer Engineering, Kiev, Ukraine, 2017: 870–875. DOI: 10.1109/UKRCON.2017.8100374
- [2] NELSON J K, HAZEN M U, GUPTA M R. Global Optimization for Multiple Transmitter Localization [C]. 2006 IEEE Military Communications Conference, Washington DC, USA, 2006: 1–7. DOI: 10.1109/MILCOM.2006.302280
- [3] WERNER J, HAKKARAINEN A, VALKAMA M. Cramer-Rao Bounds for Hybrid RSS-DOA Based Emitter Location and Transmit Power Estimation in Cognitive Radio Systems [C]. 38th IEEE Vehicular Technology Conference, Philadelphia, USA, 2013: 1–7. DOI: 10.1109/VTCFall.2013.6692149
- [4] AKKAŞ M A. Nano-Sensor Capacity and SNR Calculation According to Transmit Power Estimation for Body-Centric Nano-Communications [C]. 3rd IEEE International Symposium on Wireless Systems Within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems, Offenburg, Germany, 2017: 51–55. DOI: 10.1109/IDAACS-SWS.2016.7805785
- [5] CHEN X, GONG Y, LIN X H, et al. High Precision RF Transmit Power Calibration Based on Recursive Least Squares Estimation [C]. IEEE International Conference on Communication Systems, Shenzhen, China, 2017: 1–6. DOI: 10.1109/ICCS.2016.7833569
- [6] ZAFER M Z, KO B J, HO I W-H. Transmit Power Estimation Using Spatially Diverse Measurements Under Wireless Fading [J]. IEEE/ACM Transactions on Networking, 2010, 18(4): 1171–1180. DOI:10.1109/TNET.2009.2039801
- [7] POLITIS C, MALEKI S, DUNCAN J M, et al. SDR Implementation of a Testbed for Real-Time Interference Detection With Signal Cancellation [J]. IEEE Access, 2018, 6(99): 20807–20821. DOI: 10.1109/ACCESS.2018.2825885
- [8] ZHANG Y, LIANG J, JIANG S, et al. A Localization Method for Underwater Wireless Sensor Networks Based on Mobility Prediction and Particle Swarm Optimization Algorithms [J]. Sensors, 2016, 16(2): 212. DOI: 10.3390/s16020212
- [9] ZHAO H M, WU Y N. An Improved Positioning Algorithm Based on PSO and PLS [C]. International Conference on Advances in Mechanical Engineering and Industrial Informatics, Zhengzhou, China, 2015: 237–243. DOI: 10.2991/ameii-15.2015.47
- [10] XU X P, LI J, ZHOU M C, et al. Accelerated Two-Stage Particle Swarm Optimization for Clustering Not-Well-Separated Data [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018. DOI: 10.1109/TSMC.2018.2839618
- [11] ZHANG J Q, LU S Y, ZANG D, et al. Integrating Particle Swarm Optimization with Stochastic Point Location Method in Noisy Environment [C]. IEEE International Conference on Systems, Man, and Cybernetics, Budapest, Hungary, 2016. DOI: 10.1109/SMC.2016.7844544
- [12] ZHANG L B, LV H P, TAN D P, et al. Adaptive Quantum Genetic Algorithm for Task Sequence Planning of Complex Assembly Systems [J]. Electronics Letters, 2018, 54(14): 870–872. DOI: 10.1049/el.2018.0609
- [13] BHARATHI C, REKHA D, VIJAYAKUMAR V. Genetic Algorithm Based Demand Side Management for Smart Grid [J]. Wireless Personal Communications, 2017, 93(2): 481–502. DOI: 10.1007/s11277-017-3959-z
- [14] AFZALIRAD M, SHAFIPOUR M. Design of an Efficient Genetic Algorithm for Resource-Constrained Unrelated Parallel Machine Scheduling Problem with Machine Eligibility Restrictions [J]. Journal of Intelligent Manufacturing, 2018, 29(2): 423–437. DOI: 10.1007/s10845-015-1117-6
- [15] WU X L, WU S M. An Elitist Quantum-Inspired Evolutionary Algorithm for the Flexible Job-Shop Scheduling Problem [J]. Journal of Intelligent Manufacturing, 2017, 28(6): 1441–1457. DOI: 10.1007/s10845-015-1060-6
- [16] HAGER T, MELLOULI A, MASMOUDI F. A Multi-Objective Genetic Algorithm for Assembly Line Resource Assignment and Balancing Problem of Type 2 (ALRABP-2) [J]. Journal of Intelligent Manufacturing, 2017, 28(2): 371–385. DOI: 10.1007/s10845-014-0984-6
- [17] ZANG W K, REN L Y, JIANG Z N, et al. Modified Kernel-Based Intuitionistic Fuzzy C-Means Clustering Method Using DNA Genetic Algorithm [J]. Journal of Software Engineering, 2017, 11(2): 172–182. DOI: 10.3923/jse.2017.172.182
- [18] YANG M-D, LIN M-D, LIN Y-H, et al. Multiobjective Optimization Design of Green Building Envelope Material Using a Non-Dominated Sorting Genetic Algorithm [J]. Applied Thermal Engineering, 2017, 111: 1255–1264. DOI: 10.1016/j.applthermaleng.2016.01.015
- [19] LIU X B, GUAN Y L, KOH S N, et al. Low-Complexity Single-Channel Blind Separation of Co-Frequency Coded Signals [J]. IEEE Communications Letters, 2018, 22(5): 990–993. DOI: 10.1109/LCOMM.2018.2805332
- [20] HAN X, ZHANG Y, ABANA M A, et al. A Blind Signal Detection Scheme for Co-Channel Interference Cooperative Systems [C]. 23rd IEEE International Conference on Telecommunications, Thessaloniki, Greece, 2016: 1–5. DOI: 10.1109/ICT.2016.7500405

Biographies

LIU Shen received his B.S. from Zhengzhou University (ZZU), China in 2016. He received his M.S. in electronic and communication engineering from Guilin University of Electronic Technology, China in 2019. He studied and worked in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences and Shenzhen Institute of Radio Testing & Tech, China from 2017 to 2018. He is currently an engineer in Monolithic Power System (MPS), China. His research interests include indoor localization and software defined radio when he was a student.

QIN Yuannian received his B.S. from University of Electronic Science and Technology, China in 1994. In 2005, he received the title of senior experimental teacher. He is currently employed as a professor. He is the director of the Communication Experimental Center. The Communication Experimental Center has been awarded the title of the Experimental Teaching Demonstration Center of Guangxi Universities. As the main participant, he participated in 6 provincial and ministerial scientific research projects, presided over one Guangxi Natural Fund project, one Guangxi science and technology development project, three horizontal scientific research projects, 14 horizontal scientific research projects as the main personnel, published 22 papers, and participated in the compilation and publication of the textbook Mobile Communications.

LI Xiaofan received her B.S. and Ph.D. degrees from Beijing University of Posts and Telecommunication, China in 2007 and 2012. From 2010 to 2011, she studied in University of Washington, USA as an exchanged Ph.D. student. She joined the State Radio Monitoring Center and Testing Center (SRTC) in 2012 and was transferred to SRTC Shenzhen Lab from 2013. She is now an associate professor in Jinan University, Zhuhai, China. Her research interests include interference analysis among different radio systems, testing and evaluation methods for innovative radio technologies, cooperative communication, cognitive radio, internet of things, radio management strategy, etc.

ZHAO Yubin (zhaoyubin 2001@gmail.com) received his B.S. and M.S. from Beijing University of Posts and Telecommunications, China in 2007 and 2010. He received his Ph.D. degree in computer science in 2014 from Freie Universitat Berlin (FU Berlin), Germany. He has been an assistant professor in Center for Cloud Computing, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China, since 2014. He serves as the reviewers for many scientific journals and session chairs for several conferences. He co-organized the workshop in ICCS 2019. His current research interests include indoor localization and target tracking, wireless power transfer and mobile edge computing.

XU Chengzhong received the Ph.D. degree from the University of Hong Kong in 1993. He is a full professor in the Department of Computer and Information Science, Faculty of Science and Technology, State Key Laboratory of IoT for Smart City, University of Macau, China. His research interests include networked computing systems and applications. He is an IEEE Fellow due to contribution in resource management in parallel and distributed computing.

ZTE Communications Guidelines for Authors

Remit of Journal

ZTE Communications publishes original theoretical papers, research findings, and surveys on a broad range of communications topics, including communications and information system design, optical fiber and electro-optical engineering, microwave technology, radio wave propagation, antenna engineering, electromagnetics, signal and image processing, and power engineering. The journal is designed to be an integrated forum for university academics and industry researchers from around the world.

Manuscript Preparation

Manuscripts must be typed in English and submitted electronically in MS Word (or compatible) format. The word length is approximately 3000 to 8000, and no more than 8 figures or tables should be included. Authors are requested to submit mathematical material and graphics in an editable format.

Abstract and Keywords

Each manuscript must include an abstract of approximately 150 words written as a single paragraph. The abstract should not include mathematics or references and should not be repeated verbatim in the introduction. The abstract should be a self-contained overview of the aims, methods, experimental results, and significance of research outlined in the paper. Five carefully chosen keywords must be provided with the abstract.

References

Manuscripts must be referenced at a level that conforms to international academic standards. All references must be numbered sequentially in-text and listed in corresponding order at the end of the paper. References that are not cited in-text should not be included in the reference list. References must be complete and formatted according to ZTE Communications Editorial Style. A minimum of 10 references should be provided. Footnotes should be avoided or kept to a minimum.

Copyright and Declaration

Authors are responsible for obtaining permission to reproduce any material for which they do not hold copyright. Permission to reproduce any part of this publication for commercial use must be obtained in advance from the editorial office of *ZTE Communications*. Authors agree that a) the manuscript is a product of research conducted by themselves and the stated co-authors, b) the manuscript has not been published elsewhere in its submitted form, c) the manuscript is not currently being considered for publication elsewhere. If the paper is an adaptation of a speech or presentation, acknowledgement of this is required within the paper. The number of co-authors should not exceed five.

Content and Structure

ZTE Communications seeks to publish original content that may build on existing literature in any field of communications. Authors should not dedicate a disproportionate amount of a paper to fundamental background, historical overviews, or chronologies that may be sufficiently dealt with by references. Authors are also requested to avoid the overuse of bullet points when structuring papers. The conclusion should include a commentary on the significance/future implications of the research as well as an overview of the material presented.

Peer Review and Editing

All manuscripts will be subject to a two-stage anonymous peer review as well as copyediting, and formatting. Authors may be asked to revise parts of a manuscript prior to publication.

Biographical Information

All authors are requested to provide a brief biography (approx. 100 words) that includes email address, educational background, career experience, research interests, awards, and publications.

Acknowledgements and Funding

A manuscript based on funded research must clearly state the program name, funding body, and grant number. Individuals who contributed to the manuscript should be acknowledged in a brief statement.

Address for Submission

<http://mc03.manuscriptcentral.com/ztecom>

ZTE COMMUNICATIONS

中兴通讯技术(英文版)

ZTE Communications has been indexed in the following databases:

- Abstract Journal
- Cambridge Scientific Abstracts (CSA)
- China Science and Technology Journal Database
- Chinese Journal Fulltext Databases
- Index of Copernicus
- Inspec
- Ulrich's Periodicals Directory
- Wanfang Data

ZTE COMMUNICATIONS

Vol. 17 No. 3 (Issue 67)

Quarterly

First English Issue Published in 2003

Supervised by:

Anhui Publishing Group

Sponsored by:

Time Publishing and Media Co., Ltd.

Shenzhen Guangyu Aerospace Industry Co., Ltd.

Published by:

Anhui Science & Technology Publishing House

Edited and Circulated (Home and Abroad) by:

Magazine House of ZTE Communications

Staff Members:

Editor-in-Chief: WANG Xiyu

Associate Editor-in-chief: JIANG Xianjun

Executive Associate Editor-in-Chief: HUANG Xinming

Editor-in-Charge: ZHU Li

Editors: XU Ye and LU Dan

Producer: ZHA Hongqin

Circulation Executive: WANG Pingping

Assistant: WANG Kun

Editorial Correspondence:

Add: 12F Kaixuan Building, 329 Jinzhai Road,
Hefei 230061, P. R. China

Tel: +86-551-65533356

Email: magazine@zte.com.cn

Online Submission: <https://mc03.manuscriptcentral.com/ztecom>

Annual Subscription: RMB 80

Printed by:

Hefei Tiancai Color Printing Company

Publication Date: September 25, 2019

Publication Licenses: ISSN 1673-5188
CN 34-1294/TN