

Yue Wang
Yuyang Liu
Jiaqi Zou
Mengyao Huo *Editors*

Signal and Information Processing, Networking and Computers

Proceedings of the 10th International
Conference on Signal and Information
Processing, Networking and Computers
(ICSINC)



Series Editors

Leopoldo Angrisani, *Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy*

Marco Arteaga, *Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico*

Bijaya Ketan Panigrahi, *Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India*

Samarjit Chakraborty, *Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany*

Jiming Chen, *Zhejiang University, Hangzhou, Zhejiang, China*

Shanben Chen, *Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China*

Tan Kay Chen, *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*

Rüdiger Dillmann, *Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany*

Haibin Duan, *Beijing University of Aeronautics and Astronautics, Beijing, China*

Gianluigi Ferrari, *Università di Parma, Parma, Italy*

Manuel Ferre, *Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain*

Sandra Hirche, *Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany*

Faryar Jabbari, *Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA*

Limin Jia, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*

Janusz Kacprzyk, *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Alaa Khamis, *German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt*

Torsten Kroeger, *Stanford University, Stanford, CA, USA*

Yong Li, *Hunan University, Changsha, Hunan, China*

Qilian Liang, *Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA*

Ferran Martín, *Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*

Tan Cher Ming, *College of Engineering, Nanyang Technological University, Singapore, Singapore*

Wolfgang Minker, *Institute of Information Technology, University of Ulm, Ulm, Germany*

Pradeep Misra, *Department of Electrical Engineering, Wright State University, Dayton, OH, USA*

Sebastian Möller, *Quality and Usability Laboratory, TU Berlin, Berlin, Germany*

Subhas Mukhopadhyay, *School of Engineering and Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand*

Cun-Zheng Ning, *Electrical Engineering, Arizona State University, Tempe, AZ, USA*

Toyoaki Nishida, *Graduate School of Informatics, Kyoto University, Kyoto, Japan*

Luca Oneto, *Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy*

Federica Pascucci, *Dipartimento di Ingegneria, Università degli Studi Roma Tre, Roma, Italy*

Yong Qin, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*

Gan Woon Seng, *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore*

Joachim Speidel, *Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany*

Germano Veiga, *Campus da FEUP, INESC Porto, Porto, Portugal*

Haitao Wu, *Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China*

Walter Zamboni, *DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy*

Junjie James Zhang, *Charlotte, NC, USA*

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Yue Wang · Yuyang Liu · Jiaqi Zou ·
Mengyao Huo
Editors

Signal and Information Processing, Networking and Computers

Proceedings of the 10th International
Conference on Signal and Information
Processing, Networking and Computers
(ICSINC)

Editors

Yue Wang
China Academy of Space Technology
Beijing, China

Jiaqi Zou
Beijing University of Posts
and Telecommunications
Beijing, China

Yuyang Liu
Beijing University of Posts
and Telecommunications
Beijing, China

Mengyao Huo
Beijing University of Posts
and Telecommunications
Beijing, China

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-19-9967-3

ISBN 978-981-19-9968-0 (eBook)

<https://doi.org/10.1007/978-981-19-9968-0>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

It is our great honor to welcome you to the 10th International Conference on Signal and Information Processing, Network and Computers (ICSINC 2022). The ICSINC 2022 Committee has been monitoring the evolving COVID-19 pandemic. We have decided to delay the 2022 edition of this conference from July to September.

ICSINC 2022 provides a forum for researchers, engineers, and industry experts to discuss recent development, new ideas and breakthrough in signal and information processing schemes, computer theory, space technologies, big data, and so on.

ICSINC 2022 received 314 papers submitted by authors, in which 150 papers were accepted and included in the final conference proceedings. The accepted papers were presented and discussed in 27 regular technical sessions and 2 workshops.

On behalf of the ICSINC 2022 committee, we would like to express our sincere appreciation to the TPC members and reviewers for their tremendous efforts. Especially, we appreciate all the sponsors for their generous support and advice, including Springer, China Unicom, HuaCeXinTong company. Finally, we would also like to thank all the authors for their excellent work and cooperation.

Songlin Sun
Yue Wang
Changbo Zhu
ICSINC 2022 General Chairs

Committee Members

International Steering Committee

Songlin Sun	Beijing University of Posts and Telecommunications, China
Ju Liu	Shandong University, China
Chenwei Wang	DOCOMO Innovations (DoCoMo USA Labs), USA
Jiaxun Zhang	China Academy of Space Technology, China
Takeo Fujii	The University of Electro-Communications, Japan
Hongwu Li	China Unicom Research Institute, China
Xiaoyu Ye	China Unicom Research Institute, China
Changbo Zhu	China Unicom QingHai Branch, China
Fang Wang	China Unicom QingHai Branch, China

General Co-chairs

Songlin Sun	Beijing University of Posts and Telecommunications, China
Yue Wang	China Academy of Space Technology, China
Changbo Zhu	China Unicom QingHai Branch, China

Technical Program Committee Chairs

Xinzhou Cheng	China Unicom Research Institute, China
Qiang Zhang	China Unicom QingHai Branch, China

Publication Chair

Yuyang Liu	Beijing University of Posts and Telecommunications, China
------------	--

Exhibits Chair

Lianxin Xin

China Unicom QingHai Branch, China

Local Service Committee

Zhijuan Tao

China Unicom QingHai Branch, China

Wei Shang

China Unicom QingHai Branch, China

Sponsors

Springer



China Unicom



HuaCeXinTong



Contents

Information Technology Session 1

A New Multi-instance Learning Algorithm Integrated with TF-IDF for Entity Relation Extraction in Electronic Medical Records	3
<i>Yiyang Xiong, Yajuan Qiao, Shilei Dong, Xuezhi Zhang, and Hua Tan</i>	
A Deep Learning Algorithm Using Feature Engineering to Adjust Attention Mechanisms and Neural Network for Cloud Security Detection	11
<i>Yiyang Xiong, Yajuan Qiao, Shilei Dong, Xuezhi Zhang, and Hua Tan</i>	
The Design and Implementation of a Ranging System Based on the DSSS	19
<i>Yingpeng Cao, Fu Li, Shuyi Qin, Wei Zhang, and Zihong Yuan</i>	
Application of Grey Approximation Ideal Solution Ranking Methods in Optimal Selection of Satellite Initial Orbits	27
<i>Yang Lei, Wang Fan, and Bo Zhenyong</i>	
Filter and Closed Loop Control Technology of Piezoelectric Drivers	39
<i>Qian Lu, Bin Ren, and Le Zhang</i>	
A Novel Back-Feed Double-Layer Microstrip Antenna	48
<i>Pengfei Zhao, Dabao Wang, Mingxuan Wu, Xiangwei Ning, Xiaojuan Li, and Jinyuan Ma</i>	
A Full-Digital Test Method for Satellite-Borne Software	55
<i>Pan Xin, Jiao Rong-Hui, and He Jing</i>	
3D Warehousing: Enabling Intelligent Warehousing Visualization Based on Three.js	63
<i>Wenqian Shan, Wenbo Wan, Anyi Chen, Leichao Ren, Jiande Sun, Meng Fang, and Yantong Zhan</i>	
Deep Learning Neural Networks with Auto-adjustable Attention Mechanism for Server Fault Diagnosis Under Log Data	72
<i>Yiyang Xiong, Yajuan Qiao, Shilei Dong, Xuezhi Zhang, and Hua Tan</i>	
A Hierarchical Classification Algorithm of Spacecraft Telemetry Data Based on Time Series Characteristics	80
<i>Peng Wan, Tang Li, Da Tang, and Qianqian Zhang</i>	

A Saliency-Transformer Combined Knowledge Distillation Guided Network for Infrared Small Target Detection	88
<i>Wei Zhang, Wenquan Feng, Menghao Li, Shuchang Lyu, and Ting-Bing Xu</i>	
Research on Cloud-Edge Collaboration Based on Object Storage for Internet of Things	96
<i>Qiu hong Zheng, Dan Liu, Yun Shen, and Peng Ding</i>	
Industrial Defect Detection System Based on Edge-Cloud Collaboration and Task Scheduling Technology	104
<i>Yuying Xue, Yun Shen, and Huibin Duan</i>	
Actuator Grouping Optimization Method of Antenna Reflector Based-on Ant Colony Algorithm	113
<i>Yan Du, Songjing Ma, Tao Ma, Xiangshuai Song, and Jiayou Zhang</i>	
A Method to Measure the Quality of a Cloud Network Across Multiple Resource Pools	122
<i>Su Yue, Chen Ruihao, Li Wei, Zhao Weibo, and Ma Fei</i>	
A New Cloud Computing Deployment Model: Proprietary Cloud	130
<i>Weibo Zhao, Su Yue, Ma Fei, Ruihao Chen, and Li Wei</i>	
Overview of the Development of Secure Access Service Edge	138
<i>Ruihao Chen, Su Yue, Weibo Zhao, Ma Fei, and Li Wei</i>	
Optimization and Numerical Investigation of Micro-pin-Fin Structure on Heat Sink with Checkerboard Nozzles	146
<i>Huajie Lv, Kehan He, and Xing Ju</i>	
A Method of Design and Evaluation of X-ray Communication in Black-Out Area	155
<i>Fangyuan Xia, Bo Wang, Guoting Zhang, Yabo Yuan, Jian Wu, Yingjun Zhang, Yao Li, Furui Zhang, Zhenkun Tan, Yun Du, and Tong Su</i>	
A Robust DOA Estimation Method Based on Auxiliary Sensors and Power Inversion	161
<i>Yihao Song, Puming Huang, Ledu Qiu, Shuai Li, and Ming Li</i>	
Diffusion Convolution Graph Attention Network for Spatial-Temporal Prediction	170
<i>Xiang Yin, Lei Wu, Yanqiang Zhang, Yanni Han, and Kun Zhai</i>	

Fast Estimation Technology of Orbit Information for Non-cooperative Space Targets Based on AREKF Filtering Theory 179
Zhuo Zhang, Gang Liu, Henghai Fan, Dongmei Kuang, Anliang Li, and Ruixi Gaoya

Information Technology Session 2

Design of Intelligent Power Supply and Distribution Test System for Small Satellite Based on Distributed Structure 189
Yuexin Hu, Jing Zhang, Guiying Zhang, Guangjie Ren, and Hui Yao

Design of Small Satellite Power Supply Interface Test System Based on PXI Bus 198
Jianfeng Dai, Zhi Yang, Haichao Wu, and Yiming Cheng

Research on Method of On-Orbit Evolution for Aerospace Control Software 207
XiaoFeng Li, XiaoGang Dong, Bin Gu, Ming Zhao, RuiMing Zhong, Yi Li, and Jing Wang

Research on Automatic Assembly Technology for Spacecraft System 219
Zejian Chen, Jizhi Yang, Jianping Xu, Yangcheng Zhang, Yongliang Li, and Yi Yue

Design and Research of Multi-user Distributed Configuration Management Based on Zookeeper 228
Ming Zhang, Zhaojian Shen, Bin Yin, Li Cui, and Fan Xu

A Complete Ensemble Local Mean Decomposition and Its Application in Doppler Radar Vital Signs Monitoring System 236
Meng Zhang, Zhibin Yu, Pang Rong, and Gao Yuan

The Key Research of High Speed Camera Based on Multiple Core CMOS Sensors 245
Yanzhang Li, Hong Zeng, Guangsen Liu, Jun Zhu, and Xiaoling Che

Research on Cargo Level Optimization Based on Multi-objective Optimization Algorithm 253
Lin Shi, Yanan Tan, Hongling Chen, Wenbo Wan, Yunhua Wang, and Leichao Ren

The Application of Single Pulse Output Module Integrated by FPGA in On-Board Electronic Products 260
Sun Xiaofeng, Sun Hao, Xing Weiwei, and Wangbo

Design of Steady Speed Control System for Space Borne Scanning Loads	267
<i>Yong Zhou, Jie Pang, Fuqiang Liu, Chao Ma, and Chao Dong</i>	
Research on Slotting Optimization Based on MOEA/D	276
<i>Hongling Chen, Yanyan Tan, Lin Shi, Wenbo Wan, Leichao Ren, and Yunhua Wang</i>	
An Improved Density Peak Clustering Algorithm Based on Gravity Peak	284
<i>Hui Han and Rui Zhang</i>	
A DPC Based Recommendation Algorithm for Internship Positions	292
<i>Rui Zhang, Lingyun Bi, and Tao Du</i>	
Optimal Method of Air Cargo Loading Under Multi-constraint Conditions	300
<i>Yuheng Lu, Chunyun Dong, Meng Nan, Xiaolong Chen, and Yufan Wei</i>	
Research on Software Synthesis Method for Spacecraft Control System	309
<i>Yi Li, Xiaogang Dong, Xiaofeng Li, Bin Gu, Xingsong Zhao, Yanxia Qi, and Bo Yu</i>	
Prediction and Analysis of Infectious Disease Visit Data Based on DPC	319
<i>Rui Zhang, Yuhui Song, and Tao Du</i>	
Research on the Method of Improving the Accuracy of MRP Calculation Through Big Data in Aerospace Enterprise	328
<i>Wen Liu, Ruihua Li, and Yuzhu Li</i>	
Design and Application of Bad Block Management Strategy for On-Board Solid-State Memory	334
<i>Hongjun Ma, Jianhong Xiao, Dayang Zhao, and Yifeng He</i>	
An Automated Scoring System for Photoshop Course in Secondary Vocational Colleges	341
<i>Peng Liu, Zhiyan Wang, Xiufang Liu, and Wenbo Wan</i>	
Discussion on the Application of Contributing Student Pedagogy in Vocational Education “Computer Network Technology”	349
<i>Fanjie Lv and Jiande Sun</i>	
Research on NURBS Interpolation Algorithm Based on Newton Iteration and Adams Equation	355
<i>Yunsen Wang, Lin Li, Chengzhi Ma, SiPei Shao, and Yangchuang Cao</i>	

Fake AP Detection Technology Based on Improved DBSCAN Algorithm 362
Zhaoyuan Mei, Chenwei Wang, Lvxin Xu, and Songlin Sun

Vehicles Location Estimation and Prediction in Integrated Sensing
and Communications: Sensing Vehicles as Extended Targets 371
Guanyu Zhang, Zhilei Ling, and Songlin Sun

Research on Vehicle and Cargo Matching of Electric Materials Based
on Weed Optimization Algorithm 380
Xiao Yu, Dedi Huang, Yanlong Gao, Jun Zhu, and Bohan Ren

Improved MSFLA-Based Scheduling Method of Electric Power
Emergency Materials 389
Tiezheng Wang, Xiao Yu, Ming An, and Liang Shi

A Survey of Class Activation Mapping for the Interpretability
of Convolution Neural Networks 399
Mingwei He, Bohan Li, and Songlin Sun

Robust Multi-person Reconstruction in Right Spatial Arrangement
from In-the-Wild Scenes 408
Wei Wang, Ronghui Zhang, Hai Huang, and XiaoJun Jing

Research on the Curriculum System of Electronic Information
Engineering for the Certification of Engineering Education 417
*Songlin Sun, Jiaqi Zou, Zhilei Ling, Hai Huang, Tiankui Zhang,
Caili Guo, and Changchuan Yin*

Multimedia and Communication Networks

An S Band Antenna Using Meta-surface for Satellite TT & C Test 425
*Xiangwei Ning, Fanhui Zhou, Liang An, Shuo Feng, Pengfei Zhao,
and Aixin Chen*

Design and Implementation of a Test and Verification System
for Spaceborne Synthetic Aperture Radar 430
An Liang, Liang Jian, Yang Li, Hao Zhiya, and Zhongjiang Yu

An Optimized Scheme for Telecommand Transmission 437
Liu Bo, Weining Hao, Hongguang Wang, and Weisong Jia

Transmission Efficiency Improvement of Ka-Band VCM
Satellite-to-Ground Data Transmission System of LEO Remote Sensing
Satellite Based on DVB-S2X 443
Zhongguo Wang and Dabao Wang

Anti-interference Control for On-Orbit Servicing Spacecraft Formation Under Communication Resource Limitation	452
<i>Tao Wang, Yingchun Zhang, Heli Gao, Jie liu, and Hongchen Jiao</i>	
FPGA-Based High-Speed Remote Sensing Satellite Image Data Transmission System	460
<i>Yifeng He, Peng Sun, Qian Liu, Hao Zhang, and Dayang Zhao</i>	
Application of Frequency Domain Filtering in Edge Detection	471
<i>Bo Li and Hongyan He</i>	
Research on the Anti-interference of a Satellite Telecommand Interface	478
<i>Shifeng Gao, Jingang Wang, Luyuan Wang, Zhenyu Wu, and Duanguo Si</i>	
Intelligent Propagation Prediction Model for Wireless Radio Channel Based on CNN	486
<i>Yajuan Qiao, Yiyang Xiong, Shilei Dong, Xuezhi Zhang, and Hua Tan</i>	
Research on Application Efficiency Analysis Method of Remote Sensing Satellite for Emergency	494
<i>Gao Han, Bai Zhaoguang, Che Xiaoling, Zhao Liang, Wu Bin, Wang Chao, Zhu Jun, and Bai Yuchen</i>	
Research on Blind Restoration Algorithm of Motion Blurred Remote Sensing Image	505
<i>Hanwen Yang, Yanran Liu, Li Luan, Yunsen Wang, and Haoxuan Wang</i>	
A Fast Algorithm of CU Block Division Based on Frequency Domain on VVC	513
<i>Zihao Liu, Lei Chen, and Chenggang Xu</i>	
Optimization of AVS3 Intra Derived-Tree Mode Algorithm Based on DCT Coefficients	520
<i>Chenggang Xu, Fuyun Kang, Yi Wu, and Lei Chen</i>	
Integrated Sensing and Communications with OTFS: Application in V2I Communications	528
<i>Zhilei Ling, Yuxin Wu, Guanyu Zhang, and Songlin Sun</i>	
Space Technology Session 1	
Fault Analysis and Evaluation of Control Moment Gyroscope on Orbit	541
<i>Jinfeng Zhou, Xi Chen, and Boneng Tan</i>	

Precision Orbit Determination for LEO Based on BDS Navigation Data 551
Chao Li, Xiusong Ye, Hong Ma, Shouming Sun, and Yang Yang

Research and Practice of Electrostatic Grounding Technology
of Aerospace Electronic Products Based on Intelligent Active Protection
Concept 560
Ying Zhu, Dong Wang, Ke Li, Wei Qiao, Yijian Zheng, and Wenze Qi

Investigation on Feasibility of Covert TT&C Scheme Using Satellite
Payload Channel 567
Jie Chen, Haiming Qi, and Fengchun Wang

A Satellite Imaging Stability Monitoring Method Based on Lunar
Calibration Technology 574
Yong-chang Li, Qiang Cong, and Shun Yao

Design and Verification of mK Temperature Fluctuation Control
for Gravity Gradiometer 581
Wei Liu, Yupeng Zhou, and Tinghao Li

A Perigee Orbit Change Method for Geostationary Satellite 590
ShuHua Zhang, Heng Yao, JianPing Li, and ZhongMou Lei

Optimal Design of Power Drive Unit Layout Based on Improved
Discrete Particle Swarm Optimization 598
Yufan Wei, Chunyun Dong, Meng Nan, Xiaolong Chen, and Yuheng Lu

Preliminary Discussion on the Application of Digital Thread
in the Model-Based Spacecraft Development 606
Lingfeng Zhao, Ziwei Wu, Lingyun Cheng, and Guoxiong Zhan

Study on the Construction Method of Spacecraft System Model Based
on Meta-model 614
Jiangdong Ruan, C. G. Li, Yuxuan Li, and Xinwu Chen

The Design and Implementation of Magnetic Control of Ultra Quiet
Control System in Gravity Satellite 622
Bin Guan, Yiwu Liu, Qirui Liu, Yan Li, and Jiakun Fan

Intercomparison of On-Orbit Consistency of Radiometric Calibration
Between Two Sensor Calibration Spectrometers of HY Satellite 630
Huiting Gao, Yue Ma, and Qingjun Song

Loose Formation Keeping Control for Low-Earth Orbit Satellite Cluster Considering the Interaction Between Cluster Satellites	638
<i>Huang He, Yang Xu, Zhang Qi, and Jing Baohong</i>	
Research on Data Fusion Architecture of GNC Subsystem of China Space Station	647
<i>Jingsong Li, Jing Wang, Haixin Yu, Xiaofeng Li, Ruiming Zhong, Xiaogang Dong, Zhaohui Chen, and Junchun Yang</i>	
Application in Emergency Tasks Schedule with the GFDM-1 Satellite	656
<i>Mingliang Liu, Hu Qiu, and Ziwei Li</i>	
Emergency Scheduling Process Design for ZY303 Satellite	663
<i>Ziwei Li, Mingliang Liu, and Hu Qiu</i>	
Research on the Application of TDRSS for Manned Spacecraft	673
<i>Da Tang, Tang Li, Peng Wan, Zhisheng Wang, and Yushu Feng</i>	
Modeling and Analysis of Satellite with a Large Inertia Rotating Load	681
<i>Guiming Li, Shixuan Liu, Zhihui Li, Xinyan Wu, Kui Zou, and Rui Liu</i>	
Research on Satellite-to-Ground Integration Joint Mission Planning System for Mega-constellation Multi-user Remote Sensing Information Resource Guarantee	689
<i>Shaoyu Zhang, Xu Yang, Ruolan Zhang, Keqiang Xia, Xiaobo Hui, Haipeng Liu, and Xing Wang</i>	
Analysis of the Influence of Satellite Drift Angle Correction Datum on Imaging Quality	696
<i>Chao Wang, Yuze Cao, Han Gao, Cheng Yan, Hongzhi Zhao, Keli Zhang, Jun Zhu, and Lili Wang</i>	
Application and Process of Inclination Adjustment of Satellites in Sun Synchronous Orbit	705
<i>Shenggang Wang and Yuan Jun</i>	
Analysis Method for Isolation Index of Transceiver Link of Space Ground TT&C Equipment	713
<i>Jingyu Zhao, Jianxiao Zou, Shuo Wang, Yalong Yan, and Ben Su</i>	
Satellite Momentum Wheel Life Prediction Method Based on PSO-LSSVM	723
<i>Zhang Chao, Wenjie Ma, and Jiexuan Song</i>	

Design of Equivalent Simulation Test System for Universal Power Supply and Distribution of Space Station Multi-aircraft 731
Xu-zhen Jing and Peng Ying

Design and Verification of a Geographic Information Prediction Scheme for Autonomous Mission Planning of Satellites 740
Rina Wu, Jie Liu, Tao Zhang, Jianmin Zhou, and Chao Chen

Space Technology Session 2

Bit Synchronization Verification Strategy for High Orbit Navigation Receiver 749
Mengdan Cao, Yu Chen, Yukui Zhou, Zhaoqiang Cheng, Yu Peng, and Dongsheng Shi

Real Time Dynamic Adjustment Calculation Method of Camera Integration Time Based on Spaceborne Digital Elevation Model 759
Linfeng Xing, Chao Xue, Xin Guan, and Gaojian Lv

A Time Sequence Design for Improving the Output Frequency and Accuracy of Star Sensor Under High Dynamic Conditions 766
Rui Liu, GuiMing Li, ZhiHui Li, and JianFu Zhang

Research on the Measurement Method of Inter-satellite Clock Difference and Distance of Formation Satellites 772
Yang Liu, Yujie Liu, and Sufang Chen

Research on Asteroid Detection Radar Based on High Precision Signal Delay Estimation 780
Sufang Chen, Mengna Jia, Yang Liu, and Dong Liu

Review of Micro-vibration Modeling, Suppression, and Measurement for Spacecraft with High-Precision 787
Meng Ge, Kuai Yu, Yongsheng Wu, Dong Wang, Guangyuan Wang, and Chongzhou Yu

Satellite Gravity Gradient Measurement Technology Based on Superconducting Suspension 797
Ming Li, Yue-xin Hu, and Ming-xue Shao

The Mathematic Analytical Results of Self Excited Flyback Converter with Double Switches 804
Shan Li, Dinghao Sun, Liwei Wang, Xuhui Liu, and Xudong Wang

Anomaly Interpretation Method Based on Data Mining for Remote Sensing Satellite Payload	814
<i>Jia You and Chengzhi Lu</i>	
A Study on Determination Method of Safety Band in the Translational Closing Section of Rendezvous and Docking	823
<i>Honghui He and Changqing Chen</i>	
Prospect Analysis of Satellite Application of Long-Distance Wireless Energy Transmission System	829
<i>Sen Wang, Xiaofei Li, Xiaochen Zhang, Yudan Liu, and Bingxin Zhao</i>	
Design of Multi-dimensional Control and Multi-subsystem Cooperative Operation Algorithm for Control Subsystem	838
<i>Hong Guan, Jian Cai, HeLong Liu, Yong Li, and Zefeng Ma</i>	
Research on Attitude Determination of Spacecraft Two-Stage Composite Control System	844
<i>Gaojian Lv, Chao Xue, Xin Guan, and Linfeng Xing</i>	
Design of Dual Robot Collaborative Assembly Scheme for Aerospace Products	850
<i>Lijian Zhang, Dingwen Pan, and Ruiqin Hu</i>	
Study on the Application of Robot-Assisted Assembly in Satellite Board Assembly	858
<i>Tiecheng Qiu, Qian Zhang, and Liang Guo</i>	
3D Model-Based Design Method for Complex Cable Network Three-Dimensional Marking of Spacecraft	866
<i>Qian Zhang, Tiecheng Qiu, and Yu Fu</i>	
Study on the Optimization of Temperature Measurement Points for Small Satellite	873
<i>Zhijia Li, Dawei Li, and Zhiming Xu</i>	
Design of Model-Based Spacecraft Flight Control Support System	881
<i>Wang Xiaoming, Gao Hongtao, Fan Daliang, Qu Xiaoyu, and Zhang Xiaogong</i>	

Big Data Workshop Session 1

Research and Application of 5G Massive MIMO Antenna Weight Intelligent Optimization Based on 4G/5G Coordination 893
Zixiang Di, Xinzhou Cheng, Jiajia Zhu, Yi Li, Lexi Xu, Jinjian Qiao, Lu Zhi, and Wenzhe Wang

A Method of 5G Core Network Service Fault Diagnosis 902
Shiyu Zhou, Yong Wang, Xiqing Liu, Xinzhou Cheng, Zhenqiao Zhao, and Tian Xiao

Research on User Behavior Credibility Evaluation Model in Trusted Network 911
Anshun Zhou, Suimin Wang, Mingde Huo, Jianzhi Wang, Yuwen Huo, Shihan Fu, and Lexi Xu

Research and Application of Key Technologies of Intelligent Manufacturing Based on 5G Vehicle 920
Shangyu Tang, Mingde Huo, Yan Zhang, Xin Zhao, Yuwen Hou, Ying Ji, and Guoyu Zhou

Research on Key Technologies of 5GC Network Intelligent Operation Under the Background of Digital Transformation 929
Zhenqiao Zhao, Xiqing Liu, Yong Wang, Jie Miao, Shiyu Zhou, Xinzhou Cheng, Lexi Xu, Xin He, and Icy Yan

Research on Location Algorithm of Mobile Network Based on Hidden Markov Model 938
Bei Li, Wei Zhao, Kai Zhou, Lexi Xu, Guanghai Liu, Tian Xiao, Chen Cheng, Lu Zhi, Wei Zhang, and Ziwei Zhu

Research on the Development of 5G Messaging Service for Operators 947
Mengni Chen, Ziyuan Zhu, MuYun, Jianying Guo, Yimeng Song, and Sainan Hou

Investigation of 5G Terminal Users with Dual SIMs 954
Jiajun Li, Bawei Pei, Yukun Liu, and Fengwei Chen

Research on Deep Packet Inspection for Driving Digital Operation 961
Miao Sun, Qiang Zhang, and Gan Han

Research on the Operation Model of Operator Computing Products 968
Yimeng Song, Fan Shi, Mengni Chen, Yun Mu, and Jianying Guo

Big Data Processing Mode Based on Cross Data Source SQL Engine 975
Yilong Li, Qiang Zhang, Bin Han, and Jianru Wang

Online Machine Learning-Based Quality Difference Identification and Prediction Prevention for Broadband Users 981
Wei Deng, Lin Yuan, Ling Zhen, Xina Zhang, Jingzhu Chen, and Xiuyuan He

Achieving Privacy-Preserving Diagnosis with Federated Learning in LEO Satellite Constellation 990
Qinglei Kong, Zhidi Lin, Feng Yin, Lexi Xu, Xinzhou Cheng, Shuguang Cui, and Xiaoyu Ye

Cellular Network Coverage Hole Detection and Diagnosis Method Using WaveCluster 999
Zijing Yang, Lexi Xu, Feibi Lyu, Lixia Liu, Jiajia Zhu, Kun Chao, and Xinzhou Cheng

A Novel Method of Trace Message Decoding and the Performance Analysis 1007
Shengli Guo, Jinghui Li, Xiaodong Cao, Chao Wang, Chuntao Song, Xinzhou Cheng, Runsha Dong, Tianyi Wang, and Xijuan Liu

Study on Cost Difference Between Peak-Valley Pricing and Flat Pricing 1014
Zeyi Yang, Yang Liu, Feng Luo, Xiaoxuan Du, Zetao Xu, Pengcheng Liu, Ao Shen, and Yuan Fang

Three-Dimensional Integrated Base Station Rental Fee Benchmark Evaluation 1022
Zetao Xu, Pengcheng Liu, Yang Zhang, Shuaizhong Pan, Zeyi Yang, Yang Liu, Feng Luo, and Xiaoxuan Du

The Energy Saving Measurement System and Method of Main Base Station Communication Equipment 1030
Yang Liu, Zeyi Yang, Feng Luo, Xiaoxuan Du, Xidian Wang, Zetao Xu, Ao Shen, and Yuan Fang

Energy Saving Technology and Parameter Research of 5G Terminal 1038
Yuan Fang, Jiandi Luo, Yun Lu, Jian Hu, Ao Shen, Pengcheng Liu, Yao Cen, Zetao Xu, Jimin Ling, and Qintian Wang

Research on Energy Consumption Modelling of 5G Wireless Communication Main Equipment Based on GBRT Algorithm 1046
Yao Cen, Yuan Fang, Zeyi Yang, Yunlong Liu, Sheng Zhang, Zetao Xu, Ao Shen, and Yang Liu

Research of Voice Speech Quality Evaluation Based on Trace Data 1054
Lu Zhi, Xinzhou Cheng, Jiajia Zhu, Lexi Xu, Bei Li, Zixiang Di, Liang Liu, Jinjian Qiao, and Zhaoning Wang

Analysis of the Impact of COVID-19 on the Development of National ICT Industry 1062
Yang He, Cheng Feng, Quan Yu, Rui Zhou, and Ziyuan Zhao

Research on Low Latency and High Reliability Wireless Technologies and Solutions 1070
Yi Li, Qingliang Long, Zixiang Di, Yuchao Jin, Yuting Zheng, Lexi Xu, Xinzhou Cheng, and Bei Li

Resource Multiplexing Schemes of URLLC and eMBB Under Multi-service Coexistence Scenario Based on Management Aspects 1078
Yuchao Jin, Yi Li, Yuting Zheng, Deyi Li, Xinzhou Cheng, Qingliang Long, Lexi Xu, and Tian Xiao

Big Data Workshop Session 2

5G Construction Efficiency Enhancement Based on LSTM Algorithm 1089
Tian Xiao, Qingliang Long, Lexi Xu, Guanghai Liu, Zixiang Di, Bei Li, Zhaoning Wang, Shiyu Zhou, and Fei Xue

Business Circle Attraction Based on DPI 1097
Wei Zhang, Yuhui Han, Qingqing Zhang, Tianyi Wang, Chen Cheng, Lexi Xu, Xinzhou Cheng, and Bei Li

A Novel Method of Obtaining Location Information 1105
Shengli Guo, Lexi Xu, Xiaodong Cao, Chao Wang, Chuntao Song, Xinzhou Cheng, Runsha Dong, Tianyi Wang, Jinghui Li, and Xijuan Liu

Research on Network Quality Intelligent Monitoring Technology Based on Prophet Model 1112
Hongbo Shen, Baoyou Wang, Saibin Yao, Jiucheng Huang, Zhanqiang Liu, Lei Chen, and Jie Tian

Uplink Tx Switching Application Optimization Based on XGBoost Algorithm 1121
Tao Huang, Tian Xiao, Yao Wei, Lexi Xu, Ji Lan, Pengxiang Li, Jiansheng Liang, and Xinyan Wang

Research and Application of Key Technologies in 5G Mobile Network E2E Slicing	1130
<i>Rui Xiao, Jihua Li, Sirui Zhong, Xinzhou Cheng, Xiaoyu Ye, and Xinrui Yang</i>	
Research on 5GtoB Service Monitoring and Prediction Scheme in the Full Scenario of ‘Smart Winter Olympics’	1138
<i>Rui Xiao, Wei Zeng, Lei Wang, and Xueting Zhang</i>	
Research on Handling User Complaints in 5G Key Scenarios Based on Reinforcement Learning	1147
<i>Wei Zeng, Jihua Li, Weiye Xia, Xueting Zhang, and Ulrich Kleber</i>	
Research on 5G Networking and Massive MIMO Intelligent Optimization Method Based on Big Data and AI for Winter Olympics Venues	1155
<i>Wei Zeng, Jun Fan, Haotian Wu, Andrey Krendzel, and Rui Xiao</i>	
Users’ Perception Monitoring Operation System Based on Big Data Algorithm for Highway Scenarios	1163
<i>Chao Liu, Bei Li, Hui Pan, Li Xu, Xufeng Hang, Zhenwei Jiang, Baoyou Wang, and Ziwei Zhu</i>	
Weak Coverage Analysis Method for Mobile Networks Based on Machine Learning	1171
<i>Jinjian Qiao, Jiajia Zhu, Xinzhou Cheng, Lexi Xu, Ning Meng, Lijun Cheng, Jinyu Zhai, Zixiang Di, and Fred Dong</i>	
Indoor Positioning Based on Enhanced 5G Fingerprint Positioning Algorithm	1179
<i>Li Xu, Saibin Yao, Sibing Rao, Qiuyue Hu, Chao Liu, and Haiyun Zhu</i>	
Root Cause Analysis Based on Trace for Mobile Network Problem	1185
<i>Liang Liu, Xinzhou Cheng, Jiajia Zhu, Lexi Xu, Songbai Liang, Lijun Cheng, Jinyu Zhai, and Fred Dong</i>	
A Solution of 5G Multi-beam Optimization Based on Big-Data	1193
<i>Xufeng Hang, Yi Zeng, Baoyou Wang, Chao Liu, and Fengli Dai</i>	
A Microservice-Based Visual Orchestration Platform for 5G IoT Applications	1201
<i>Zhaoning Wang, Jiajia Zhu, Xinzhou Cheng, Jinjian Qiao, Feibi Lyu, Tian Xiao, Xu Han, and Tao Lyu</i>	

Comprehensive Analysis Scheme of Video Service Based on XDR 1209
Lijuan Cao, Xin He, Yuwei Jia, Kun Chao, Yunyun Wang, Lexi Xu, Chen Cheng, Heng Zhang, Yuchao Jin, and Yi Li

5G Close Loop Proactive Optimization Using Network Data Analysis Function 1217
Feibi Lyu, Jiajia Zhu, Zhaoning Wang, Jinjian Qiao, Liang Liu, Zixiang Di, Chen Cheng, Tian Xiao, and Yuan Zhang

Evolution of Intelligent Telecom Networks in 5G Era 1225
Runsha Dong, Xiaodong Cao, Han Liu, Lexi Xu, Xin He, Chen Cheng, Shengli Guo, and Xinzhou Cheng

Research and Application of 5G Super Uplink Technology in Plateau 1234
Lixia Liu, Weijun Chen, Xun Zhu, and Xinjie Hou

Author Index 1241

Information Technology Session 1



A New Multi-instance Learning Algorithm Integrated with TF-IDF for Entity Relation Extraction in Electronic Medical Records

Yiyang Xiong[✉], Yajuan Qiao, Shilei Dong, Xuezhi Zhang, and Hua Tan

China Telecom Corporation Limited Research Institute, Beijing, China
xiongyy1@chinatelecom.cn

Abstract. Electronic medical record (EMR) is the core data which generated in the process of clinical treatment. By extracting the record text's entity relationship, we can catch a large amount of information closely related to the patient's condition. Deep learning is widely used in the field of entity relationship extraction in EMR. Among them, seq-to-seq is the most popular network structure for text sequence, but it can't distinguish the influence between different words. Therefore, multi-instance learning which are improved by attention mechanism can adjust the weight of words is proposed and achieved outstanding results. However, the reason for the creation of the attention mechanism parameter is always puzzling: it lacks interpretability and it is difficult for us to understand the origin of the numbers. Term frequency-inverse document frequency (TF-IDF) is an algorithm to calculate the importance of words according to their occurrence frequency, so we think integrating it into the construction of neural network may increase model's interpretability. In this paper, we propose a new multi-instance learning algorithm, which integrates the TF-IDF into the construction of deep learning model. Compared with the network without TF-IDF, our model achieves an improvement of nearly 2% and the convergence rate is reduced by 5% on average.

Keywords: Multi-instance learning · TF-IDF · Relation extraction

1 Introduction

1.1 Background

In daily medical activities, doctors, nurses and other medical personnel usually use the medical information system to generate multiple types of medical records which is called electronic medical records (EMR). EMR records the diagnosis and treatment process and medication suggestions in detail, which has high research value [1]. With gradual improvement of medical informatization, EMR is becoming more and more popular, and gradually plays a more important role in clinical diagnosis and treatment.

In EMR natural language processing, entity relation extraction is the core tasks. It refers to obtain the relationship between known entities. For example, here is a text: doctors treat patients. This text has a pair of entities of doctors and patients, and the entity relationship is 'treat' (the relationship of 'treat' needs to be defined in advance).

Deep learning not only has strong ability, but also can solve problems end-to-end, so it become mainstream method of EMR entity relation extraction. Deep learning applied to EMR can be roughly divided into supervised learning algorithm and unsupervised learning algorithm. Supervised learning algorithm usually adopts convolutional neural network which represented by textCNN [2] and PCNN [3], it can obtain excellent performance [4], but these methods need a large number of positive samples, so the cost of large-scale learning is high. Semi supervised learning can make up for this problem because it can quickly obtain a large number of annotation data. The commonly used semi supervised learning include remote supervision learning [5] and multi-instance learning [6]. Among them, multi-instance learning introduces attention mechanism, so the weight of context can be adjusted, and it achieves better results and become the most popular algorithm.

1.2 Our Innovation

The introduction of attention mechanism greatly increases the processing ability of neural network for long text [7]. Transformer, which abandons neural network and adopts attention mechanism completely [8], even leads the research direction of natural language processing. Because the text length of EMR is generally long, it is particularly effective. However, the initial parameters of attention mechanism are often set randomly, which brings two problems: 1. The algorithm has great randomness in convergence. 2. Insufficient interpretability, so it is difficult to realize improvement in the future. In addition, the embedding layer of neural network lacks meaningful annotation of data: it usually only has location information. So, if we can set the initial parameters of neural network and attention mechanism through the understanding of the text, can we create a better model?

TF-IDF is a very mature technology in the field of text search [9]. It can judge the importance of a word in the text through word frequency. In this paper, we introduce its calculation results into the initial parameters of model, and propose a new multi-instance learning algorithm for EMR. By doing this, we can deliver neural network the importance of each word before training starts. Compared with the network without integration, the convergence speed of our proposed algorithm has been improved by more than 5%, and the accuracy has also been improved.

2 Model Architecture

2.1 Overview

When encountering deep learning problems, the ideal state is that the data is either black or white, such as whether a picture is a face. However, in practice, the samples we have are difficult to meet the situation above, especially for the problem of entity relationship extraction: when any two entities appear at the same time, the relationship is usually not unique. For example, there are two sentences: doctors treat nurses and nurses cooperate with doctors. Nurses and doctors have different relationships in these two sentences: treatment and cooperation. Remote supervision method and the multi-instance learning belong to semi supervised learning which can solve the problem above.

Since multi-instance learning can significantly reduce the impact of false tagging in Distant Supervision method [10], we adopt its framework.

The multi-instance learning algorithm makes a hypothesis: if an entity pair has a certain relationship in the knowledge base. Then, in the data containing the entity pair, at least one sentence expresses this relationship. Unlike ordinary supervised learning, which aims at a single sample, multi-instance learning is bag oriented, and a bag contains multiple samples. It has two definitions: 1. If one sample in a bag is a positive sample, the whole bag is marked as a positive bag. 2. If a bag is all negative samples, then the bag is negative bag. Let me give you an example to understand this. For example, we can judge whether there is gold ore from the excavated ore. If there is gold in the trial excavated ore, it is considered that there is gold ore here. If none of the trial excavated ore contains gold, it is considered that there is no gold ore here.

Therefore, for the problem of electronic medical record, we adopt multi-instance learning algorithm. Our multi-instance learning algorithm consists of two parts: neural network part and attention mechanism part. Our innovation integrates the results of TF-IDF calculation into the construction of two modules. Through our innovation, our model has the ability to transmit the importance of each word to the neural network from the perspective of word frequency, which quickly improves the convergence speed and slightly improves the accuracy. Next, we will introduce the construction of the model in detail.

2.2 Neural Network

In terms of neural network, we adopt Piecewise Convolutional Neural Network (PCNN) as basic framework. The TF-IDF PCNN model we proposed has following framework. Different colors in the figure represent the results of different convolution kernels. In addition, we use blue to represent the starting position of the entity.

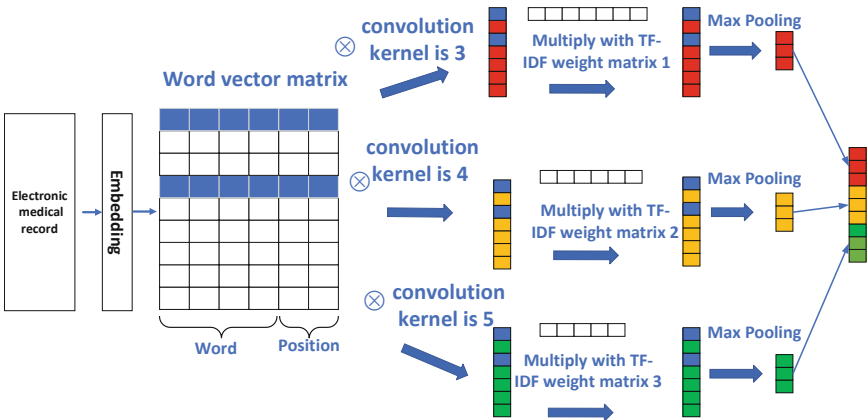


Fig. 1. The figure shows the structure of the neural network. It consists of five parts: word embedding, convolution, TF-IDF weight adjustment, Max pooling and splicing.

As shown in Fig. 1, We use a common sentence in the electronic medical record to describe how to deal with this example and convert it into sentence vector. Next, we will elaborate the processing process of each step through the example.

Word embedding. We first need to convert the text into a word vector that can be understood by computer and neural network. We take a common Chinese text in electronic medical record as an example. Here is a Chinese sentence: “医生和护士相互协作”. Since it consists of 9 rows (each row represents a word), and the number of columns is set to 6. This text is transformed into a $9 * 6$ matrix.

There are two kinds of word embedding methods. One is that word vector can be obtained in other corpora in advance, and the other is obtained by network training as unknown parameters [11]. These two methods have their own advantages. The first method can use other corpora to get more a priori knowledge. The second method can grasp the features associated with the current task better. In order to make good use of the advantages of both, our model adopts a dual channel method: one of them is embedded in pre trained words and will not change; The other is initialized in the same way, but will be changed with the training process. Please note that since we use the method of piecewise convolution, we need to mark the position of the entity and record it in the word vector (the position part of the vector). The starting position of doctors and nurses is marked here and highlighted in blue in Fig. 1. It can be described as the following formula.

$$S = R^{r*c} \quad (1)$$

R is the length of the sentence (number of words), C is the width of the convolution kernel, and the size of C is determined by position coding and word embedding.

Convolution. The biggest feature of convolution in natural language processing is the width of convolution kernel is the same as that of sentence matrix, and it can only move in the height direction, which reflects words are the smallest constituent unit of sentences. Therefore, in the example given in this paper, the width of convolution kernel is 6. In terms of length, since convolution kernels of different sizes can learn complementary features from each other [12], we set up three convolution kernels with lengths of 3, 4 and 5, and obtained convolution results respectively. The contents can be described by the following formula.

$$C \in R^{r+w-1}, C_j = wq_{j-w+1:j} \quad s + w - 1 \geq j \geq 1, \quad (2)$$

W is the size of convolution. Since we use three convolution kernels, so j is equal to three and the convolution output is:

$$C = \{C_1, C_2, C_3\}, C_{i3} = w_i q_{3-w+1:3} \quad (3)$$

TF-IDF adjustment. Here is an important embodiment of our TF-IDF in innovation. In the construction of word vector, each word is embedded according to the way of pre training or neural network convergence, and the neural network treats them equally. However, we should understand that the information behind each entity is actually different. If a word appears infrequently in the corpus and suddenly appears in our sentences,

it may bring more information and deserve our more attention. TF-IDF reflects this fact. Its formula consists of three parts:

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (4)$$

TF (Term Frequency) indicates the frequency of the entry in the text, where $n_{i,j}$ indicates the number of times the entry t_i appears in the document d_j , and $TF_{i,j}$ indicates the frequency of the entry t_i in the document d_j .

$$IDF_i = \log \frac{|D|}{1 + |j : t_i \in d_j|} \quad (5)$$

IDF (inverse document frequency) indicates the prevalence of keywords, where $|D|$ indicates the number of all documents and $|j : t_i \in d_j|$ indicates the number of documents containing the term t_i .

$$TF - IDF = TF * IDF \quad (6)$$

TF-IDF is obtained by multiplying formulas (1) and formulas (2). We can infer that TF-IDF tends to filter out common words and pay attention to the words that appear frequently in the text and are rare in the corpus.

We calculate TF-IDF value of each word according to word frequency and combine them to form TF-IDF vector, which can reflect the importance of each word in the sentence. We multiply it with the result of convolution:

$$C_t = TF * IDF * C \quad (7)$$

Now, we can tell the neural network which words need more attention.

Max pooling and splicing. Finally, we use pooling layer to reduce the complexity of the system, and splice the results of different convolution kernel calculations. A complete sentence vector is now complete.

2.3 Attention Mechanism

Multi-Instance Learning usually produces a lot of false data, we need to reduce the impact of noise. In addition, sentences deserve different attention due to its word's frequency. Based on the problems above, we propose a new attention mechanism improved by TF-IDF. The specific framework is shown in Fig. 2. Our attention mechanism is sentence-level.

We creatively calculate the contribution of each sentence to the text through TF-IDF, which means that the model has an initial parameter with physical meaning. In addition, the data oriented by multi-instance learning has many wrong labels, so it is also necessary to reduce the impact of these wrong labels through the attention mechanism. The formula of the attention mechanism is as follows:

$$\beta = \frac{\exp(e_i)}{\sum_k \exp(e_k)} \quad (8)$$

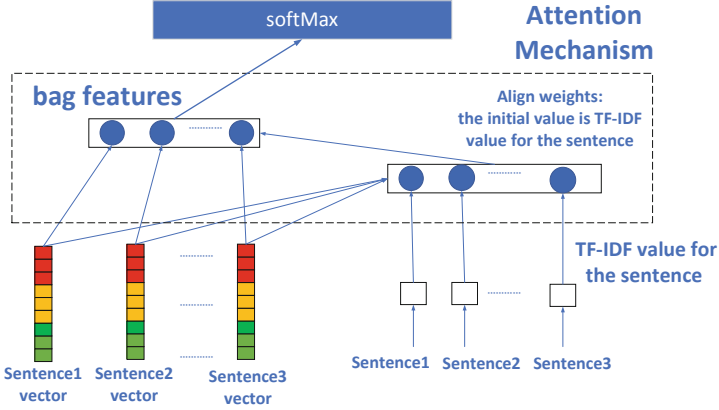


Fig. 2. The figure describes our sentence-level attention mechanism. Its initial parameters are calculated by TF-IDF. Its calculation will get the final classification result through SoftMax.

Among them, e_i is called a query function, which scores the matching degree between the input sentence x_i and the prediction relationship r and its specific expression is as follows, TF-IDF determines its initial value, where A is a weighted diagonal matrix:

$$e_i = x_i r * TF - IDF * A \quad (9)$$

Finally, we define the output of SoftMax according to [13]:

$$p(r|S) = \frac{\exp(o_r)}{\sum_{k=1}^{n_r} \exp(o_k)} \quad (10)$$

Among them, n_r is the total number of relations in the corpus and o is the output of neural network.

3 Experiment and Analysis

We are committed to the informatization of the medical industry. Therefore, the experiment of this paper is oriented to electronic medical records. Related work is usually based on the English data set. The difference of this paper is that the experiment is carried out for the Chinese text. The data set we use is from Ruijin Hospital in Shanghai. The content of the dataset is diabetes clinic information and related academic papers provided by Ruijin hospital. It has the entity information that has been tagged, so we can extract entity relationship directly. There are ten kinds of entity relations defined in the dataset. Five of them are related to diseases: Test_Disease, Symptom_Disease, Treatment_Disease, Drug_Disease, Anatomy_Disease. The remaining categories are related to drugs: Frequency_Drug, Duration_Drug, Amount_Drug, Method_Drug, SideEff_Drug.

Our experiment evaluates the model ability by using Precision, Recall and F1, we compare our proposed algorithm with multi-instance learning algorithm (PCNN + ATT),