



信息通信领域产学研合作特色期刊

第三届全国期刊奖百种重点期刊 | 中国科技核心期刊

ISSN 1009-6868

CN 34-1228/TN

CODEN ZTJHAY

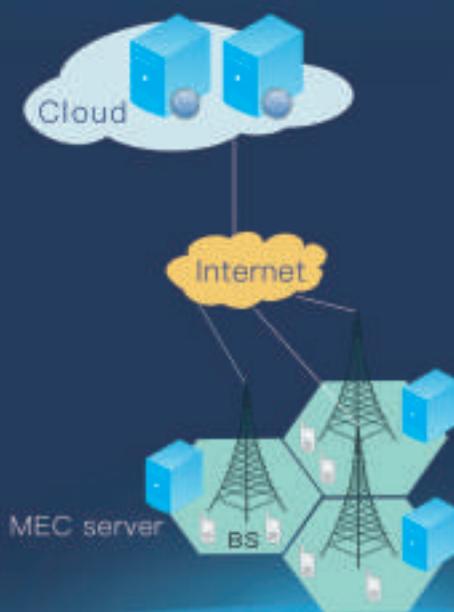
# 中兴通讯技术

## ZTE TECHNOLOGY JOURNAL

tech.zte.com.cn

2018年4月 • 第2期

### 专题：大数据智能化无线网络技术



中兴通讯技术

ZTE TECHNOLOGY JOURNAL

二〇一八年四月

第二十四卷第二期

总第一三九期



# 《中兴通讯技术》第7届编辑委员会委员名单

**顾问** 侯为贵（中兴通讯股份有限公司创始人） | 钟义信（北京邮电大学教授） | 陈锡生（南京邮电大学教授）

**主任** 陆建华（中国科学院院士,清华大学教授）

**副主任** 赵先明（中兴通讯股份有限公司总裁） | 糜正琨（南京邮电大学教授）

**副主任** 马建国（广东工业大学教授） | 陈前斌（重庆邮电大学副校长）

## 编委（按姓名拼音排序）

- |                             |                                      |
|-----------------------------|--------------------------------------|
| <b>曹淑敏</b> 北京航空航天大学党委书记     | <b>唐雄燕</b> 中国联通网络技术研究院首席专家           |
| <b>陈建平</b> 上海交通大学教授         | <b>童晓渝</b> 中电科软件信息服务有限公司副总经理         |
| <b>陈杰</b> 中兴通讯股份有限公司高级副总裁   | <b>王京</b> 清华大学教授                     |
| <b>陈前斌</b> 重庆邮电大学副校长        | <b>王文东</b> 北京邮电大学教授                  |
| <b>葛建华</b> 西安电子科技大学教授       | <b>王喜瑜</b> 中兴通讯股份有限公司副总裁             |
| <b>管海兵</b> 上海交通大学教授         | <b>王翔</b> 中兴通讯股份有限公司副总裁              |
| <b>洪波</b> 中兴发展股份有限公司总裁      | <b>卫国</b> 中国科学技术大学教授                 |
| <b>洪伟</b> 东南大学教授            | <b>吴春明</b> 浙江大学教授                    |
| <b>纪越峰</b> 北京邮电大学教授         | <b>邬贺铨</b> 中国工程院院士                   |
| <b>蒋林涛</b> 中国信息通信研究院科技委主任   | <b>徐安士</b> 北京大学教授                    |
| <b>李尔平</b> 浙江大学教授           | <b>续合元</b> 中国信息通信研究院副总工              |
| <b>李红滨</b> 北京大学教授           | <b>徐慧俊</b> 中兴通讯股份有限公司执行副总裁           |
| <b>李建东</b> 西安电子科技大学副校长      | <b>薛一波</b> 清华大学教授                    |
| <b>李军</b> 清华大学教授            | <b>杨义先</b> 北京邮电大学教授                  |
| <b>李乐民</b> 中国工程院院士,电子科技大学教授 | <b>杨震</b> 南京邮电大学校长                   |
| <b>李融林</b> 华南理工大学教授         | <b>尤肖虎</b> 东南大学教授                    |
| <b>李少谦</b> 电子科技大学教授         | <b>张宏科</b> 北京交通大学教授                  |
| <b>李星</b> 清华大学教授            | <b>张平</b> 北京邮电大学教授                   |
| <b>刘建伟</b> 北京航空航天大学教授       | <b>张云勇</b> 中国联通研究院院长                 |
| <b>陆建华</b> 中国科学院院士,清华大学教授   | <b>赵慧玲</b> 中国通信标准化协会网络与业务能力技术工作委员会主席 |
| <b>马建国</b> 广东工业大学教授         | <b>赵先明</b> 中兴通讯股份有限公司总裁              |
| <b>孟洛明</b> 北京邮电大学教授         | <b>郑纬民</b> 清华大学教授                    |
| <b>糜正琨</b> 南京邮电大学教授         | <b>钟章队</b> 北京交通大学教授                  |
| <b>庞胜清</b> 中兴通讯股份有限公司执行副总裁  | <b>周亮</b> 南京邮电大学教授                   |
| <b>孙知信</b> 南京邮电大学教授         | <b>朱近康</b> 中国科学技术大学教授                |
| <b>谈振辉</b> 北京交通大学教授         |                                      |



信息通信领域产学研合作特色期刊  
 第三届国家期刊奖百种重点期刊  
 中国科技核心期刊  
 工信部优秀科技期刊  
 中国五大文献数据库收录期刊  
 ISSN 1009-6868  
 CN 34-1228/TN  
 1995年创刊

### 办刊宗旨

以人为本,荟萃通信技术领域精英  
 迎接挑战,把握世界通信技术动态  
 立即行动,求解通信发展疑难课题  
 励精图治,促进民族信息产业崛起

# Contents 目次

中兴通讯技术 总第139期 第24卷 第2期 2018年4月

## 专题:大数据智能化无线网络技术

- 02 大数据驱动的“人工智能”无线网络 ..... 张琰,盛敏,李建东
- 06 大数据驱动的无线网络资源管理及控制 ..... 刘媛妮,赵国锋
- 10 基于强化学习的无线网络智能接入控制技术 ..... 严牧,孙耀,冯钢
- 15 移动边缘计算中数据缓存和计算迁移的智能优化技术 ..... 汪海霞,赵志峰,张宏纲
- 19 TD-LTE网络中大气波导干扰的分析与预测 ..... 孙天宇,周婷,杨咏
- 25 基于数据驱动深度学习方法的无线信道均衡 ..... 杨咏,李扬,周明拓
- 30 mMTC网络中基于空口流量的入侵检测 ..... 卢楠,杜清河,任品毅
- 38 5G网络人工智能化的基本框架和关键技术 ..... 王威丽,何小强,唐伦

## 专家论坛

- 43 智能物联网技术和应用的发展趋势 ..... 杨咏

## 企业视界

- 47 免调度非正交多址技术及其接收机设计 ..... 邱刚,田力,王沙,袁志锋

## 技术广角

- 54 新一代无线定位技术研究与发展趋势分析 ..... 陈诗军,王慧强,陈大伟
- 59 软件定义天地一体化网络:架构、技术及挑战 ..... 许方敏,仝宗健,赵成林,秦智超

期刊基本参数:CN 34-1228/TN\*1995\*b\*16\*64\*zh\*P\* ¥ 20.00\*15000\*12\*2018-04

# Contents 目次

ZTE TECHNOLOGY JOURNAL Vol. 24 No. 2 Apr. 2018

## Special Topic: Big Data-Driven Intelligent Wireless Network

- 02 Big Data-Driven "Artificial Intelligence" Wireless Network ..... ZHANG Yan, SHENG Min, LI Jiandong
- 06 Resource Management and Control of Big Data-Driven Wireless Network ..... LIU Yuanni, ZHAO Guofeng
- 10 The Intelligent Access Control Mechanisms in Wireless Network Based on Reinforcement Learning ..... YAN Mu, SUN Yao, FENG Gang
- 15 Intelligent Optimization of Data Caching and Computation Offloading in Mobile Edge Computing System ..... WANG Haixia, ZHAO Zhifeng, ZHANG Honggang
- 19 Analysis and Prediction of Atmospheric Duct Interference in TD-LTE Networks ..... SUN Tianyu, ZHOU Ting, YANG Yang
- 25 The Wireless Channel Equalization Based on the Data-Driven Deep Learning Method ..... YANG Yang, LI Yang, ZHOU Ming Tuo
- 30 Air-Interface Traffic-Load Based Intrusion Detection over mMTC Networks ..... LU Nan, DU Qinghe, REN Pinyi
- 38 5G Networks Based on Artificial Intelligence: Basic Framework and Key Techniques ..... WANG Weili, HE Xiaoqiang, TANG Lun

## Expert Forum

- 43 Trends of Technologies and Applications for Intelligent IoT ..... YANG Yang

## Enterprise View

- 47 On Grant-Free NOMA Technology and the Associated Receiver Design ..... QIU Gang, TIAN LI, WANG Sha, YUAN Zhifeng

## Technology Perspective

- 54 New Generation Wireless Locations Technology: Research and Development Trend ..... CHEN Shijun, WANG Huiqiang, CHEN Dawei
- 59 Architecture, Technology and Challenges of Software Defined Integrated Space and Terrestrial Network ..... XU Fangmin, TONG Zongjian, ZHAO Chenglin, QIN Zhichao

## 敬告读者

本刊享有所发表文章的版权,包括英文版、电子版、网络版和优先数字出版版权,所支付的稿酬已经包含上述各版本的费用。

未经本刊许可,不得以任何形式全文转载本刊内容;如部分引用本刊内容,须注明该内容出自本刊。

## 2018年第1—6期专题

### 1 5G 承载网技术和优化组网

张云勇 中国联通研究院院长  
徐雷 中国联通研究院高级工程师

### 2 大数据智能化无线网络技术

陈前斌 重庆邮电大学副校长

### 3 毫米波与太赫兹通信技术

洪伟 东南大学教授  
王海明 东南大学教授

### 4 5G 回传网络光电子器件技术

孙笑晨 中兴光电子技术有限公司技术总监  
徐勇积 中兴光电子技术有限公司总经理

### 5 可再生能源供电的无线通信

牛志升 清华大学教授

### 6 区块链技术及其物联网应用

刘建伟 北京航空航天大学教授

# 专题名称:大数据智能化无线网络技术

## 策划人简介



陈前斌

重庆邮电大学二级教授、博士生导师,享受国务院政府特殊津贴专家,现为重庆邮电大学副校长、移动通信技术重庆市重点实验室主任、中国通信学会会士、重庆市电子学会副理事长兼秘书长;获得国家科技进步二等奖、重庆市技术发明一等奖、重庆市科技进步一等奖等奖项;主持国家和省部级重点项目等各类科研项目 20 余项;获得授权国家发明专利 92 项(采用或转让 27 项),发表学术论文 100 余篇。

## 内容导读

为了使以 5G 为代表的无线网络更加自动化和智能化,目前国际电信联盟 (ITU) 正式成立了“未来网络(包括 5G)-机器学习焦点组”,重点研究机器学习、人工智能在包含第 5 代移动通信(5G)系统的未来网络中的应用。为了实现灵活、高效、高质量的管理、服务、运营的“自动化”,欧洲电信标准化协会(ETSI)发布了一份名为《自动化下一代网络中的网络和服务操作的必要性和益处》白皮书,核心目标是基于大数据、人工智能实现 5G 网络与服务管理、运营的自动化。未来移动通信论坛发布了《无线大数据与智慧 5G》白皮书,提出“智慧 5G”这一全新概念,认为利用人工智能与无线大数据技术,5G 网络能够具备一定的自动化和智能化的能力。

大数据、人工智能与无线网络技术日趋融合,以 5G 为代表的无线网络设计及管理将随着大数据和人工智能技术的应用而得到极大提升,基于无线网络数据驱动,结合数据挖掘、统计分析、深度学习、通信理论和专家系统等诸多技术的人工智能技术,为网络管理及网络性能的优化提供智能化和自动化的可能,有助于提高无线通信网络的设计、运营和优化,能够从业务体验、用户感受、网络质量、网络效率和网络成本等各个方面自主优化网络并提升网络性能。

大数据智能化无线网络技术的发展前景究竟如何?所面临的问题与挑战究竟怎样解决?采用了哪些关键技术和方案?为了反映大数据、人工智能在无线网络领域的研究现状及最新研究成果,展示大数据、人工智能面临的理论和技术上的新挑战,揭示大数据和人工智能的研究热点及研究方向,本期专题从大数据驱动的“人工智能”无线网络、大数据驱动的无线网络资源管理及控制、无线网络智能接入控制、移动边缘计算中数据缓存和计算迁移的智能优化技术、基于数据驱动深度学习方法的无线信道均衡、人工智能化的基本框架等,分别在多个方面阐述了大数据、人工智能领域的研究成果和经验,希望能给读者提供有益的启示和参考。在此,对各位作者的积极支持和辛勤工作表示衷心的感谢!

陈前斌

2018 年 3 月 22 日



# 大数据驱动的“人工智能”无线网络

## Big Data-Driven "Artificial Intelligence" Wireless Network

张琰/ZHANG Yan

盛敏/SHENG Min

李建东/LI Jiandong

(西安电子科技大学, 陕西 西安 710071)  
(Xidian University, Xi'an 710071, China)

无线网络是大数据应用的重要载体之一;另一方面,无线网络的规模化、密集化发展也将产生海量的运维数据。以某区域运营商现网情况为例,无线基站每日操作维护中心(OMC)产生的各类数据将近数百吉比特,但目前运营商对于该类数据的应用率仅约30%左右。因此,如何充分挖掘和利用网络运维大数据来支撑无线网络的智能化已经成为未来无线网络发展的关键技术问题。

本文从“人”(用户特征获取与分析)、“工”(工作参数配置与调整)、“智”(智慧资源管控与优化)、“能”(承载能力提升与进化)4个方面阐述了支撑无线网络智能化发展的功能架构,如图1所示。同时,结合不同功能和应用场景的特点,提出了大数据分析以及人工智能在无线网络中的典型用例,从而实现网络资源与用户需求的精准匹配。本文旨在给大家提供一种启示性研究思路,使得大数据分析以及人工智能在未来无线通信网络中得到广泛关注,支撑未

收稿日期: 2018-01-15

网络出版日期: 2018-03-17

基金项目: 国家科技重大专项(2017ZX03001010-004); 国家自然科学基金(61725103)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0002-004

**摘要:** 提出了以用户特征获取与分析、工作参数配置与调整、智慧资源管控与优化、承载能力提升与进化为支撑的智能无线网络功能架构。以网络资源与用户需求的精准匹配为目标,研究了大数据分析以及无线网络人工智能特性间的耦合关系。给出了大数据分析以及人工智能在无线网络中的典型用例,并初步验证了其有效性,同时展望了智能无线网络的发展前景和方向。

**关键词:** 大数据分析; 人工智能; 无线资源管理; 服务质量(QoS)

**Abstract:** In this paper, a smart wireless network architecture based on acquisition and analysis of user characteristics, configuration and adjustment of work parameters, intelligent resource management and optimization, and capacity improvement and evolution is proposed. With the accurate matching of network resources and user requirements, the coupling between big data analysis and artificial intelligence of wireless network is studied. The typical use cases of big data analysis and artificial intelligence in wireless networks are given. The validity of this model is verified, and the development prospect and direction of smart wireless networks are also forecasted.

**Keywords:** big data analysis; artificial intelligence; wireless resource management; quality of service (QoS)

来人与人、人与物、物与物的全面互联需求。

### 1 “人”: 用户特征获取与分析

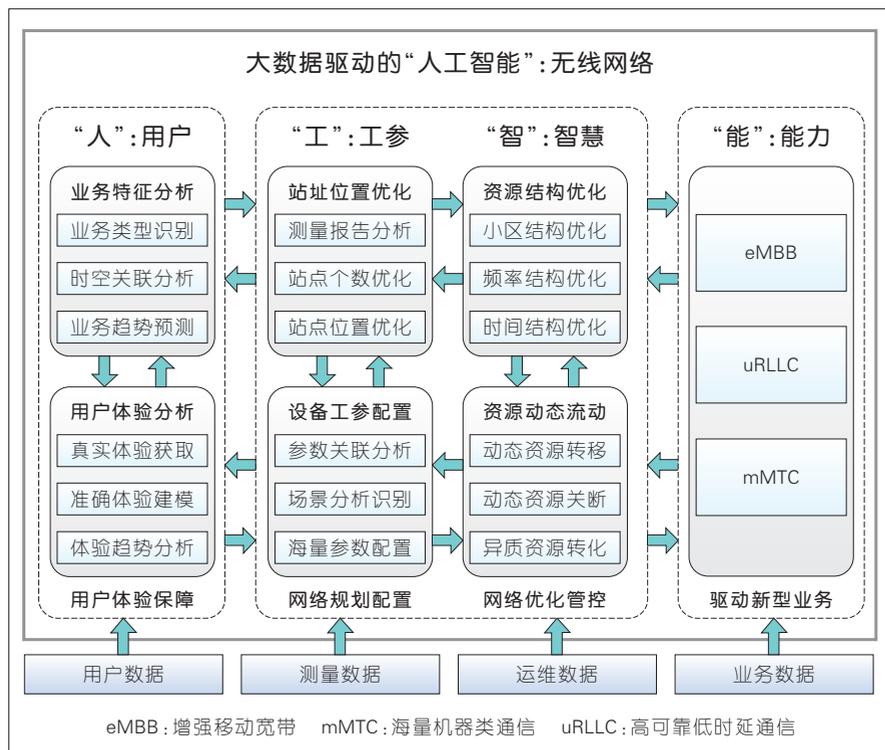
用户是无线网络服务的核心对象,实时、准确地获取用户的业务信息同时对用户的体验信息进行有效建模、评估是保证无线网络智能运行的基础,与此同时掌握用户的行为特征、业务特征也有助于网络的资源配置与管理。

#### 1.1 业务特征分析

无线网络的性能与业务模型紧密相关,而移动互联网新兴业务的不断涌现给无线网络的运维提出了挑

战,传统无线网络是通过采集核心网信令进行业务深度解析并检测出业务类别。然而其业务识别种类及实时感知控制能力尚不能满足未来智能化无线网络的应用需求。通过大数据分析手段建立终端、接入网、核心网、业务提供商的全链条业务识别与分析体系,可有效改善整网的业务保障能力,例如:可改善时延敏感类业务的用户体验,提升高价值业务感知,限制低价值业务的资源占用等。

另一方面,随着数据分析手段和计算能力的提升,使得利用业务特征数据建立大量用户及业务的时空关联特征成为可能。例如:在文献[1]中作者建立了贝耶斯网络结构学习框架,利用观测样本数据来确定业务的



▲图1 无线网络智能化功能架构

空间统计分布及依赖关系,进而可以进行小区间业务的合理推理与趋势预测。基于业务数据的时空关联关系可为网络资源分配与预留提供重要依据。

## 1.2 用户体验分析

用户体验是判断智能无线网络优劣的基本准则。现如今,各类移动数字设备层出不穷,社会各个领域的信息化不断加深,因此,存在大量与人类领域、技术领域、经济领域、环境领域相关的复杂因素影响移动用户的业务体验,分析的维度以及关联关系复杂度不断提升。此外,研究表明:85%以上的用户并不愿意反馈自己的业务体验,导致用户的实时准确信息难以获取,使得用户体验的分析变得更加困难。因此,通过用户的客观行为建立用户的真实体验模型成为了极具吸引力的新方向。

以移动视频业务为例,用户的客观行为表征参数,如放弃率、观看时长、快进频率等具有较好的可测性,

可在用户无感知的条件下为用户体验的建模提供基础行为数据。在此基础上,可以结合移动视频应用环境来进一步优化用户体验建模的准确性。例如:根据全球定位系统(GPS)信息和基站信息确定用户所处位置;利用终端的麦克采集环境噪音强度;利用终端感光器采集环境的光强;根据时钟信息确定时间等来判断用户体验业务时的周边环境,将此类数据与用户客观行为数据进行联合处理可以有效地获得用户的一些真实业务体验。

传统的无线网络中的数据报表往往是单维的、无关联的,无法描述网络的运行细节,更无法体现用户的真实感知,因此通过用户体验与业务特征的智能分析为网络的智慧运营打下了良好的基础。

## 2 “工”：工作参数配置与调整

为了应对数千倍容量需求,无线网络的基本形态将是超高密度布设

的小基站(热点/家庭基站/微微小区)+微小区+宏小区,并与无线局域网网络互补共赢、异构融合。因此,站址位置的选取以及基站工作参数的配置复杂度将成指数增长,通过引入大数据与人工智能的处理手段,可有效减轻网规网优人员的负担。

### 2.1 站址位置优化

传统无线网络规划通常需要考虑网络结构、业务分布、网络覆盖、无线传播环境、干扰规避、边界协调等方面的要求,目前已有针对2G、3G和长期演进(LTE)系统所量身打造的网规网优系统,比如中兴通讯的ZTE UniPOS无线网络规划系统和ZTE UniPOS无线网络优化系统,但传统网络规划问题站址个数和备选位置往往相对较少,随着微站和微微站的大规模商用使得网络站址选取与优化复杂度不断提升,与此同时家庭基站的引入使得蜂窝基站的布设位置不再受运营商精准控制,这也无形中增加了无线网络规划的难度。因此,面向5G超密集基站部署问题传统的测量与优化方法将不能很好地满足应用需求。通过粒子群优化、量子粒子群优化等智能算法的引入,可有效解决站址个数、站址位置的双重优化问题,同时提升网络规划的自主性、最优性以及鲁棒性。

### 2.2 设备工参配置

网络的密集布设使网络需管理和配置的参数大幅增长,例如:在2G网络中,单节点需要配置和优化的参数大约为500个;3G网络需要配置和优化的参数达1000个;4G则有1500个参数,而在下一代网络中,单节点<sup>[1]</sup>可能需要配置和优化达2000个参数,与此同时,各个工作参数间的关联关系更为复杂。

针对无线网络巨量工作参数配置复杂度指数增长挑战,我们需要构建多维工作参数的网络化自配置架构,设计分层分区的低开销分配与

配置方案。例如:可根据网络规模与密度,将覆盖区域划分为紧密相接的弹性虚拟小区,根据最大化空间复用原则为每个弹性虚拟小区分配局部最优的工参配置表。基站可根据自身地理位置与密集小区工参配置表进行映射,快速配置多维工作参数,从而减少多个密集区域配置参数及配置开销,且仅需低频次网络化的联动交互,就可实现网络工作参数的无冲突分配和配置,使得网络资源分配和配置开销不随网络规模和密度增大而显著增长。此外,网络化的自配置架构同时可以支撑网络参数的在线调整,为网络资源的按需流动提供了基础。

### 3 “智”:智慧资源管控与优化

智慧的资源管控是智能化无线网络的核心,其目标是通过无线资源的优化分配与动态流转来匹配用户的需求。

#### 3.1 资源结构优化

网络资源结构的优化是在相对较大的时间尺度上为网络提供优化的资源结构,主要包含小区结构优化、频率结构优化、时间结构优化。

##### (1) 小区结构优化

由于基站布设密度、用户密度以及业务密度的不断提升,网络干扰成为了制约网络服务质量提升的关键因素,尤其是小区边缘用户服务质量更是无法很好地得到满足,通过统一网络化资源管控架构可以实现物理小区的合并与拆分,也就是在网络部署后可以实现网络结构的动态调整来适应不同的业务分布。因此,可以通过将小区结构优化与业务特征分析相结合,构建最佳的小区结构。

##### (2) 频率结构优化

当前的无线网络有 D、E、F 等多个频段可以使用,未来 5G 在 3.5 GHz 以及更高频率将会有更多可供使用的频段,大量接入点的频点配置及其

形成的网络频率结构自优化将成为关键。文献[3]中将该问题建模为一个非合作速率最大化博弈模型,并设计了一种完全分布式的频点分配算法,在不需要大量信息交互的条件下实现了即插即用的自主频率优化。

##### (3) 时间结构优化

由于业务的多样性和时变性导致上下行业务存在不均衡的现象,而在时分双工(TDD)系统中可通过动态调整上下行子帧的配比来动态适应网络中业务的变化。然而,不同接入点使用不同的子帧配比也将引入交叉子帧干扰,从而降低了网络的整体性能,因此需要整网的数据统计和分析来实现最优的子帧配比。文献[4]、[5]中作者提出了一种基于动态 TDD 的匹配网络业务分布的最优资源分配方法,该方法首先根据网络中小区的业务需求将小区进行聚类,推导了不同业务负载对蜂窝链路成功传输概率的影响。以最小化每类小区的服务时间为目标,求解出了与每类小区上下行业务需求最优匹配的上下行子帧配置因子,实现了网络业务与资源的最优匹配。

#### 3.2 资源动态流动

资源动态流动是在小时间尺度上智能地为用户提供最佳资源,主要包含动态资源转移、动态资源关断、异质资源转化。

##### (1) 动态资源转移

根据无线资源细粒度调度和自适应调整的需求,可在优化的资源结构基础上,通过资源状态交互,发现并利用空间隔离区域小区间的隐藏关联关系,将多个相互独立小区的接入资源进行聚合,形成网络化的群聚资源。根据用户业务时空变化特性,综合考虑密集区域内站址的空间分布与频率资源的耦合度,以密集区域内接入用户数最大化为目标,动态构建最佳小区形状及动态区域结构,进行资源的分配和网络化加载,使用户在新构建的区域小区平滑流转,实

现了资源加载过程的快速收敛和区域间的负载均衡,有效地降低系统的阻塞率,提升了系统容量。

##### (2) 动态资源关断

在优化网络运行状态的过程中,除了系统吞吐量,网络传输过程中的能量效率同样是一个值得关注的问题。在 5G 密集组网场景中除了较大时间尺度上对密集接入点进行关断/开启外,载波关断技术也作为一个节能策略被广泛研究,例如:在分时期演进(TD-LTE)网络中,当业务负载较轻时,将一帧中的下行子帧配置成一个或多个多播/组播单频网络(MBSFN)子帧,可以实现更多的符号关断。通过动态配置 MBSFN 区域内的小区,并分配 MBSFN 区域间的多播帧配置,可以实现多小区资源的联动和干扰的有机规避,从而在节省基站功率消耗的同时提升了网络吞吐量,进而可以支持更多用户。

##### (3) 异质资源转化

在密集接入场景下无线局域网与蜂窝网络的深度融合已成为必然趋势。两种网络资源在时、空、频多个维度均存在差异。因此需要对异质资源配置进行实时优化,以适应无线信道状态的随机变化,匹配不同类型用户的业务特征,文献[6]中作者提出了一种两种异质网络的联合资源转化方法,充分利用不同网络的业务负载与传输能力的异质化特性,细粒度地缓解网络间干扰,保障用户需求。此外,计算资源与存储资源的优化分配及其与通信资源间的有机转化也是目前研究的热点。

### 4 “能”:承载能力提升与进化

支持增强移动宽带业务,如无线虚拟现实/增强现实(VR/AR);超可靠低时延业务,如车联网控制信息;海量机器类通信(mMTC),如物联网业务,是 5G 无线网络承载能力提升的重要目标。以数据为驱动力,我们突破了网络化智能资源管控技

术,并在第3届全国互联网+大赛主会场——西电远望谷体育馆构建了“繁星”密集无线通信试验系统,系统通过认知资源弹性配置、区域资源链式联动、跨域资源聚合协同,实现网络连接密度提升与无线全息业务承载的新突破,保障了互联网+大赛的顺利圆满进行,为用户提供了“身临其境”的极速无线业务体验,如图2所示。

无线网络是数据传输的载体,同时数据的不断积累也成为无线网络智能化水平不断的先决条件,未来无线网络将成为不断具备自配置、自优化、自愈合的智能化系统<sup>[7-9]</sup>,甚至将具备自进化的能力,如图3所示。

## 5 结束语

大数据与人工智能已经在一些规则的、有边界的问题处理中取得了比较好的应用效果,如围棋、医疗等领域,然而针对支持万物互联的无线网络是否有惊人的优势还在不断探索中;但海量数据与智能算法一定会成为无线网络进一步发展的重要基础。通过在用户特征分析、工参优化配置、智能资源管控、承载能力提升等方面引入人工智能可以帮助人们在系统布设、自动化运营和优化等各个核心环节做出更明智的决策,能进一步满足并催生新兴业务。

### 参考文献

- [1] HAN W J, SANG H Y, SHENG M, LI J D. Efficient Learning of Statistical Primary Patterns via Bayesian Network[C]// IEEE International Conference on Communications. USA:IEEE, 2015 :4871-4876. DOI: 10.1109/ICC.2015.7249094
- [2] IMRAN A, ZOHA A, A ABU-DAYYA A. Challenges in 5G: How to Empower SON with Big Data for Enabling 5G [J]. IEEE Network, 2014, 28(6): 27-33. DOI: 10.1109/MNET.2014.6963801
- [3] XU C, SHENG M, WANG X J. Distributed Subchannel Allocation for Interference Mitigation in OFDMA Femtocells: A Utility-based Learning Approach[J]. IEEE



▲图2 “繁星”密集无线通信系统部署、测试与业务体验

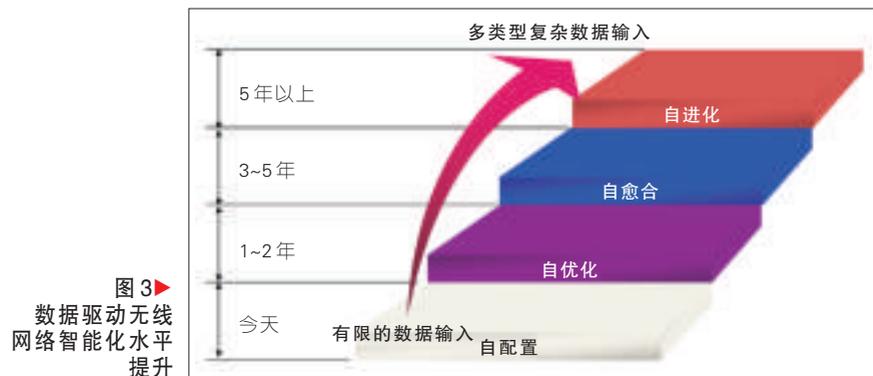


图3  
数据驱动无线网络智能化水平提升

- Transactions on Vehicular Technology, 2015, 64(6): 2463-2475. DOI: 10.1109/TVT.2014.2344434
- [4] SUN H G, SHENG M, MATTHIAS W, et al. Traffic Adaptation for Small Cell Networks with Dynamic TDD[C]//Proceedings of 2016 IEEE Global Communications Conference. USA:IEEE, 2016
  - [5] SUN H G, SHENG M, MATTHIAS W, et al. Traffic Adaptation and Energy Efficiency for Small Cell Networks With Dynamic TDD[J]. IEEE Journal on Selected Areas in Communications, 2016, 34: 3234-3251. DOI: 10.1109/JSAC.2016.2600442
  - [6] WANG L, SHENG M, ZHANG Y, et al. IM-Torch: Interference Mitigation via Traffic Offloading in Macro/Femtocell+WiFi HetNets [C]//2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications. USA:IEEE, 2013: 1705-1709. DOI: 10.1109/PIMRC.2013.6666417
  - [7] KHATIB E J, BARCO R, MUNOZ P, et al. Self-Healing in Mobile Networks with Big Data [J]. IEEE Communications Magazine, 2016, 54(1): 114-120. DOI: 10.1109/MCOM.2016.7378435
  - [8] GOMEZ-ANDRADES A, MUNOZ P, SERRANO I. Automatic Root Cause Analysis for LTE Networks Based on Unsupervised Techniques[J]. IEEE Transactions on Vehicular Technology, 2016, 65(4): 2369-2386
  - [9] BARCO R, LAZARO P, MUNOZ P. A Unified Framework for Self-Healing in Wireless Networks[J]. IEEE Communications Magazine, 2012, 50(12): 134-142. DOI: 10.1109/MCOM.2012.6384463

### 作者简介



张琰,西安电子科技大学副教授;主要研究领域为无线网络资源分配、智能网络;先后主持和参加基金项目10余项,获得国家技术发明奖二等奖2项;已发表论文20余篇,其中被SCI/EI检索10余篇。



盛敏,西安电子科技大学教授、博士生导师,教育部长江学者特聘教授,国家杰出青年科技基金获得者;主要研究方向为智能宽带无线通信、认知无线网络、大规模自组织网络等;先后主持和参加基金项目20余项,获得国家技术发明奖二等奖2项;已发表论文100余篇,其中被SCI/EI检索80余篇。



李建东,西安电子科技大学教授、博士生导师,教育部长江学者特聘教授,国家杰出青年科技基金获得者;主要研究方向为智能宽带无线通信、认知无线网络、大规模自组织网络以及无线网络的干扰管理等;先后主持和参加基金项目20余项,获得国家技术发明奖二等奖2项;已发表论文200余篇,其中被SCI/EI检索100余篇。

# 大数据驱动的无线网络资源管理及控制

## Resource Management and Control of Big Data-Driven Wireless Network

刘媛妮/LIU Yuanni  
赵国锋/ZHAO Guofeng

(重庆邮电大学, 重庆 400065)  
(Chongqing University of Posts and  
Telecommunications, Chongqing 400065,  
China)

### 1 无线网络资源管理面临的问题

为了保证构建于无线网络之上的各种应用能够给用户带来较好的用户感知体验,研究者利用动态频谱接入、无线覆盖测量、信道估计等技术来对频谱资源进行管理和规划,提高无线资源的利用率。然而,随着无线移动通信技术的快速发展,人们对无线频谱资源日益增加的需求,造成了无线频谱拥塞。为此,需要构建面向全域的无线网络资源管理池,通过对无线网络资源信息的数据收集、分析,实现对资源的统一管理和控制。然而,现有的技术在应对上述任务时主要面临以下3方面的问题:

首先,泛在化的数据急剧增长,增加了数据收集的范围及复杂度。传统的无线网络数据收集<sup>[1-3]</sup>涉及到的方法有人工路测、专门的无线传感器网络的构建等。这些方法在一定程度上能够实现一定范围内的数据收集,但也有自身的局限性:如人工路测方法通常会导致数据收集成本

收稿日期: 2018-01-16  
网络出版日期: 2018-03-23  
基金项目: 国家自然科学基金(61501075)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0006-004

**摘要:** 认为将大数据技术融入到无线网络资源管理中,有利于统筹分析大范围、多应用之间以及应用内部的无线资源需求,使无线网络资源管理变得更加灵活、便捷、高效。提出大数据驱动的无线网络资源管理及控制框架:首先利用移动群智感知(MCS)技术方法进行无线网络全域范围内资源信息的收集;其次利用数据虚拟化方法技术,构建全网的无线资源统一视图并提供不同粒度及角度的数据服务;最后利用软件定义网络(SDN)思想方式,研究面向应用全生命周期的无线资源管理控制机制,实现对应用的个性化资源管理与控制。

**关键词:** 无线网络; 大数据; 移动群智感知; 数据虚拟化; 网络资源管理

**Abstract:** It is believed that the integration of big data technology into wireless network resource management will facilitate the co-ordination analysis of wireless resource requirements in a wide range of multiple applications and the internal of applications, making wireless network resource management more flexible, convenient and efficient. In this paper, a big data-driven wireless network resource management and control framework is proposed. First, mobile crowd sensing (MCS) technology is used to collect the whole-scale wireless network resource information. Second, data virtualization (DV) is used to build a unified wireless resource view, and to provide data services with different granularities and angles. Finally, in order to realize the personalized resource management and control of applications, the software defined network (SDN) method is used in the radio resource management control mechanism oriented to the whole application life cycle.

**Key words:** wireless network; big data; mobile crowd sensing; data virtualization; network resource management

较高,并且通过该方式收集的无线网络环境数据存在盲区<sup>[4]</sup>,且灵活性较差;而构建专门的无线传感器网络需要在无线网络全域范围内部署大量的网络基础设施,增加了数据收集及网络维护的成本。

其次,多源异构的数据特性为数据的分析管理带来新的挑战。现有大部分的数据分析系统往往面向单一的上层应用,多应用之间无法实现某些相同底层数据的复用,造成数据的重复收集、分析及浪费。即便是面向多源异构数据处理的数据仓库在

构建全域的无线网络资源池的需求下也变得捉襟见肘,不具有可持续性:数据源的数量太过多样化,使得数据在可用性、质量和格式方面的差异太大,导致数据仓库需要不断地进行完整的数据提取、转换、加载等操作;另一方面,网络管理及控制的需求逐渐加快,日趋专业化,难以为每个需求开发定制专门的提取、转换、加载流程。

最后,无线网络资源的管理及控制主要表现为单向的管理及优化,更多的是从网络侧的角度考虑资源的

分配及控制,无法针对具体应用从用户的感知体验考虑用户的需求变化对资源变化的需求。即大部分的网络管理及控制策略考虑网络侧资源的使用情况,较少考虑用户侧应用需求变化导致的资源变化,缺乏相应的反馈机制

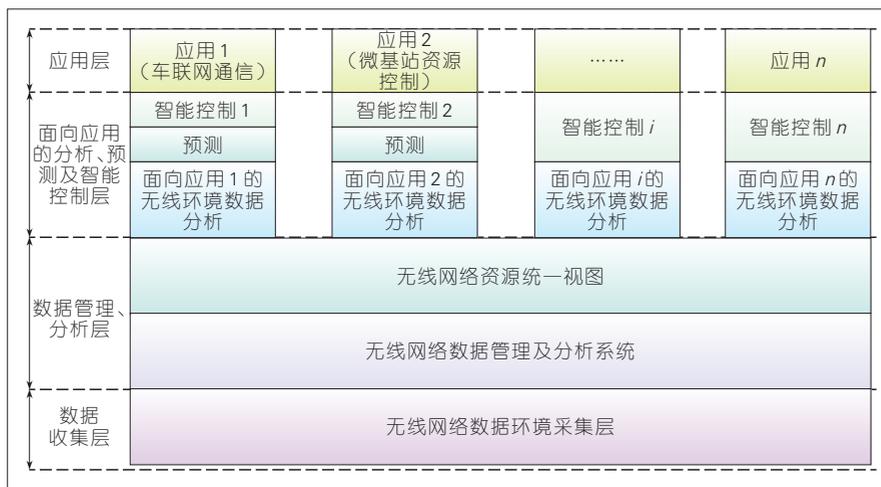
面对这些挑战,解决的关键是如何实现无线网络环境全局信息数据的收集,构建统一的无线网络资源管理框架,实现面向无线网络全域范围内的数据搜集、分析以及面向应用的资源管理及控制。在对泛在的、多源、异构的无线网络数据收集的基础上,通过数据的互通共享、融合分析,并针对无线网络上层应用中面临的需求,为上层应用提供动态、按需、自动化的数据存储、处理、分析、可视化等服务,实现大数据驱动的无线网络资源管理及控制。

## 2 大数据驱动的无线网络资源管理及控制的内涵及挑战

大数据驱动的无线网络资源管理及控制,就是在实现无线网络全域内数据收集的基础上,将无线网络数据进行融合分析、互通共享,构建统一的无线网络资源池,这些资源包括频率、时间、功率、编码、网络设施、资源等,并以数据服务的形式向上层应用提供查询、分配等功能。其次,对构建在无线网络之上各种应用的资源使用情况(如反映用户感知体验指标的相关数据)进行收集及反馈,从而使不同类型、不同需求的无线网络应用业务在兼顾整体资源分配的前提下能够获得按需服务,最大化无线资源使用效率。

如图1所示,这种大数据驱动的无线网络资源管理及控制主要涉及以下3方面的关键技术:

(1) 泛在、灵活的无线网络数据收集方法,即构建无线网络统一资源管理池需要收集的信息包括网络侧



▲ 图1 大数据驱动的无线网络资源管理参考模型

及用户侧两方面的数据。网络侧的数据包括能够反映无线网络设备、资源、性能、状态、拓扑业务/内容等多个方面。用户侧数据则是直接能够反映用户业务感知体验的数据,如数据视频业务浏览过程中能够反映用户感知体验的首包时延、卡顿次数、平均下载时间等。

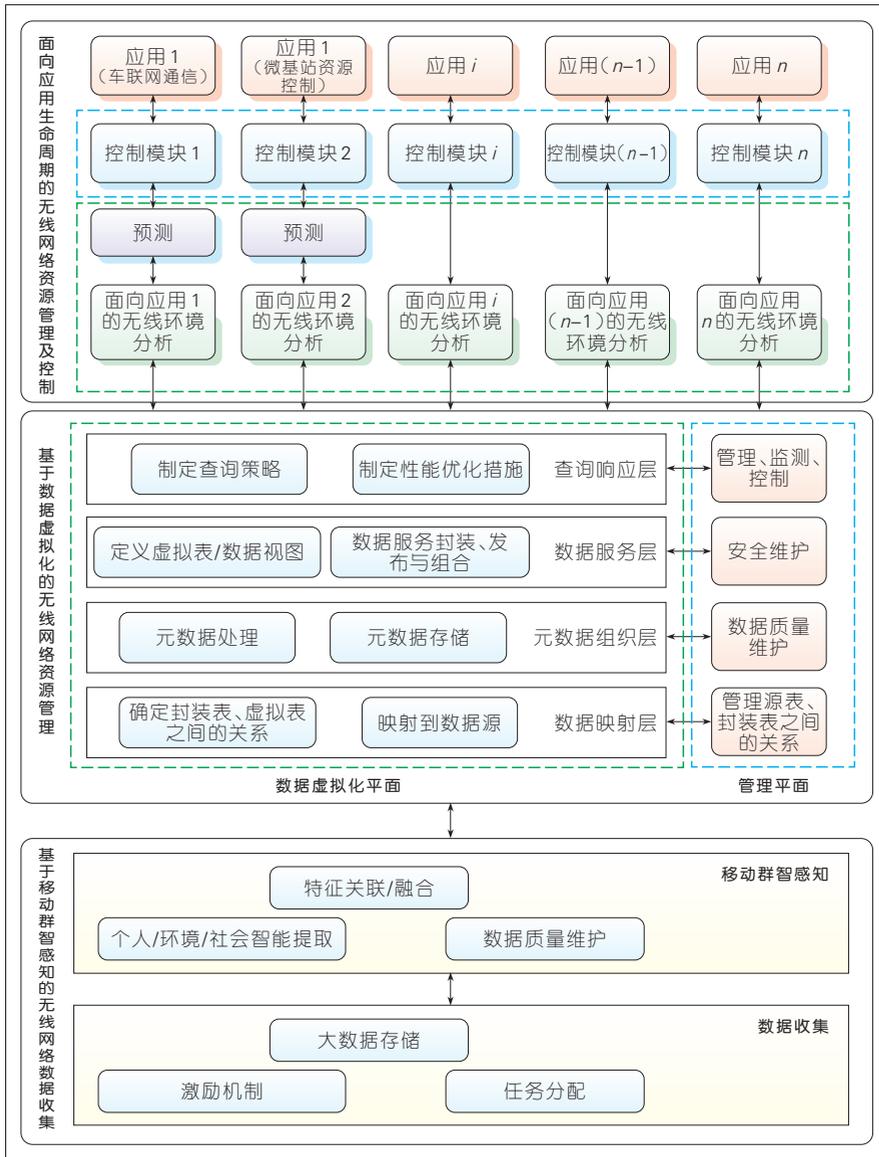
(2) 面向多源、异构的无线网络数据的管理机制,即构建无线网络全局资源的统一视图,需要屏蔽各个数据源的差别,使用户能够通过统一的结构请求访问这些资源数据而无需关注各个数据源的具体结构。另一方面,底层数据提供者对数据结构、数据格式等方面的改变,都不会对数据使用者产生影响。这就使上层的数据使用者不需要亲自执行查询、访问并聚合多个数据源的数据即可得到自己想要的信息。为实现大数据驱动的无线资源管理,需要至少具备两个功能:能够根据本地数据源的“特性和能力”主动对外发布相关的数据服务;当外界的数据消费者请求相关的数据服务时,数据管理层能够根据请求提取相关的数据,并根据外界数据标准进行一定的规范变换,然后提交给外界的请求对象。

(3) 面向应用的无线网络资源管理及控制,即面向应用的无线网络资源管理及控制不仅针对无线网络上

的各种应用进行资源的统一管理及分配,而且需要根据业务对资源的动态需求情况,实现面向应用的个性化、动态资源管理。也就是说,该层面一方面应该能够掌控无线网络环境全局的资源,兼顾各种应用对资源的需求并实现资源的统筹分配;另一方面需要针对每一个特定的应用,构建对应的资源管理及控制模块,对单个应用的资源使用情况进行实施监测控制,并根据应用对资源的实际需求情况实现动态调整。

## 3 大数据驱动的无线网络资源管理及控制系统的参考框架

为了满足上述的要求,本文提出了基于大数据驱动的无线网络资源管理及控制框架(如图2所示),自底向上分别是基于移动群智感知<sup>[4]</sup>的无线网络数据收集层,基于数据虚拟化的无线网络管理及分析层<sup>[5]</sup>,面向应用全生命周期的无线网络资源管理、控制层,最上层为构建于无线网络之上的各种应用。基于大数据驱动的无线网络管理及控制首先在数据收集层利用移动群智感知的方法实现大规模、多样化的无线网络数据收集,其次利用数据虚拟化的方法构建多源、异构数据的统一视图,并将资



▲图2 大数据驱动的无线网络资源管理及控制框架

源以数据服务的形式向外发布,网络资源管理及控制层面利用软件定义网络(SDN)的思想方式实现对上层应用的资源使用情况的个性化管控及反馈。下面我们将对各个层面进行详细阐述。

### 3.1 基于移动群智感知的无线网络数据收集

移动群智感知是指利用普通用户的移动设备作为基本的感知单元,通过移动互联网进行有意识或无意识的协作,实现感知任务分发以及感

知数据收集,完成大规模的、复杂的社会感知活动。移动群智感知作为一种新型的、“以人为中心”的感知模式对未来的普适计算有重要的意义。利用移动群智感知技术进行无线数据收集的优势在于:节点的移动性、泛在性以及能够通过多种方式(GSM、3G、4G、Wi-Fi等)连接到云端对数据进行存储、计算、集中化的特性。利用移动群智感知的方法进行无线网络环境数据的收集,一方面有效利用了移动端碎片化的计算、存储资源,另一方面大大降低了构建数据

收集网络基础设施的费用。

移动群智感知的生命周期<sup>[6]</sup>可划分感知任务的发布、感知用户招募、用户感知活动的进行以及感知数据上报4个阶段。因此,在该层面主要涉及以下相关问题:(1)如何对感知任务进行合理的粒度划分,使得功能单一的感知设备之间能够通过协作方式完成感知任务,最大化利用移动感知设备的碎片化资源;(2)如何在感知成本有限的情况下选择能够实现最大化覆盖的感知节点;(3)如何设置、如何构建合理的激励机制,提高用户参与感知活动的积极性;(4)如何均衡感知数据的精度与感知数据数量之间的关系,在保证感知数据质量的同时能够控制感知数据收集的规模以减轻后续的存储、计算、分析成本。

### 3.2 基于数据虚拟化的无线网络大数据管理及分析

数据虚拟化是一种虚拟化技术,对于来自多个异构数据源的数据可以隐藏真实数据的复杂性。数据虚拟化通过为上层应用程序提供一种透明的数据访问机制,可以显著地将应用程序与数据的复杂性分离。具体的方式是将每个“异构数据获取于处理描述包”封装为一个数据服务单元,这样就可以将局域网范围内的一个或者多个数据源,抽象为一组数据服务单元。其中每个数据服务单元相当于正对数据源的一个预处理查询,不仅可以被执行,而且可以被发布。使用数据虚拟化对数据进行管理,可以使数据模型的逻辑从数据的物理位置和源格式的细节中抽象出来。这意味着:使用数据的上层应用和用户不需要知道数据的物理存储方式和位置,使用底层的数据处理就可以在任何时候选择合并、移动、转换或存储数据,而不会对上游数据的工具 and 用户产生任何影响。这样避免数据的重复收集,有效地实现数据复用。

通过数据虚拟化技术提供统一的无线网络资源视图,将底层数据能够提供的功能以服务的形式进行定义、注册、权限管理等,从而实现对底层数据功能的模块化,以及针对上层控制层面触发的复杂事件进行数据服务组合的功能。根据数据虚拟化对数据的处理流程,可将数据虚拟化层面自底向上分解为数据映射层、元数据组织层、数据服务层、查询响应层。在该层面需要考虑的问题主要包括:(1)数据虚拟化层面的实施方案,即以何种架构形式实现数据虚拟化架构;(2)如何建立原始数据、映射表、封装表之间的关系,保证数据源向虚拟表格之间的正确映射;(3)如何实现元数据的抽取,并建立结构合理的元数据组织模型,实现元数据的高效存储,提高数据索引的效率;(4)如何创建相应的数据服务模块,实现对元数据的调用;(5)如何制定最佳的查询处理策略,提高查询的效率等。

### 3.3 面向应用全生命周期的无线网络资源管理及控制

传统面向应用的无线资源分析直接对数据进行“调用、分析、计算”的模式,难以区分不同用户(数据使用者、数据提供者)的权限,无法在保证数据隐私的同时扩大数据的使用范围。研究面向应用的无线网络资源管理及控制机制,根据底层的数据分析结果,构建相应的资源管理策略,并以SDN控制的方式实现对无线应用资源分析,并根据用户侧收集到的数据的分析结果,实现对无线资源的动态反馈及调整。

面向应用全生命周期的无线网络资源管理及控制层面需要重点研究的内容包括:(1)如何定义无线网络环境中各种应用的事件流,并建立复杂事件与底层数据服务的对应关系,实现面向应用的无线资源数据分析的功能;(2)如何构建基于软件定义的、面向应用的无线资源控制机制,通过调用基于数据虚拟化的无线资源分析模块的结果制定对无线资源的宏观统筹方案以及面向应用的资源调整机制,并以软件定义的方式实现对应用的资源控制。

## 4 结束语

面对无线网络应用不断增多而导致的用户对无线网络感知体验要求不断增加的新形势,需要实现面向应用全生命周期的无线网络资源管控,以提高无线网络资源的利用率,满足用户对体验质量(QoE)的需求。大数据驱动的无线网络资源管理及控制从数据在应用中流动的过程出发,通过对无线网络环境相关信息的收集、分析,制定资源管理分配策略,对无线应用所需要的资源进行分配,并通过对当前资源分配策略下用户的感知体验信息的收集、分析,进而反馈到资源管控层面,最终以SDN的方式实现面向应用的个性化、动态化的资源管理及分配。另外,随着大数据、人工智能等技术的发展,对网络资源及用户感知体验的分析将会更具洞察力,应用内部业务流的定义及监测控制也将会变得更加智能。可以预见:在不久的将来,这些技术将会更深入地被融入到无线网络资源管控中,未来的无线网络资源

管理将会变得更加灵活、便捷、智能、高效。

#### 参考文献

- [1] ABDALLAH M N, DYAB W, SARKAR T K, et al. Electromagnetic Macro Model for Analysis of Propagation Path Loss in Cellular Networks[C]//Antennas and Propagation Society International Symposium (APSURSI). USA: IEEE, 2014: 947-948. DOI: 10.1109/APS.2014.6904802
- [2] 钟志华. 无线网络优化的路测[J]. 科技传播, 2017,9(07):43-44
- [3] LIANG H W, HO C H, CHEN L S, et al. Coverage Hole Detection in Cellular Networks with Deterministic Propagation Model[C]//Intelligent Green Building and Smart Grid (IGBSG), 2016 2nd International Conference on. USA: IEEE, 2016: 1-6. DOI: 10.1109/IGBSG.2016.7539428
- [4] GANTI R K, YE F, LEI H. Mobile Crowdsensing: Current State and Future Challenges[J]. IEEE Communications Magazine, 2011, 49(11):32-39. DOI: 10.1109/MCOM.2011.6069707
- [5] GARETH MORGAN. Data Virtualisation on Rise as ETL Alternative for Data Integration [J]. Computer Weekly, 2012,4(8):15-18
- [6] ZHANG D Q, WANG L Y, XIONG H Y, et al. 4W1H in Mobile Crowd Sensing[J]. IEEE Communications Magazine, 2014, 52(8):42-48. DOI: 10.1109/MCOM.2014.6871668

#### 作者简介



刘媛妮,重庆邮电大学副教授;主要研究领域为移动群智感知网络、物联网安全、数据虚拟化;先后主持和参加基金项目10余项;已发表论文20余篇,其中被SCI/EI检索10余篇。



赵国锋,重庆邮电大学教授、博士生导师;目前主要研究领域为未来互联网、移动互联网、网络安全等;先后主持和参加国拨经费项目20余项;获得重庆市科技进步二等奖和三等奖各1项;发表论文80余篇,其中SCI、EI收录40余篇,出版著作1部,获得发明专利授权4项。

利授权4项。

# 基于强化学习的无线网络智能接入控制技术

## The Intelligent Access Control Mechanisms in Wireless Network Based on Reinforcement Learning

严牧/YAN Mu  
孙耀/SUN Yao  
冯钢/FENG Gang

(电子科技大学, 四川 成都 611731)  
(University of Electronic Science and  
Technology of China, Chengdu 611731,  
China)

当今社会已经迈入信息经济时代, 信息技术已成为推动经济结构向多样化消费和低能耗高效发展的重要驱动力。据思科公司预测, 到2019年全球移动数据总流量将增长至每月24.3 EB, 接近2000年全球互联网总流量的200倍<sup>[1]</sup>。另据全球移动通信系统(GSM)协会分析<sup>[2]</sup>, 到2020年全球支撑物联网的机器对机器通信(M2M)连接数将达到9.8亿, 接近2000年全球M2M连接数的14倍。无线通信网络在面临无线资源趋于枯竭的同时, 正在经历着前所未有的高增速无线服务需求与低效率无线服务供给之间的矛盾。

未来无线通信将利用复杂异构网络来支持多样化应用场景, 包括连续广域覆盖、热点高容量、高可靠低时延以及低功耗巨连接等。由于用

收稿日期: 2018-01-18  
网络出版日期: 2018-03-22  
基金项目: 国家自然科学基金(61631005、61471089); 中央高校基本科研基金(ZYGX2015Z005)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0010-005

**摘要:** 介绍了无线网络中的强化学习算法, 认为由于强化学习算法与环境交互并动态决策的特点, 其对复杂网络环境有着较强的适应能力; 然后针对无线网络中的强化学习方法的应用场景做了概述, 并给出了两个基于强化学习的无线接入技术案例: 毫米波技术的切换技术和 Multi-RAT 接入技术。可以看到: 智能的无线接入技术由于具备充分挖掘和扩展无线网络资源的潜力, 能够显著提高无线网络用户的体验。

**关键词:** 未来无线网络; 切换; 接入控制; 强化学习

**Abstract:** In this paper, the application of reinforcement learning in wireless network is briefly introduced. Due to the characteristics of interacting with environment and dynamic decision making, reinforcement learning algorithm has strong adaptability to complex network environment. Then the application scenarios of reinforcement learning method in wireless network are summarized, and two cases of wireless access technology based on reinforcement learning are given: handoff policy of mmWave HetNets and multi-rat access control. Intelligent access control of wireless network is powerful in exploiting wireless network resources, which can improve the quality of experiences of mobile users.

**Keywords:** future wireless network; handoff; access control; reinforcement learning

户终端性能和业务需求的不同, 用户体验质量(QoE)在不同通信场景也存在极大的差异性。出于成本和兼容性的考虑, 未来无线网络将长期处于多网共存的状态, 包括2G、3G、4G、5G、Wi-Fi等, 由于不同网络利用不同的无线接入技术, 因而形成了接入技术的差异性。同时, 为了进一步提升网络的容量, 需要在传统接入站点的基础上引入 Micro、Pico、终端直通(D2D)、移动自组织(Adhoc)及小蜂窝等接入站点, 因而形成了对网络的重叠异构覆盖。网络的高密度部署和多网络共存使得复杂异构网络

下的无线干扰环境变得更加复杂, 并对无线接入网的资源调度和控制管理提出了更高的要求。

传统的无线接入技术在“网络-频谱”的静态匹配关系下对网络进行规划设计和资源配置。设备的接入往往基于某一参数(如信号强弱、区域位置)选择单一接入网络和固定接入站点。由于复杂异构网络中海量用户行为的随机性, 不同网络的业务需求呈现出极大的时空动态变化特性。静态的“网络-频谱”匹配使得网络容量无法满足变化的网络业务需求, 大大地限制了无线网络的接入能

力,并导致用户接入体验差等问题。

为根本性地提高无线网络接入能力,必须打破传统的无线资源管理和接入控制的僵化机制,研究智能的无线接入理论与技术,充分挖掘和扩展无线网络资源的利用潜力,显著提高无线网络用户的体验。在无线网络中,由于用户行为以及网络的动态性和复杂性,使得接入控制和资源分配是非常具备挑战性的<sup>[5]</sup>。人工智能(AI)技术,比如机器学习,赋予计算机分析环境并解决问题的能力,并提供了一种有效的方法来处理动态性高、复杂度明显的问题<sup>[6]</sup>。

## 1 强化学习在无线网络中的应用

强化学习是一种在非确定环境下做决策的强劲的工具<sup>[5]</sup>。Google Deepmind最近所研发的AlphaGo以及AlphaGo Zero所使用的强化学习在围棋这类动态性明显、环境信息复杂的博弈游戏中表现良好<sup>[6]</sup>,并且取得较好的成绩。在异构网络接入控制的过程当中,由于网络的动态性导致了决策过程也必然是动态性的,我们需要主体和环境进行频繁交互、感知,从而智能化地协调用户和基站的决策行为。因此,强化学习由于其所具备的特点被我们利用到异构网络的决策过程中也是顺其自然的。

### 1.1 强化学习的分类

(1)根据强化算法是否依赖模型可以分为基于模型的强化学习算法和无模型的强化学习算法。这两类算法的共同点是通过与环境交互获得数据,不同点是利用数据的方式不同。基于模型的强化学习算法利用与环境交互得到的数据学习系统或者环境模型,再基于模型进行决策。无模型的强化学习算法则是直接利用与环境交互获得的数据改善自身的行为。两类方法各有优缺点:一般来讲基于模型的效率比无模型要高,因为智能体可以利用环境信息;但是

有些无法建立模型的任务只能利用无模型强化学习算法,因此无模型强化学习算法更具备通用性。

(2)根据策略的更新和学习方法,强化学习算法可分为基于值函数的强化学习算法、基于直接策略搜索的强化学习算法以及 Actor-Critic (AC)的方法。所谓基于值函数的强化学习方法是指学习值函数,最终的策略根据值函数贪婪得到。也就是说,任意状态下,值函数最大的动作为当前最优策略。基于直接策略搜索的强化学习算法,一般是将策略参数化,学习实现目标的最优参数。基于AC的方法则是联合使用值函数和直接策略搜索。

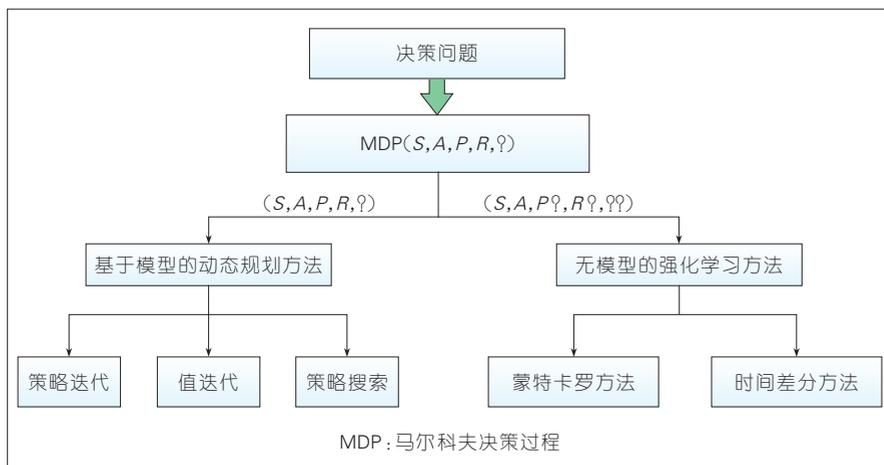
(3)根据环境返回的回报函数是否已知,强化学习算法可以分为正向强化学习和逆向强化学习。在强化学习中,回报函数是人为指定的,回报函数指定的强化学习算法称为正向强化学习。很多时候,回报无法人为指定,如无人机的特效表演,这时可以通过机器学习的方法由函数自己学出来回报。

### 1.2 强化学习在无线网络中的应用

考虑到无线网络的特殊应用场景,在基于图1的分类下,我们进一步按照无线网络的特点对强化学习进行分类,如图2所示。首先由于受限于网络中有限的频谱资源,用户总

是以竞争的关系接入到网络中,那么资源调度、小区切换等考虑用户QoE的问题往往可以建模成一个多主体马尔科夫决策过程(MDP);然后考虑到网络状态空间变化基于时间的连续性或离散性,可将网络决策过程建模为连续时间或者离散时间MDP,连续时间MDP需要决策做到快速反应,尽量做到在线学习;再者,基于网络动作空间的连续性或离散性,有分别基于策略迭代和值迭代的强化学习方法;最后考虑到传统的强化学习方法利用到网络环境中的一些不足,我们可以和深度学习结合起来做一个改进。

根据做决策的时序先后,我们可以把网络中接入用户的决策分为基于多主体的序贯博弈过程或同时博弈过程,如图3所示。具体来说,由于普通的强化学习本身就是基于MDP建模,并且解决的是序贯博弈的问题。为了解决同时博弈的问题,我们可以采用Nash Q-learning算法<sup>[7]</sup>。在Nash Q-learning的算法中,所有的决策主体在同一个决策时间从一个随机的决策开始去尝试学习它们的最优Q-value。为了达到这样的目的,每一个主体都通过其他主体的Q-value来更新自己的决策,直到达到纳什均衡点。例如:在文献[8]中,作者在认知无线mesh网络中考虑在尽可能保证主用户的服务质量(QoS)



▲ 图1 基于决策过程的强化学习算法分类

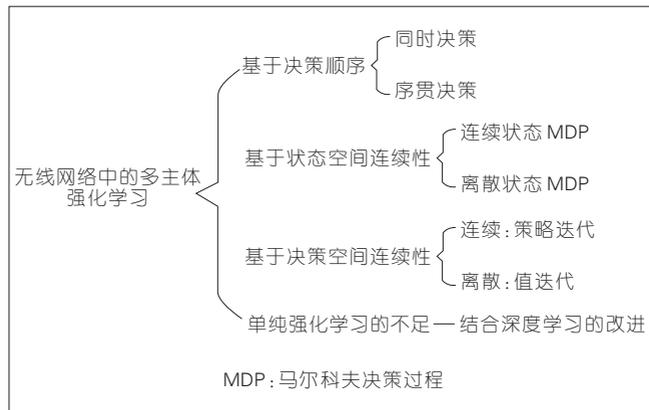


图2 无线网络中的多主体强化学习

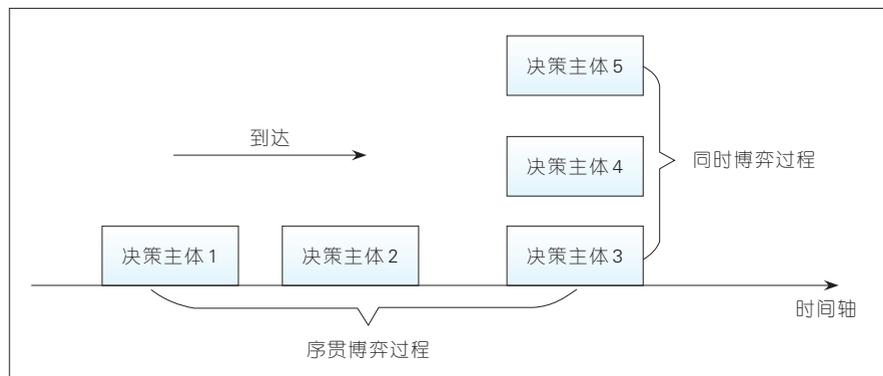


图3 基于决策时间的博弈过程

条件下,为同时接入的次级用户分配功率资源和频谱资源。考虑到次级用户之间的竞争关系(博弈关系),采用了基于多主体的 Nash Q-learning 算法,并得到较好的结果。在决策空间集较小,并且主体数量较少的情况下, Nash Q-learning 是一种很好的用于解决多主体同时博弈的算法。

在无线网络中,经常存在动作(决策)空间过大的现象,例如:在时频资源块分配问题中或者在长期演进(LTE)中非连续接收(DRX) cycle 长度的设置问题中。如果我们把频谱资源或者 cycle 的长度范围划分为较小的决策单元,那么就会使得策略空间异常大,会消耗大量的计算资源。如果我们能通过策略迭代用更平滑的手段去搜索最优策略,会显著增加学习效率,更加贴合无线网络中需求快速决策的特点。

基于状态空间在时间上的连续性或离散性,我们可以把 MDP 建模

成连续时间上的 MDP 或离散时间上的 MDP。连续时间 MDP 是基于时间序列连续的马尔科夫过程,其依然具备马尔科夫性。连续时间 MDP 和离散时间 MDP 区别在于时间指标参数从离散的  $T=\{0,1,2,\dots\}$  改为连续的实数  $T=\{t|t \geq 0\}$ 。当我们考虑小时间尺度上的网络问题,由于用户流的不间断涌入,信道质量的无规律变化等都会造成网络状态的频繁波动。因此快速决策就变得尤为重要。这里基于连续空间较好的算法是 AC 算法。AC 较好地平衡了值迭代和策略迭代这两种方法。例如:文献[9]中,作者考虑把基于流量变化下的基站开关操作建模为一个连续状态的 MDP。考虑到用户的接入流量是一个连续变化的过程,那么整个网络的状态也相应具有很强的动态性和连续性。所使用的 AC 算法在该工作中不仅加快了学习速率,TD-error 还具备预测的功能性。

无线网络中,强化学习还可以和深度学习结合起来使用,两者各有优缺点。强化学习本身由于状态空间过大导致学习时间较长(维度诅咒),在复杂的无线网络环境中,由于网络状态复杂,单纯的强化学习由于算法收敛过慢并不是十分贴合。基于神经网络的深度学习方法,可以利用历史数据对下一时刻的用户行为或者网络状态进行预测。但是,尽管深度学习能够提供较为精准的趋势分析和模式识别,也很难推导出与数据完全匹配的分布函数,在无线网络中带来决策上的明显失误,使得数据失去其应用价值。此外,为了及时保存和处理蜂窝网络数据,基站作为中心控制器需要存储大量的蜂窝网络数据,需要消耗大量的存储和计算资源。因此,我们可以将深度学习利用起来为小时间尺度上的网络决策提供先验信息,从而加速强化学习算法的收敛速度。

## 2 智能化接入控制案例分析

我们考虑两种智能化的接入控制技术作为案例研究:(1)针对毫米波异构蜂窝网我们提出了一种基于机器学习的智能切换策略,在保证用户服务质量的前提下,减少不必要的切换次数。针对单个用户,在强化学习方法中采用基于置信区间上界(UCB)算法的基站选择策略,可以降低某个用户的切换次数。(2)我们考虑将不同的 QoS 需求的用户接入到蜂窝网和 Wi-Fi 共存的异构网络中。为了在复杂和动态环境中最大化系统吞吐量并且同时满足用户 QoS 需求,我们利用基于多主体强化学习的智能多无线电接入技术,通过动态感知网络环境,来为每个用户分配相应的信道资源。

### 2.1 基于毫米波技术的智能切换技术

#### (1) 强化学习的奖励函数设计

由于与切换次数相关的奖励函数很难反映,下面我们通过巧妙地设

计奖励函数,来达到最小化  $H_n$  的目的。我们定义奖励函数为:

$$R_n^k(t) = \int_t^{t_n^k} r_n^k(t) dt \quad (1)$$

其物理意义是:如果用户(UE)  $n$  选择切换至基站(BS)  $k$ ,在发生下次切换的这段时间内所传数据的总量。其中  $t_n^k$  表示的是发生下次切换的时间。可以证明:对 UE  $n$  而言,最小化总的切换次数  $H_n$  等价于解决上述所提出的具有公式(1)的回报函数的强化学习模型。

### (2) 估计期望收益

对服务类型相同的 UE,比如服务类型均为  $C_n$  的 UE,我们让这些 UE 共用一个类回报函数,计作  $R_{C_n}^k(T_{C_n}^k)$ ,用  $C_n$  类中的每个 UE 的  $R_n^k(t)$  来共同更新类回报函数  $R_{C_n}^k(T_{C_n}^k)$ 。具体的更新定义为:

$$\check{R}_{C_n}^k(T_{C_n}^k + 1) = \frac{T_{C_n}^k \check{R}_{C_n}^k(T_{C_n}^k) + R_n^k(t)}{T_{C_n}^k + 1}$$

其中,  $T_{C_n}^k$  代表当前 BS  $k$  被服务类型为  $C_n$  的用户选中的次数。我们用类回报函数  $R_{C_n}^k(T_{C_n}^k)$  来作为 BS  $k$  关于服务类型为  $C_n$  用户的回报函数值就避免了对于某个单个用户而言,无法及时更新  $R_n^k(t)$  的问题。我们用公式(2)中的方法去估计回报函数的期望。

$$E[R_n^k(t)] = \begin{cases} \check{R}_{C_n}^k(T_{C_n}^k), & \text{if } n \in C_n, \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

由于处在同一服务类型的用户切换准则相近,在经过一段时间的学习之后,回报函数期望的估计值具有较高的准确性。

### (3) 基站选择策略

由于强化学习中的定理——探索和利用,我们不能总是选择当前回报函数期望值最大的基站进行切换。通常,我们用 Regret 来衡量强化学习中的一个策略的优劣程度。Regret 是指所采取的策略与最优策略之间的差距。在我们的这个问题中,

UE  $n$  在策略  $\pi$  下在执行了  $W$  次切换后的 Regret 可以表示为:

$$\text{Regret}_{\pi}(W) = \sum_{\omega(t)=1}^W [R_{\pi^*}(t_n^k) - R_{\pi}(t_n^k)] \quad (3)$$

其中,  $R_{\pi^*}(t_n^k)$  代表采用最优策略  $\pi^*$  在时刻  $t_n^k$  所获得的回报。在文献[9]中已经证明所能达到的最优 Regret 是关于切换  $W$  次数呈对数数量级的。算法 UCB 已经被证明:无论何种形式的 reward 函数都可以实现对数量级的 Regret。UCB 算法的选择策略为:agent 每次选择机器  $j^*$ ,其中  $j^*$  的计算方式为:

$$j^* = \arg \max_j (\bar{x}_j + \sqrt{\frac{2 \ln W}{W_j}}) \quad (4)$$

其中,  $\bar{x}_j$  为机器  $j$  所获得的平均回报值,  $W_j$  代表到目前为止机器  $j$  被选中的次数,而  $W$  表示目前为止总的执行决策的次数。

基于 UCB 算法,我们提出了一种目标基站选择策略。我们将基站  $j$  的 index 设为  $\check{R}_{C_n}^k(T_{C_n}^k) + l \sqrt{\frac{2 \ln H_n}{T_{C_n}^k}}$ ,其中  $l = \max_{k \in A_n, c \in C} \check{R}_{C_n}^k(T_{C_n}^k)$ ,  $H_n$  代表目前为止 UE  $n$  所发生的总的切换次数。因此,基于 UCB 算法,我们提出了 SMART-S 的基站选择策略。UE  $n$  在发生切换后,在可行基站集合  $A_n$  中

选择 BS  $k^*$  进行切换,其中:

$$k^* = \arg \max_k (\check{R}_{C_n}^k(T_{C_n}^k) + l \sqrt{\frac{2 \ln H_n}{T_{C_n}^k}}) \quad (5)$$

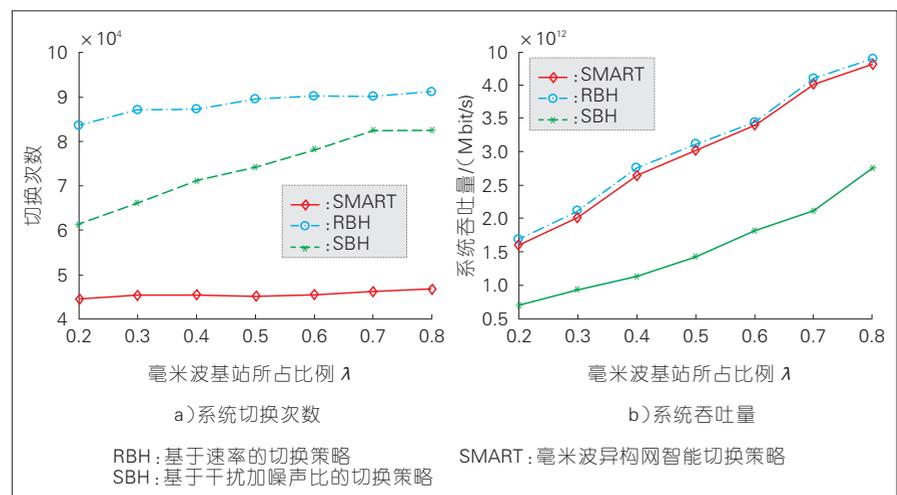
我们考察了毫米波异构网中智能(SMART)切换策略下的性能,并与下面两种传统策略进行了对比:基于速率的切换策略(RBH)是每次用户发生切换时总是选择当前可以提供最大传输速率的基站进行切换;基于干扰加噪声比(SINR)的切换策略(SBH)是用户总是选择可以提供最大信号 SINR 的基站进行切换。图4代表了这3种切换策略下的系统总的切换次数/系统吞吐量与毫米波小基站(mm-FBS)所占比例  $\lambda$  之间的关系。通过图4可以看出:我们可以通过较小的系统吞吐量的损失而带来较明显的切换次数的降低。

## 2.2 Multi-RAT 智能接入技术

为了在复杂和动态环境中最大化系统吞吐量并且同时满足用户 QoS 需求,我们利用基于多主体强化学习方法的智能多无线电接入(SARA)技术,通过动态感知网络环境,来为每个用户分配相应的信道资源。

### (1) 场景描述

我们研究的场景是蜂窝网小基站(SBS)和 Wi-Fi 热点共存的场景。LTE 下行执行正交频分多址的传输



▲ 图4 mm-FBS所占比例  $\lambda$  和系统切换次数/系统吞吐量之间关系图

方式(OFDMA),其频谱资源包含很多的时频资源块(RB),又叫做子信道。在传输的过程中,非连续波段的频谱可以利用传输数据流。为了保护正在进行的会话流,我们假设新到的业务流必须在没有多余频谱资源的情况下进行等待。基站作为中心控制器是能够获取全局的网络信息,包括用户的QoS需求和网络环境信息。由于网络的动态性和跨无线电技术(RAT)的资源调度复杂特性,多无线电技术的聚合需要更加智能化的技术支撑。

(2) 基于多主体强化学习的Multi-RAT接入机制

多无线电接入过程是一个多主体的随机过程<sup>[9]</sup>。在多主体的环境中,我们可以观测到其他所有主体所做的决策已经反馈的回报值。基于该多主体的随机过程,和图5提出的两层决策框架,无线电/信道选择过程(RSP)和资源分配过程(RAP)中分别存在着同时博弈和序贯博弈的过程。我们采取Nash Q-learning算法<sup>[9]</sup>以及蒙特卡洛树搜索(MCTS)方法<sup>[10]</sup>来解决这两个博弈的相关问题。

我们把接入过程建模成一个基于半马尔科夫(SMDP)的强化学习模型。具体来说,在我们的工作中有两个决策阶段,如图5所示:第1阶段为RSP,该阶段的目的在于尽可能地去避免碰撞和乱序情况的发生,从而压缩决策空间。当我们的算法收敛后,我们就开始第2阶段——RAP,在该阶段中,基于有限的网络资源和多样的用户喜好,我们考虑去使用有限的信道资源为用户提供合适的服务,并且使得系统平均吞吐量最大化。在这一阶段中,我们假设在蒙特卡洛树搜索中,每一个节点 $s$ 包含了 $\{r(s,a),N(s,a),Q(s,a)\}$ 的信息,其中 $r(s,a)$ 是即时的奖赏值用来衡量该资源分配决策的好坏, $N(s,a)$ 是节点的被访问次数, $Q(s,a)$ 是该节点的Q-value。在决策的搜索过程中,用到了上界信心树搜索(UCT)<sup>[11]</sup>方法。每个节点所

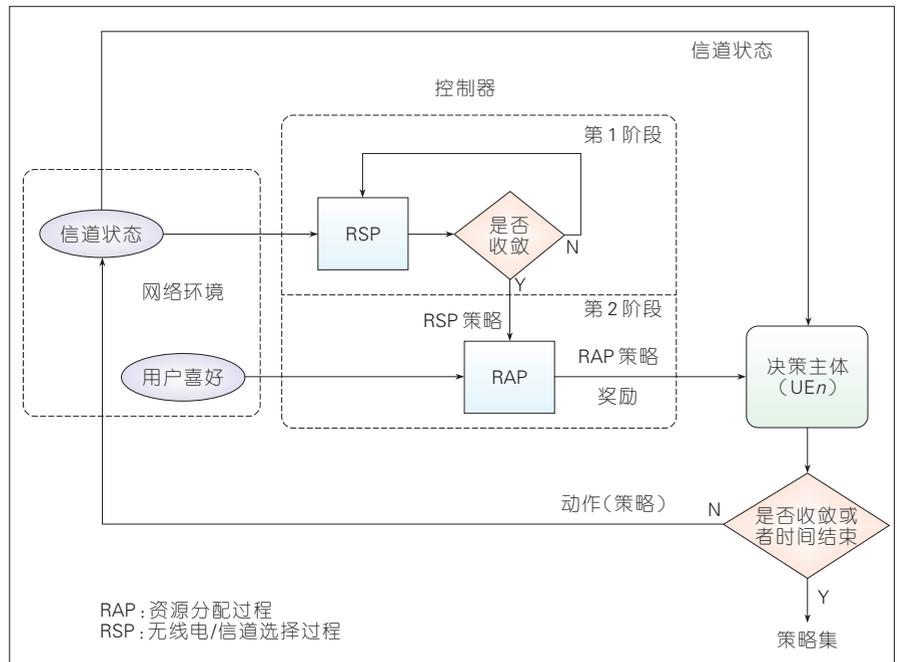
需要满足的是单个用户流的QoS需求,根节点所需要满足的是整个系统的吞吐量的最大化。

我们使用了下面的一些调度技术作为比较:多载体的比例公平调度算法(PFSMTS)<sup>[12]</sup>;LTE作为辅助传输的算法(LAA):在该算法中,Wi-Fi作为流量优先卸载的频段,LTE作为辅助频段;在线学习(OLA):对SARA中的用户进行流式处理。

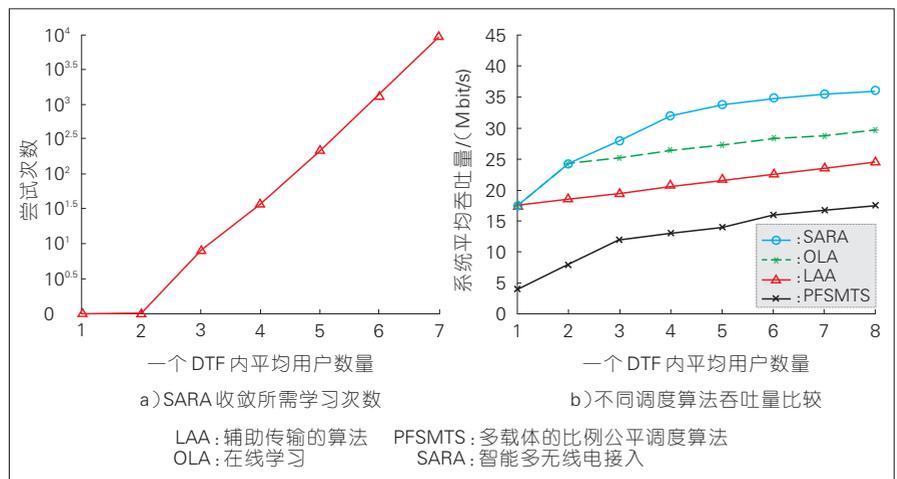
从仿真图我们得到的结论分别是:如图6a)所示,SARA所需的收敛

学习次数随着用户数量的增加而增加,复杂度也随之上升。考虑到短时间尺度调度特性,我们可以设置在短时间内进行资源调度,这样相应进入用户数量也较少,算法收敛较快,网络性能容易被满足。如图6b)所示,SARA的系统吞吐量性能明显要高于其他的调度算法(当用户数量大于3的时候),这意味着SARA这样的智能化的LTE-WiFi聚合方式可以在动态的环境中明显地提高系统资源的

➔ 下转第46页



▲ 图5 基于SMDP的两层决策架构



▲ 图6 SARA性能仿真结果

# 移动边缘计算中数据缓存和计算迁移的智能优化技术

## Intelligent Optimization of Data Caching and Computation Offloading in Mobile Edge Computing System

汪海霞/WANG Haixia  
赵志峰/ZHAO Zhifeng  
张宏纲/ZHANG Honggang

(浙江大学智能通信与网络安全研究所,  
杭州 310007)  
(Institute of Intelligent Communication  
and Network Security, Zhejiang  
University, Hangzhou 310007, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0015-004

**摘要:** 针对移动边缘计算(MEC),设计了计算迁移和数据缓存的联合优化模型,并基于改进的遗传算法(GA)对该模型的时延优化特性进行了优化。仿真表明:提出的方案比传统方案更能有效地降低用户请求的时延,对移动边缘网络的部署和应用具有一定的参考意义。

**关键词:** 移动边缘计算; 计算迁移; 数据缓存; 遗传算法

**Abstract:** In this paper, a combined optimization model for computing migration and data caching is designed. Particularly, an improved genetic algorithm (GA) is put forward to analyze the time delay optimization of the proposed model, which surely has certain significant values to the deployment and applications of mobile edge network.

**Keywords:** mobile edge computing; computation offloading; data caching; genetic algorithm

随着移动互联网和物联网(IoT)的快速发展以及各种新型业务(如虚拟现实(VR)、增强现实(AR)和视频会议等)的不断涌现<sup>[1]</sup>,用户对网络服务质量(QoS)的要求越来越高。因此,为了有效解决移动互联网发展带来的高负荷、低时延等要求的挑战,移动边缘计算(MEC)概念得以提出,并得到了学术界和产业界的广泛支持,被认为是下一代网络的关键技术之一<sup>[2-6]</sup>。

通过将相关的计算任务和请求数据迁移到近端(本地)MEC服务器上,可以减少网络设备能耗和传输时延,并大大提高用户体验。一般来说,计算迁移的关键技术就是充分利用MEC的优点(比如:缩短执行时延,降低能耗等),为此业务卸载过程决定合适的决策过程。文献[7]的工

作表明:将MEC与云计算相结合可以减少迁移任务的相关时延。文献[8]中,作者认为可以通过最小化任务的执行时延来实现最佳的迁移决策。在文献[9]中,作者提出了一种通过最小化能耗来进行决策的方法,其中优化问题被表述为一定约束条件下的马尔可夫决策过程(MDP)。此外,文献[10]和[11]分别分析了移动终端的能量消耗与多用户系统的相关时延间的关系。在文献[12]中,作者考虑在单个宏小区的微基站(BS)上进行数据缓存,通过最小化宏基站服务的请求数量来优化缓存方案。文献[13-18]中,作者也都研究分析了基站缓存中的存储分配问题。

虽然MEC具有较强的计算和数

据传输能力,但相对于现在急速增加的移动终端数量与业务,MEC的相关资源仍然十分有限。本文基于MEC与数据中心(DC)的相互配合机制,根据排队理论分析了每个任务的平均执行时延<sup>[9]</sup>,然后提出了一定缓存空间约束下的时延最小化问题,并通过改进的遗传算法解决该优化问题,从而有效降低了用户请求时延,提高了缓存性能和效率。

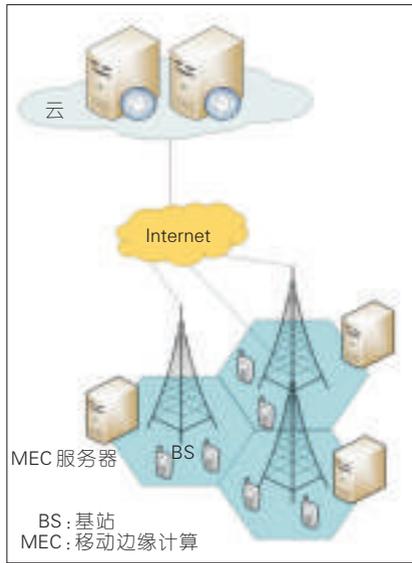
### 1 系统模型

如图1所示,MEC系统有 $B$ 个BS(每个BS配备有一个MEC服务器)和 $M$ 个用户设备(UE)。在这里我们可以把MEC服务器看作是一个具有计算和存储能力的小型数据中心,通过

收稿日期: 2018-01-12

网络出版日期: 2018-03-17

基金项目: 中兴通讯研究项目“基于认知SDN的5G-RAN技术研究”资助



▲图1 MEC系统

部署与边缘交换机关联的虚拟机,将相关的计算任务迁移在 MEC 服务器上,也可以将用户所请求的内容数据缓存在 MEC 服务器上。假设来自 UE 的业务请求的到达过程遵循泊松分布,则 UE  $i$  的请求速率被表示为  $\lambda_i$ 。 $c_i$  为 UE  $i$  所请求的计算任务,  $d_i$  表示 UE  $i$  所要访问的数据量,  $e_j$  表示 MEC 服务器  $j$  的数据缓存容量。把 BS 覆盖范围内每个 UE 当做 M / M / 1 排队系统进行处理,则 MEC 服务  $j$  和 UE  $i$  的服务速率分别可以定义为  $\mu_j$  和  $\theta_j$ 。

通常情况下,移动终端发送请求,然后由 BS 接受处理或将请求发送往远端云端进行处理,这个过程涉及 BS 和云,以及 UE 和 BS 之间的传输。为了简单起见,将 BS 与云之间传输的单位时延表示成  $e_{MC}$ ,将 UE 与 BS 传输过程中的单位时延表示为  $e_{UM}$ 。用变量  $x_{ij}$  表示服务器  $j$  是否缓存了用户  $i$  所请求的内容,如果服务器内缓存了用户请求的内容,则变量  $x_{ij}$  为 1,否则为 0。同样地,如果将用户  $i$  产生的计算任务迁移到 MEC 服务器  $j$  上进行处理,变量  $y_{ij}$  为 1;反之,若进行本地处理,则为 0。

根据排队论<sup>[9]</sup>,我们可以计算得出服务器  $j$  处理计算任务产生的时

延,即:

$$C_j^M = y_{ij} c_i \lambda_i / (\mu_j - \sum_{i=1}^M y_{ij} \lambda_i) \quad (1)$$

其中,  $T_i^{UM} = \sum_{j \in B} e_{UM} (y_{ij} c_i + d_i)$  是迁移到服务器的总计算任务量,注意  $\mu_j - \sum_{i=1}^M y_{ij} \lambda_i > 0$ 。

同样,由用户  $i$  进行本地处理相关任务的时延是:

$$C_i^U = (1 - \sum_{j \in B} y_{ij}) c_i \lambda_i / (\theta_i - (1 - \sum_{j \in B} y_{ij}) \lambda_i) \quad (2)$$

其中,  $(1 - \sum_{j \in B} y_{ij}) c_i \lambda_i$  是由用户本地处理的计算任务量,并需要满足  $\theta_i - (1 - \sum_{j \in B} y_{ij}) \lambda_i > 0$ 。

以上是对用户产生的计算任务,在不同的处理方式下计算的执行时延。对于在数据传输过程中产生的传输时延,如果 MEC 服务器缓存了用户所请求的数据,则用户和服务器之间的传输时延表示为  $T_i^{UM} = \sum_{j \in B} e_{UM} (y_{ij} c_i + d_i)$ ;若服务器缓存中没有用户所请求的内容数据,则服务器和数据中心之间传输数据所得的时延为  $T_i^{MC} = \sum_{j \in B} e_{MC} (1 - x_{ij}) d_i$ ,其中  $1 - x_{ij} = 0$  意味着用户所请求的内容缓存在服务器里,而无需访问数据中心。

综上所述,所有用户产生的总时延为请求的执行时延和传输时延的总和,即:

$$T = \sum_{i=1}^M C_j^M + C_i^U + T_i^{UM} + T_i^{MC} \quad (3)$$

其优化问题就可以表述如下:

$$\text{Minimize } T$$

$$\text{s.t. } \sum_{i=1}^M d_i x_{ij} \leq \varepsilon_j, \quad \forall j \in B, \forall i \in M \quad (4)$$

$$\mu_j - \sum_{i=1}^M y_{ij} \lambda_i > 0, \quad \forall j \in B, \forall i \in M \quad (5)$$

$$\theta_i - (1 - \sum_{j \in B} y_{ij}) \lambda_i > 0, \quad \forall j \in B, \forall i \in M \quad (6)$$

$$x_{ij} \in \{0, 1\} \quad \forall j \in B, \forall i \in M \quad (7)$$

$$y_{ij} \in \{0, 1\} \quad \forall j \in B, \forall i \in M \quad (8)$$

其中,约束条件(4)限制了缓存

在服务器的数据总量不应该超过服务器的总容量。其他的约束条件上文都有提到,不再赘述。

## 2 改进的遗传算法

遗传算法(GA)<sup>[20]</sup>是模拟自然界生物进化过程与机制求解优化问题的一类自适应的启发式智能搜索算法。该算法通过将初始个体进行选择、交叉和变异等操作来更新种群。其最主要的特征是种群搜索策略和种群个体之间的信息交换,非常适用于处理传统方法不容易解决的非线性以及复杂的问题,其应用领域非常广。模拟退火算法(SA)<sup>[21]</sup>是一种贪心算法,它在搜索过程中加入了随机因子,并且以一定的概率来接受比较差的解,这样就有可能会避免陷入局部最优。

为了更加有效地解决上述优化问题,与传统的遗传算法相比,我们提出了如下改进遗传算法(算法1):

(1)在交叉过程中,对新的种群进行了模拟退火操作,使得一些适应值较低的个体也有一定的概率遗传到下一代,这样可以有效防止算法收敛到局部最优,并使算法更加稳定。

(2)在种群进行交叉操作和变异操作之后,分别计算新种群里面个体的适应度函数。这样就可以防止一些优秀的个体在种群进化过程中被遗弃掉。

在改进的算法1中,种群个体是指上述优化问题中的变量  $\{x_{ij}, y_{ij}\}$ ,即前  $M$  维的值为  $x_{ij}$  的取值,后  $M$  维的值为  $y_{ij}$  的取值,所以个体是一个用 0 和 1 填充的  $2 \times M$  维的向量。值得注意的是:在所改进的算法中给出的可行解都满足约束方程(4)、(5)和(6),适应度函数就是所要优化的目标函数。

### 算法1 改进的遗传算法

初始化种群 Pop;  
根据公式(3)计算适应度;

将 Pop 中的最优个体进行存储;  
 While ( $i <$  最大迭代次数) do  
   根据交叉概率对种群 Pop 进行交叉得到新种群 Pop1;  
   计算种群 Pop1 的适应度;  
   并将种群 Pop1 中的最优个体与初始化存储的进行比较, 保存最优的个体;  
   根据变异概率对种群 Pop1 进行变异得到新种群 Pop2;  
   计算种群 Pop2 适应度并进行比较, 保存最优个体;  
   用模拟退火算法更新种群 Pop2;  
   计算适应度并保存最优的个体;  
 End  
 Return 最优个体以及对应的适应度函数

### 3 仿真设计与分析

为了评估本文提出的基于改进遗传算法的联合计算和缓存优化方法, 我们将其与其他两种策略进行了性能对比: 一种是不考虑边缘的计算和存储资源的传统策略, 即本地执行所有的计算任务并且所有请求的数据都存储在数据中心; 另一种是随机策略, 它虽然考虑了移动边缘网络的相关资源, 但采用随机方式来决定是在终端还是在服务器上执行计算任务, 并且数据也是随机存储在边缘服务器或数据中心。随机策略在一定程度上牺牲了服务质量, 从而降低了数据缓存过程的复杂性。

在仿真过程中, 除非文中明确说明, 否则所有主要参数值均基于表 1 进行设置。

图 2 描述了当用户数  $M = 200$  时, 在不同策略下总时延随着服务器缓存容量 (从 500 增加到 3 500) 的变化关系。可以观察到传统策略完全不受服务器缓存容量的变化, 当缓存容量从 500 增加到 1 000 时, 我们提出的策略的总时延从 510.2 ms 急剧下降到 320.1 ms。这是由于随着服务器的缓存容量的增加, 通过改进的遗传

算法的自适应调节使得数据传输时间变短。但当容量达到一定的阈值之后, 由于用户请求数量是一定的, 所以总时延减少的较为缓慢。

图 3 描述了当用户数  $M$  从 50 增加到 400 时, 不同策略的总时延变化情况。我们可以观察到: 与其他策略相比, 我们所提出的策略实现了较低的延迟, 并且可以减少超过 50% 的执行延迟。特别是当用户数量很少时,

边缘服务器有足够的资源用较少的时间来处理大部分任务。随机策略利用了边缘网络的计算和存储资源, 因此其性能优于传统策略。另外, 不同策略下的总时延间的差距也随着用户数目的增加而增大, 意味着当 MEC 系统的规模较大时, 所提出的策略具有较大的性能提升。这是由于提出的策略可以根据用户请求数据大小的不同来提前缓存, 从而使得数

▼表 1 仿真参数设置

参数	值	参数	值
$M$	[50,500]	$B$	[1,5]
$d_i$	[1,20]/MB	$c_i$	[1,5]/MB
$\lambda_i$	[1,10]/s	$\mu_j$	[1.5,2] $\times 10^3$
$\theta_j$	[0.5,1] $\times 10^3$	$e_{MC}$	1.5 s/MB
$e_{UM}$	1 s/MB	$\varepsilon_j$	1 000/MB

图 2  
不同策略下总时延与 MEC 服务器缓存容量的关系

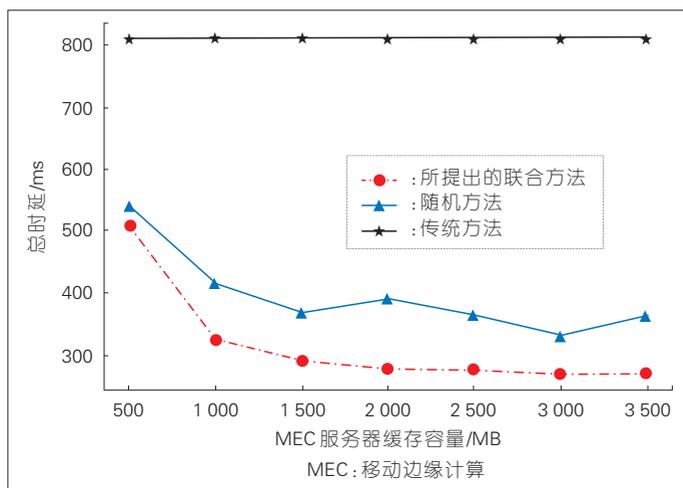
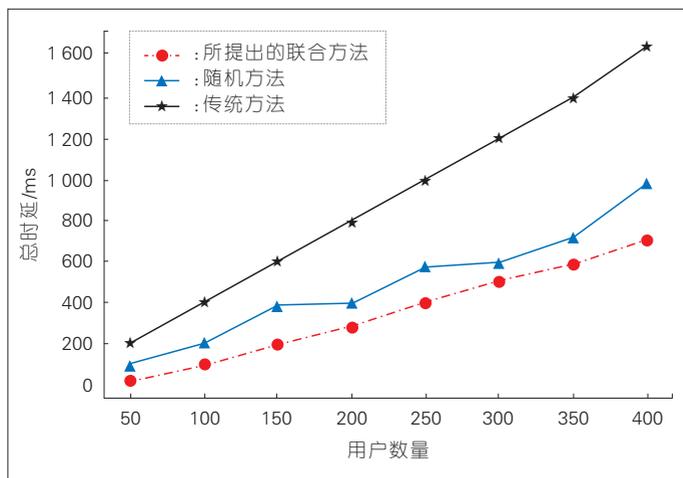


图 3  
不同策略下总时延与用户数的关系



据的传输时延较低。

## 4 结束语

在 MEC 中, 计算迁移和数据缓存是很重要的特性, 缓存决策策略的优劣直接影响移动边缘计算系统的性能。本文中我们通过联合分析执行时延和传输时延, 用改进的遗传算法来建立优化的计算迁移和数据缓存模型, 有效提高了缓存空间的使用效率, 性能方面也有较大程度的提升。

### 参考文献

- [1] MACH P, BECVAR Z. Mobile Edge Computing: A Survey on Architecture and Computation Offloading [J]. IEEE Communications Surveys & Tutorials, 2017, (99): 1-1
- [2] WANG X, CHEN M, TALEB T, et al. Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G Systems [J]. IEEE Communications Magazine, 2014, 52(2):131-139. DOI: 10.1109/MCOM.2014.6736753
- [3] European Telecommunications Standards Institute (ETSI). Mobile Edge Computing Introductory Technical White Paper[R], 2014
- [4] HU Y C, PATEL M, SABELLA D, et al. Mobile Edge Computing: A Key Technology towards 5G[J]. ETSI White Paper, 2015,11(11):1-16
- [5] KUMAR K, LIU J, LU Y H, et al. A Survey of Computation Offloading for Mobile Systems [J]. Mobile Networks and Applications, 2013, 18(1):129-140
- [6] BASTUG E, BENNIS M, DEBBAH M. Living on the Edge: The Role of Proactive Caching in 5G Wireless Networks[J]. IEEE Communications Magazine, 2014, 52(8):82-89. DOI: 10.1109/MCOM.2014.6871674
- [7] RIMAL B P, VAN D P, MAIER M. Mobile-Edge Computing vs. Centralized Cloud Computing in Fiber-Wireless Access Networks[C]//2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). USA:IEEE,2016: 991-996
- [8] LIU J, MAO Y, ZHANG J, et al. Delay-Optimal Computation Task Scheduling for Mobile-Edge Computing Systems[C]//2016 IEEE International Symposium on Information Theory (ISIT).USA:IEEE, 2016: 1451-1455. DOI: 10.1109/ISIT.2016.7541539
- [9] KAMOUN M, LABIDI W, SARKISS M. Joint Resource Allocation and Offloading Strategies in Cloud Enabled Cellular Networks [C]//2015 IEEE International Conference on Communications (ICC).USA:IEEE, 2015: 5529-5534
- [10] JIANG Z, MAO S. Energy Delay Tradeoff in Cloud Offloading for Multi-Core Mobile Devices[J].IEEE Access, 2015, 3:2306-2316. DOI: 10.1109/ACCESS.2015.2499300
- [11] KWAK J, KIM Y, LEE J, et al. Dream: Dynamic Resource and Task Allocation for Energy Minimization in Mobile Cloud Systems[J].IEEE Journal on Selected Areas in Communications, 2015, 33(12):2510-2523. DOI: 10.1109/JSAC.2015.2478718
- [12] POULARAKIS K, IOSIFIDIS G, TASSIULAS L. Approximation Algorithms for Mobile Data Caching in Small Cell Networks[J].IEEE Transactions on Communications, 2014, 62(10):3665-3677. DOI: 10.1109/TCOMM.2014.2351796
- [13] GU J, WANG W, HUANG A, et al. Proactive Storage at Caching-Enable Base Stations in Cellular Networks[C]//2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC). USA:IEEE, 2013:1543-1547. DOI: 10.1109/PIMRC.2013.6666387
- [14] BASTU E, BENNIS M, KOUNTOURIS M, et al. Cache-Enabled Small Cell Networks: Modeling and Tradeoffs[C]// Wireless Communications Systems (ISWCS), 2014 11th International Symposium on.USA:IEEE, 2015. DOI: 10.1109/ISWCS.2014.6933434
- [15] BLASCO P, GUNDUZ D. Learning-Based Optimization of Cache Content in A Small Cell Base Station[C]//2014 IEEE International Conference on Communications (ICC). USA:IEEE, 2014: 1897-1903. DOI: 10.1109/ICC.2014.6883600
- [16] GOLREZAEI N, SHANMUGAM K, DIMAKIS A G, et al. Femtocaching: Wireless Video Content Delivery through Distributed Caching Helpers[C]// Proceedings of IEEE INFOCOM 2012.USA:IEEE, 2012: 1107-1115. DOI: 10.1109/TIT.2013.2281606
- [17] POULARAKIS K, IOSIFIDIS G, SOURLAS V, et al. Multicast-aware Caching for Small Cell Networks[C]//2014 IEEE Wireless Communications and Networking Conference (WCNC). USA:IEEE, 2014:2300-2305. DOI: 10.1109/WCNC.2014.6952688
- [18] POULARAKIS K, IOSIFIDIS G, TASSIULAS L. Approximation Caching and Routing Algorithms for Massive Mobile Data Delivery [C]//2013 IEEE Global Communications Conference (GLOBECOM).USA: IEEE, 2013. DOI: 10.1109/GLOCOM.2013.6831621
- [19] COOPER R B. Introduction to Queueing Theory[M]. New York: North Holland, 1981
- [20] DAVIS L. Handbook of Genetic Algorithms [R]. Van Nostrand Reinhol, 1991
- [21] KIRKPATRICK S, GELATT C D, VECCHI M P, et al. Optimization by Simulated Annealing [J]. Science, 1983, 220(4598):671-680

### 作者简介



**汪海霞**, 浙江大学信息与电子工程学院硕士研究生在读; 主要研究方向为移动边缘计算、智能数据缓存技术。



**赵志峰**, 浙江大学信息与电子工程学院副教授; 主要研究方向为认知无线电、无线 mesh 网络、移动边缘计算和 SDN 在无线通信中的应用; 已发表论文 100 余篇, 申请发明专利 20 余项。



**张宏纲**, 浙江大学信息与电子工程学院教授; 主要研究方向为移动通信、人工智能、机器学习、认知无线电、绿色通信和下一代异构蜂窝网络架构; 已发表论文 200 余篇, 提出 IEEE 802.15 国际标准提案达 16 项, 并拥有 3 项国际专利。

# TD-LTE 网络中大气波导干扰的分析与预测

## Analysis and Prediction of Atmospheric Duct Interference in TD-LTE Networks

孙天宇/SUN Tianyu<sup>1,3</sup>  
周婷/ZHOU Ting<sup>2</sup>  
杨旸/YANG Yang<sup>1</sup>

(1. 中国科学院上海微系统与信息技术研究所, 上海 200050;  
2. 中国科学院上海高等研究院, 上海 201210;  
3. 中国科学院大学, 北京 100049)  
(1. Shanghai Institute of Microsyst & Information Technology, Chinese Academy of Sciences, Shanghai 200050, China;  
2. Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China;  
3. University of Chinese Academy of Sciences, Beijing 100049, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0019-006

**摘要:** 使用来自于江苏移动的实时网络侧数据来分析大气波导干扰(ADI)的特征, 同时结合网络侧数据与气象数据, 使用两种机器学习方法对 ADI 强度进行预测, 并相互比较。仿真结果表明: 使用机器学习可以获得不错的 ADI 预测效果, 当训练样本达到 40 000 条时, 准确率与召回率分别可以达到 72% 与 75% 以上。

**关键词:** 时分复用长期演进(TD-LTE); 大气波导; 机器学习; 干扰预测

**Abstract:** In this paper, the big data of network-side from the current operated network of China Mobile is used to analyze the characteristics of atmospheric duct interference (ADI). Combining network side data with meteorological data, two machine learning methods are used to predict the ADI intensity, and are compared with each other. The simulation results show that machine learning can achieve good ADI prediction effect. When the training sample reaches 40 000, the accuracy and recall rate can reach 72% and 75% respectively.

**Keywords:** time division-long term evolution (TD-LTE); atmospheric ducts; machine learning; interference prediction

大气波导是一种发生在低层大气中的物理现象, 在大气波导中电磁波反射系数随高度呈现负梯度, 使得电磁信号被约束在大气波导层内传播, 并比在非波导层中经历更少的衰减。1963年, BEAN 和 DUTTON<sup>[1]</sup> 首先设定了大气电磁反射率的公式并证明了大气波导的存在。

文献[2]中作者指出波导层可能导致一系列的影响, 比如: 传输的损失、探空雷达的高度误差等。在雷达系统中, 波导层的存在会减少探测范围<sup>[3]</sup>。表面波导会导致气象雷达自动降雨率计算的错误<sup>[4]</sup>。在文献[5]中, ORAIZI 和 HOSSEINZADEH 研究了大气波导对于正交频分复用(OFDM)数字广播系统的影响, 他们论证了大气波导会增加电磁波的强度从而产生

干扰并且降低频率复用距离。

一些迹象表明: 大气波导会在现行的大规模时分双工(TDD)网络中造成干扰, 譬如时分同步码分多址(TD-SCDMA)、时分复用长期演进(TD-LTE)甚至 5G 的 TDD 网络。这是因为对于 TDD 系统而言, 上行和下行传输使用的是相同的频率。来自其他基站的下行信号可能对本地的上行信号产生同频干扰(CCI), 也就是所谓的上下行干扰。所以, TDD 网络必须在上下行信号转换之间设计出一段空白时隙作为保护时隙(GP)来避免上下行干扰。但是在大气波导发生的情况下, 来自远端基站的下行信号传播时较正常情况下会经历很少的衰减, 从而干扰本地的上行信号。我们用 ADI 来表示这种大气波

导干扰。目前在业界 CCI 消除技术已经被广泛的研究, 而最广泛使用的技术就是小区内干扰消除协调(ICIC)。文献[6]中作者提出了一种基于时域 ICIC 的功率控制方法。文献[7]中作者提出一种机制, 可以有效消除全球微波互联接入(WiMAX)系统中小区边界用户的 CCI。但上述这些方法大都侧重于用户侧的 CCI 消除, 并不适用于 ADI 的场景。TDD 系统中的大气波导干扰吸引了来自学术和工业界的大量关注。通过来自中国移动的实时网络侧数据, 我们发现在江苏省徐州市, ADI 干扰最严重的情况下, 高达 27.6% 的小区收到的上行干扰大于 -90 dBm, 此时干扰完全淹没了有用信号导致小区内无法正常通信。ADI 严重影响了 TD-

收稿日期: 2018-01-15  
网络出版日期: 2018-03-28

LTE 系统的性能,因此 ADI 问题亟需被分析与解决。

目前业界存在 2 种主流的方法来检测与估计大气波导:(1)使用探空雷达测量温度、气压等物理量并根据公式来计算修正后的大气电磁折射率<sup>[2]</sup>;(2)使用海面杂波雷达预测海洋电磁折射扇区面<sup>[3]</sup>。方法 1 测量难度大而且花费较高,而方法 2 只适用于海洋场景。此外,这些方法只关注了大气波导本身,一些对于 ADI 关键问题的解答,譬如 ADI 的发生时间或如何鉴别及预测 ADI,依旧没有被解答或被深入研究。除了对干扰的分析,对于运营商来说,预测 ADI 的发生以提高网络性能是十分重要的。

在这篇文章中,我们创造性地使用来自运营商的实时网络侧数据来进行 ADI 预测。由于 ADI 产生的随机性,我们无法从帧结构上解决。因此,成功地预测 ADI 的发生是解决大气波导干扰的关键。ADI 干扰产生的原因复杂繁多,传统的建模方法并不适用。所以,我们结合了网络侧数据与气象数据来进行预测,并使用了支持向量机(SVM),通过将干扰问题转化为一个分类问题实现预测。此外,考虑到关联特征与预测结果之间的相关性,我们还使用因子分解机(FM)算法,将干扰问题转化为一个线性加二阶组合特征的回归问题。该工作填补了大规模 TDD 网络中 ADI 研究的空白,且帮助移动运营商在大气波导场景下提高网络性能。

## 1 ADI 的概述

### 1.1 大气波导现象

文献[2]中表明:大气中异常的电磁信号传播条件来源于大气折射率的变化,折射率如公式(1):

$$N = \frac{77.6}{T} \left( p + \frac{4810e}{T} \right) \quad (1)$$

这里  $N$  是无量纲的折射系数,  $T$  是绝对温度单位为  $K$ ,  $p$  是气压单位为  $hPa$ ,  $e$  是水气压单位为  $hPa$ 。考

虑到地球曲率的影响,修正后的折射系数  $M$  为:

$$M = N + \frac{z}{R_e \times 10^6} \quad (2)$$

$z$  是距地表的高度,  $R_e$  是地球的半径。在正常情况下  $M$  随着高度的增加而增加。当  $M$  具有负梯度的时候,对应的高度上就会产生大气波导,电磁信号就会被捕获而进入这段高度之中。

图 1 展示了几种异常的大气状况。

### 1.2 ADI 对 TD-LTE 网络的影响

在 ADI 的场景下,信号的传播会经历更少的衰减。从而,来自于超过 GP 最大保护距离外基站的信号有可能对本地的上行信号造成同频干扰。目前,业界对于 TDD 系统同频干扰的研究大都集中在用户侧的下行同频干扰消除。这些技术是建立在准确的信道估计或邻近基站的协同的基础上,并不适用于 ADI 的场景。同时,在无线传感网中广泛使用的定位技术<sup>[4]</sup>也不适用于 ADI 场景,这是因为 TDD 系统无法将远端基站与本地用户的信号分离开来,从而我们就无法根据接收信号的强度来确定干扰源。

图 2 展示了大气波导存在时上下行信号的碰撞情况,来自远端基站的下行信号传输超过了 GP 对应的最大距离从而有部分与本地的上行信号

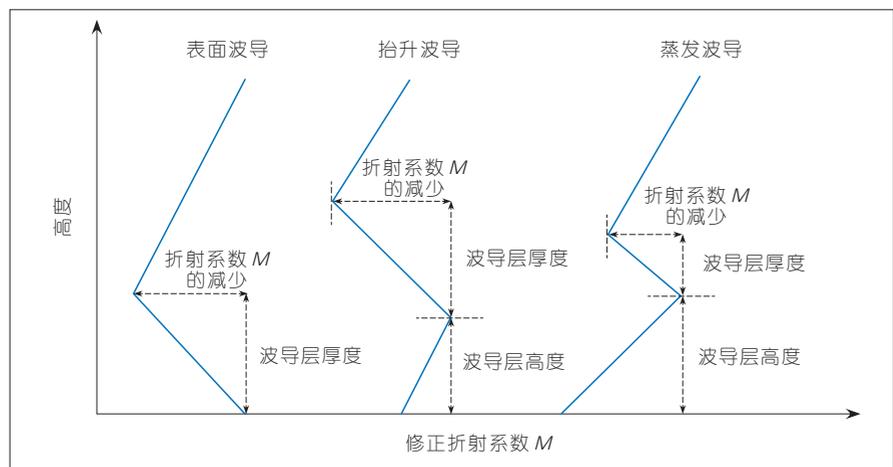
形成重叠,产生同频干扰。ADI 对大规模组网的 TD-LTE 网络会造成严重的影响。根据中国移动的报告,在 2016 年 8 月 25 日,在江苏省徐州市,高达 27.6% 的小区就曾受到了功率大于  $-90$  dBm 的 ADI 干扰,在这种情况下有用信号完全被干扰淹没,小区的正常通信受到阻碍。因此,有效地预测 ADI 的发生具有很高的研究价值。

### 1.3 TD-LTE 网络中大气波导干扰的特征

我们使用来自江苏移动 2016 年 5 月 16 日的网络侧数据绘制了干扰地图。图 3 的 a)、b)、c)和 d)分别展示 1:00 am、7:00 am、13:00 pm 和 19:00 pm 时刻全江苏省的 ADI 干扰情况。

在图 3a)中,同频干扰正在逐渐增强,并在图 3b)中达到最强,之后在图 3c)、d)中干扰水平持续降低,并保持稳定,低于  $-90$  dBm。这些迹象都表明在 1:00 am ~ 7:00 am 这段时间内,ADI 干扰现象很有可能发生。从气象学上可以解释为:在午夜至清晨的这段时间内,地面的温度下降较快,地表的低层大气容易产生一个逆温层,即随着高度的上升大气温度也是上升的。根据公式(1),修正折射率  $M$  会在该范围内具有一个负梯度,从而导致大气波导层的产生。

此外在图 3b)中的中北部受到的干扰强度会大于南部。考虑到江苏



▲ 图 1 异常大气条件下修正折射系数  $M$  随高度变化的示例

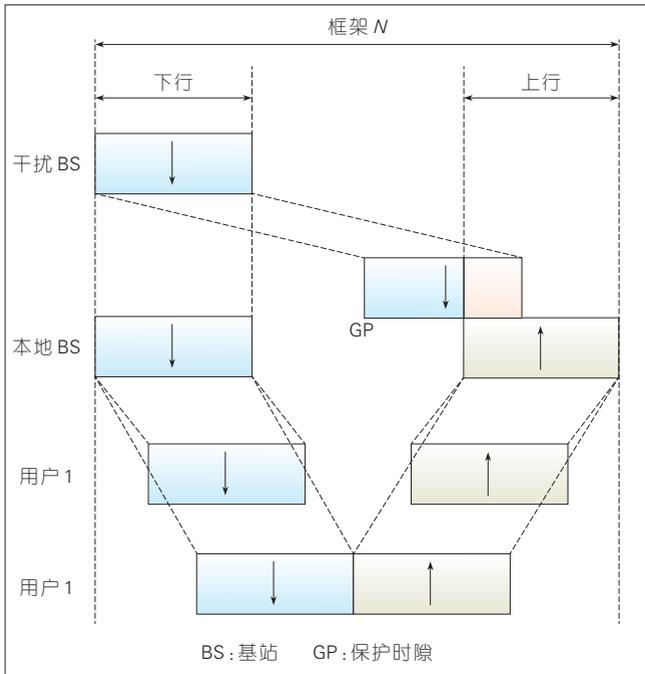


图 2 大气波导存在时上下行信号碰撞情况

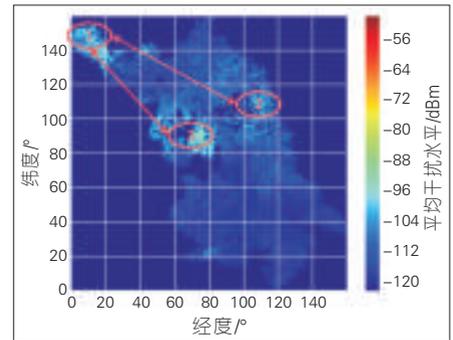


图 4 ADI 互易性质说明

发现 1、2 和 3 号区域的干扰都很强。与 1:00 am 时的情况相比, 这 3 个地区的干扰明显都同时增强了。此外图 5 给出了这 3 个地区的受干扰的基站

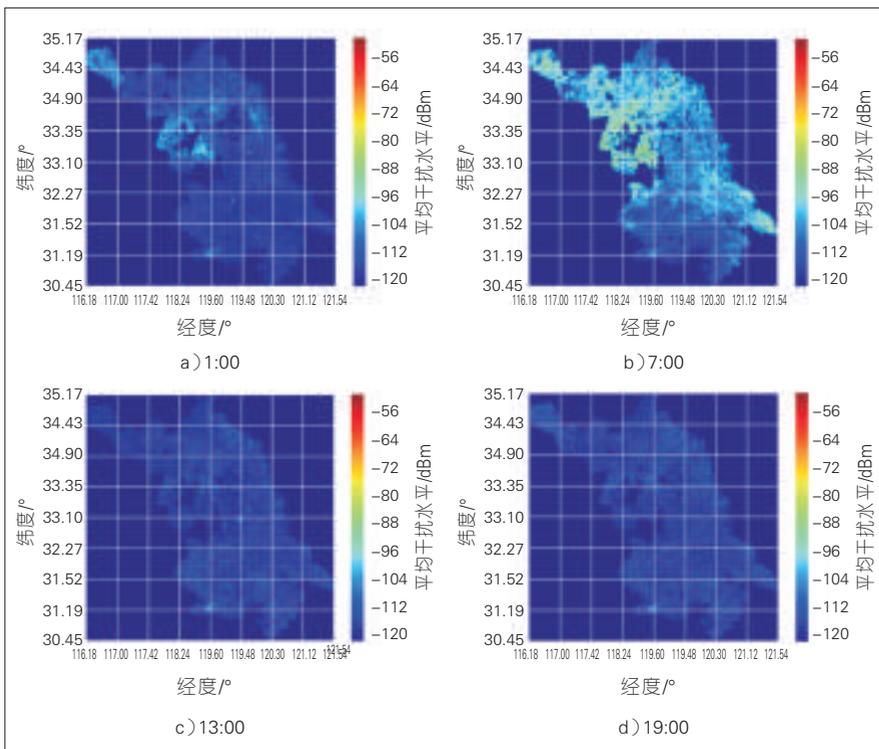


图 3 大气波导干扰地图

省的实际地理情况, 中北部农村较多, 地势开阔平坦, 利于波导层的展开; 而在城市之中, 高楼较多, 干扰信号易被阻断, 从而干扰的水平会降低。由此我们可以得出: ADI 具有随

时间变化的特点, 而且易于发生在凌晨到清晨这段时间; 乡村地区通常有更大的概率受到 ADI 影响。

图 4 展示的是 2016 年 5 月 16 日 2:00 am 的干扰情况。我们可以清楚

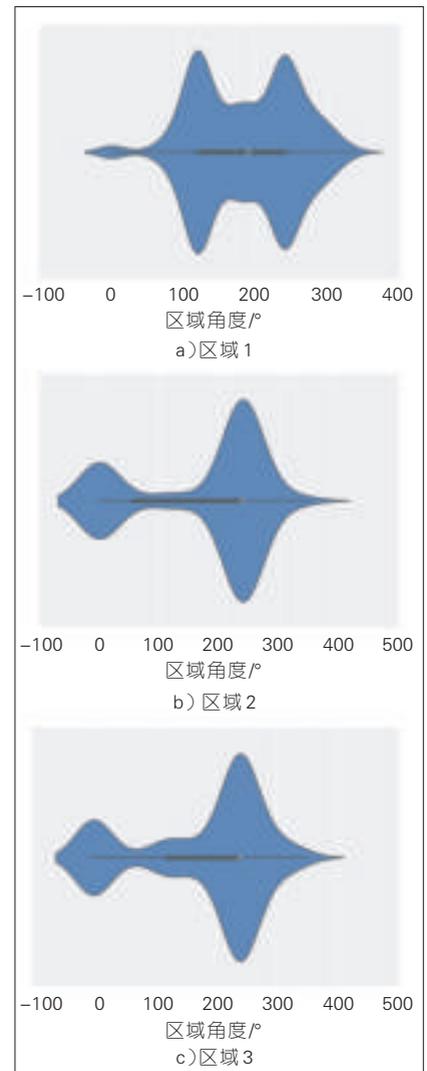


图 5 受 ADI 干扰基站数量在扇区角上的分布

数量在基站扇区角上的分布。我们可以明显发现:区域1的干扰来自于120°和240°的方向,120°正好分别指向区域2和区域3;而区域2和区域3的干扰主要来自240°方向,恰好是区域1的方向。至于区域1来自于240°方向的干扰,极有可能来源于江苏省的西北方向的省份,譬如安徽与河南。图4和图5说明了ADI的方向特性以及互易性质。

上述的种种现象表明:ADI干扰的发生与气象数据和网络侧数据是存在着关联性的,因此我们也可以使用这些数据来对ADI的发生进行有效预测。

## 2 基于机器学习的ADI预测方法

### 2.1 基于SVM的分类器训练

SVM是一种二分类的机器学习算法<sup>[10]</sup>。分类器意味着当待分类数据输入时,我们可以自动判别它属于哪一类。在我们的问题中,输入是处理过的基站侧数据与气象数据;而输出是1或者-1,分别代表基站是否会被ADI影响。SVM的主要思想是在样本空间中寻找一个鲁棒性最好的划分超平面,使得正负样本间的距离最大。划分超平面可以用 $\omega^T x + b = 0$ 表示。假设我们有训练样本集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ 。其中, $y_i \in \{1, -1\}$ 是我们训练样本的标签。在实际中我们根据运营商的需求设定阈值,将干扰大于阈值的样本标记为1,代表此基站受到ADI的影响;将干扰小于阈值的样本标记为-1,代表正常的基站。 $x_i$ 是样本的特征,它由基站侧数据与气象数据通过处理生成。我们使用训练数据集训练出 $\omega$ 和 $b$ ,从而获得一个线性分类器: $f(x) = \text{sign}(\omega^T x + b)$ 。

SVM的优化问题可以写为:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \max(1 - y_i(\omega^T x_i + b), 0) \quad (3)$$

其中, $C$ 是常数, $C \sum_{i=1}^m \max(1 - y_i(\omega^T x_i + b), 0)$ 是正则项,我们使用它来避免过拟合与线性不可分的情况。公式(3)是一个二次规划的优化问题,求解它的时间复杂度是 $O(n^2)$ ,当训练样本量过大会消耗大量的训练时间。文献[11]中,作者提出了一种最小序列优化的方法,避免了对海森矩阵求逆;但该方法侧重于减少空间复杂度,对时间复杂度的优化效果不大。利用文献[12]中提出的乘子交替法(ADMM)框架可以分布式训练,我们可以将时间复杂度降低为 $O(n^2/k^2)$ ,其中 $k$ 为训练数据的分区数。

### 2.2 基于因子分解机的回归预测

SVM是一种广义线性模型。显然,在线性模型下,样本中的特征分别单独对最后的标签形成影响,相互之间并不形成关联。而由公式(1)、(2)我们可以发现:ADI的形成受多种气象特征相互影响,即关联特征与标签之间是存在正相关性的。因此,引入两个特征的组合是非常有意义的。

目前主要有2种手段得到组合特征:(1)通过数据分析加人工构造来进行人工特征工程;(2)通过模型做组合特征的学习。方法1需要专家知识,在我们的问题中并不适用;而方法2常用的模型有因子分解机(FM)、神经网络等。

多项式模型是包含特征组合的最直接的一种模型,其表达式如公式(4):

$$y = \omega_0 + \sum_{i=1}^n \omega_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \omega_{ij} x_i x_j \quad (4)$$

此时,我们的问题变为了一个回归问题。对应于2.1节的分类问题,我们令 $y_i$ 表示接收到的ADI干扰强度,单位为dBm。直观地看公式(4),前两项是线性回归模型的表达式,而最后一项则是二阶特征交叉项,即组

合特征项。这类问题可以使用一般的最小二乘规划来求解参数。

然而在ADI问题中,存在着许多类别变量,如基站扇区角、天线下倾角等。这些一维的特征经过One-Hot编码后维数会大大增加,使得训练数据变得稀疏。在这种情况下,只有当 $x_i x_j$ 同时不为0才能有效,从而学习到的 $\omega_{ij}$ 不是充分统计量结果,会导致 $\omega_{ij}$ 不准确。

文献[13]中提出的FM方法通过引入隐向量内积来代替交叉项系数,使得 $x_m x_i$ 与 $x_i x_j$ 的参数不再相互独立,从而使得我们可以在样本稀疏的情况下合理地估计参数。改进后的表达式可以写为:

$$y = \omega_0 + \sum_{i=1}^n \omega_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j \quad (5)$$

其中, $\langle v_i, v_j \rangle$ 表示两个长度为 $k$ 的向量的内积,即:

$$\langle v_i, v_j \rangle = \sum_{f=1}^k v_{i,f} \cdot v_{j,f} \quad (6)$$

这相当于对我们使用 $k$ 个描述特征的因子 $v_i = (v_{i,1}, \dots, v_{i,k})$ ,来表示每一维特征分量 $x_i$ 。这样做的好处是:所有满足存在某个 $j \neq i$ ,使得 $x_i x_j \neq 0$ 的样本都可以用来学习隐向量 $v_i$ ,在很大程度上避免了数据稀疏造成参数估计不准确的影响。

## 3 仿真结果

### 3.1 仿真环境

我们通过仿真来验证和比较两种机器学习预测算法的性能,仿真环境配置见表1。

### 3.2 预测结果

仿真在不同大小的训练集下进行,样本数分别设为2 000、4 000、10 000、20 000、40 000。每次训练完,我们按照训练集20%的大小选取测试集进行测试。我们使用了K-近邻<sup>[14]</sup>(KNN)算法作为比较基准。在

▼表 1 仿真环境配置

操作系统	ubuntu 16.04.2 LTS
编程语言	python2.7
工具库	scikit-learn v0.20.dev0
气象数据来源	中国气象数据网 <a href="http://data.cma.cn/">http://data.cma.cn/</a>
网络侧数据	来自江苏省移动共计 24 4375 个基站,数据维度包含基站的经纬度、时间、扇区角、频段、天线高度、ADI 干扰的强度
气象数据	包含了气温、气压、水汽压这 3 个式(1)中的关键参数,以及其他气象上的物理量

ADI: 大气波导干扰

KNN 算法中,测试样本的标签由样本空间中离它最近的  $k$  个样本投票决定。显然 KNN 是一种懒惰的方法,它没有任何的学习过程,只是存储数据并查找。当我们不使用任何学习策略的时候,这显然是一种最为本能的预估方法。

我们使用准确率和召回率来评估最后的训练结果:

$$\text{accuracy} = \frac{n_{\text{correct}}}{n_{\text{all}}}, \text{recall} = \frac{n1_{\text{correct}}}{n1_{\text{all}}} \quad (7)$$

其中,  $n_{\text{all}}$  表示总测试样本数,  $n1_{\text{all}}$  表示标记为 1,也就是被干扰的测试样本数。  $n_{\text{correct}}$  和  $n1_{\text{correct}}$  分别表示总测试样本和标记为 1 的样本中被判断正确的样本数。召回率反映了有多少被干扰的基站会被成功预测出来,对于运营商提高用户体验来说,是一

个重要的指标。

图 6(a)、b) 分别展示了 3 种方法在准确率与召回率上的仿真结果。当样本数较少时, KNN 的性能要大大好于 SVM 与 FM,这是因为训练数据集不充分导致 SVM 与 FM 学习不充分。而随着训练样本数的提升, SVM 与 FM 性能提升迅速,因为它们充分学习了训练数据中的特征;而 KNN 因为本身不具备学习能力,性能增长放缓。我们可以发现:当训练样本数大于 18 000, FM、SVM 的性能都要大大优于 KNN。其次, FM 的性能一直优于 SVM,这正是由于 FM 使用了关联特征,提高了自身的泛化性能。

#### 4 结束语

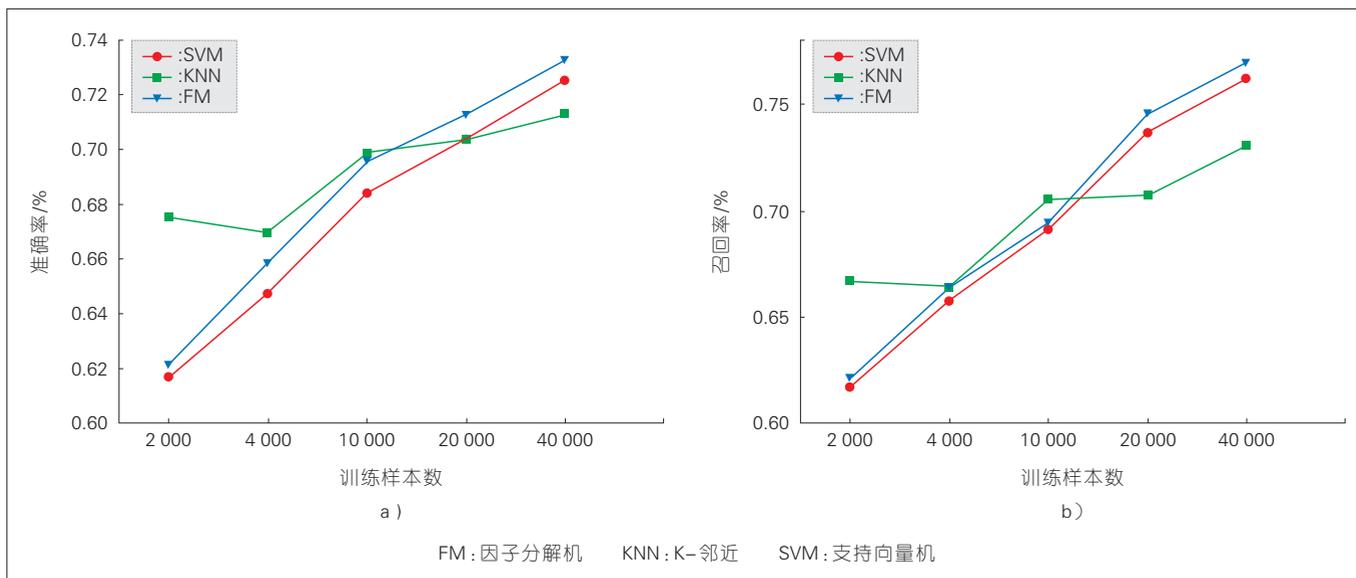
在本文中,我们分析与预测了大规模组网 TD-LTE 系统中的 ADI 问

题。不同于传统气象学中大气波导的检测方法,我们使用了机器学习并结合运营商提供的实时网络侧数据。同时,我们也给出了 ADI 在物理帧结构上的成因,并利用干扰地图等分析手段得出了 ADI 在发生时段与空间分布上的一些特征。

我们首先提出了一种基于 SVM 二元分类器的 ADI 预测方法,创新性地结合了实时网络侧大数据与气象数据。仿真结果表明:随着训练样本数的增长,准确率与召回率都会增长。当训练样本数到达 40 000 时, SVM 的准确率与召回率分别达到 72% 与 75.5%,均显著好于 KNN。

在此基础上,我们提出了一种基于 FM 回归预测 ADI 强度的方法。在线性回归的基础上引入了关联特征,从而提高泛化性能。仿真结果表明:当训练样本数到达 40 000 时, FM 的准确率与召回率分别达到 73.6% 与 76.4%,要优于 SVM。

我们将这篇文章作为 ADI 研究的一个出发点。在未来的工作中仍然有很多的问题亟待解决:(1) ADI 的统计特征与如何对它建模值得研究;(2) 如何精确定位 ADI 的干扰源? 目前来自于运营商的数据在扇区角方面还比较粗糙,只包含 3 个方向的信



▲图 6 各预测方法准确率与召回率性能比较

息,通过场地测量可以获得更为精确的扇区角信息从而帮助进行 ADI 干扰源的定位。

### 致谢

本文的研究工作和完稿得到了江苏移动研究中心的龚淑蕾和上海无线通信研究中心的张武雄与许晖的帮助和支持,我们表示衷心感谢。

### 参考文献

- [1] BEAN B R, DUTTON E, FRANK V, et al. A Radio-Meteorological Study[C]// Antennas and Propagation Society International Symposium. USA:IEEE, 1963:56-61
- [2] MENTES S S, KAYMAZ Z. Investigation of Surface Duct Conditions over Istanbul, Turkey [J]. Journal of Applied Meteorology and Climatology, 2007, 46(3): 318-337.DOI: 10.1175/JAM2452.1
- [3] ANDERSON K D. Radar Detection of Low-Altitude Targets in a Maritime Environment [J]. IEEE Transactions on Antennas and Propagation, 1995, 43(6): 609-613.DOI: 10.1109/8.387177
- [4] GERSTOFT P, ROGERS L T, HODGKISS W S, et al. Refractivity from Clutter Using Global Environmental Parameters[C]//IEEE 2001 International Geoscience and Remote Sensing Symposium. USA:IEEE, 2001,6: 2746-2748
- [5] ORAIZI H, HOSSEINZADEH S. The Effect of Atmospheric Duct on Modern OFDM-Based Digital Broadcasting Systems[C]// 33rd European Microwave Conference. USA: IEEE, 2003: 747-750. DOI: 10.1109/EUMA.2003.341061
- [6] LU W, FAN Q, LI Z, et al. Power Control Based Time-Domain Inter-Cell Interference Coordination Scheme in DSCNs[C]//IEEE International Conference on Communications (ICC). IEEE: USA, 2016.DOI: 10.1109/ICC.2016.7511467
- [7] XU W, SEZGINER S. Co-channel Interference Cancellation in Reuse-1 Deployments of WiMAX System[C]//IEEE Wireless Communications and Networking Conference (WCNC). USA: IEEE, 2012:342-346
- [8] LI X H, HE Z S, HE J X, et al. Design of an Ocean Atmospheric Duct Signal Processor [C]// International Symposium on Intelligent Signal Processing and Communication Systems. USA:IEEE, 2010:1-4
- [9] MAO G, ANDERSON B D O, FIDAN B. Path Loss Exponent Estimation for Wireless Sensor Network Localization[J]. Computer Networks, 2007, 51(10): 2467-2483.DOI: 10.1016/j.comnet.2006.11.007
- [10] BURGESS S, VANDENBERGHE L.A Tutorial on Support Vector Machines for Pattern Recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(1):121-167
- [11] PLATT J. A Fast Algorithm for Training Support Vector Machines[J]. Journal of Information Technology, 1998, 2(5):1-28
- [12] BOYD S, PARIKH N, CHU E, et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers[J]. Foundations and Trends in Machine Learning, 2011, 3(1): 1-122
- [13] RENDLE S. Factorization Machines with libFM[J]. Acm Transactions on Intelligent Systems & Technology, 2012, 3(3):1-22. DOI: 10.1145/2168752.2168771
- [14] LAROSE D T. Discovering Knowledge in Data: An Introduction to Data Mining[M]. USA: Wiley-Interscience, 2004

### 作者简介



孙天宇,中国科学院上海微系统与信息技术研究所硕士在读;主要研究方向是机器学习在无线通信中的各种应用。



杨婷,中国科学院上海微系统与信息技术研究所博士毕业,中国科学院上海高等研究院新媒体无线技术研究中心研究员;主要研究方向为异构网络中的资源分配与智能网络;曾获 2015 年中国通信协会技术创新一等奖、上海科学技术进步二等奖;已发表论文 20 余篇,同时有 40 余项专利被批准或正在受理。



杨励,中国科学院上海微系统与信息技术研究所研究员、学位委员会副主任,中国科学院无线传感网与通信重点实验室主任,上海无线通信研究中心主任,上海雾计算实验室联合主任,国际雾计算产学研联盟大中华区主任,中国通信学会学术工作委员会委员等;研究领域包括无线传感器网络(物联网)、新一代移动通信系统(5G)、雾计算与网络技术、开放无线测试验证平台等;承担了国家科技重大专项(03 专项)、国家“863”计划、国家自然科学基金重点等一系列前沿课题研究;获选 IEEE Fellow、中组部“万人计划”、上海市“千人计划”、中科院“百人计划”、上海市“优秀学术带头人”和“领军人才”等荣誉;发表论文 150 余篇,申报专利 80 余项。

# 基于数据驱动深度学习方法的无线信道均衡

## The Wireless Channel Equalization Based on Data-Driven Deep Learning Method

杨旸/YANG Yang<sup>1,2</sup>

李扬/LI Yang<sup>1</sup>

周明拓/ZHOU Mingtuo<sup>2</sup>

(1. 中国科学院上海微系统与信息技术研究所, 上海 201899;

2. 上海无线通信研究中心, 上海 201210)

(1. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 201899, China;

2. Shanghai Research Center for Wireless Communications, Shanghai 201210, China)

在数字无线通信系统中, 二进制符号通过色散信道传输, 导致产生时延扩展和符号间串扰 (ISI); 而 ISI 的存在阻碍了频率带宽的有效利用和系统性能的提高<sup>[1]</sup>。无线信道基本上可以用一个复数值有限脉冲响应 (FIR) 滤波器来表示, 而信道输出是滤波器抽头权重的线性组合, 并且被噪声污染。信道均衡则是基于信道观测来抵消 ISI 和噪声的影响, 从而重建传输序列。通常, 传输信道会受到线性和非线性失真的影响, 现实中往往将信道建模为一个非线性系统。传统的线性均衡算法, 如递归最小二乘 (RLS)<sup>[2]</sup> 算法, 在信道非线性特性强烈及多径丰富的情况下无法保证较低的误码率。

人工神经网络的强大非线性拟合能力近来在无线信道均衡领域受到了较多关注。文献[3-4]中, 作者分

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0025-005

**摘要:** 无线信道均衡可以被看成将接收端符号恢复成发射符号集中某个符号的问题; 而无线通信系统中的许多恢复过程可以被认为是通过学习一组具有良好的概率包络和相干时间的随机滤波器来克服信号的线性混合、旋转、时移、缩放以及卷积等特性。具体地, 使用卷积神经网络 (CNN) 来学习这些滤波器, 然后将学习到的滤波器送入后续的循环神经网络进行时域建模, 最后对信号进行分类。实验显示: 卷积-循环神经网络 (CRNN) 均衡器与传统的递归最小二乘滤波器 (RLS)、多层感知机滤波器 (MLP) 在达到相同误码率 (SER) 情况下好 2 ~ 4 dB。

**关键词:** 信道均衡; 无线通信; 深度学习; 神经网络

**Abstract:** Channel equalization can be viewed as a task that classifies or reconstructs the received signal as a symbol from the transmitting symbol set at the receiver. Many recovery processes in wireless communication systems can be considered to overcome linear mixing, rotation, time-shift, scaling and convolution by learning a set of random filters with good probabilistic envelope and coherent time. Concretely, convolutional neural network (CNN) is used to learn these filters, which are send into the subsequent recurrent neural network (RNN) for temporal modeling, and finally the signals are classified. Experimental results show that our convolutional recurrent neural network-based (CRNN) equalizer outperforms the recursive least square (RLS) and multi-layer perceptron network (MLP) equalizers by average 2 to 4 dB with the same symbol error rate (SER).

**Keywords:** channel equalization; wireless communications; deep learning; neural network

别使用了不同结构的神经网络进行均衡, 并与传统的信道均衡器做对比, 发现神经网络算法能达到更低的误比特率。文献[5]中作者提出一种基于自适应决策反馈的信道均衡器, 其中使用了多层感知机 (MLP) 神经网络结构, 来解决通信卫星的信道均衡问题。由于无线信号由复数值表示, 文献[6]中作者提出一种基于复数值的多层感知机神经网络, 并提出一种基于复数值的反向传播算法, 使得其更适用于实际系统。

受时变信道的影响, 无线信号属于时间序列数据, 但是上述信道均衡器并未将这一特性考虑在内。因此, 本文中我们提出一种基于卷积-循环神经网络 (CRNN) 的信道均衡器, 其特点是既考虑了信道的非线性又兼顾了信道的时变性。我们利用卷积神经网络的平移不变性来学习像传统滤波器抽头权重那样的匹配滤波器, 然后再将学习到的滤波器送入接下来的循环神经网络进行时间建模, 继而就可以用于对接收到的信号进

行分类。

具体地,我们使用自制的8×8并行多输入多输出(MIMO)通信系统收集实际的收发信号,构建成用于训练我们神经网络的数据集。特别地,收发端仅使用一根天线,从而组成一个单发单收系统,信号发送采用正交相移键控(QPSK)调制。据我们所知,我们是第一个采用CRNN结构进行信道均衡的<sup>[7]</sup>。我们的主要贡献总结如下:

(1)通过将信道均衡建模成一个条件概率分布学习问题,从而可以使用神经网络算法进行求解。

(2)基于问题建模,我们提出一种基于CRNN的信道均衡器,在其中既考虑了信道的非线性失真又兼顾了信道的时变性。

我们将所提出的CRNN均衡器与传统的RLS均衡器,以及MLP均衡器在误码(SER)率性能上进行了比较。大量实验结果显示:我们的CRNN均衡器要比其他另外两个均衡器在达到相同SER时好2~4 dB。

### 1 问题建模

无线信道均衡可以表示成如图1所示的条件概率分布学习问题。具体地,我们假设发送符号集为S,当中包含K个符号,为方便建模,我们以自然数1~K来表示这K个符号。对于给定的接收信号 $y^{(i)}$ ,信道均衡的作用就是将其识别或者还原为发

送符号集的某符号 $x^{(i)}=k$ ,通过计算对应于每个发送符号的条件概率(也即整个符号集的条件概率分布),我们可以得出条件概率最大的符号即为 $y^{(i)}$ 最可能的对应发送符号。在这当中, $y^{(i)}$ 由有限个时延副本构成:

$$y^{(i)} = [y^{(i)}, y^{(i-1)}, \dots, y^{(i-N+1)}]^T \quad (1)$$

其中,N为信号经历的多径个数,假设其可在均衡前提前估计出。同时,条件概率可以表示为:

$$P_{X|Y=y} = P(X=x|Y=y) \quad (2)$$

然后,我们可以用一个神经网络作为近似函数来参数化这个条件概率,以如下方式:

$$f_w: R^N \rightarrow R^K \quad (3)$$

$f_w$ 即以 $w=(w_1, \dots, w_K)^T$ 为参数的神经网络近似函数,作用是将一个N维输入映射为一个K维输出。 $w$ 为N×K的矩阵,为神经网络各神经元之间连接的权重。具体地, $f_w$ 可以由下面的函数形式将条件概率分布参数化:

$$P_{X|Y=y^{(i)}} = f_w(y^{(i)}) = \frac{e^{x_j}}{\sum_{j=1}^K e^{x_j}} \in [0, 1] \quad (4)$$

其中, $x=(w_1^T y^{(i)}, \dots, w_K^T y^{(i)})^T$ ,值得注意的是此函数处处可微,分母项将K个条件概率归一化。 $w$ 控制条件概率分布的表现,我们通过调节 $w$ 即神经网络的连接权重来优化基于条件概率分布的符号预测准确率,即信道均

衡性能。

基于以上推导,目前我们需要一个标准对神经网络调整参数后的预测性能进行评估。具体地,我们引入交叉熵损失函数进行这项评价,其表示如下:

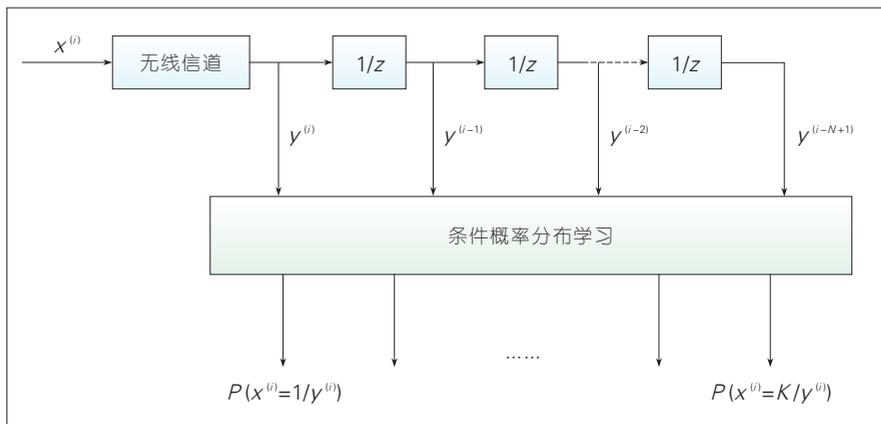
$$E(w) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^K 1\{x^{(i)}=j\} \log \frac{e^{w_j^T y^{(i)}}}{\sum_{l=1}^K e^{w_l^T y^{(i)}}} \quad (5)$$

其中,m为训练集样本的个数,一个训练样本可表示为 $(y^{(i)}, x^{(i)})$ 。 $y^{(i)}$ 为接收符号, $x^{(i)}$ 为该接收信号对应的真实发送符号,在机器学习的语境中则为样本对应的标签。 $1\{\cdot\}$ 为指示函数,表示在大括号里的假设若为真则返回1,否则返回0。该函数对一组神经网络权重 $w$ 所导致的预测误差进行计算,我们利用收集到的训练样本集合通过调节 $w$ 最小化这一函数从而提高预测准确率。

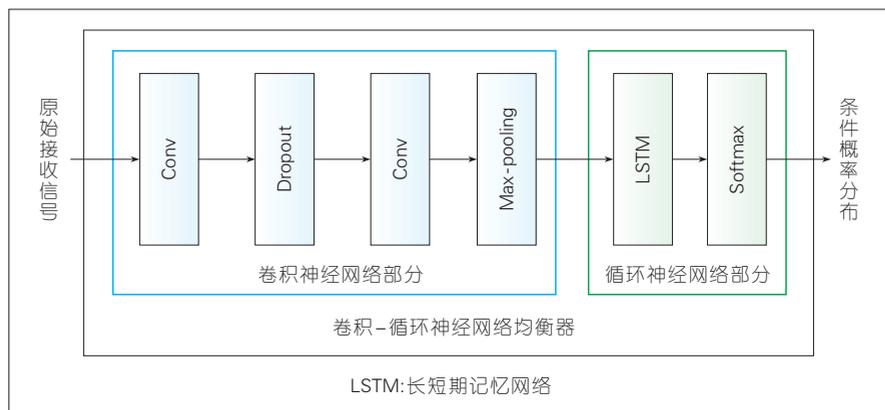
### 2 CRNN 无线信道均衡器

基于以上问题建模,我们提出一种流水线式神经网络均衡器,由2部分构成:第1部分为一个卷积神经网络;第2部分为一个循环神经网络。神经网络的原理详细可见文献[8],本文由于篇幅所限将不深入讨论。

我们所提出的CRNN均衡器如图2所示,除输入和输出之外,其为一个5层神经网络结构。输入层为一个2×N的接收符号,分别由N个时延副本的实部和虚部构成。输入层后接第1个卷积层Conv,卷积层是卷积神经网络中最关键的环节,用于提取输入样本的特征<sup>[9]</sup>。每个卷积层都由若干个卷积核组成,每个卷积核都是一个特定长、宽的2D滤波器,在计算机中由一个随机初始化的2D权重矩阵表示。这些卷积核会以一定的步长遍历整个上层输入进行卷积操作,生成特征图,即我们前文提到的匹配滤波器。神经网络下一层为Dropout,表示上层有多少比例的神经元与下一层进行连接,由1个0~1之间的参数控制。该层的作用是适当降低神



▲图1 将信道均衡建模成一个条件概率学习问题



▲图2 CRNN均衡器

网络的表达能力,降低神经网络过拟合的可能性。接下来的卷积层的作用与第1个卷积层相同,但是进一步学习更加抽象地表达,即可能得到的是有利于识别信号的幅度、相位、多普勒频移等特征。卷积神经网络的最后一层是1个Max-pooling层,具体操作类似于卷积核,但仅提取对应面积内最大的权重值,一方面用于减少连接数,降低复杂度,另一方面可以获得局部相似性特征。我们使用卷积神经网络的原因在于无线通信系统中许多信号识别或还原是通过学习一组具有良好的概率包络和相干时间的随机滤波器来克服信号的线性混合、旋转、时移、缩放以及卷积等特性。那么经过上述一系列操作之后,卷积神经网络就学习出了这些滤波器。

下一步我们将这些学习到的匹配滤波器送入由长短期记忆(LSTM)单元<sup>[10]</sup>构成的循环神经网络用于时间建模,以克服无线信道的时变特性。具体地,LSTM单元通过记忆上一时刻的状态并把这一状态加入当前状态的计算,来进行时间依赖学习。需要注意的是:LSTM层的权重在每一时刻都是共享的。最后,我们使用Softmax对结果进行归一化并得到发送符号的条件概率分布。

至此,我们可以开始使用大量预先收集的收发数据训练所提出的CRNN均衡器。下一步,我们将讨论

如何获取这些所需数据。

### 3 数据获取

如图3所示,我们采用自研的8×8 MIMO系统进行数据采集,为了保证测试数据的真实性和均衡算法的可靠性,保证采集数据的多样化,我们在多个典型室内通信场景下采集无线收发数据。在采集实验中信号发射端采用正交相移键控(QPSK)调制模式,发送端以100 Mbit/s的码片速率连续发送4 096个QPSK码片200次。采集的数据平均分布于信噪比为-10~20 dB的区间内,且为简单地验证起见,采集过程中我们仅使用单发单收设置。值得注意的是:在特定信噪比下收集的数据将单独构成子数据集,我们的CRNN训练将由每个不同的子数据集单独训练并筛选出



▲图3 数据采集设备与环境

表现最好的模型。具体地,每个子数据集包含11万对收发数据,其中10万对数据用于训练CRNN,其余数据用于对模型进行测试。

## 4 模型训练及测试结果

### 4.1 超参数优化

在第3节中,我们指出通过调整神经网络的连接权重 $w$ ,就可以优化神经网络均衡器对于接收信号的预测结果。除了连接权重这些参数,类似于神经网络隐藏层个数、特定隐藏层的神经元个数、神经网络进行训练时的梯度下降学习率等需要人工试验的参数,我们称之为超参数。因此,均衡器最终的预测效果除了依赖于神经网络的连接权重,也取决于神经网络的各个超参数。而如何确定超参数的值,以及采取何种操作使得神经网络的预测效果最好的过程被称为超参数优化。

在实际工程中,我们通常使用网格搜索法进行超参数优化,也即采用穷搜法逐个试验各个人工设定的超参数组合,从而确定一组最佳超参数组合。本文中,我们将固定一些超参数从而避免搜索空间过大,导致调节效率低下。

除批数量外,表1中选取固定的超参数基本为我们的CRNN均衡器的结构参数。批数量是利用数据集

▼表 1 选取固定的超参数

输入层大小	2 × 12
第 1 层卷积大小	2 × 4 × 64
第 2 层卷积大小	2 × 4 × 32
Max-pooling 大小	2 × 2
LSTM 层大小	100
输出层大小	4
批数量	1 024

LSTM: 长短期记忆网络

进行最小化损失函数的随机梯度下降操作时需要确定的参数,它决定了在每次梯度下降时可同时参与更新权重  $w$  所应选取的数据个数。

选取了固定的超参数之后,我们需要对剩下的超参数进行优化。具体地,我们将优化学习率  $\gamma$ 、Dropout 层神经元连接丢弃比率、训练批次等 3 个超参数。如图 4 所示,在训练过程中交叉熵损失在当学习率  $\gamma=1 \times 10^{-2}$  或  $\gamma=1 \times 10^{-3}$  时下降。虽然不够明显,但  $\gamma=1 \times 10^{-3}$  时交叉熵损失函数收敛到一个更低的值。我们还可以看到:当  $\gamma=1 \times 10^{-1}$  时,交叉熵损失一直处于振荡状态并未明显下降。这说明此学习率过大,对于梯度下降步长过大导致无法找到正确的使损失函数下降的方向。我们将训练集划分为互不相交的训练集和验证集,与之对应得到的准确率分别为训练准确率和验证准确率。验证准确率指示了模型的泛化性能,因为验证集中的数据在训练集中从未遇到过。从而,我们可以利用验证准确率来判断该超参数设置是否真正有效。图 5 表示当我们设置不同的神经元连接丢失比率时,CRNN 均衡器预测准确率的表现,其中实线为训练准确率,虚线为验证准确率。可以看出:当 Dropout 层神经元连接丢弃比率为 0 时,我们得到更高的训练准确率和验证准确率,这说明先前的超参数设置导致的模型训练结果并未过拟合我们的训练集。因此,针对本文实验而言,Dropout 层并无必要加入 CRNN 均衡器结构中。最后一个需要

优化的超参数为训练批次,其定义为训练集被完全迭代的次数,而训练过程遍历完一次训练集称为训练一个批次。当训练批次较大时,不同的训练批次导致的验证准确率差别很小,因此我们使用验证误差来评价超参数的影响。如图 6 所示,训练批次对于准确率的增益是有上限的。那么我们需要找到最小的上限,以防止过度训练。在工程实验中,我们可以使用训练提前停止方法,即监测控制训练批次对于准确率增益的影响,当相邻若干次准确率增长非常小时,我们认为训练批次达到了它对于准确率增益的上限,这时便可以提前终止训练。在我们的实验中,当训练批次为 40 时,验证误差最小,也即验证准确

率最大。另外,需要强调的是:上述超参数优化是在信噪比为 15 dB 的数据集下进行的。我们在线下对比发现:信噪比为 15 dB 时的数据集训练出的 CRNN 均衡器的泛化性能最好,因此我们在其他信噪比下也使用信噪比为 15 dB 的数据集模型。

#### 4.2 误码率性能

误码率(SER)是信道均衡器的一个重要性能指标,我们也实现了 RLS 均衡器和 MLP 均衡器并将它们与我们训练得到的 CRNN 均衡器对比 SER 性能。

图 6 显示了 3 种信道均衡器在信噪比为 -10 ~ 20 dB 时的 SER 性能,显然当信噪比低于 0 dB 时,它们均无法

图 4 ▶  
训练过程中不同  
学习率下的交叉熵  
损失趋势

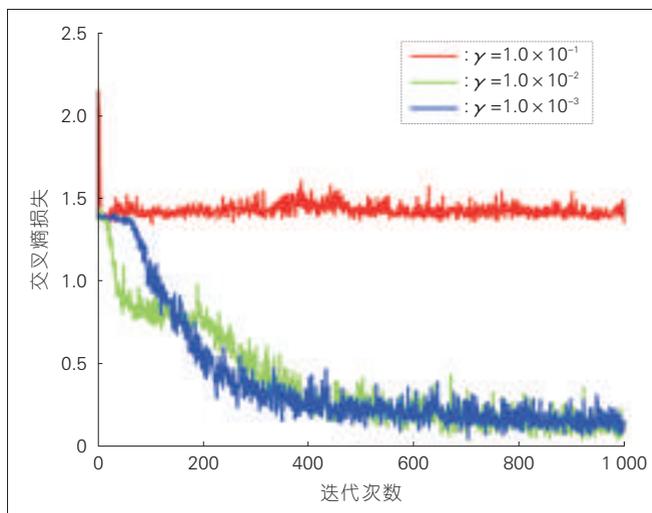
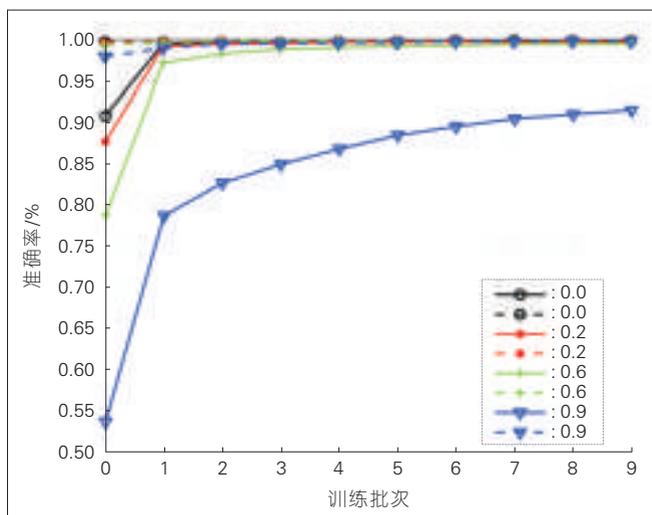


图 5 ▶  
Dropout 层神经元连  
接丢弃比率对训练、  
验证准确率的影响



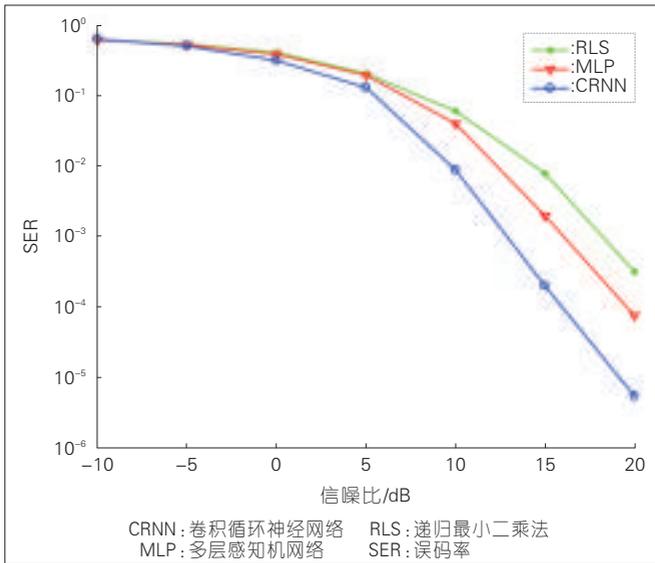


图6  
不同信噪比下CRNN  
与RLS、MLP的SER  
性能对比

符合性能要求,此时CRNN基本只能从噪声中学习而无法学习到有效的信道特征。当信噪比在0~10 dB时,CRNN均衡器在达到相同SER情况下要比RLS、MLP均衡器平均高2 dB。而当信噪比在10~15 dB时,CRNN均衡器在达到相同SER情况下比另外两个均衡器平均高4 dB,这意味着CRNN在信号特征明显时能比传统均衡器学习到更多的信道特征。

## 5 结束语

本文中我们提出了一种基于CRNN的算法,来解决非线性时变信道的均衡问题,并在实际系统中进行了验证。具体地,我们使用卷积神经网络来学习匹配滤波器,然后将学习到的滤波器送入后续的循环神经网络进行时域建模,最后对信号进行分类。实验表明:相比RLS、MLP算法,CRNN均衡算法能够更好地处理真实通信环境中的无线信道均衡问题。

### 参考文献

[1] LYU X, FENG W, SHI R, et al. Artificial Neural Networkbased Nonlinear Channel

- Equalization: A Soft-Output Perspective[C]// Proceedings of the International Conference on Telecommunications (ICT). USA:IEEE, 2015: 243-248.DOI: 10.1109/ICT.2015.7124690
- [2] WANG Y. Channel Equalization Using a Robust Recursive Least-Squares Adaptive-Filtering Algorithm[C]//Proceedings of the IEEE International Conference on Computer and Information Technology (ICCI). USA: IEEE, 2012: 135-138.DOI: 10.1109/CIT.2012.49
- [3] BURSE K, YADAY R N, SHRIVASTAVA S C. Channel Equalization Using Neural Networks: A Review[J].IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),2010,40(3): 352-357.DOI: 10.1109/TSMCC.2009.2038279
- [4] PATRA J C, BENG POH W, CHAUDHARI N S, et al. Nonlinear Channel Equalization with QAM Signal Using Chebyshev Artificial Neural Network[C]// Proceedings of the IEEE International Joint Conference on Neural Networks. USA:IEEE, 2005,(5):3214-3219. DOI: 10.1109/IJCNN.2005.1556442
- [5] CHANG P, WANG B. Adaptive Decision Feedback Equalization for Digital Satellite Channels Using Multilayer Neural Networks [J].IEEE Journal on Selected Areas in Communications, 1995,13(2): 316-324
- [6] HUANG R C, CHEN M S. Adaptive Equalization Using Complex-Valued Multilayered Neural Network Based on the Extended Kalman Filter[C]// Proceedings of the International Conference on Signal Processing. USA:IEEE, 2000:519-524
- [7] LI Y, CHEN M H, YANG Y, M.T.et al. Convolutional Recurrent Neural Network-Based Channel Equalization: An Experimental

Study[C]// Proceedings of the Asia-Pacific Conference on Communications (APCC). USA:IEEE, 2017. DOI: 10.23919/APCC.2017.8304090

- [8] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep Learning[M]. USA:MIT Press, 2016
- [9] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-Based Learning Applied to Document Recognition[C]//Proceedings of the IEEE. USA:IEEE, 1998, 86(11):2278-2324
- [10] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780

## 作者简介



杨旸,中国科学院上海微系统与信息技术研究所研究员、学位委员会副主任,中国科学院无线传感网与通信重点实验室主任,上海无线通信研究中心主任,上海雾计算实验室联合主任,国际雾计算产学研联盟大中华区主任,中国通信学会学术工作委员会委员等;研究领域包括无线传感器网络(物联网)、新一代移动通信系统(5G)、雾计算与网络技术、开放无线测试验证平台等;承担了国家科技重大专项(03专项)、国家“863”计划、国家自然科学基金重点等一系列前沿课题研究;获选IEEE Fellow、中组部“万人计划”、上海市“千人计划”、中科院“百人计划”、上海市“优秀学术带头人”和“领军人才”等荣誉;发表论文150余篇,申报专利80余项。



李扬,中国科学院上海微系统与信息技术研究所博士研究生在读;主要研究方向为机器学习/深度学习在无线通信系统中的应用、强化学习、基于Wi-Fi的室内定位、数据分析等。



周明拓,上海无线通信研究中心计算通信部部长,曾在包括ICC、GLOBECOM、PIMRC在内的40多个国际会议担任技术计划委员会成员、财务主席、当地安排主席和会议主席等;研究方向包括雾计算、5G、无线传感器网络、工业物联网、工业4.0和网络物理系统等;发表论文70余篇,出版书籍6部。

# mMTC网络中基于空口流量的入侵检测

## Air-Interface Traffic-Load Based Intrusion Detection over mMTC Networks

卢楠/LU Nan  
杜清河/DU Qinghe  
任品毅/REN Pinyi

(西安交通大学, 陕西 西安 710049)  
(Xi'an Jiaotong University, Xi'an  
710049, China)

机器类通信(MTC)是指不需要或只需要很少的人工干预下机器之间的通信。海量机器类通信(mMTC)面向物联网低成本节点泛在信息交换,成为5G通信系统的主要场景之一。目前,已经有50亿机器通信终端连接无线网络,到2020年这个数字预计将达到500亿<sup>[1]</sup>。机器通信网络有众多应用,如:自动驾驶、智慧医疗、智能测量、家居管理、智慧城市<sup>[2]</sup>。

安全性保障是大规模机器网络的重要任务之一。然而,大规模机器通信网络往往要求节点具有低成本特性,这一要求也削弱了节点的安全防护能力。安全问题有可能阻碍机器通信的发展甚至危害机器通信的各种应用。所以,机器通信中的安全性问题目前已经吸引了越来越多的研究<sup>[2-3]</sup>。大规模机器通信网络中的安全问题可以分为以下几类:物理攻击、配置攻击、协议攻击。机器通信网络中的安全保障机制也可以在网

收稿日期:2018-01-13  
网络出版日期:2018-03-22

基金项目:国家自然科学基金(61431011、61671371);陕西省重点研发计划重点项目(2017ZDXM-GY-012);中央高校基本科研业务费专项资金

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868(2018)02-0030-08

**摘要:** 提出基于空口负载特征学习的入侵检测体系与方法。基站通过分析海量机器类通信(mMTC)节点随机接入过程中的空口信号,可以智能化学习接入负载特征。在此基础上,结合常态流量负载统计信息,设计了入侵攻击检测的框架与实时检测方案。分析与仿真结果表明所提方法可以较准确地跟踪接入负载变化。与基准方案相比,可获得较高的检测概率和较短的检测时间。方案不依赖于高层安全协议,可基于底层信号实现快速入侵检测,为未来的物联网(IoT)安全防护提供了新思路与参考方案。

**关键词:** 入侵检测; MTC网络; 随机接入; 最大似然检测

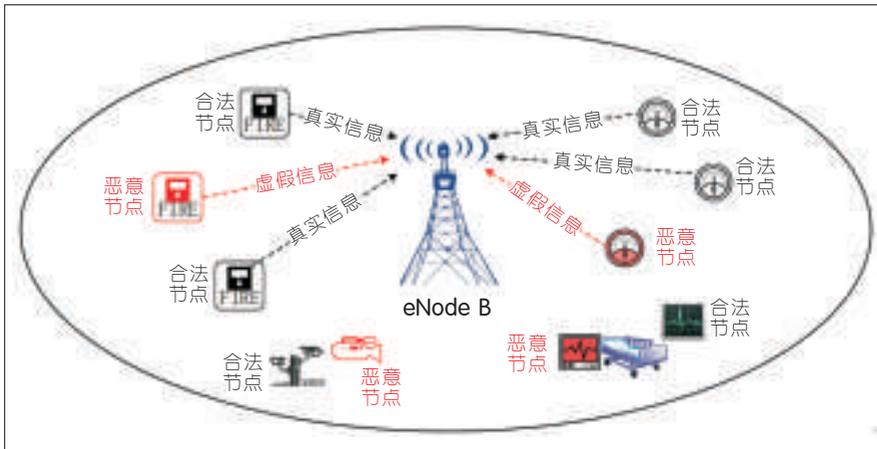
**Abstract:** In this paper, an air-interface traffic-load based intrusion detection approach is proposed. The base station can intelligently learn the traffic-load features by analyzing the air-interference signal in the massive machine type communications (mMTC) nodes' random access procedure. With the help of the statistic information under the normal case, the framework of intrusion and attack detection for massive machine type communications (MTC) networks is established and a real time detection scheme is designed. The performance analysis and simulation results demonstrate that our scheme can well track the arrival process with high accuracy, and outperform the baseline schemes in terms of the detection probability and the detection time. Our low layer signal based approach can make an agile intrusion detection and does not depend on security protocol applied on the high layer, which provides novel thinking and a reference scheme for the security enhancement in future Internet of things (IoT).

**Key words:** intrusion detection; MTC networks; random access; maximum likelihood detection

络协议栈的各层进行配置,例如:鉴权机制、加密算法、安全路由协议等。在mMTC网络中,入侵攻击是一种典型的安全隐患。如图1所示,在某些情况下,机器通信网络可能被一些恶意终端入侵,这些恶意终端接入后可以像正常的终端一样传递信息。恶意终端不会以瘫痪网络为目的进行攻击,而是在隐藏自己的同时发送虚假的感知信息,以达到其他恶意目的。目前,通信领域已经开发了一些手段来防止恶意终端入侵网

络。其中,第3代合作伙伴计划(3GPP)推进了先进包交换鉴权密钥协议的标准化<sup>[4]</sup>。此外,也有大量基于鉴权和密钥协议来保障网络安全性的研究<sup>[5-6]</sup>。

上述大部分机制在网络上层增强安全性,这需要一定的信号开销和复杂的信任管理机制。所以,它们难以有效适应面向5G的大规模低成本机器通信网络的发展。作为上层安全机制的补充,我们通过研究发现可通过观测空口流量负载变化进行入



▲图1 系统模型

入侵检测。根据这一思路,我们提出了一种基于空口流量负载学习的入侵检测方法。本方法利用入侵攻击和正常状态网络流量的差异性来判断是否有入侵发生。已经有文献利用空口流量来研究小区网络中流量特征分类问题以提高服务可见性<sup>[7]</sup>,或者设计自适应算法改变随机接入的有关参数以均衡网络负载<sup>[8]</sup>;而尚未有利用空口流量进行安全入侵攻击检测的手段。

为了提升大规模机器通信网络的安全性,我们给出了 mMTC 网络的入侵模型并设计了入侵检测方法。本方法不同于传统上层安全保障机制,而是利用了空口流量负载特征进行入侵攻击发现。本文方法包含 2 个部分:首先,基于机器类型终端随机接入过程中碰撞与成功状态进行流量负载估计;在此基础上,结合流量负载规律建立了入侵检测框架及实时检测方法。本方法可以作为现有安全协议的补充。同时,该方案利用底层信息,不会造成新的信令开销,并能有效地减小攻击发现时间,有望为未来物联网(IoT)安全入侵防护提供理论基础及参考方案。

## 1 mMTC 系统模型

我们考虑如下的 mMTC 系统,该系统由一个基站和在其覆盖范围内的大量 MTC 终端组成。正常情况下,

所有的 MTC 终端都是注册过的合法用户,其数量用  $N$  表示;但是在某些时候,一定数量的恶意 MTC 终端会在未经许可的情况下进入该网络,并且成功通过了鉴权机制。这些恶意终端会发送错误信息来扰乱 IoT 系统的常规运转,或者占用系统的时频资源,比如物理上行/下行共享信道(PUSCH/PDSCH),从而进一步达到不法的恶意目的。为了更好地隐蔽自己,恶意终端不会采取强烈或者易被察觉的攻击,比如:拒绝服务(DoS)攻击。当恶意终端的数量相对较小时,mMTC 网络不会被明显影响到;反之,虚假或错误信息传播形成规模,造成极大危害。我们定义可以接受的最大恶意终端的数量为  $N_1$ ,如果恶意终端的数量超过  $N_1$ ,则认为 mMTC 网络发生了入侵行为。我们用  $H_0$  和  $H_1$  分别代表假设:入侵未发生和发生。mMTC 网络中的合法 MTC 终端与恶意 MTC 终端数量的和定义为  $N_0$ 。那么,我们的检测问题可以描述为:

$$\begin{cases} H_0: N_0 \leq N + N_1 \\ H_1: N_0 > N + N_1 \end{cases} \quad (1)$$

目前,5G mMTC 网络仍然处于标准化的初始阶段,具体协议尚未确定。因此,本文中我们暂时遵循长期演进(LTE)网络规范<sup>[10]</sup>。LTE 网络中的机器类型网络规范的核心特征包

括 Beta 分布到达模型、访问类别限制(ACB)机制,及 4 次握手接入协议。

### 1.1 流量模型

在网络中有  $N$  个注册过的合法 MTC 终端。文献[9]给出了两种流量模型:模型 1 可以视为 MTC 终端在一段时间内均匀地接入网络,比如非同步模式;模型 2 可以视为大量 MTC 终端以高度同步的模式接入网络,比如一次断电后的接入。考虑到网络流量的突发性,我们采用文献[9]的模型 2 来描述本网络中的合法 MTC 终端的到达过程。具体说来,MTC 终端在  $t$  时刻发送接入请求的数量满足概率密度函数  $g(t)$ ,其中  $g(t)$  服从 Beta 分布,如下:

$$g(t) = \frac{t^{\alpha-1}(T-t)^{\beta-1}}{T_A^{\alpha+\beta-1} \text{Beta}(\alpha, \beta)} \quad (2)$$

$$\alpha > 0, \beta > 0, 0 \leq t \leq T_A$$

其中,  $T_A$  是时间长度,  $\text{Beta}(\alpha, \beta)$  是 Beta 函数<sup>[10]</sup>,对  $g(t)$  在时间上积分可以求出在第  $i$  次接入中的到达终端数  $A_{[i]}$ ,下标“ $[i]$ ”表示第  $i$  个时隙。在我们的模型中,恶意 MTC 终端以一种最隐蔽的方式存在于网络中。也就是说,它们有和合法终端同样的到达过程、时间起点和接入过程。

### 1.2 接入控制

在本系统模型中,时间划分为时隙,每个时隙由下标“ $[i]$ ”索引,且 MTC 终端遵循 LTE 网络中的 ACB 机制<sup>[11]</sup>。在每个时隙开始前,eNodeB 广播 ACB 因子  $p$ 。在每一个时隙,每个准备接入的 MTC 终端生成一个 0 和 1 之间的随机数  $q$ 。如果  $q$  小于  $p$ ,该终端则通过 ACB 过程,进入基于竞争的随机接入。否则,该终端退避一段时间,时间长度为随机变量  $T_1$ ,由公式(3)给出:

$$T_1 = (a_0 + b_0 \times \text{rand}) \times T_0 \quad (3)$$

其中,  $\text{rand}$  表示在区间  $[0, 1]$  中产生的均匀随机数,  $a_0$  和  $b_0$  是正实数,  $T_0$

是退避时间参数。在  $T_1$  S后,被退避的终端重新开始ACB过程。我们定义在第  $i$  次接入机会时的准备接入终端数为  $D_{[i]}$ ,它是在该时刻新到达的终端数、被退避至该时刻的终端数和在上一次接入中被碰撞的终端数之和。另外,在第  $i$  次接入机会时,通过ACB过程的终端数为  $M_{[i]}$ 。

图2给出了入侵检测方案的框图,图2的上半部分是接入控制和随机接入过程的示意图,下半部分是入侵检测过程的示意图,图中各符号省略了下角标。

### 1.3 随机接入过程

所有通过ACB过程的终端都需经过随机接入过程来传输它们的数据。实际中有2种接入模式:适用于高优先级终端的非竞争模式和适用于普通终端的时隙化竞争模式。在本文中,我们考虑采用竞争模式,其适用于存在大量普通终端的一般物联网。LTE网络下的竞争接入模式包含4个阶段<sup>[12]</sup>:第1阶段,每个终端从所有可选导频信号(Preamble)中随机选择一个,并在当前时隙通过物理随机接入信道发送该导频(在网络中,假设一共有  $K$  个可用导频信号,它们之间两两正交,典型的持续时间为  $1\text{ ms}$ );第2阶段,eNodeB对每个被选择导频进行回应,发送随机接入响应消息(RAR),每个RAR包含对应于某一导频的资源块分配命令;第3阶段,每个终端根据自己在第1阶段

中发送的导频检索RAR中信息,并在得到的物理上行共享信道上传输连接请求信息;第4阶段,eNodeB向数据包被成功解码的终端发送竞争解决方案消息。

对于在第1阶段选择相同导频信号的终端,其传输可能发生碰撞;但由于发射的信号相同,基站通常也可能正确接收。然而,即使在第1阶段不发生碰撞,在第3阶段用户发送连接请求数据的时候,传输的数据包在同一资源块并且不同用户的信号不同。此时碰撞不可避免,不能被eNodeB成功解码。这些导频碰撞的终端将在下一次随机接入机会时从ACB过程开始它们的接入过程。对于每一个终端,随机接入机会每  $T_m$  秒出现一次,通常  $T_m$  为  $0.005$ 。

在本文中,为了便于分析并更好地关注如何入侵检测,我们采用简化的握手传输模型,即如果用户选择同一导频,那么则假设会发生碰撞。发生碰撞的用户在下一个接入机会开始时仍可进行接入竞争。这里需要指出:恶意终端因为需要尽可能伪装成合法节点而仅散播虚假或错误信息,所以它们也会遵循协议的随机接入与退避策略,从而避免基站很容易发现它们的非常规接入行为。

## 2 基于空口流量的入侵检测方案

我们的目标是估计网络中MTC终端的到达过程,并由此判断是否出

现了异常流量,也就是MTC网络是否被入侵。如图2所示,本入侵检测方案包括两部分:系统状态估计即到达过程的估计和实时的入侵检测判决。下面分别描述这两部分。

### 2.1 到达过程的估计算法

我们假设在每次随机接入过程中,eNodeB知道空闲导频、只被一个MTC终端占用的导频和被多个MTC终端选择的导频的数量。这3类导频的数量分别定义为  $A, B, C$ 。假设基站掌握这3类导频数量的合理性在于:对于仅有只被一个MTC终端占用的导频,基站可以正确检测到并统计数量;对于被多个MTC终端选择同一导频而发生碰撞的情况,基站可以检测到较强的信号能量但不能正确译码请求数据包,从而可以区分这一类导频并统计数量;对于空闲导频,基站将仅观测到很低的能量,进而也可区分这类导频并统计数量。我们将导频状态向量  $S$  定义为  $(A, B, C)$ 。 $M$  的最大后验(MAP)估计可以由公式(4)得到:

$$\hat{M} = \arg \max_{0 \leq m \leq N} \{ \Pr(M=m | A=a, B=b, C=c) \} \quad (4)$$

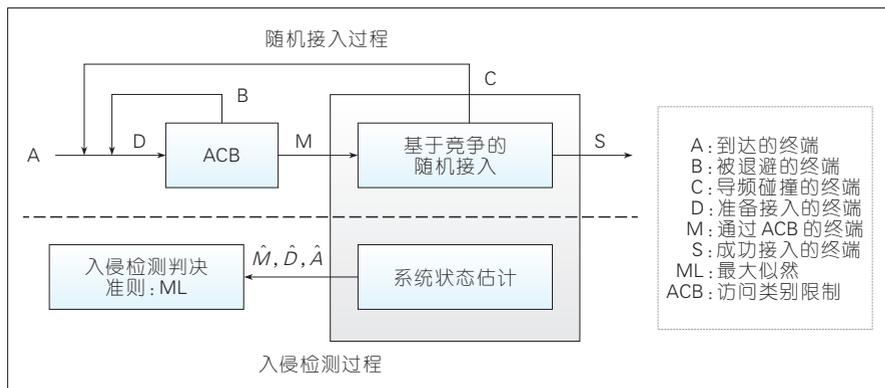
其中,  $a, b, c = 0, 1, \dots, K$ 。在用户独立选择的条件下,最大后验估计退化为最大似然(ML)估计。 $M$  的最大似然估计由公式(5)给出:

$$\hat{M} = \arg \max_{0 \leq m \leq N} \{ \Pr(A=a, B=b, C=c | M=m) \} \quad (5)$$

在  $A+B+C=K$  的约束下,  $S$  一共有  $(K+1)(K+2)/2$  种可能。对于  $(A, B, C)$ , 它的状态索引定义为:

$$n = \frac{1}{2}(2K-A+3)A+B+1 \quad (6)$$

我们可以假设每次接入竞争机会中用户的选择是独立的,但是顺序发生。这样,我们可以采用马尔科夫链模型来完成最大似然估计中条件概率的计算。具体的概率转移矩阵及相关计算参见我们的前期工作<sup>[13]</sup>。为了本文的完整性,我们将过程做如下简述。每次接入机会中用户逐个



▲图2 入侵检测方案框图

选择导频的概率转移矩阵定义为  $P$ 。其中,  $P_{n_1, n_2}$  代表  $S$  从  $n_1$  状态转移到  $n_2$  状态的概率。对于当前状态  $n_1$  (其状态矢量描述记为  $(A, \mathcal{B}, C)$ ) 在添加一个导频后, 转移为状态  $(A-1, \mathcal{B}+1, C)$ ,  $(A, \mathcal{B}, C)$ ,  $(A+1, \mathcal{B}, C-1)$  的概率分别为  $A/K$ ,  $\mathcal{B}/K$ ,  $C/K$ 。因此有:

$$P_{n_1, n_2} = \begin{cases} \frac{A}{K}, & \text{if } n_2 = \frac{1}{2}(2K-A+4)(A-1) + \mathcal{B} + 2 \\ \frac{\mathcal{B}}{K}, & \text{if } n_2 = \frac{1}{2}(2K-A+3)A + \mathcal{B} + 1 \\ \frac{C}{K}, & \text{if } n_2 = \frac{1}{2}(2K-A+2)(A+1) + \mathcal{B} + 1 \\ 0, & \text{others} \end{cases} \quad (7)$$

构造出概率转移矩阵  $P$  后, 我们可以得到第  $m$  步后的概率分布向量  $\pi(m)$  为:

$$\pi(m) = \pi(0)P^m \quad (8)$$

其中,  $\pi(0)$  是初始概率的分布向量。对于第  $n$  种状态, 公式(5)可以写为:

$$\hat{M} = \arg \max_{0 \leq m \leq N} \{\pi(m)_n\} \quad (9)$$

其中, 下标  $n$  代表向量的第  $n$  个元素。由公式(9)可以求出  $M$  的最大似然估计。 $M$  的取值范围是  $[0, N]$ , 但在实际系统中, 遍历所有的可能性会引入较大的复杂度。因此, 在本文仿真中, 我们设定一个  $M$  的上限  $M_{\max}$ , 仅在  $[0, M_{\max}]$  范围内考察负载大小。这里  $M_{\max}$  的取值远远大于每一次竞争的节点平均数目, 因此不会对方案性能产生显著影响。

上述方法完成了对于通过 ACB 的终端数的估计, 下面我们需要对网络内的到达节点数进行估计。如果网络中目前有  $D$  个准备接入的 MTC 终端, 它们首先需要进行 ACB 过程, 通过 ACB 过程的 MTC 终端数为  $M$ ,  $M$  服从二项分布, 试验次数为  $D$ , 概率为  $p$ 。那么, 已知  $\hat{M}$  后,  $D$  的最大似然估计值为:

$$\hat{D} = \frac{\hat{M}}{p} \quad (10)$$

在第  $i$  个时隙, 准备接入的 MTC 终端由 3 部分组成: 由公式(2)定义的新到达的 MTC 终端, 在  $i-1$  时隙碰

撞的 MTC 终端和退避到第  $i$  个时隙的 MTC 终端。所以, 第  $i$  个时隙的新到达 MTC 终端的数量可以由式(11)计算:

$$\hat{A}_{[i]} = \hat{D}_{[i]} - \bar{B}_{[i]} - \hat{C}_{[i-1]} \quad (11)$$

其中,

$$\hat{C}_{[i-1]} = \hat{M}_{[i-1]} - \mathcal{B}_{[i-1]} \quad (12)$$

在公式(12)中,  $\mathcal{B}_{[i-1]}$  是  $i-1$  时隙的成功接入 MTC 终端的数量, 它等于随机接入过程中第 4 步竞争释放消息的数量。另外,  $\bar{B}_{[i]}$  是  $B_{[i]}$  的期望, 由公式(3)我们得出:

$$\bar{B}_{[i]} = \begin{cases} 0, & i \leq l \\ \frac{1}{r} \sum_{s=1}^{i-l} D_{[s]} (1-p_{[s]}), & l < i \leq l+r \\ \frac{1}{r} \sum_{s=1}^r D_{[i-l-s]} (1-p_{[i-l-s]}), & i > l+r \end{cases} \quad (13)$$

其中,  $l = a_0 T_0 / T_m$ ,  $r = b_0 T_0 / T_m$ 。

## 2.2 基于最大似然准则的实时检测算法

我们根据得到的当前时隙及过去时隙的到达节点数的估计值, 基于最大似然准则估计网络内的总节点数并做出判决。在我们的实时入侵检测算法中, 我们每  $\lambda$  个时隙做一次判决, 比如当  $\lambda = 20$ , 随机接入信道 (RACH) 的周期是 5 ms, 我们每 0.1 s 做一次判决。 $\hat{A}_k$  表示在第  $k-1$  次判决和第  $k$  次判决间的到达 MTC 终端的数量, 其中, 下标 “ $k$ ” 表示第  $k$  次判决。对于第  $k$  次判决, 我们已经得到了过去所有  $k \times \lambda$  个时隙的到达终端的估计值。我们的流量模型是一个概率密度函数遵循 Beta 分布的随机到达过程。在过去的  $k \times \lambda$  个时隙中, 总的到达 MTC 终端数量服从二项分布, 二项分布的参数为  $N_0$  和  $\tilde{g}_k$ ,  $\tilde{g}_k$  代表  $g(t)$  的累计分布函数, 如公式(14):

$$\tilde{g}_k = \int_{0_0}^{t_0} g(t) dt \quad (14)$$

至此, 我们可以得到网络内的总节点数  $N_0$  的最大似然估计为:

$$\hat{N}_0 = \frac{\sum_{z=1}^k \hat{A}_z}{\tilde{g}_k} \quad (15)$$

如果  $\hat{N}_0$  大于合法终端数和允许的最大恶意终端数之和, eNodeB 会发送一个警告信息; 否则, eNodeB 会发送一个空白信息, 我们则认为目前的 MTC 网络处于正常情况。我们得到入侵检测问题的最大似然判决如下:

$$\hat{N}_0 \stackrel{H_1}{\geq} N + N_1 \quad (16)$$

当累积的终端到达数量较小时, 到达过程的随机性可能导致较高的错误检测概率。为了降低错误检测概率, 我们设置  $k_0$  作为启动最大似然判决的门限。  $k_0$  时刻需要满足错误检测概率小于  $\delta$  的条件,  $\delta$  是一个很小的概率值。第  $k$  次判决时的错误检测概率  $P_{F_i}$  由式(17)给出:

$$P_{F_i} = \sum_{x=\tilde{g}_k(N+N_1)}^N \binom{N}{x} \tilde{g}_k^x (1-\tilde{g}_k)^{N-x} \quad (17)$$

可以利用 De Moivre-Laplace 定理化简公式(17)。De Moivre-Laplace 定理是中心极限定理的特殊形式, 它指出在一定条件下正态分布可以作为二项分布的近似。因此可以将公式(17)写为:

$$P_{F_i} \approx Q\left(\frac{N\tilde{g}_k}{\sqrt{N\tilde{g}_k(1-\tilde{g}_k)}}\right) - Q\left(\frac{N(1-\tilde{g}_k)}{\sqrt{N\tilde{g}_k(1-\tilde{g}_k)}}\right) \quad (18)$$

其中,  $Q(\cdot)$  是 Q 函数。给定  $N$ ,  $N_1$  和  $\delta$ , 利用式(18)可以由二分法求出  $\tilde{g}_k$  和  $k_0$ 。假定  $N = 30000$ ,  $N_1/N = 5\%$ , 当  $\tilde{g}_k = 3\%, 5\%, 10\%$  时, 可以得到  $P_{F_i} = 6.39\%, 2.35\%, 0.19\%$ 。当启动最大似然判决时, 累积的到达节点数需要大于  $N\tilde{g}_k$ 。

## 3 入侵检测方案性能分析

### 3.1 估计算法的跟踪性能分析

2.1 节所述到达过程的估计算法第 1 步是依据最大似然准则估计  $M$ 。定义  $M$  的误差为:  $\Delta M = M - \hat{M}$ 。在不引起歧义的情况下, 我们省略下

标“ $[i]$ ”，当涉及下标“ $[i-1]$ ”时，则不会省略角标。对于  $M$  个通过 ACB 的终端随机选择导频，导频状态概率分布向量为： $\pi(M)$ ，若  $\pi(M)$  中第  $k$  个元素不为零，则对应的第  $k$  个导频状态的概率为  $[\pi(M)]_k$ 。对于第  $k$  个导频状态，它的最大似然估计为：

$$\hat{M}_k = \arg \max_{0 \leq m \leq N} \{[\pi(m)]_k\} \quad (19)$$

则  $M$  的估计误差的期望值  $E\{\Delta M\}$  可以写为：

$$E\{\Delta M\} = \sum_{k=1}^{(K+1)(K+2)/2} [\pi(M)]_k * (\hat{M}_k - M) \quad (20)$$

其中， $E\{\cdot\}$  表示期望。由式(20)，可以计算不同  $M$  下的  $E\{\Delta M\}$ 。

2.1 节所述到达过程的估计算法第 2 步是根据式(11)计算  $\hat{A}$ ，定义  $A$  的估计误差为： $\Delta A = A - \hat{A}$ 。由公式(11)可知  $\Delta A$  由  $D$ 、 $B$ 、 $C$  的估计误差构成，且有如下关系：

$$\Delta A = \Delta D - \Delta B - \Delta C \quad (21)$$

其中， $\Delta D = D - \hat{D}$ ， $\Delta B = B - \hat{B}$ ， $\Delta C = C - \hat{C}$ 。 $D$  的估计误差由 2 部分构成： $M$  的估计误差和二项分布参数的最大似然估计值和实际值间的误差。经过化简， $\Delta D$  的期望为：

$$E\{\Delta D\} = \frac{E\{\Delta M\}}{p} \quad (22)$$

$B$  的估计误差是  $B$  的期望值和实际值间的误差，所以有  $\Delta B$  的期望为 0。经过化简， $C$  的估计误差等于上一时隙  $M$  的估计误差，所以我们得到  $\Delta C_{[i]}$  的期望为：

$$E\{\Delta C_{[i]}\} = E\{\Delta M_{[i-1]}\} \quad (23)$$

综合式(21)、(22)和(23)，我们可以得到  $\Delta A$  的期望：

$$E\{\Delta A\} = \frac{E\{\Delta M\}}{p} - E\{\Delta M_{[i-1]}\} \quad (24)$$

公式(24)说明： $A$  的估计误差的期望只和  $M$  及  $M_{[i-1]}$  的估计误差的期望有关，而  $E\{\Delta M\}$  和  $E\{\Delta M_{[i-1]}\}$  可

以由公式(20)得到。

### 3.2 实时检测算法的性能分析

我们引入对比方案 1，利用它可以得出本文方案成功检测概率和错误检测概率的下界。对比方案 1，它和本文方案的区别在于：对比方案 1 中 eNodeB 利用到达过程的估计值只在第 10 秒进行判决。所以，对比方案 1 称为保守的基于空口流量的入侵检测方案（简称为保守空口检测方案）。对于对比方案 1，第 10 秒时  $N_0$  的估值为：

$$\hat{N}_0^{b1} = \sum_{i=1}^{T/T_n} \hat{A}_{[i]} \quad (25)$$

引入对比方案 1 后，我们可以得到对比方案 1 和本文所提方案下成功检测概率  $P_D$  和错误检测概率  $P_F$  的关系为：

$$P_D > P_D^{b1}, P_F > P_F^{b1} \quad (26)$$

其中，上角标  $b1$  表示对比方案 1 下的相应变量。如前所述， $A$  的估计误差的期望可以由公式(24)计算出。但是， $A$  的估计误差难以计算。为了便于分析，我们假设  $N_0$  的估计误差是一个均值为 0、方差为  $\sigma^2$  的高斯随机变量。则  $\hat{N}_0^{b1}$  可以写为：

$$\hat{N}_0^{b1} = N_0 + n, \quad (27)$$

其中， $n \sim Norm(0, \sigma^2)$ ， $Norm(\mu, \nu)$  表示均值为  $\mu$ 、方差为  $\nu$  的高斯随机变量。当  $N_0 = N + N_1 + 0.5\%N$  时，我们可以得到：

$$\frac{\hat{N}_0^{b1}}{N} \sim Norm\left(1.005 + \frac{N_1}{N}, \frac{\sigma^2}{N}\right) \quad (28)$$

对比方案 1 下的成功检测概率为：

$$P_D^{b1} = 1 - Q\left(\frac{0.005}{\sqrt{\sigma^2/N}}\right) \quad (29)$$

同理可得，当  $N_0 = N + 0.5\%N$  时，对比方案 1 下的错误检测概率为：

$$P_F^{b1} = Q\left(\frac{N_1/N - 0.005}{\sqrt{\sigma^2/N}}\right) \quad (30)$$

当  $\sigma^2$  取不同值，可计算出对比方案 1 的成功检测概率和错误检测概率作为本方法成功检测概率和错误检测概率的下界。例如：当  $N = 30000$ ， $N_1/N = 5\%$  时， $\sigma^2$  分别取 2，4，6，则  $P_D^{b1}$  分别为 72.99%，66.75%，63.82%， $P_F^{b1}$  分别为 0%，0%，0.07%。

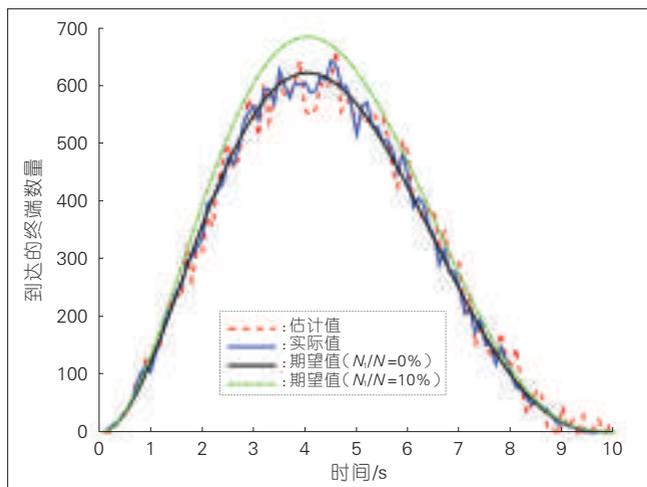
## 4 仿真评估

本节我们利用仿真来评估我们的入侵检测方案的性能。我们采用文献[10]中的仿真参数。仿真中假设在单个小区中有 30000 个 MTC 终端需要进行数据传输，随机接入请求符合 Beta 分布，其中  $\alpha = 3$ ， $\beta = 4$ ， $T = 10s$ 。ACB 过程的参数为  $a_0 = 0.7$ ， $b_0 = 0.6$ ， $T_0 = 4s$ 。物理随机接入信道的配置索引为 6，这意味着随机接入信道每隔 5 ms 出现一次，带宽则为 180 kHz。我们假设一次随机接入中可用的导频总数  $K$  为 54，最大似然判决的平滑因子  $\lambda$  设置为 20。

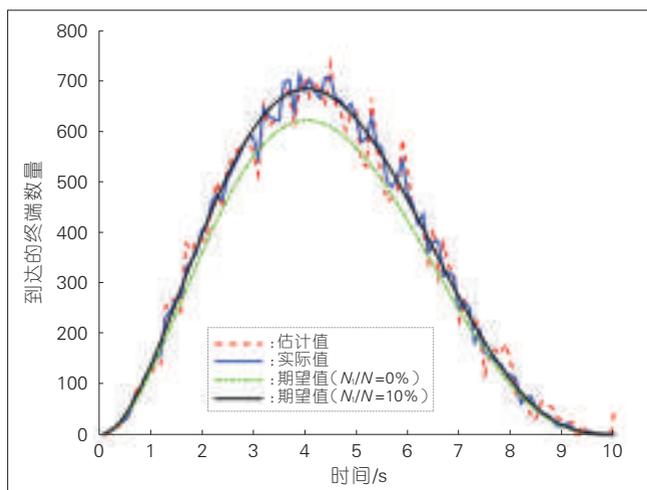
### 4.1 估计算法的跟踪性能

本文方案的第 1 步是基于机器类型终端随机接入过程中碰撞与成功状态进行流量负载估计，所以估计算法的性能对最终的检测效果影响很大。为了评估 2.1 节到达过程的估计算法的性能，我们在图 3 和图 4 中分别给出了在不同情况下到达过程的实际值和估计值。如图 3 所示，正常情况下估计值和实际值之间的误差很小，可见我们的估计算法的跟踪性能是很理想的，这为我们的入侵检测方案实现较低的错误检测概率提供基础。同样地，观察图 4，可以看到我们的估计算法的跟踪性能在入侵发生时也是很理想的，图中的估计值曲线和正常情况下的期望值曲线差异明显，这为我们的入侵检测算法实现较高的检测概率和较短的检测时间提供基础。

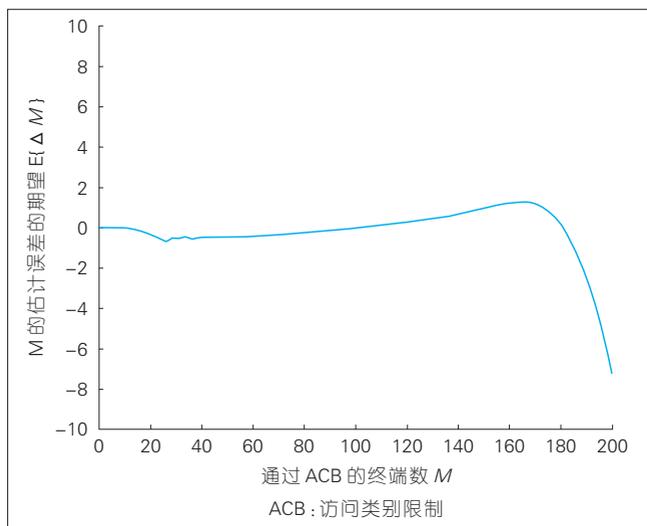
图 5 给出了不同的  $M$  取值下  $M$  的估计误差的期望值  $E\{\Delta M\}$  的变化



◀图3  
正常情况下到达过程的估计情况



◀图4  
入侵时到达过程的估计情况



◀图5  
M的估计误差的期望 E{ΔM} 随 M 的变化情况

情况,可以看出在不同的  $M$  取值下  $E\{\Delta M\}$  近似为 0,说明我们的估计算法在不同的负载情况下都有准确的

估计。注意到,当  $M$  接近 200 时估计误差的期望值较大。这是因为为了降低算法的复杂度,估计算法考察的

每次竞争中的最大节点数为  $M_{\max}$  (参见 2.1 节)。所以,在逼近边界条件的时候,估计的误差较大。

#### 4.2 入侵检测方案的性能

我们比较了本文方案和其他 3 种对比方案下的检测时间、成功检测概率、错误检测概率。3 个对比方案中,eNodeB 利用到达过程的估计值或者导频碰撞概率的观测值在第 10 秒进行判决,具体方案如下:

(1)保守的基于空口流量的入侵检测方案:详见 3.2 节。

(2)基于导频平均碰撞概率的入侵检测方案(简称为平均碰撞概率检测):

$$\bar{\eta} \underset{H_0}{\overset{H_1}{\geq}} \bar{\eta}_1 \quad (31)$$

其中,  $\bar{\eta}$  是 10 s 内观测到的导频平均碰撞概率,  $\bar{\eta}_1$  是当假设  $H_1$  为真时的导频平均碰撞概率。

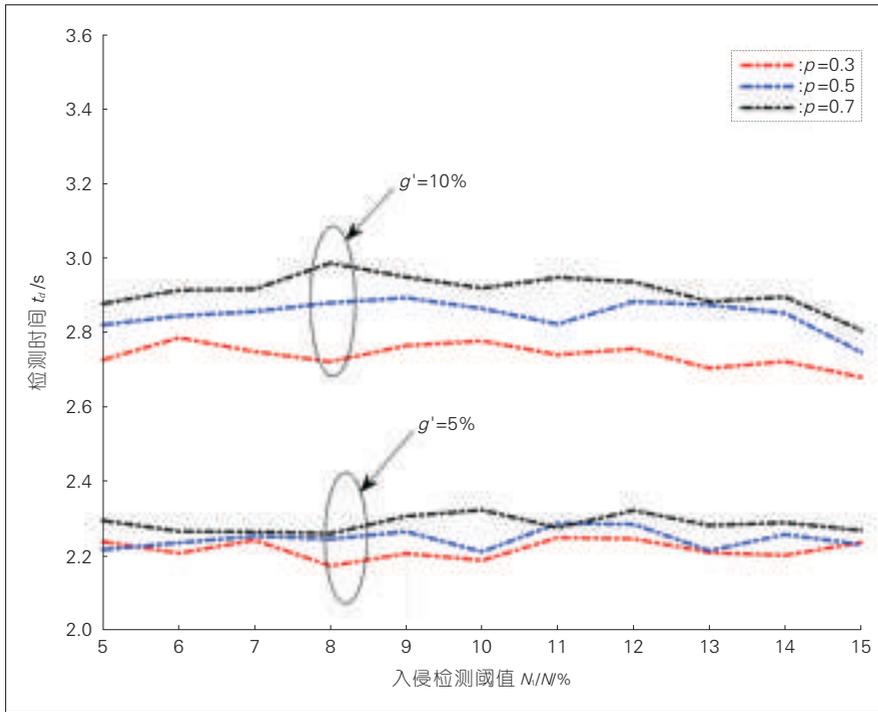
(3)基于导频平均碰撞概率偏差的入侵检测方案(简称为平均碰撞概率偏差检测):

$$\bar{\eta} \underset{H_0}{\overset{H_1}{\geq}} (1 + \varepsilon) \bar{\eta}_0 \quad (32)$$

其中,  $\bar{\eta}_0$  是当假设  $H_0$  为真时的导频平均碰撞概率,  $\varepsilon$  是允许的最大导频碰撞概率偏差因子。

这 3 个对比方案都是在累积了大量的随机接入信息后进行判决,所以它们的检测时间均为 10 s,而本文方案是低于 10 s 的。对比方案 2 和 3 利用了导频碰撞概率,所以它们只能用在 ACB 退避因子  $p$  是常数的情况。值得注意的是:在仿真中,我们假设入侵发生时的 MTC 终端总数  $N_0$  为  $N \times (1 + N_1/N + 0.5\%)$ ,作为所有  $N_0 > N + N_1$  的典型值。另外,我们假设正常情况下的 MTC 终端总数  $N_0$  等于  $N$ ,作为所有  $N_0 < N + N_1$  的典型值。

图 6 给出了检测时间  $t_d$  随系统被入侵的检测阈值(简称为入侵检测阈值)  $N_1/N$  变化的情况,其中退避因子  $p$  取不同值。如图 6 所示,在同一  $\bar{g}_i$  (图中简称为  $g'$ ) 下对于所有的



▲图6 检测时间  $t_d$  随入侵检测阈值  $N_i/N$  的变化情况

$N_i/N$ , 本方案下的检测时间是基本相同的。这是因为对于不同的  $N_i/N$ , 我们设置了相应的恶意终端数量  $N_1$ , 所以不同  $N_i/N$  下检测的难度相当。另外, 也可以看到对于不同的  $p$ , 本方法具有稳定的性能。对于不同  $\tilde{g}_{k_0}$ , 由于启动最大似然判决的时刻不同, 所以平均检测时间不同。如前文所述, 3种对比方案的检测时间为 10 s, 所以在图 6 中省略了。图 7 给出了成功检测概率  $P_D$  随系统被入侵的阈值  $N_i/N$  变化的情况。本文可以实现 95% 以上的成功检测概率。对比方案 1 的检测概率低于本文方案是因为本方案是实时检测, 而对比方案 1 只在第 10 秒检测。对比方案 2 的检测概率较低, 是因为碰撞概率不能很好地描述少量入侵者存在时的空口流量变化。对比方案 3 中, 随着横坐标增大检测概率上升, 这是因为随横坐标增加, 入侵者数量增多, 碰撞概率增加。

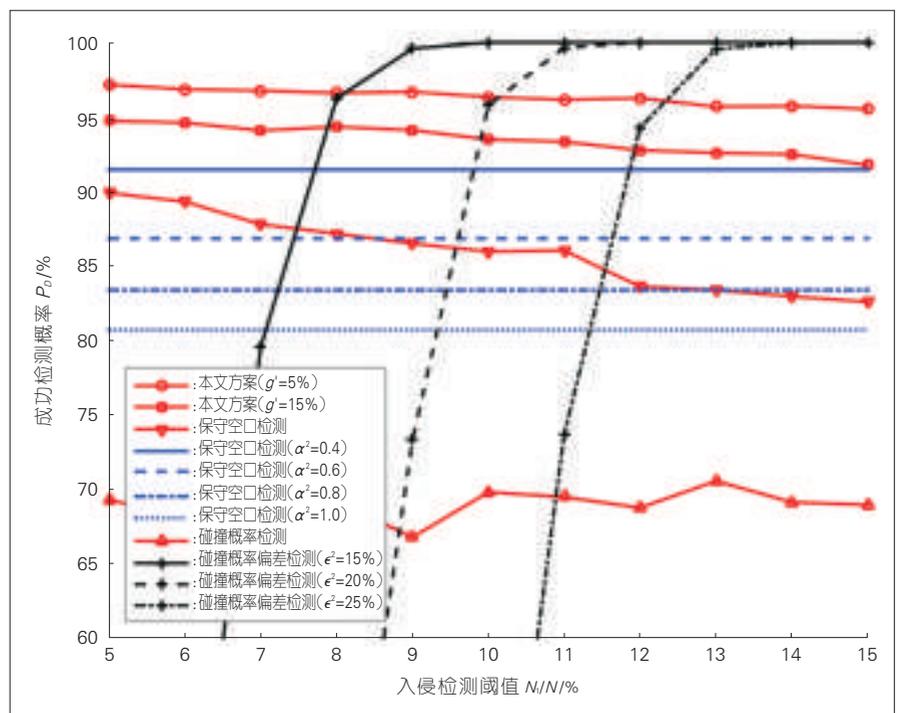
图 8 给出了入侵检测阈值  $N_i/N$  变化时错误检测概率  $P_F$  的变化情况。当  $N_i/N$  增加时,  $P_F$  迅速下降。

这是因为对于固定的  $N_0$ , 更大的入侵检测阈值意味着更松弛的安全要求。为了衡量本方案的稳健性, 我们引入参数  $\delta$  作为系统中 MTC 终端总

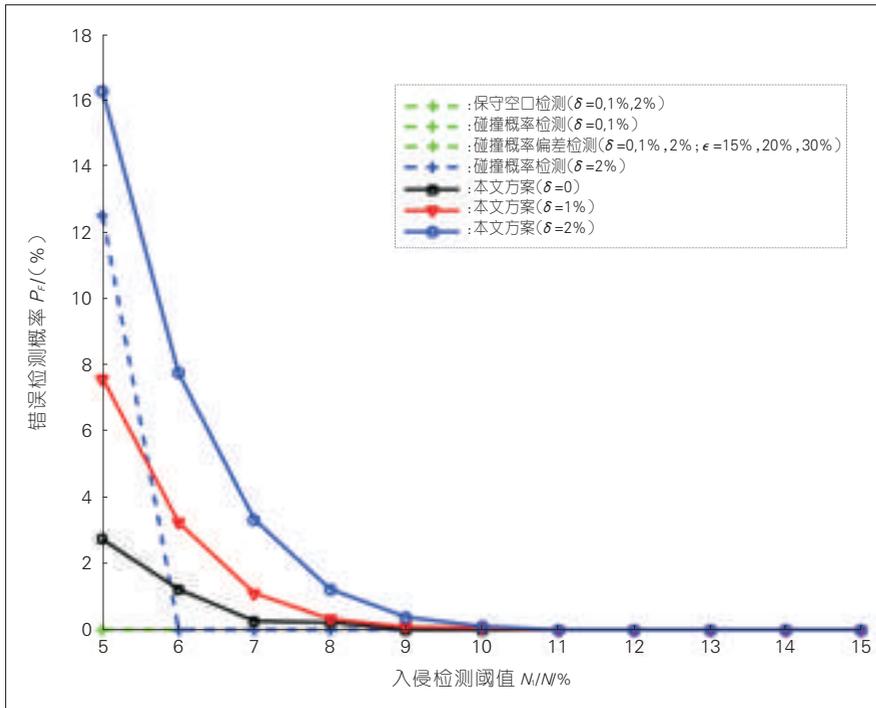
数  $N_0$  的误差系数, 系统中 MTC 终端总数  $N_0$  等于  $(1+\delta)N$ 。当  $\delta=1\%, 2\%$  时,  $P_F$  在大多数  $N_i/N$  下低于 5%, 可见本方案在  $N_0$  有一定误差时的错误检测概率也是较低的。对于少数  $P_F$  大于 5% 的情况, 此时的  $N_i/N$  较小, 例如  $N_i/N=5\%$ , 这意味着网络的安全性要求是较高的。当有一定数量的恶意 MTC 终端出现时, 即使未达到阈值, 此时发出警报也是有益的。从图 8 中还可以看出: 3 个对比方案在大多数参数设置下,  $P_F$  总是 0, 这是因为这 3 个对比方案都很保守, 只在第 10 秒判决, 相应的错误概率会很低。但是, 对比方案 2 在  $\delta=2\%$  时错误检测概率较高, 这是因为它直接利用导频碰撞概率进行判决, 所以对  $N_0$  的误差较敏感。

### 5 结束语

本文提出了基于空口负载特征学习的入侵检测体系与方法。基站通过分析 mMTC 节点随机接入过程中的空口信号, 可以智能化学习接入负载特征。在此基础上, 结合常态流



▲图7 成功检测概率  $P_D$  随入侵检测阈值  $N_i/N$  的变化情况



▲图8 错误检测概率  $P_f$  随入侵检测阈值  $N_i/N$  的变化情况

量负载统计信息,我们设计了入侵攻击检测的框架与实时检测方案。分析与仿真结果表明本文所提方法可以较准确地跟踪接入负载变化。与基准方案相比,可获得较高的检测概率和较短的检测时间。本文方案可以作为现有安全协议的补充,同时不会造成新的信号开销,可以用于低成本 mMTC 终端的智能管理和未来 IoT 安全防护的参考方案。下一步,我们将关注如何整合更多、更深入的信息来服务于 mMTC 和物联网中的入侵检测。

#### 参考文献

- [1] Ericsson. More than 50 Billion Connected Devices[R]. 2011
- [2] BARKI A, BOUABDALLAH A, GHAROUT S, et al. M2M Security: Challenges and Solutions [J]. IEEE Communications Surveys & Tutorials, 2016, 18(2):1241–1254.DOI: 10.1109/COMST.2016.2515516
- [3] CHENG Y, NASLUND M, SELANDER G, et al. Privacy in Machine-to-Machine Communications A State-of-the-Art Survey [C]//2012 IEEE International Conference on Communication Systems (ICCS). USA:IEEE, 2012:75–79.DOI: 10.1109/ICCS.2012.6406112
- [4] LAI C, LU R, ZHENG D, et al. Toward Secure

- Large-scale Machine-to-Machine Communications in 3GPP Networks: Challenges and Solutions[J].IEEE Communications Magazine, 2015, 53(12): 12–19.DOI: 10.1109/MCOM.2015.7355579
- [5] 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; 3GPP System Architecture Evolution (SAE); Security architecture (Rel. 12); 3GPP TS 33.401 V12.13.0[S]. 2014
- [6] LI J, WEN M, ZHANG T. Group-Based Authentication and Key Agreement With Dynamic Policy Updating for MTC in LTE-A Networks [J]. IEEE Internet of Things Journal, 2016, 3(3): 408–417.DOI: 10.1109/JIOT.2015.2495321
- [7] CAO J, LI H, MA M. GAHAP: A Group-Based Anonymity Handover Authentication Protocol for MTC in LTE-A Networks[C]//2015 IEEE International Conference on Communications (ICC). USA:IEEE, 2015: 3020–3025.DOI: 10.1109/ICC.2015.7248787
- [8] LANER M, SVOBODA P, RUPP M. Detecting M2M Traffic in Mobile Cellular Networks[C]//IWSSIP 2014 Proceedings. Croatia, 2014: 159–162
- [9] HE H, DU Q, SONG H, et al. Traffic-Aware ACB Scheme for Massive Access in Machine-to-Machine Networks[C]//2015 IEEE International Conference on Communications (ICC). USA:IEEE, 2015:617–622
- [10] 3rd Generation Partnership Project, Study on RAN Improvements for Machine-type Communications; (Rel. 11)[S]. 3GPP TR 37.868, V1.1.0.0. 2011
- [11] GUPTA A K, Nadarajah S. Handbook of Beta Distribution and Its Applications[J]. Biometrics, 2004, 62 (1): 309–310

- [12] PHUYAL U, KOC A T, FONG M H et al. Controlling Access Overload and Signaling Congestion in M2M Networks[C]//2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, USA:IEEE, 2012:591–595
- [13] 3rd Generation Partnership Project; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification, Third-Generation Partnership Project, Sophia Antipolis Cedex, France[S]. TS 36.321 V12.7.0. 2015
- [14] HE H, DU Q, SONG H, et al. Traffic-Aware ACB Scheme for Massive Access in Machine-to-Machine Networks[C]//2015 IEEE International Conference on Communications (ICC). USA:IEEE, 2015: 617–622.DOI: 10.1109/ICC.2015.7248390
- [15] DAVID T, VISWANATH P. Fundamentals of Wireless Communication[M]. Beijing: Posts & Telecom Press, 2009

#### 作者简介



卢楠,西安交通大学研究生在读;主要研究领域为无线网络中的物理层安全技术;在SCI期刊及IEEE通信领域会议上发表多篇论文;获得授权发明专利1项。



杜清河,西安交通大学副教授;主要研究领域为5G系统关键技术、无线网络物理层安全技术、无线网大数据与机器学习技术、无线多媒体传输技术、无线通信信道建模与仿真技术等;先后主持和参加国家自然科学基金项目、“863”项目、国家科技重大专项项目10余项;获得2项科技成果奖,2007年获IEEE GLOBECOM最佳论文奖,2017年获《China Communications》最佳论文奖;已发表论文100余篇,其中被SCI检索期刊论文50余篇,获授权专利10项。



任毅,西安交通大学教授;主要研究领域为无线网络物理层安全技术、5G系统关键技术、无线网大数据与机器学习技术、无线认知网络、无线自组织网络;先后主持和参加国家自然科学基金项目、“863”项目、国家科技重大专项项目30余项;2010年获IEICE通信协会最佳论文奖,2017年获《China Communications》最佳论文奖,并获得2项科技成果奖;已发表论文150余篇,其中被SCI检索期刊论文60余篇,获授权专利30余项。

# 5G网络人工智能化的基本框架和关键技术

## 5G Networks Based on Artificial Intelligence: Basic Framework and Key Techniques

王威丽/WANG Weili  
何小强/HE Xiaoqiang  
唐伦/TANG Lun

(重庆邮电大学移动通信重点实验室,  
重庆 400065)  
(Key Laboratory of Mobile  
Communication, Chongqing University of  
Posts and Telecommunications,  
Chongqing 400065, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0038-005

**摘要:** 认为未来的5G蜂窝网络需要满足大量服务和应用的不同需求。随着网络配置复杂性的提高及新型业务的不断涌现,5G网络的标准化进程亟需新的技术支持。通过对5G网络新特征的分析,将人工智能与5G网络相结合,设计了一种智能化5G网络框架,该框架由环境舱、智能中心及网络策略3部分组成。同时,归纳了当前人工智能技术在业务预测、网络切片、无线资源分配以及资源共享领域的应用,并总结了实现智能化5G网络的机遇和挑战。

**关键词:** 5G蜂窝网络;人工智能;新特征;智能化框架

**Abstract:** The future 5G cellular network is required to support a large number of services and applications. With the increasing complexity of network configuration and the continuous emergence of new-type services, the standardization process of 5G networks urgently calls for new technical support. Based on the new features of 5G network, an intelligent 5G framework is designed by combining artificial intelligence with 5G network. This framework is composed of three parts: environment capsule, intelligent center and network strategy. Then, the current applications of artificial intelligence technology in traffic forecasting, network slicing, wireless resource allocation and resource sharing are summarized. Finally, the opportunities and challenges of realizing intelligent 5G networks are summarized.

**Keywords:** 5G cellular networks; artificial intelligence; fresh features; intelligent framework

### 1 智能化5G蜂窝网络

第5代移动通信(5G)网络采用了复杂的无线传输技术和无线网络架构,5G将是融合、协同的多制式共存的异构网络。从技术上看,将存在多层、多无线接入技术的共存,导致网络结构非常复杂,各种无线接入技术内部和各种覆盖能力的网络节点之间的关系错综复杂,网络的部署、管理、维护将成为一个极具挑战性的工作。为了降低网络部署、运营维护复杂度和成本,提高网络运维质量,未来5G系统还须具备充分的灵活性,具有网络自感知、自调整等智能化能力,以应对未来移动信息社会难以预计的快速变化。

#### 1.1 5G网络新特征

在2G时代,网络中需要配置的参数为50个,到3G时代,数量增加到100个,而4G时代需要配置的参数

已经到达1 500个。按照这个趋势,5G时代需要配置的参数将可能会达到2 000个<sup>[1]</sup>。然而,当前4G网络的管理和配置依然采用手动/半自动的方式,这种方式不仅成本高且效率低下,给5G网络的性能实现带来了巨大的挑战。因此,为了高效地管理和配置5G网络,自组织性将会成为未来5G网络的重要特征。

随着人们对通信网络依赖性的增加,新的业务类型(如:增强移动宽带、高可靠低时延通信、大规模物联网)不断涌现,5G时代将面临不断改变的业务类型模式带来的挑战<sup>[2]</sup>。在

这种情况下,5G蜂窝网络需要具备自主识别新业务类型、高效的资源调度机制、按需定制相应网络切片的一些功能。

目前,为了使5G系统更加自动化和智慧化,国际电信联盟(ITU)正式成立了“ITU面向包括5G在内的未来网络的‘机器学习’焦点组,重点研究机器学习、人工智能在包含5G系统的未来网络中的应用<sup>[3]</sup>。此外,面向将来的5G网络与服务,为了实现灵活、高效、高质量的管理、服务、运营的“自动化”,欧洲电信标准化协会(ETSI)发布了一份名为《自动化下一

代网络中的网络和服务操作的必要性和益处》的白皮书<sup>[4]</sup>,核心的目标是实现 5G 网络与服务管理、运营的自动化。

## 1.2 人工智能

人工智能的最终目标是建立一个类似于人类思维活动的系统模型。因此,人工智能的实现主要在于构建出来的操作系统能否根据系统的“思维活动”采取理想的行动<sup>[5]</sup>。人工智能领域处理的问题主要包括感知、挖掘、预测以及推理。

- 感知:具有感知能力的智能体对外部环境进行监测;
- 挖掘:对感知到的外部信息进行分类和分析;
- 预测:基于系统经验获得概率模型;
- 推理:具有认知功能的智能体根据外部模型做出推断性结论。

从实践角度来说,设计能够自主思考的系统对 5G 网络的意义重大,例如:5G 网络可以利用感知技术进行网络异常检测以实现网络的自修复,利用挖掘技术对网络业务进行分类分析,利用预测技术预测用户的移动趋势和业务量变化以及利用推理技术配置一系列的参数以更好地适应业务等。人工智能技术不仅具备和环境交互的能力且可以根据不同的环境条件自动执行合理的行动,因此各行各业的人工智能化已经成为社会的必然发展趋势。

## 1.3 5G 网络人工智能化的重要意义

(1) 5G 网络需要通过智能决策,管理种类繁多的资源和动态变化的业务流量。在早期以语音为主的通信网络中,流量模型易于预测,这一时期的流量需求管理也相对较为简单。然而,智能设备的出现使得无线流量模型在维度和粒度上变得更加复杂。为了使网络在面对不同的业务和流量需求时依然能够保证最佳用户体验,策略控制系统需要进行异

常复杂的处理。以网络功能虚拟化为例,必须使其核心决策算法能够自动匹配当前的无线、用户以及流量条件,以实现计算资源的动态分配。而在这方面,人工智能是最佳候选技术,可以为当前的无线系统提供更敏捷和健壮的复杂决策能力。

(2) 5G 网络需要通过自动化,在提高效率的同时降低成本。如何低成本、高效率地运营日益复杂的网络是当前面临的一项重大挑战。目前 2G、3G、4G 网络的信息数据基本上是通过路测、用户投诉记录或操作维护中心(OMC)的报告来获取的,这种数据获取方法效率较低,已经不能满足 5G 网络对于低时延和实时跟踪来提高资源利用率的需求。为了实现 5G 网络自动化,需要对用户域(包括用户分布、用户需求等)、网络域(网络负载、拥塞状态等)和无线域(频谱利用率、链路质量等)的动态网络状况有全面充分的了解。因此,智能感知技术是实现 5G 网络自动化的一个重要条件。

(3) 5G 网络需要根据业务特征按需提供服务。提高网络的资源利用率是满足日益增加的网络业务需求的必要条件。当前的移动网络采用一种网络架构服务所有业务类型的模式。由于网络的单一性,网络中的所有用户只能采用相同的带宽消耗模式,不仅每一类业务类型的特定性能需求无法得到保证,还会大大降低网络资源利用率。5G 网络切片的出现使得用户可以根据特定需求定制针对性服务<sup>[6]</sup>,而为了实现网络切片的灵活调用,切片的创建、部署和管理都将离不开智能化技术。

## 2 智能 5G 网络架构

为了提升当前网络的决策能力,实现网络自动化,并能为特定业务按需建立网络切片,以达到灵活地使用和管理网络的目的,有必要让蜂窝网络具备感知环境的能力,对不确定因素进行学习,设计相应行为,合理地

配置网络。人工智能主要就是解决怎样对变量进行学习,预测未来的状态,通过和环境交互找到潜在解决方案的方法。因此,5G 时代的蜂窝网络可以利用人工智能技术和环境进行交互以改变网络的运作模式,实现智能的 5G 时代。

本文提出的智能 5G 网络架构主要包括 3 个部分:环境舱、智能中心和网络策略。智能 5G 网络架构能够从真实网络环境中读取、观察和分析各类网络信息,同时也能采取相应的行动策略对网络环境加以控制。

### 2.1 环境舱

环境舱是智能 5G 网络架构中唯一可以和真实网络环境直接交流的部分,它主要执行以下两种任务:

(1) 为了维持网络运作秩序,环境舱需要存储从真实环境中观察到的网络状态信息,包括网络拓扑信息、业务请求信息以及资源使用情况等,并负责将这些信息批量传送到智能中心。

(2) 为了保证向智能中心传送的网络信息的实时性,环境舱还需要将网络策略采取的行动信息(包括切片配置、用户关联以及资源分配等)传送到真实网络环境中,以便网络环境实时更新状态信息。

### 2.2 智能中心

智能中心是智能 5G 网络架构的核心部分,智能 5G 网络框架如图 1 所示。通过利用人工智能中的感知、挖掘、预测以及推理方法,对从环境舱中获得的信息进行处理,例如:进行拓扑感知、业务预测以及请求分析等。处理的结果最终会发送给网络策略模块,并做出相应的决策。这些结果可能包括业务分析报告(如资源配置建议)、用户控制信息(如用户服务优先级)以及网络配置通知(如需要调整的参数)等。

以网络切片为例,人工智能利用感知模块获取网络拓扑信息,接着利

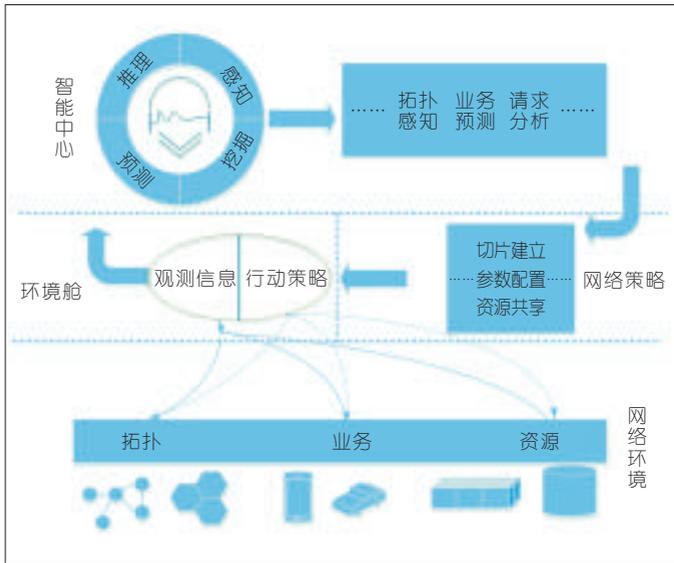


图1  
智能5G网络框架

用挖掘模块对搜集的信息进行分析和处理,并基于历史物理网络节点资源的使用情况对节点的资源需求量进行预测,然后利用推理模块获取网络切片的虚拟网络功能可能的部署方案。

### 2.3 网络策略

根据智能中心传递来的各类信息形成最终的网络策略。由网络策略决定最终采取的各类行动,如给某类请求业务建立切片、网络参数的配置和调整以及虚拟网络映射等。最终的决策行动会经环境舱反馈给真实网络环境,网络环境则会根据最新行动更新其状态信息。

以网络切片的建立为例,网络策略会根据由智能中心传来的业务分析报告来决定分配给每个切片的资源类别以及数量,同时,网络策略会将最终的资源分配策略传递给环境舱,而环境舱会反馈到真实网络中以及及时更新网络资源相关的整体使用情况。

## 3 5G网络智能化的关键技术

这一部分将从4个方面来详细分析当前的无线网络智能化方案,主要包括业务预测、网络切片、无线资源

分配以及资源共享。

### 3.1 业务预测

随着用户业务需求量的增加,业务预测已经成为5G网络监督和管理的关键部分。精确的业务预测需要通过追踪数据流的变化来建立实际的网络业务模型。由于网络中诸多非线性因素的影响,5G网络的业务变化将在时域上呈现较大的不规则性,从而使得传统的线性回归技术已经不再适用于预测和分析当前网络的业务变化情况。基于现有的研究,表1总结了几种可用于智能5G网络的业务预测方法<sup>[7]</sup>:实时方法、时序分析法、反向传播<sup>[8]</sup>及支持向量机,并分析了各自的优点和缺点。

### 3.2 网络切片

当前“一体适用”的网络架构会

表1 智能5G网络的业务预测方法

方法	方法描述	优缺点
实时方法	系统将最新观测到的业务情况作为下一周期的业务量	只适用于各周期间业务波动较平稳的情况
时序分析	用数据报告中的相邻值来表征连续的测量结果	模型泛化能力差,不能可靠地预测观察报告之外的系统状态
反向传播	学习采样由输入层提供,经过中间层的神经元传输到输出层,反向传播算法在隐藏的中间层执行	此方法不仅可以识别非线性模式,且预测精确度较高,可以缩短训练过程所需的时间
支持向量机	优化支持向量找出训练数据中的错误,训练数据由3个参数(流速、体积、密度)组成,用这3个参数来刻画业务流	即使当业务状态变化幅度较大时,此方法依然具有很好的性能。此外,此方法还可以解决非线性问题

造成不同类型业务的需求冲突,从而影响用户体验。网络切片技术的出现,使得5G网络可以根据不同业务的特定需求自行租用共享的物理网络基础设施,用以构建多个逻辑独立的网络。网络切片提供了一个网络即服务(NaaS)模型,能够根据动态的业务和应用需求灵活地分配和重新分配网络资源,从而为不同的5G通信场景定制网络切片。为了实现网络切片的灵活调度,文献[9]将人工智能方法运用到5G网络切片中设计了代理商实体,并且通过3个基本模块的联合调用来提高网络资源利用率。智能5G网络切片框架具体如图2所示。

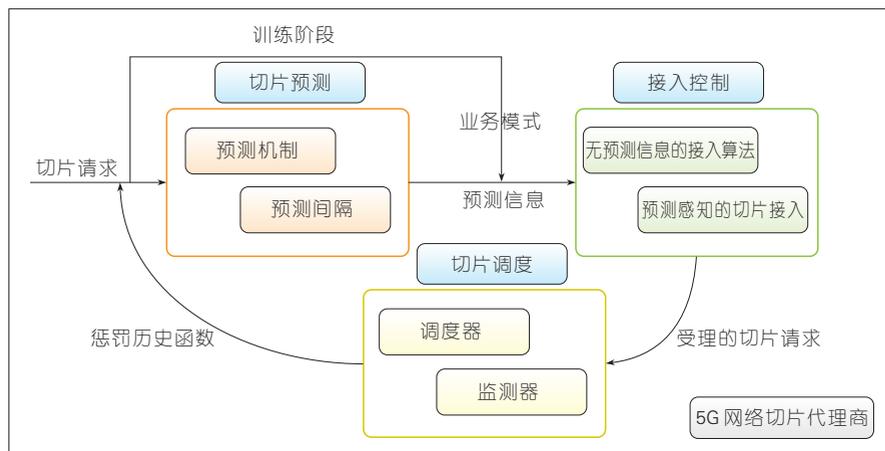
- 预测模块:分析网络切片的业务模式,使用 Holt-Winters 预测方法分析并预测未来的切片业务请求,并将预测信息提供给接入控制模块;

- 接入控制模块:使用接入控制策略来选择下一个时间窗口内可被授权的网络切片请求;

- 切片调度模块:接收来自接入控制模块的受理切片请求列表,并给网络切片分配物理资源;监测(通过使用惩罚历史函数)服务水平协议的违反情况。

### 3.3 无线资源分配

为了使5G网络能够根据不断变化的业务需求为用户进行动态资源分配,多种基于人工智能的算法已经被应用到无线资源管理领域中,表2对几种相关算法进行了总结并分析了其相应的特点,具体包括:遗传算



▲图2 智能5G网络切片框架图

▼表2 无线资源分配方法

方法	方法描述	特点
遗传算法	根据建立的优化目标,全局搜索最优解	能解决非线性优化问题;巨大的全局搜索容量;较强的健壮性
多臂赌博机	通过调整分配给各个竞争者的资源比例,在最大化本身的收益和满足其他设备的收益之间折衷	适用于多个竞争者的自适应决策问题;分布式优化方法
Q学习	通过在特定的状态执行能够使累积收益最大的行动得到最优的资源分配策略	适用于系统模型未知和网络和环境随机变化的情况

法<sup>[10]</sup>、多臂赌博机<sup>[11]</sup>和Q学习算法<sup>[12]</sup>。

### 3.4 资源共享

网络功能虚拟化技术允许底层物理网络经营者向服务提供商租出部分基础设施,服务提供商利用租用的基础设施资源创建自己的虚拟网络以向终端用户提供端到端服务。虚拟网络由一系列的虚拟节点和链路组成,这些虚拟节点和链路由底层物理网络的路径和节点支撑。高效的底层物理资源共享过程主要分为两步:第一步是虚拟网络映射,将虚拟网络节点和链路映射为底层物理节点和路径;第二步是在虚拟网络的生命周期内为虚拟节点和链路动态分配资源。为了实现对虚拟网络的高效管理,文献[13]利用人工神经网络(ANN)技术设计了一种自动化系统为虚拟网络提供自适应的资源分配。将底层物理网络的节点和路径用ANN进行模拟,ANN的输入为网络资源状态,输出为资源共享方案。文献[13]还采用了一种类似于增强学

习的误差函数对ANN的输出结果进行满意度评估,以完成ANN的在线训练过程。

文献[14]提出了一个基于图形神经网络(GNN)来预测虚拟化网络功能(VNFs)对资源的需求,该GNN模型表示由VNF及其连接的相邻VNFs形成的星型拓扑,通过将前馈神经网络(FNN)函数应用于不同的GNN层来计算VNF对资源的需求,从而实现对于NFV中的资源动态感知和管理。

## 4 研究展望

### 4.1 智能化5G网络的机遇

人工智能给5G网络的实现带来了更多的机会。首先,面对超负荷的蜂窝网络数据量,人工智能可以利用这些数据预测可能发生的事件,并在此基础上实现网络资源的动态分配和参数的动态配置;其次,随着网络的异构化,人工智能技术能够分析新兴的业务请求并为其匹配最合适的接入点以满足用户的性能需求;最

后,人工智能技术可以使5G网络对于系统警报响应和安全威胁防护都变得更加敏捷,从而使网络能够更好地应对意料之外的网络状况。

人工智能算法包含领域较广且种类繁多,其中机器学习方法(如Q学习和支持向量机)、深度学习方法(如卷积神经网络和循环神经网络)、神经网络(如反向传播算法、玻尔兹曼机等)、群智能算法(如蚁群算法)以及进化算法(如经验竞争算法)都可以用来改善5G网络性能。

在上述技术中,神经网络和深度学习目前受到了广泛关注。一般来说,神经网络由许多神经元和神经元之间的加权连接组成,其中的神经元可以抽象为变量而加权值可以视为相关参数。神经网络可以利用学习技术实现合理配置以保证相关应用可以根据一组输入值得到理想的输出值,通过学习算法迭代调整当前所有神经元对之间的连接权重值,使神经网络的输出值和输入值达到最佳匹配。当前,神经网络已经在认知无线网络的状态估计和预测方面表现出了良好的性能,而为了让神经网络算法在5G网络中发挥其自适应性及分散控制的优势,此项技术正在被逐渐运用到NFV的虚拟网络映射和资源共享领域,然而当前神经网络技术在5G网络中的应用依然处于初步阶段。

进化算法也将因其复杂度低和收敛性快的优势成为下一代无线网络的重要实现技术,例如:遗传算法就可以通过模仿自然选择过程解决网络优化问题。相比之下,机器学习算法的执行主要包含两个阶段:训练阶段和测试阶段,其中训练阶段的复杂性要比测试阶段高得多,由于当前对5G网络在能效和灵活性方面的严格要求,对于小型的移动终端将只执行其机器学习算法中的测试阶段。

### 4.2 5G网络智能化挑战

将人工智能技术运用到5G网络

中的好处显而易见,但其中仍然存在一些值得思考的问题:首先,在5G时代,网络数据是一把双刃剑,虽然海量的数据给人工智能的训练和使用提供了宝贵的基础,但找到一个简单有效的模型来匹配这些数据却相当困难;其次,为了及时处理蜂窝网络中的数据,需要网络具备巨大的存储和计算资源,且可能会威胁到信息安全;最后,人工智能算法的使用必须以数据的集中处理为前提。所有这些因素均会对网络实体的计算容量和运营成本带来巨大的负担。

## 5 结束语

本文首先介绍了5G网络的新特征以及人工智能的基本概念,在此基础上阐述了将人工智能技术运用到5G网络中的重要理论和实践意义。其次,本文设计了一种智能化5G网络架构,此架构主要由3部分组成:和真实网络环境进行实时交互的环境舱;利用人工智能技术对从环境舱传来的信息进行处理智能中心;根据智能中心的处理结果采取行动的网络策略。这种智能化5G网络架构能够通过和网络环境的交互解决5G网络中的各类决策性问题。进一步,本文对当前人工智能技术在业务预测、网络切片、无线资源分配以及资源共享方面的应用作了简单总结并分析了各种算法的特点和适用场景。最后,本文分析了人工智能技术运用到5G网络中的机会和挑战。虽然5G网络的人工智能化进程依然存在重重阻碍,但我们相信人工智能技

术将会翻开未来移动通信网络的新篇章。

### 参考文献

- [1] IMRAN A, ZOHA A. Challenges in 5G: How to Empower SON with Big Data for Enabling 5G [J]. Network IEEE, 2014, 28(6):27-33. DOI: 10.1109/MNET.2014.6963801
- [2] 任驰,马瑞涛.网络切片:构建定制化的5G网络[J].中兴通讯技术,2017,24(1):26-30. DOI: 10.3969/j.issn.1009-6868.2018.01.006
- [3] ITU Focus Group on Machine Learning for Future Networks including 5G [EB/OL]. (2017-11-06)[2018-03-19]. <http://www.itu.int/en/ITU-T/focusgroups/ml5g/Pages/default.aspx>.
- [4] ETSI Launches Zero Touch Network and Service Management Group [EB/OL]. (2017-12-14)[2018-03-19]. <http://www.etsi.org/news-events/news>
- [5] LI D. Artificial Intelligence in the Rising Wave of Deep Learning: The Historical Path and Future Outlook [Perspectives][J]. IEEE Signal Processing Magazine, 2018, 35(1):180-177. DOI: 10.1109/MSP.2017.2762725
- [6] 5G Network Architecture: A High Level Perspective [EB/OL]. (2017-07-21)[2018-03-19]. <http://www.huawei.com/minisite/5g/img>
- [7] YAN X B, LU T, LI Y J, et al. Research on Event Prediction in Time-Series Data[C]// International Conference on Machine Learning and Cybernetics. USA:IEEE, 2004, (5):2874-2878. DOI: 10.1109/ICMLC.2004.1378522
- [8] PAN X, LEE B, ZHANG C. A Comparison of Neural Network Backpropagation Algorithms for Electricity Load Forecasting[C]// IEEE International Workshop on Intelligent Energy Systems. USA: IEEE, 2014:22-27. DOI: 10.1109/IWIES.2013.6698556
- [9] DARIO B, M G, ALBERT B. Mobile Traffic Forecasting for Maximizing 5G Network Slicing Resource Utilization [C]// IEEE INFOCOM. USA:IEEE, 2017. DOI: 10.1109/INFOCOM.2017.8057230
- [10] TSENG L Y, CHEN S C. Two-Phase Genetic Local Search Algorithm for the Multimode Resource-Constrained Project Scheduling Problem [J]. IEEE Transactions on Evolutionary Computation, 2009, 13(4): 848-857. DOI: 10.1109/TEVC.2008.2011991
- [11] MAGHSUDI S, STAMCZAK S. Channel Selection for Network-Assisted D2D Communication via No-Regret Bandit

Learning With Calibrated Forecasting [J]. IEEE Transactions on Wireless Communications, 2015, 14(3):1309-1322. DOI: 10.1109/TWC.2014.2365803

- [12] ONIRETI O, ZOHA A, MOYSEN J, et al. A Cell Outage Management Framework for Dense Heterogeneous Networks[J]. IEEE Transactions on Vehicular Technology, 2016, 65(4):2097-2113. DOI: 10.1109/TVT.2015.2431371
- [13] MIJUMBI R, GORRICO J L, SERRAT J, et al. Neural Network-Based Autonomous Allocation of Resources in Virtual Networks [C]// European Conference on Networks and Communications. USA: IEEE, 2014:1-6. DOI: 10.1109/EuCNC.2014.6882668
- [14] MIJUMBI R, HASIJA S, DAVY S, et al. Topology-Aware Prediction of Virtual Network Function Resource Requirements [J]. IEEE Transactions on Network and Service Management, 2017, 14(1): 106-120. DOI: 10.1109/TNSM.2017.2666781

### 作者简介



王威丽,重庆邮电大学博士生在读;主要研究方向为5G网络切片、人工智能算法等。



何小强,重庆邮电大学博士生在读;主要研究方向为新一代无线通信网络、人工智能算法研究、网络切片、网络功能虚拟化等。



唐伦,重庆邮电大学教授、博士生导师;主要研究方向为下一代移动通信网络、网络资源分配、网络切片、网络功能虚拟化等;已发表论文100余篇,其中被SCI/EI检索50余篇。

# 智能物联网技术和应用的发展趋势

## Trends of Technologies and Applications for Intelligent IoT

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0043-004

**摘要:** 提出跨垂直行业的、基于智能物联网(IoT)技术和应用的“一横一纵”的系统架构和发展趋势,即基于通用处理器(GPP)的共享计算平台和基于开源软件的应用开发环境将联合构建新一代智能物联网(IoT)的系统架构和应用基础。这一通用、普适的软硬件系统架构将助力IoT技术和应用从数据收集到信息提取再到知识创造的快速演进和创新飞跃,实现未来智能IoT服务的宏大使命。

**关键词:** IoT; GPP; 软件定义一切(SDX); 系统架构

**Abstract:** In this paper, the system architecture and development trend of the “general purpose processor (GPP)-horizontal and software defined everything (SDX)-vertical” in vertical industries is proposed, which is on the basis of intelligent Internet of things (IoT). It means that the GPP-based shared computing platform and open source software-based application development environment will be jointed to establish the system architecture and application foundation of next generation intelligent IoT. This universal software and hardware architecture will facilitate IoT technology and applications to rapidly innovate from data collection, information extraction to knowledge creation, which can achieve the future mission of intelligent IoT.

**Key words:** IoT; GPP; SDX; system architecture

杨旸/YANG Yang<sup>1,2</sup>

(1. 中国科学院上海微系统与信息技术研究所, 上海 201899;  
2. 上海无线通信研究中心, 上海 201210)  
(1. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 201899, China;  
2. Shanghai Research Center for Wireless Communications, Shanghai 201210, China)

得了不少重要的成果,但在实际部署过程中,多数垂直行业的物联网应用以封闭方案为主流,这一方面造成了系统软件与硬件的兼容性不强;另一方面,导致了不同物联网应用之间信息的交互性不佳,具体主要包含以下3个方面:

(1) 系统应用。物联网中嵌入式感知/控制设备采用不同的软/硬件平台,部署于无法自由交换信息的多种异构网络环境,专用的业务系统和管理平台难以兼容互通。

(2) 技术研发。物联网的用户参与度低,应用开发门槛高,开发周期长;平台架构的耦合度高,扩展性互通性差,基础设施的部署及维护成本高;异构网络与终端多样化;物联网各种资源缺少统一的描述模型,无法对资源进行统一的协同处理与协作调度。

(3) 架构设计。物联网系统呈现出竖井化的信息孤岛特征,仅适合需求明确的大规模垂直行业应用场景;系统基础设施架构和服务能力难以分享和重用,第三方资源也难以被低成本地集成进系统,使得物联网很难实现大规模的应用和推广。

为更好地推进物联网产业进步,

### 1 应用驱动促进了物联网技术的快速发展

近年来,经过国家与地方的大力推广,以及产业链的成熟,物联网(IoT)<sup>[1-2]</sup>的发展与应用已逐步进入快车道,促进了中国各行业应用的信息化与智能化。从系统工程角度来看,IoT具有将感知、控制和短距离无线通信能力的传感设备嵌入到各类器件与终端中,拓展系统自动获取物理世界信息的作用直径和能力;在传统以人类语义为核心的信息网络基

础上,引入物体贡献数据的新型信息维度,形成物理世界与人类信息领域融合的抽象数字生态系统;将对物理世界的感知引入现有的网络与信息系统,并对系统整体性能进行智能优化处理,兼容并包现有体系,共同形成对人类意义重大的信息网络体系。从技术演进角度来看,以应用驱动的物联网架构设计是物联网发展的关键,随着中国牵头制定的全球首个物联网顶层架构国际标准正式通过国际标准草案(DIS)投票<sup>[3]</sup>,NB-IoT标准化工作的完成<sup>[4-5]</sup>以及5G<sup>[6]</sup>商用的展望,将对物联网未来发展产生深远的影响。

尽管物联网产业化初见成效,取

收稿日期: 2018-01-20  
网络出版日期: 2018-03-29

突破物联网发展瓶颈,顺应物联网应用驱动的需求,本文提出跨垂直行业的智能物联网技术和应用的“一横一纵”系统架构和发展趋势,联合构建兼具灵活性与普适性的新一代智能物联网的系统架构和应用基础,助力物联网未来飞跃发展。

## 2 物联网应用的过去、现在和演进态势

### 2.1 过去:数据收集

自从“感知中国”计划拉开物联网发展的帷幕以来,中国的物联网事业取得了蓬勃的发展。早期物联网应用的研发与部署,多侧重于信息的感知,其数据采集技术主要涉及传感器、射频识别(RFID)技术、多媒体信息采集、微机电系统(MEMS)、条码和实时定位等技术。感知信息的组网通信技术主要实现传感器、RFID等数据采集技术所获取数据的短距离传输、自组织组网。感知层传输技术包括有线和无线方式:有线方式有现场总线、公共交换电话网络(PSTN)等传输技术;无线方式有RFID、Wi-Fi、超宽带、短距通信(NFC)等传输技术,此外还包括多传感器对数据的协同信息处理技术。

对于感知信息的传输主要通过移动通信网、互联网、卫星网、广电网、行业专网等,而针对数据的应用基本上停留在信息的存储、统计与发布层面。在这一阶段,其技术发展以刚性需求为主,重点是感知功能、应用性能,通过机器到机器(M2M)连接,很少跨领域信息交互。

### 2.2 现在:信息提取

随着物联网应用范围的不断拓展,以及物联网系统的广泛部署,尤其是“云计算”与“大数据”等新技术的层出不穷,推动了物联网技术的不断进步。中科院上海微系统与信息技术研究所作为主要参与单位,承接并实施的“新一代宽带无线移动通信

网”重大专项“面向南水北调工程安全的传感器网络技术研发”中,针对南水北调中线干线工程“三个安全”的重大需求,通过在长达1400 km的南水北调中线干线工程上,部署10万余只各种不同类型的传感器,有效地保障了“三个安全”(工程安全、供水安全、人身安全)的重大需求<sup>[1]</sup>。同时,结合前端感知信息的提取,有效地实现了异常状态预警、防入侵监测等关键功能,进一步证实了“传感器、数据、平台”等关键技术在大规模传感器网络应用中的可行性,推动了“三位一体”模式成为物联网应用发展的现实标准。

### 2.3 演进:态势和挑战

得益于前几年的海量感知信息积累,公众应用正在慢慢起步。这种公众应用以弹性需求为主,对公众网络依赖程度高,广泛利用现有各种终端。面对日益增长的物联网应用需求,各种平台纷纷涌现出来,譬如:阿里、京东分别推出各自的物联网平台,尽管这些企业级的平台可以提供共性服务的支撑,但往往难以实现不同平台之间的融合,而且缺乏专家系统支持,从发展的角度来看,无法满足物联网未来的快速增长。实际项目和工程实践中,每一个垂直行业都希望开发和掌控自己的平台和应用,造成垂直行业的物联网应用以封闭方案为主流,系统呈现竖井化信息孤岛特征,无法实现跨行业的大规模推广,而且在这种需求下,随着行业物联网部署规模的不断扩大,将造成这种竖井化信息孤岛越来越大。

如何实现跨平台的信息融合,如何减少并消除行业竖井化信息孤岛,将是未来智能物联网需要面对的挑战。纵观物联网的发展可以看出:数据收集是基础,信息提取是重点,知识创造是核心,跨界融合是关键。可以预见:未来在海量感知终端部署与海量物联网信息存储的基础上,物联网的发展必将从信息化与智能化过

渡为知识化;构建具有灵活性与普适性的软/硬件平台实现跨行业、跨平台信息融合,已成为智能物联网未来发展的迫切需求。

## 3 智能物联网技术和应用的发展趋势

尽管目前物联网处于规模化发展阶段,需求各异且信息孤立,由于政策的推动以及利益最大化的导向,未来垂直行业内部以及行业间的信息交互将成为必然趋势。同时,未来物联网技术的发展,不是对现有技术的颠覆性革命,而是通过对现有技术的综合运用,进一步结合先进技术(如:人工智能、机器学习等),融合现有技术实现全新模式转变来实现智能物联网。只有这样,才能有效地实现智能物联网的知识创造,实现智能物联网与现有的网络系统的平滑升级态势,催生出一系列新的框架与技术。

鉴于此,我们提出基于通用处理器的共享计算平台和基于开源软件的应用开发环境的“一横一纵”系统架构和发展趋势(如图1所示),着力于解决当前物联网领域的市场碎片化、应用规模小的问题,突破垂直行业应用端到端的封闭方案占据产品主流,行业壁垒较高,整体解决方案成本高等局限。

### 3.1 GPP:基于通用处理器的共享计算平台

基于通用处理器的共享计算平台从系统应用的角度出发,提取不同物联网业务和系统的共性元素;再根据共性元素的技术特征分类,以模块化方式实现共性元素子类,并依据具体业务需求选择不同子类组合,构成物联网应用的共性计算平台,以提供各种协同计算、设备管理、数据采集交换和共享管理、用户管理、权限管理、应用数据挖掘、专家系统、应用接口等共性功能和服务;然后以统一的开放硬件计算平台结合定制化的软



图1  
跨行业应用的智能  
物联网系统架构

件,适应多样化的应用需求,实现不同底层技术方案的组合和集成,并按照统一的模式平台化管理和发布各类信息。

基于通用处理器的共享计算平台体系设计,分离了物联网不同应用的共性技术特点和差异性,为解决应用场景多样化的矛盾提供了有效的思路,有利于规范物联网从感知互动设备到网络传输再到应用服务的明确产业链发展,在统一的标准体系下推进物联网的研究、开发、集成和应用。同时,也更好地体现不同行业应用在物联网共性平台统一架构下的融合衍生、集成创新,为多种产业应用提供科学而全面的服务。

基于通用处理器的共享计算平台的技术特征如下:

(1)缩短技术研发周期,降低开发成本

该技术可以帮助科研院所与小微企业快速开发应用系统,降低开发成本,缩短产品的上市时间。采用共享计算平台提供的稳定可靠的共性功能和服务,并对应用层和感知层提供标准接口,方便进行二次开发和功能扩展。新业务开发时,用户只需按照平台的接口开发应用功能即可,无需开发共性功能和服务,工作量大大减少,有利于短时间内快速开发出满足需求的应用系统,降低开发成本。

(2)实现数据共享、运营,消除信

息孤岛

通过应用集成平台建设物联网数据中心,制订完善的数据管理、存储、交换和共享策略和规范,使得各个系统之间的数据能够融合起来,有效解决了应用系统间的数据共享问题,从而有效消除了信息孤岛现象。同时,数据的融合和共享,使得物联网的最终运营成为可能。物联网应用集成平台是物联网“共性平台”的核心组成,依托该平台和技术,来解决物联网环境下多个异构系统之间的数据交换和共享以及应用集成的难题,为上层的智慧应用构建数据和信息资源中心。

(3)资源开放与复用,减少重复建设投入

该技术实现物联网共性的功能和服务支撑,这些功能可以被所有应用业务共享和复用,无需再次开发,从而避免了重复建设。进一步地,通过结合开源软件专属应用构建广域物联网基础架构,以集约化建设思想为主导,通过融合技术屏蔽复杂的网络环境和非规范信息资源环境,形成中国创新的物联网业务模式,以科研创新提升业内影响。

### 3.2 SDX: 基于开源软件的应用开发环境

郭贺铨院士等专家多次指出<sup>[8-10]</sup>: 我们正在进入一个软件定义的时代,

软件定义的技术本质是把原先一体化的硬件设施打破,将基础硬件虚拟化并提供标准化的基本功能,然后通过管控软件,控制其基本功能,提供更开放、灵活、智能的管控服务。物联网发展未来的机遇就是软件定义一切(SDX),其本质就是所有的信息都是数字化的——万物数字化,其实现的两个前提条件目前已基本具备:

(1)性能优异的硬件组件。硬件尺寸越来越小,功耗越来越低,散热越来越少,可靠性越来越高,加工工艺越来越熟,成本越来越低,可以更加顺利地执行软件发出的指令。

(2)突破时空限制的通信网络。通信网络不但随时随地存在,而且带宽越来越高,可靠性越来越高,建设和运营成本越来越低。这也得益于软件,因为软件定义了频率,定义了传输交换。于是,软件不但可以指挥本地硬件实现各种功能,还可以通过通信网络指挥远处的信息系统协同实现各种功能。正是在硬件和网络的大力支持下,软件才进入了定义一切的时代。

搭建的应用子集——开源软件专属应用,将为各类物联网产品增加一个通用软件。在产品物理功能尽量简单的同时,应用范围可以无限拓展,功能可以无限丰富,能力可以不断升级。

“共性平台+应用子集”的下一代智能物联网应用架构体系,将坚持标准与应用相结合、互促进的指导思想,在遵循物联网体系架构和标准的同时,以标准化推动方案规范化,确保未来物联网体系架构的灵活性和普适性,并增强方案可复制性,有利于面向全行业的应用推广。

## 4 结束语

物联网的发展正面临着关键时期,如何解决物联网应用的碎片化问题,消除信息孤岛,实现物联网平台的融合已刻不容缓。本文提出的“一横一纵”智能物联网体系架构和发展

趋势,可以有效地指导未来物联网技术的研发与系统的部署,从而有效地助力物联网技术和应用从数据收集到信息提取,并向知识创造的快速演进发展和创新飞跃。随着5G标准化的完成,以及与商用序幕的拉开,智能物联网技术与应用必将迎来全新的发展。

#### 参考文献

- [1] ITU-T. Overview of the Internet of things[S]. Y.2060. 2012
- [2] 塞缪尔·格林加德. 物联网[M]. 第1版. 北京: 中信出版社, 2016: 19-34
- [3] ISO/IEC JTC 1/SC 41, ISO/IEC 30141 ED1 [EB/OL]. (2018-01-15)[2018-02-03].http://www.iec.ch/dyn/www/f?p=103:38:15605361321219:::FSP\_ORG\_ID,FSP\_APEX\_PAGE,FSP\_PROJECT\_ID:20486,23,100717
- [4] Narrowband Internet of Things (NB-IoT); Technical Report for BS and UE Radio Transmission and Reception (Release 13): 3GPP TR 36.802[S]. 2016
- [5] Cellular System Support for Ultra-Low Complexity and Low throughput Internet of Things (CloT): 3GPP TR 45.820 V13.1.0[S]. 2015
- [6] YANG Y, XU J, GE X, et al. 5G Wireless Systems Simulation and Evaluation Techniques[M].Germany: Springer International Publishing, 2016
- [7] STEADTER T. 100,000 IoT Sensors Monitor a 1,400-Kilometer Canal in China [EB/OL]. (2018-01-15)[2018-02-03].https://spectrum.ieee.org/tech-talk/telecom/internet/a-massive-iot-sensor-network-keeps-watch-over-a-1400kilometer-canal
- [8] 郭贺栓. 物联网,计算无处不在,软件定义一切,网络包容万物,连接随手可及[EB/OL]. (2017-04-28)[2018-02-03].https://www.leiphone.com/news/201710/njWSLnuUcu8EvU8Q.html
- [9] 梅宏. 软件定义的未来——万物皆可互联,一切均可编程[EB/OL]. (2017-10-26)[2018-02-03].https://www.leiphone.com/news/201710/njWSLnuUcu8EvU8Q.html
- [10] 刘韵洁. 软件定义网络将推动社会进步[EB/OL]. (2017-09-11)[2018-02-03].http://tech.sina.com.cn/d/2017-09-12/doc-ifykufzt6333091.shtml

sina.com.cn/d/2017-09-12/doc-ifykufzt6333091.shtml

#### 作者简介



杨旸,中国科学院上海微系统与信息技术研究所研究员、学位委员会副主任,中国科学院无线传感网与通信重点实验室主任,上海无线通信研究中心主任,上海雾计算实验室联合主任,国际雾计算产学研联盟大中华区主任,中国通信学会学术工作委员会委员等;研究领域包括无线传感器网络(物联网)、新一代移动通信系统(5G)、雾计算与网络技术、开放无线测试验证平台等;承担了国家科技重大专项(03专项)、国家“863”计划、国家自然科学基金重点等一系列前沿课题研究;获选 IEEE Fellow、中组部“万人计划”、上海市“千人计划”、中科院“百人计划”、上海市“优秀学术带头人”和“领军人才”等荣誉;发表论文150余篇,申报专利80余项。

←上接第14页

利用率。

### 3 结束语

本文中,我们简要介绍了强化学习,并研究了强化学习在无线网络中的一些应用。我们给出了两个针对复杂异构无线网络、动态网络环境下的智能接入技术。从仿真结果来看:和传统的接入控制算法相比,我们提出的智能化接入技术可在增加很小代价的情况下提升较大的网络性能。

#### 参考文献

- [1] Cisco. Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014-2019[R], 2015
- [2] GSM Association. The Mobile Economy Report[R],2015
- [3] CAO B, HE F, LI Y, et al. Software Defined Virtual Wireless Network: Framework and Challenges[J]. IEEE Network, 2015;29(4): 6-12, 2015.DOI: 10.1109/MNET.2015.7166185
- [4] SIMON P. Too Big to Ignore: The Business Case for Big Data[M].British: John Wiley & Sons, 2013
- [5] LITTMAN M L. Reinforcement Learning Improves Behavior from Evaluative Feedback [J]. Nature, 2015,521(7553):445-451
- [6] 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(1):484-489
- [7] HU J, WELLMAN M P. Nash Q-Learning for General-Sum Stochastic Games [J]. Journal of Machine Learning Research, 2003, 4(6): 1039-1069
- [8] CHEN X F, ZHAO Z F, ZHANG H G. Stochastic Power Adaptation with Multi-agent Reinforcement Learning for Cognitive Wireless Mesh Networks[J]. IEEE Transactions on Mobile Computing, 2013, 12(11):2155-2166.DOI: 10.1109/TMC.2012.178
- [9] LI R, ZHAO Z, CHEN X, PALICOT J, et al. TACT: A Transfer Actor-Critic Learning Framework for Energy Saving in Cellular Radio Access Networks [J]. IEEE Transactions on Wireless Communications, 2014, 13(4):2000-2011.DOI: 10.1109/TWC.2014.022014.130840
- [10] SILVER D, HUANG A, MADDISON C J A, et al. Mastering the Game of Go with Deep Neural Networks and Tree Search[J]. Nature, 2016, 529(1):484-489
- [11] ROSIN C D. Multi-Armed Bandits with Episode Context[J]. Annals of Mathematics and Artificial Intelligence, 2011, 61(3):203-230
- [12] KIM H, KIM K, HAN Y, et al. A Proportional Fair Scheduling For multi-carrier Transmission Systems[C]//Vehicular Technology Conference. USA,2004,(1):409-413

#### 作者简介



严牧,电子科技大学博士生在读;主要研究方向为面向下一代移动通信的智能化的接入控制;已在IEEE期刊和国际会议发表论文2篇。



孙耀,电子科技大学博士生在读;主要研究方向为面向下一代移动通信的智能化用户接入和切换;已在IEEE期刊和国际会议发表论文4篇。



冯钢,电子科技大学通信与信息工程学院教授、博士生导师,网络空间安全研究中心移动通信网络安全研究所所长;主要研究方向为无线通信与技术;主持自然科学基金(NSFC)、“973”计划、“863”计划以及“新一代宽带无线移动通信网”重大专项课题等;已发表论文150余篇。

# 免调度非正交多址技术及其接收机设计

## On Grant-Free NOMA Technology and the Associated Receiver Design

邱刚/QIU Gang<sup>1</sup>

田力/TIAN LI<sup>1</sup>

王沙/WANG Sha<sup>1</sup>

袁志锋/YUAN Zhifeng<sup>1,2</sup>

(1. 中兴通讯股份有限公司, 深圳 518057;

2. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518057)

(1. ZTE Corporation, Shenzhen 518057, China;

2. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518057, China)

与长期演进(LTE)类似,第5代通信系统(5G)仍然首选正交多址作为基本的多址方式,但与此同时非正交多址(NOMA)也受到广泛的关注和研究,在第3代合作伙伴项目(3GPP) Rel.15中成功立项并有望在1~2Y内完成标准化<sup>[1]</sup>。与正交多址相比,非正交多址可以在上下行系统容量、接入能力等方面提供性能增益<sup>[2-4]</sup>。此外,上行NOMA技术天然容易与免调度接入相结合,进一步降低系统接入时延、信令开销以及终端功耗等<sup>[5-6]</sup>。

在基于免调度的接入假设下,用户终端(UE)可以自发地进行上行传输,而不需要发起调度请求等待调度,因此可以节省信令开销并降低时延和功耗。由于UE自发地进行上行传输,多个用户之间不可避免会形成竞争接入。虽然可以通过预配置或预留资源的方式避免竞争,但在一般

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0047-007

**摘要:** 主要探讨5G非正交多址(NOMA)及相应的接收机设计。NOMA在与免调度传输结合时,其收发机的设计除了需要考虑多用户干扰抑制及分集增益的最大化以外,还需要考虑实际的用户检测和识别以及存在碰撞情况下的非理想信道估计。在经典的最小均方差估计结合串行干扰消除(MMSE-SIC)接收机算法基础上,针对几种不同的免调度实现方案给予相应的解决方案。仿真结果表明:所提方案在误差传播抑制和复杂度降低等方面均有较好的性能增益,在实际系统中可以同时满足免调度及高过载的业务需求。

**关键词:** NOMA; 免调度传输; 海量机器类通信(mMTC)

**Abstract:** 5G Non-Orthogonal Multiple Access (NOMA) and the associated receiver design are discussed in this paper. When NOMA is combined with non-scheduled transmission, the suppression of multi-user interference and the maximization of diversity gain should be considered for the design of transceiver. It is also necessary to consider the realistic user detection, identification and the non-ideal channel estimation under contention-based transmission. Based on the classical minimum mean square error - successive interference cancellation (MMSE-SIC) receiver algorithm, the corresponding solutions are given for several different implementation schemes. Simulation results demonstrate that the system performance in terms of error propagation minimization and complexity reduction can be enhanced, and the true grant-free transmission and high overloading capability can be achieved.

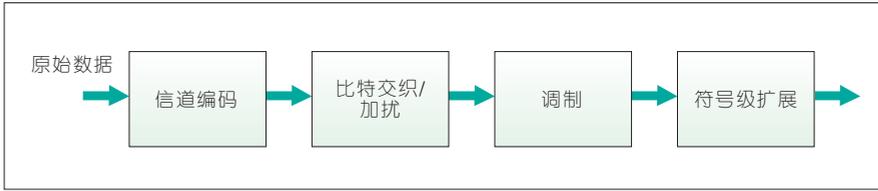
**Keywords:** NOMA; grant-free transmission; massive machine-type-communications (mMTC)

情况下由于基站对UE何时会发起免调度传输并不知情,资源利用率会相对较低,因此采用竞争接入的方式更为合适。NOMA技术的优势就在于:即使多个用户共享同一份资源并在相同的资源上进行,采用先进接收机仍可以实现多用户分离及成功解调。免调度和非正交多址的结合作为一种通用技术适用于多种场景,例如:针对海量机器类通信(mMTC)可以增加用户连接数及系统容量;对于高可靠低时延(uRLLC)业务可以降低时延及频谱效率,并能使移动宽带中的小包业务降低功耗,使初始接入

由传统的4步简化为2步(2-step RACH)等<sup>[7-8]</sup>。

一个典型的NOMA系统设计包括发射端处理和高级接收机设计。如图1所示,通常发射端是采用低码率编码、比特级交织或加扰,或符号级扩展等方式降低用户谱效,从而在复用更多用户的同时减轻用户间的干扰,进而使得接收端可以利用先进接收机进行多用户分离。典型的接收机包括基于最小均方差和串行干扰消除(MMSE-SIC)的接收机、基于匹配滤波和串行干扰消除(MF-SIC)的接收机、基于最大似然(ML)

收稿日期: 2018-01-16  
网络出版日期: 2018-04-08



▲图1 发射端处理流程

的接收机,以及基于消息传递算法(MPA)的接收机等<sup>[9-10]</sup>。接收机的选择与发射端的处理密切相关。

接收机的设计对NOMA系统尤其是其硬件实现的可行性和复杂度至关重要,在真实的免调度系统中接收机设计除了多用户检测算法的理论推导以外,还需要考虑很多实际因素,例如:非理想的用户识别、非理想信道估计、时频和频偏估计、复杂度降低等。

## 1 免调度NOMA接入技术的不同实现方案

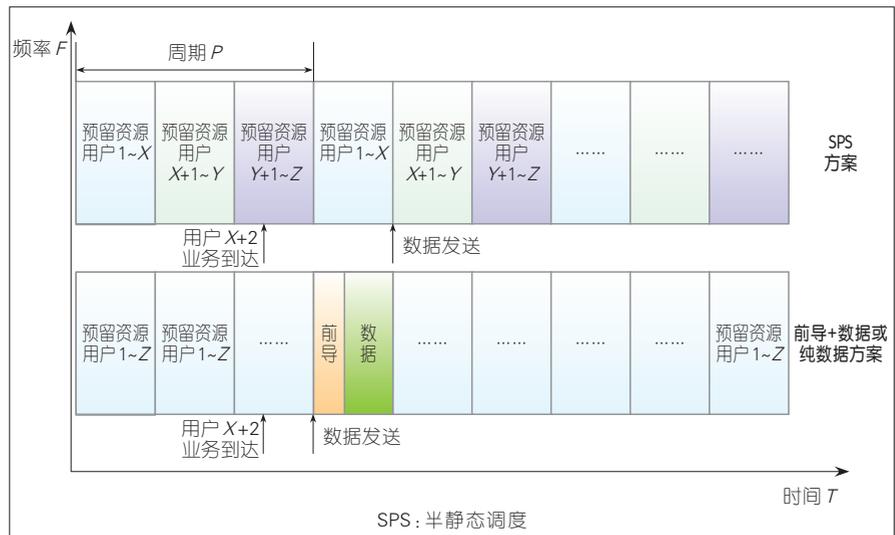
### 1.1 基于SPS的免调度方案

基于半静态调度(SPS)的免调度传输主要应用于无线资源控制(RRC)连接态下的数据传输,典型的应用有网络电话(VoIP)业务以及uRLLC业务,其主要特点为UE的时频资源是周期性预配置好的,如图2所示,因此对于每一次的上行传输不需要发起调度请求,可以节省信令开销和传输时延。基于SPS的免调度传输与NOMA技术的结合可以进一步提高资源利用率和频谱效率,尤其是针对非周期业务而言。当多个UE共享相同的物理资源时,一般可以通过预配置的参考信号(RS)资源进行用户识别。由于VoIP或是uRLLC业务的用户负载一般相对较小,因此复用在相同时频资源的UE可以配置正交的RS序列。由于每个UE资源配置的周期可以很小,所以其等待传输机会的时延也是相对可以接受的。对于节点(gNB)而言,每个时频资源上的潜在复用用户数是配置好的,但是每个发送时刻真实发起上行传输

的用户数取决于用户自己是否有缓存数据,因此接收端需要对活跃的UE进行盲检测。

当业务对接入连接数或用户负载非常高时,例如:在mMTC场景下,基于SPS的免调度传输方式则不太适用,因为:由于mMTC场景单小区需要支持的UE数目巨大(百万级),且RS序列数量有限(通常只有几个到几十个),所以很难通过周期性预配置的方式来避免所有用户的RS碰撞。另外,当潜在接入用户数目太大时,为了最小化RS碰撞的影响,SPS资源分配的周期会非常长,由此导致资源利用率过低,尤其是对于mMTC中大部分用户都是偶发小包的非周期性业务和大规模连接的情况下,会存在如下问题:

- 很难满足mMTC业务延迟需求;
- 很难进行混合自动重传请求(HARQ)重传;
- 很难支持mMTC业务的各种数据包大小;



▲图2 免调度传输的不同实现方案

• 精准的定时需求或在接收下行同步信号预分配之前的很长一段时间内UE必须处于唤醒状态,较难符合低能耗需求。

另外,对于移动用户存在小区切换状态下,相邻的多个小区需要同时给该用户预留周期性的资源,这进一步提高了每个小区潜在的接入用户数,增加了预配置的难度。

### 1.2 基于前导或RS的免调度方案

由上节分析可以推断出:当业务负载过高时UE资源间的碰撞几乎是不可避免的,此时可以采用UE随机选择资源的方式减少不必要的信令开销和等待时延,这里的资源包括时频资源和RS序列。

为了便于UE识别和信道估计,可以考虑如图2所示的前导序列加数据信道结构。前导序列可以认为是一种特殊的RS,用于识别不同的用户,例如:基于Zadoff-Chu(ZC)序列可以作为前导序列设计的出发点,其已广泛应用于物理随机接入信道(PRACH)、SRS和多用户多入多出(MU-MIMO)的解调参考信号(DMRS)中。为了克服mMTC用户同步不精准的问题,前导可以采用与数据部分不同的子载波间隔,用以增加时偏估计的范围,提升同步性能。

当用户负载较高时,为了减少碰撞概率,即避免相同时频资源上共享的多个 UE 选择同样前导序列,前导序列的资源池应该设计得足够大。值得注意的是:ZC 序列池的大小又受限于前导序列的长度。另一方面,为了保证信令开销在可接受的范围内,前导序列不能过长,因此 ZC 序列池大小是有限的。此外,多用户盲检测的复杂度与序列池大小成正相关的关系,即随着序列池大小的增长,多用户盲检测复杂度也线性增加。总之,在前导加数据的解决方案中,前导的设计是至关重要的,应该折衷考虑带宽、子载波间隔、根序列长度、循环移位等参数间的权衡。

### 1.3 基于纯数据的免调度方案

上述两种方案都依赖于基于前导或参 RS 的用户识别,由于 RS 的资源数往往是有限的,因此对于用户随机选择资源的免调度系统来说其性能会不可避免地受到碰撞的影响。另一种可以尝试的方案是:不依赖 RS 而进行盲的用户识别,其思想主要是尽可能利用数据本身和调制符号星座图统计特征来支持免调度和多用户盲检测。多用户盲检测是指当两个 UE 选择相同的物理资源时,仍可以对具有较高信号与干扰加噪声比(SINR)的强用户进行解码,主要是利用扩展码和星座图的先验信息和接收信号统计特征对强用户数据符号星座图的变化进行盲估计和盲均衡,从干扰信号中恢复强用户信号从而成功解码。此时数据中最好明确含有 UE ID,一旦数据解码成功,便可以实现 UE 识别。由于没有 RS,其初始盲信道估计的准确性也较为受限,但成功解码后的数据还可进一步用于优化信道估计,通过码字级干扰消除最小化误差传播,同时也可以用于解码具有较低 SINR 的 UE。基于纯数据的免调度解决方案的优势有:可以省去前导或 DMRS 的开销,提升系统频谱效率;由于系统性

能不受前导或参考信号资源数的限制,可以实现更高的过载能力。

## 2 接收机设计

### 2.1 传统的 MMSE-SIC

传统的码字级 MMSE-SIC 接收机主要应用于基于理想 UE 识别,即当多用户的资源(时频资源、扩频序列、RS 序列)均由基站预配置时的 NOMA 方案中。接收机处理过程如下。

假设有  $N$  个非正交用户共享同一份时频资源,其扩频因子为  $L$ ,则在接收端  $R$  个天线端口的接收信号可写为:

$$y = \sum_{i=1}^N \hat{h}_i x_i + n \quad (1)$$

其中,  $y$  是维度为  $LR \times 1$  的接收符号向量,  $[x_1, x_2, \dots, x_N]$  是每个用户在该时频资源上发送的调制符号,  $\hat{h}_i = (\hat{h}_{i,1}, \hat{h}_{i,r}, \hat{h}_{i,R})^T$  代表第  $i$  个用户的等效信道。假设信号在该时频资源内的快速衰落是恒定的,且发射功率为  $P_i$ ,每个用户每个接收天线端口的传播信道冲击响应为  $h_{i,r}$ ,第  $i$  个用户的  $L$  长扩频序列为  $s_i$ ,则等效信道可以表示为  $\hat{h}_{i,r} = \sqrt{P_i h_{i,r} s_i}$ ,另外  $n$  是指功率为  $\sigma^2$  的加性高斯白噪声。对第  $i$  个用户而言,公式(1)中的接收信号可以表示为:

$$y = \sum_{i=1}^N \hat{h}_i x_i + n = \hat{h}_i x_i + z_i \quad (2)$$

其中,  $z_i = \sum_{m \neq i}^N \hat{h}_m x_m + n$  代表噪声与第  $i$  个用户受到其他用户的干扰之和。

线性 MMSE 接收机的均衡权值由公式(3)计算得出:

$$w_i = R_{z,i}^{-1} \hat{h}_i \quad (3)$$

其中,  $z_i$  的协方差表示为:

$$R_{z,i} = \sum_{m \neq i}^N \hat{h}_m \hat{h}_m^* + \sigma^2 I_{LR} \quad (4)$$

其中,  $()^*$  代表 Hermitian 转置且  $I_{LR}$  是维度为  $LR \times LR$  单位阵。如果信道已知或已通过 RS 估计得到,则基于线性 MMSE 均衡的第  $i$  个用户的 SINR 可以由公式(5)估计得到:

$$\text{SINR}_i = \hat{h}_i^* w_i = \hat{h}_i^* R_{z,i}^{-1} \hat{h}_i \quad (5)$$

根据 MMSE-SIC 的流程,优先选择 SINR 高的用户进行依次译码。这次假设第  $i$  个用户的 SINR 最高,第  $i$  个用户的信号可由公式(6)得出:

$$\hat{x}_i = w_i^* y \quad (6)$$

在该用户信号经过信道译码和循环冗余校验(CRC)校验并译码成功后,会进行串行干扰消除(SIC)的流程,即对该用户信号进行重构,再从接收到的总信号中消除重构的信号,然后尝试对 SINR 次高的用户进行均衡和译码,以此类推直到所有用户都成功译码,或所有的剩余用户都无法译码正确。

### 2.2 传统 MMSE-SIC 的改进

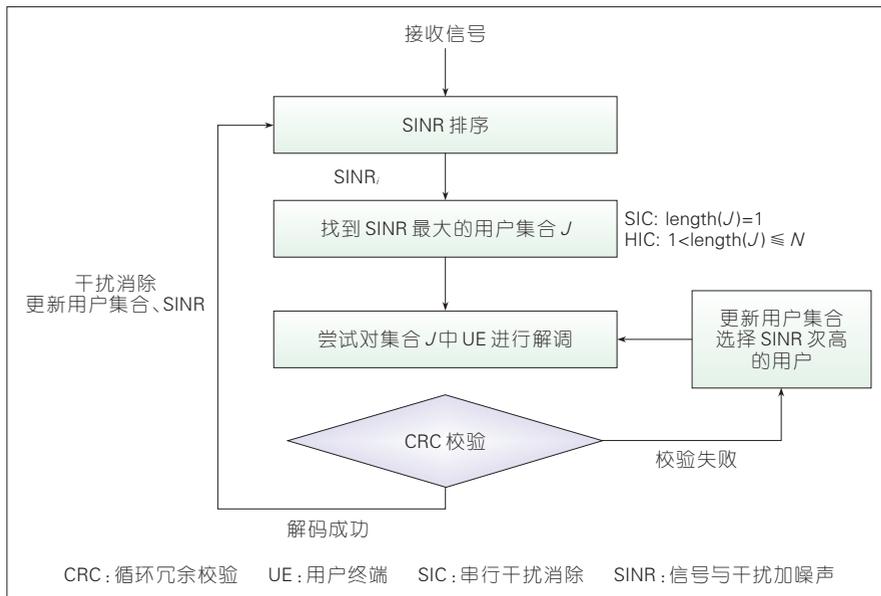
#### (1) 串并联合干扰消除(HIC)

HIC 是一种结合 SIC 和并行干扰消除的改进方案,其接收机流程如图 3 所示。与单纯的 SIC 流程相比,HIC 会根据公式(5)中统计的 SINR 的大小,尝试对具有较高 SINR 的多个用户并行进行解码,而不是逐一解码用户。在大多数情况下,HIC 可缩短处理延迟。与传统的 SIC 相比,由于每个用户有多个解码机会,所以综合译码性能可以得到提升。

#### (2) 数据辅助的信道估计改进

由于非理想的信道估计会导致干扰消除存在残余误差,即译码正确的信号无法完美重构及消除,尤其是较高功率 UE 的残留信号将对弱功率 UE 造成干扰,因此误差传播对于 SIC 接收机是至关重要的。本小节介绍一种利用译码正确的用户数据进行信道估计增强的方案。

假设第 1 个用户已经被正确解码,且  $p_1 = s_1 x_1$  是用户 1 发送的扩频后



▲ 图3 改进的MMSE-SIC接收机流程图

符号,则该用户的信道估计值可通过最小二乘(LS)法来改进:

$$\hat{h}_1 = (p_1^H p_1)^{-1} p_1^H y \quad (7)$$

其中, $y$ 是基于扩频后的接收信号。

若有多个用户的数据被成功解码,则多份数据会用于联合的信道估计增强。例如假设两个用户其扩频后符号分别为 $p_1$ 和 $p_2$ ,则改进的信道估计如下:

$$\hat{h}_{1,2} = (p_{1,2}^H p_{1,2})^{-1} p_{1,2}^H y \quad (8)$$

其中, $\hat{h}_{1,2} = (\hat{h}_1, \hat{h}_2)^T$ ,  $p_{1,2} = (p_1, p_2)$ ,  $(\cdot)^T$ 表示转置运算。随着更多的用户被成功解码,则更多用户数据的符号可用于改进信道估计。通过改进的信道估计值可以最小化SIC算法中的误差传播,提升多用户解调性能。

### (3) 复杂度降低

一般来讲,MMSE-SIC接收机的复杂度相对于最大似然估计类接收机是比较低的,它的复杂度随着UE数量呈线性增长。由上述公式可以看出:MMSE-SIC的复杂度主要源于矩阵求逆。矩阵求逆在某些情况下,如扩频长度太长或者有接收天线太多时,可能会非常高。在这种情况下,我们可以利用MMSE矩阵的复对

称性,在不降低准确性的前提下,通过增量矩阵求逆技巧诸如Sherman-Morrison公式<sup>[11]</sup>或其他的降维算法,降低矩阵求逆的复杂度。

此外,传统码字级MMSE-SIC需要计算每个符号的协方差矩阵并对每个符号进行相应的矩阵逆处理。事实上,在NOMA的典型应用场景如mMTC或URLLC场景中,用户的移动速率较低,而且包不会很大,因此大多数情况下数据包在时域的信道变化很小。公式(4)中的协方差矩阵可以通过在多个符号上进行平均得到,不需要逐符号进行求逆操作,从而可以在保证较小的损失精度下进一步减少矩阵求逆数,降低MMSE-SIC接收机复杂度。

### 2.3 基于前导/RS方案的接收机处理

与传统的基于理想用户发现的MMSE-SIC相比,在基于前导或RS的传输方案中,接收机需要额外考虑基于前导/RS的实际用户检测和识别。

对于前导和数据一次性传输的方案,其接收机流程如图4所示。首先需要完成前导序列的盲检测,其中前导序列资源池是预定义的,例如:前导序列采用ZC序列,则可以使用

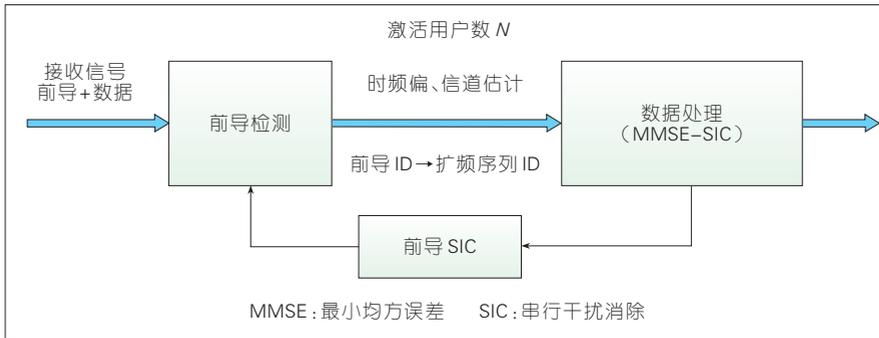
诸如根索引、循环移位(CS)、正交码(OCC)或频分复用(FDM)的相关参数用作区分用户的前导ID。可通过序列滑动窗相关,将相关峰值与一定虚警率下的门限值进行对比,检测该用户是否激活。一旦识别了用户的前导ID,则可以根据预定义的映射规则来确定该UE所使用的扩频序列。检测到的UE数、识别出的扩频序列,以及基于前导序列的信道估计将用于后续MMSE-SIC接收机的处理,该部分与图1所示的传统MMSE-SIC的数据处理过程类似。

关于免调度传输应该考虑的实际问题应包括:

(1) 虚警率。由于gNB不知道UE实际选择的前导序列,所以前导检测到的UE数可能大于实际发送的UE数。此时,gNB便会在MMSE-SIC过程中尝试解码虚警用户,这就引入了额外的复杂度。由于虚警用户的信道估计和SINR通常很低,对协方差矩阵的计算和实际发送用户的信道均衡影响几乎可以忽略不计,因此总体的接收机性能不会受影响。

(2) 漏检率。漏检率又分为2种:由低信噪比引起的漏检。在这种情况下,前导检测到的用户数小于实际发送的UE数,基站接收端不会对漏检用户进行数据译码的尝试,因而漏检用户的数据就认为是丢包了。对于其他正确检测出的用户来说,由于漏检用户的信道未知,其干扰无法通过MMSE均衡掉,所以整体性能也会有所影响。同时漏检用户的信噪比(SNR)通常也是比较低的,所以漏检对其他UE的解调性能影响不大。另一种为由碰撞引起的漏检,即如果两个或两个以上的UE选择相同的前导序列,则最多只有一个前导序列ID被检测到,这意味着将只有一个用户可能被检出。同时其相应的信道估计表现为多个UE信道响应之和,从而会导致干扰难以消除而引起系统性能的显著下降。

基于以上分析,前导或RS的碰



▲图4 基于前导的NOMA接收机流程

撞对免调度 NOMA 的传输性能有着重要的影响。一个可行的思路是通过 SIC 减轻前导序列碰撞问题,即当发生碰撞时(虽然 gNB 不知道前导是否碰撞,但 gNB 可假定检测到的前导序列总是被多个用户共享),如果其中一个强用户仍能够成功译码,则可通过如 2.2 节中所述的方法利用译码正确的数据对该用户的信道估计进行增强,然后重构并消除该用户的信号, gNB 将尝试利用相同的前导序列在残留信号中寻找之前因为碰撞被覆盖掉的漏检用户。

### 2.4 盲 MMSE-SIC 检测

纯数据解决方案的接收机过程参见图 5。由于没有用于用户识别的参考信号,所以需要数据本身进行盲信道估计、盲均衡和译码,其处理

流程与传统的 MMSE-SIC 有明显区别,详细的实现过程如下:

(1)假设每个 UE 具有  $K$  个数据符号,且假设在这些符号内每个用户的快衰落是恒定的,那么对于每个扩频后的发送数据符号,其对应的接收信号可以表示为  $y = \sum_{i=1}^N h_i s_i x_i + n$ ,其中

$h_i$  代表第  $i$  个 UE 的包含路损、阴影衰落、快速衰落效应的信道冲击响应,  $s_i$  是第  $i$  个 UE 随机选取的扩频码,  $x_i$  为第  $i$  个 UE 的调制符号,  $N$  是 UE 总数,  $n$  是加性高斯白噪声(AWGN)噪声。

(2)联立所有符号上的接收信号,表示为  $L \times K$  维的矩阵  $Y_l$ ,其中  $L$  是扩频长度,统计得到接收信号的协方差矩阵为  $R = Y_l Y_l^H / K$ 。

(3)由于不知道具体有多少个用

户存在于该时频资源上,因此需要遍历扩频序列池中的每个扩频序列,并进行联合信道均衡和解扩  $s_m^H R^{-1} y$ 。以  $s_i^H$  为例,均衡后的候选信号流见公式(9):

$$\begin{aligned} s_i^H R^{-1} y &= s_i^H (Y_l Y_l^H / K)^{-1} (h_i s_i x_i + z_i) \\ &= \frac{1}{h_i^H} s_i^H h_i (Y_l Y_l^H / K)^{-1} (h_i s_i x_i + z_i) \quad (9) \\ &= \frac{1}{h_i^H} A x_i + z_i' \end{aligned}$$

其中,  $A = s_i^H h_i (Y_l Y_l^H / K)^{-1} h_i s_i$ ,  $z_i$  和  $z_i'$  分别是解扩前后第  $i$  个用户受到的干扰。 $\frac{1}{h_i^H} A$  可以等效为对原始星座点进行缩放和相位旋转的线性处理。

(4)如果该扩频序列对应的用户是有发送数据的,则接收信号中存在相对应的数据流分量,其盲均衡后的信号星座点通常是原始信号星座点的线性失真。对于未被任何用户使用的扩频序列对应的数据流分量,则其盲均衡后的星座点不会存在明显的结构性,更像是噪声一般散乱分布。这里我们只考虑前者,对盲均衡后的星座点补偿缩放和相位旋转后便可将原始星座点周围的估计信号放回原位,  $\hat{x}_j = \frac{h_j^H}{A} s_j^H R^{-1} y$ ,并计算每个候选信号流的误差向量幅度

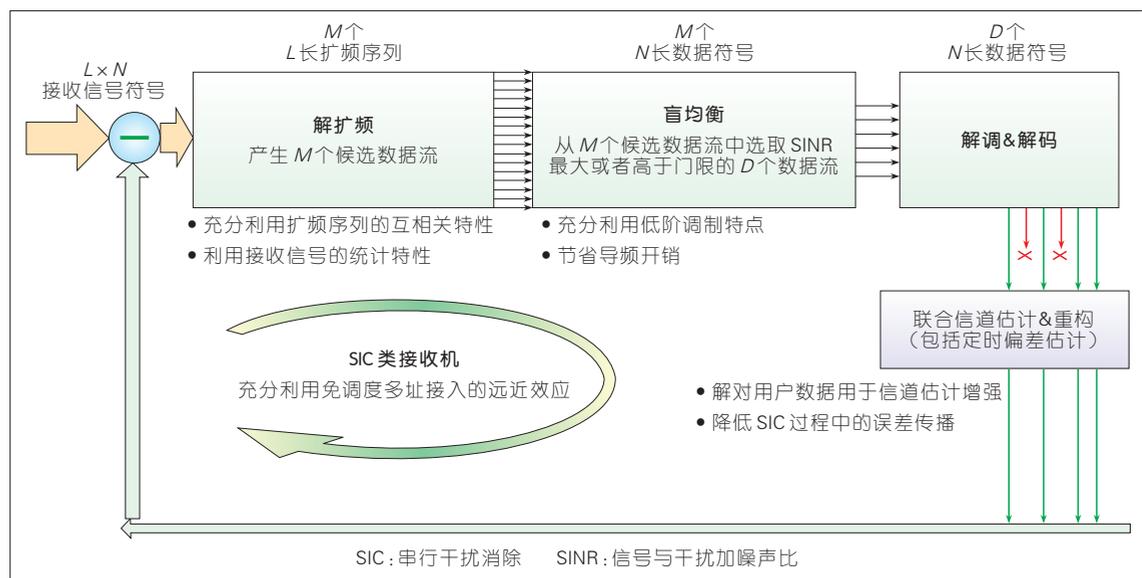


图5 纯数据方案的接收机盲检流程

(EVM)和等效 SINR。

(5)对 SINR 进行排序,选择具有最高 SINR 的多个候选信号流(考虑复杂度和可靠性,如 6 个)并尝试对其进行译码。对每个候选信号流  $\hat{x}_j$ ,还考虑到星座点相位模糊的可能性,因此有必要将信号的多种可能性输入到译码器进行解码,例如:对于二进制相移键控(BPSK)星座点,  $\hat{x}_j$  和  $-\hat{x}_j$  都应该被送去译码尝试。每个信号流是否能译码成功取决于 CRC 校验的结果。

盲 MMSE-SIC 过程的其余部分和 SIC 流程结束的判定都与传统的 MMSE-SIC 相同。类似于基于前导的解决方案,盲 MMSE-SIC 检测也应该考虑虚警率和漏检率,但由于盲信道估计主要依赖于数据本身,所以可以避免碰撞的问题。

### 3 仿真结果

我们采用多用户共享接入(MUSA)<sup>[9]</sup>作为非正交多址方案示例进行仿真对比。MUSA 为一种基于低互相关的符号级扩展方案,本例中采用扩展长度为 4 的序列,其扩展序列的元素取值于  $\{1, -1, j, -j\}$ 。图 6 中首先对比了理想信道估计下 MUSA 性能随用户负载的变化情况,其中调制方式为 QPSK,采用 Turbo 编码且码率为 1/2,接收机采用理想的 MMSE-SIC 或 MMSE-HIC。从仿真结果可以看出:在同时接入的用户数小于或等于 16 时,基于 MUSA 的非正交多址可以逼近单用户的性能,此时过载率(即用户数除以扩展序列长度)为 400%。当过载率进一步提升至 500% 即 20 个用户时,基于 MMSE-SIC 的 MUSA 会有 1.5 dB 的性能损失,而采用改进的 MMSE-HIC 接收机后性能略有提升。

需要注意的是:虽然理想信道估计下 NOMA 技术本身在支持 20 个用户时并没有明显性能损失,但是在真实信道估计下,如果仍采用基于导频的发射方案,那么由于导频端口的限制,例如:LTE 系统最多只支持 8 个导

频端口,当用户数多于 8 个时就会发生导频碰撞,其信道估计准确度难以保证,即同一资源上复用的用户数会受到较大限制。图 7 进一步给出了基于纯数据方案的 MUSA 性能。仿真假设免调度传输,调制方式为 BPSK, Turbo 编码码率为 1/2,接收机采用盲 MMSE-SIC 的检测方案,同时考虑实际的用户发现和真实信道估计。从仿真结果来看:随着用户数增多,系统的链路性能如误块率方面有所下降,但是由于纯数据的免调度传输方案并没有导频碰撞的限制,所以该仿真结果是表明其真实能支持的

用户数。

### 4 结束语

本文中我们主要讨论了 NOMA 以及其接收机设计,理论分析和评估了几种免调度 NOMA 实现方案的优缺点,并针对不同的免调度传输信道结构,基于 MMSE-SIC 接收机分别进行了针对性的优化和改进,并通过链路级仿真评估了所提方案和接收算法在用户接入过载能力方面的性能。

### 致谢

本研究得到中兴通讯袁弋非博

图 6 理想信道估计下 MUSA 方案性能对比

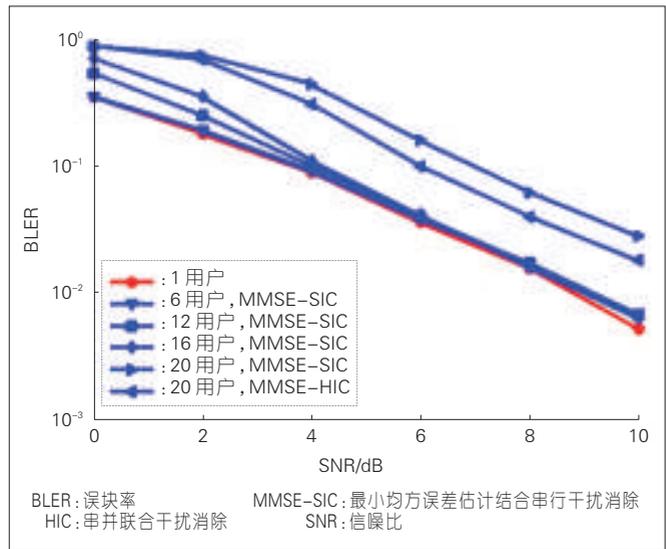
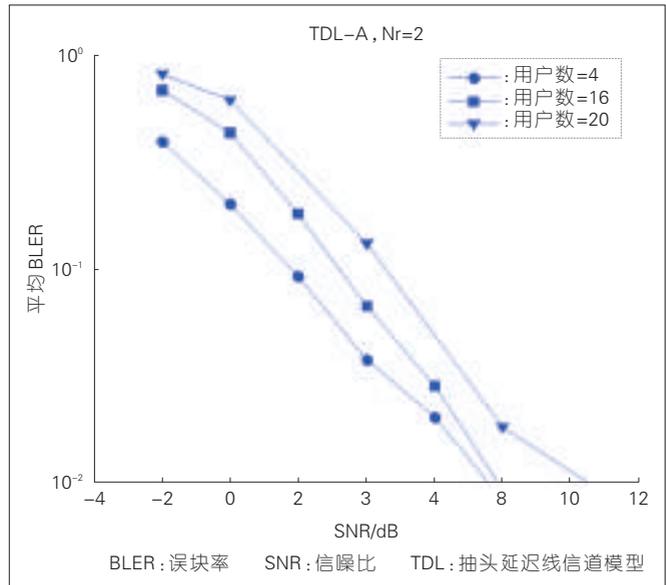


图 7 BPSK 调制下纯数据 MUSA 方案性能评估



士、曹伟博士、严春林博士、李卫敏工程师、戴建强工程师、胡宇洲工程师的大力帮助,谨致谢意!

#### 参考文献

- [1] 3GPP. SID on Non-Orthogonal Multiple Access for NR: RP-170681[S], 2017
- [2] SAITOY, KISHIYAMA Y, BENJEBBOUR A, et al. Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access[C]// IEEE 77th Vehicular Technology Conference. USA: IEEE, 2013:1-5
- [3] DAI L, WANG B, YUAN Y, et al. Non-Orthogonal Multiple Access for 5G: Solutions, Challenges, Opportunities, and Future Research Trends[J]. IEEE Communications Magazine, 2015,53(9):74-81.DOI:10.1109/MCOM.2015.7263349
- [4] 3GPP. Study on Downlink Multiuser Superposition Transmission (MUST): TR 36.859[S], 2016
- [5] YUAN Z, YU G, LI W, et al. Multi-User Shared Access for Internet of Things[C]// IEEE 83rd Vehicular Technology Conference. USA:IEEE, 2016,(1):1-5.DOI:10.1109/VTCSpring.2016.7504361
- [6] TIAN L, YAN C, LI W, et al. "On Uplink Non-Orthogonal Multiple Access for 5G: Opportunities and Challenges [J]. China Communications,2017,14(12): 142-152
- [7] BOCKELMANN C, PRATAS N, NIKOPOUR H, et al. Massive Machine-Type Communications in 5G: Physical and MAC-Layer Solutions[J]. IEEE Communications Magazine,2016,54(9): 59-65. DOI:10.1109/MCOM.2016.7565189
- [8] 3GPP. Study on New Radio (NR) Access Technology Physical Layer Aspects: TR 38.802, v0.3.2[S]
- [9] TSE D. Fundamentals of Wireless Communications[M]. British: University of California
- [10] CHENG Y S, NEELY M J, CHUGG K M. Iterative Message Passing Algorithm for Bipartite Maximum Weighted Matching[C]// IEEE International Symposium on Information Theory, USA:IEEE, 2006
- [11] SHERMANJ, MORRISON J. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix[J]. The Annals of Mathematical Statistics, 1950, 21(1): 124-127

#### 作者简介



**邱刚**,中兴通讯股份有限公司无线研究院副院长、高级工程师、专家委员会委员、首席系统架构师、LTE&5G研发总工兼产品经理,深圳市地方领军人才;主要负责4G LTE和5G的产品开发工作;作为课题组长完成多个重大专项;协同高速无线通信系统获2014年国家科学技术进步奖二等奖,第4代移动通信系统TD-LTE关键技术与应用获2016年国家科学技术进步奖特等奖,面向Pre5G的多天线/超大规模天线技术及应用获2016通信学会科技进步一等奖,并获2009—2016年度深圳市产业发展与创新人才奖。



**田力**,中兴通讯股份有限公司资深标准专家,同时担任《IEEE Transactions on Vehicular Technology》、《IEEE Access》等国际期刊审稿人,并担任IEEE Globecom 2017 workshop、WF-5G'18、EuCNC 2018等国际会议组委会委员;主要负责传播信道建模、5G物理层新空口标准、非正交多址技术等工作;参与多项国家自然科学基金及重大专项;共发表论文35篇。



**王沙**,中兴通讯股份有限公司无线算法部工程师;主要负责5G物理层新空口标准、非正交多址接入技术方面的工作。



**袁志锋**,中兴通讯股份有限公司无线算法部预研工程师;主要负责卷积码/RS码的FPGA收发系统、SOC ASIC芯片设计、新多址技术研究,研究领域涉及信息论、多址、纠错码、自适应算法、MIMO系统和高速软硬件算法;参与多项国家“863”及重大专项;发表论文10余篇。

# 新一代无线定位技术研究与发展趋势分析

## New Generation Wireless Location Technology: Research and Development Trend

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0054-005

**摘要:** 认为现有的无线定位技术无法满足移动互联网应用在定位网络覆盖范围、部署费用、定位精度以及终端兼容性等方面的需求。通过对无线定位产业的现状以及未来无线定位标准提案信息的分析,指出了通信定位一体化的发展趋势,并从运营商的角度提出了通信定位一体化网络发展的技术路线。

**关键词:** 室内定位; 位置服务; 管道化; 定位网

**Abstract:** Existing wireless positioning technologies cannot meet the Internet applications' requirements in terms of coverage, deployment costs, location accuracy and terminal compatibility. By analyzing the status of wireless positioning industry and standard initiative for future location technology, the integration trend of communication and positioning is demonstrated. Furthermore, the roadmap of communication and positioning integrated network is proposed from the perspective of communication carriers.

**Key words:** indoor positioning; location-based service; pipelining; location network

陈诗军/CHEN Shijun<sup>1,3</sup>  
王慧强/WANG Huiqiang<sup>2</sup>  
陈大伟/CHEN Dawei<sup>1</sup>

(1. 中兴通讯股份有限公司, 广东 深圳 518057;  
2. 哈尔滨工程大学, 黑龙江 哈尔滨 150001;  
3. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518057)  
(1. ZTE Corporation, Shenzhen 518057, China;  
2. Harbin Engineering University, Harbin 150001, China;  
3. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518057, China)

随着移动互联网产业的蓬勃发展,尤其是电子地图应用的大量涌现,基于位置的服务(LBS)已成为当前研究的热点<sup>[1]</sup>。针对人们日常生活中衣、食、住、行等各个方面,各种新兴服务行业不断推出嵌入LBS的移动应用程序,为终端用户提供了极大的便利<sup>[2]</sup>。此外,如车辆、运输、交通管理等传统行业为完成信息化升级,主动引入LBS,以完成相关产业及其生态链的深度改造和整合<sup>[3-4]</sup>,产生了以滴滴快车、优步、ofo共享单车等为代表的形态。可以看出,无论从新兴行业还是传统行业角度观察,LBS都是强有力的增长点。

目前,虽然以全球定位系统(GPS)为代表的定位技术的发展已经能够为用户提供满意的室外定位精度,但城市化进程的加速使人们大多数时间在室内活动,来自商业、政府、通信领域的室内定位需求与日俱增。新兴应用需求对定位服务提出了更高的覆盖要求<sup>[5]</sup>,定位网络应该不仅能够覆盖室外环境,也应该能够覆盖包括室内空间在内的特殊场景,以实现室内室外定位服务的无缝衔接。从相关产业的发展趋势可判断:能否提供有效覆盖室外环境与复杂室内环境的一体化定位服务,关乎人们生活质量乃至国民经济发展,是定位服务提供商急需解决的问题。

一体化定位服务需要在覆盖范围与定位精度两方面兼顾室内、室外

场景。与宽广的室外环境不同,室内场景空间狭窄,且随着5G时代的来临,物联网和智能化对位置服务提出了更高的要求,一般认为室内定位精度需达到3 m以下,甚至小于1 m才能够满足日常使用要求,更要求无缝定位覆盖<sup>[6]</sup>;另外,区别于室外较为简单的信道环境,室内的建筑结构、材料材质将引起复杂的多径效应,无线信号传播受到非视距(NLOS)因素影响较大,上述室内环境的特性导致传统的面向室外环境的定位技术已经不能满足室内定位的精度需求,而专门的室内定位技术往往存在覆盖范围小、部署成本高等缺点<sup>[7-8]</sup>,高精度、无缝覆盖对通信网和定位网协同发展提出了迫切的新需求。

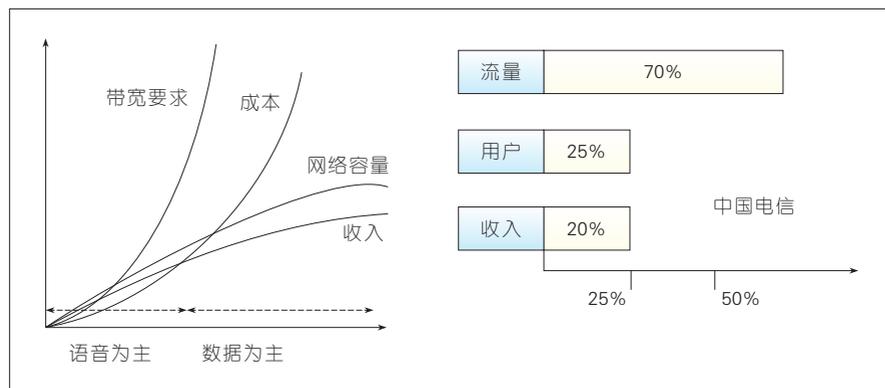
### 1 通信行业面临的问题

移动互联网时代,数据流量需求

收稿日期: 2017-07-22  
网络出版日期: 2017-09-11

激增,运营商之间、运营商和关联行业之间的“花园围墙”不复存在,整个产业价值链主导权发生转移,主要表现为运营商的管道化困境<sup>[9]</sup>,即对业务和用户控制权日益削弱。运营商成本与收入变化如图1<sup>[10]</sup>所示。为避免管道化命运,电信运营商虽然积极开展新业务,然而面临的竞争却日益加剧。一方面,运营商推出的新业务范围互相重叠,如:视频业务、应用商店业务等,在业务内容和业务模式方面缺乏新意,导致运营商之间的同质竞争越发激烈;另一方面,运营商介入IM、电子商务、手机支付等互联网应用领域,而不得不对传统互联网公司的正面交锋。

然而相应的努力收效甚微,其原因可以归纳为以下两点。首先,受传统思维影响,运营商难以主动放弃通信行业的核心地位,导致其产业规划主要以技术为导向,而忽视市场需求。但在消费者自我意识逐渐成熟,互联网思维被普遍接受的市场环境中,通信行业的发展必然会从以运营商主导向以消费者需求主导的模式转移,以用户需求为中心的服务模式是不可逆转的产业发展趋势。其次,技术的发展会使边际成本越来越高,而边际效益则随之降低,运营商不可能从已经高度发达的互联网应用领域攫取利益。因此,运营商为解决其面临的问题,必须准确发掘用户痛点,分享新技术起步阶段的研发红利,深耕蓝海市场。



▲图1 运营商管道化趋势

移动互联网浪潮为运营商带来了重大的发展机遇。以手机为代表的移动终端,已成为人们生活中不可或缺的一部分,人们的工作和生活都围绕移动终端进行,而移动终端的精准定位正在成为改善用户体验,增强应用品质的关键因素。移动终端的无缝覆盖正是运营商的传统优势领域,仅需要在传统的带内信号基础上添加必要基础设施就可以轻松完成通信定位一体化的网络,相对于GPS定位、无线传感器网络定位,在室内外融合场景下具有先天优势,且无需增添新的定位辅助设备。LBS服务可以作为新的“基础设施服务”提供给各类互联网公司使用,从在广阔的市场应用前景,从而带来新的巨大利润增长点。因此,如何针对用户对基于LBS的日常应用的使用需求,结合自身优势发展通信定位一体化网络是运营商在未来5G时代解决自身管道化问题的有效途径。

## 2 无线定位技术及存在问题

目前,针对不同的应用场景与需求存在多种特点各异的无线定位技术,下面我们对常见的技术做介绍。

### (1) 卫星定位系统

目前手机、车载导航仪等设备的基础定位信息主要是由以GPS、GLONASS、Galileo以及北斗等卫星导航系统提供<sup>[11-14]</sup>。

然而,由于导航卫星位于万米高空,信号衰落与障碍物的阻碍导致终

端在特定位置无法搜索到足够数量的定位卫星,难以满足定位的必要条件;同时,美国军方只开放C/A码供民用系统使用,其定位精度普遍在几十米到百米的范围内,该定位精度虽然能够支持道路导航类服务,但无法满足室内定位、车辆管理等应用的米级定位精度要求。此外,目前也产生了GPS、GLONASS、Galileo以及北斗的融合定位算法,但是精度一般仍保持在5m以上。

### (2) 地面伪卫星解决方案

为弥补卫星定位系统的缺陷,NextNav等公司提出了在地面布设伪卫星(PL)的补充方案。所谓伪卫星,是一种地面定位信号发射装置,发射的信号格式、带宽、电文等一般都和GPS一致<sup>[15]</sup>。伪卫星定位的基本原理与GPS相同,排除硬件设备对定位精度的影响,其准确性主要由伪卫星的几何分布决定<sup>[16]</sup>。虽然PL部署位置灵活,缺陷也与GPS相同,难以满足室内定位的需求。

### (3) 辅助GPS技术(A-GPS)技术

GPS设备利用卫星发出的无线电信号进行定位,但在实际使用过程中,地理环境和人文建筑常常会对信号造出一定的影响,使得定位信号被削弱,而不能完成定位的请求。在这种情况下,可以利用A-GPS系统完成快速定位。A-GPS以GPS为基础,是一种利用移动通信网络基站信息辅助GPS实现移动设备定位的技术,能够在各种制式的2G、3G乃至4G网络中使用。

利用A-GPS进行定位,除需要一部拥有GPS定位功能的手机外,还需要电信运营商为通信网络增加位置服务器、差分GPS基准站等设备<sup>[17]</sup>。简单地说,A-GPS是借助手机基站来为GPS提供前期位置调星服务与后期位置校准服务,以提高GPS的定位精度<sup>[18]</sup>。

### (4) Wi-Fi定位技术

基于无线局域网的解决方案是室内定位问题研究的一个重要方向,

该方案需要按照定位需求设置相应数量的Wi-Fi热点,利用基于信号强度的指纹定位技术,对移动终端进行定位<sup>[19]</sup>。首先需要建立Wi-Fi信号强度的指纹库<sup>[20]</sup>,在定位过程中终端设备将实时监测的Wi-Fi热点信号强度与指纹库中记录进行比对,确定定位结果。

#### (5) 无线射频识别(RFID)定位技术

基于RFID的定位系统通常将信号强度作为定位依据。定位系统主要由RFID标签、参考标签以及读卡器3部分组成<sup>[20]</sup>。其中,读卡器能够发出特定频率的信号,以激活休眠状态的标签,同时也能读取标签的识别码信息与信号强度。RFID标签被激活后,能够利用标签内置天线发送包含自身识别码的信息。当待定位的RFID标签处于读卡器作用范围内时,读卡器读取该标签识别码,并将其信号强度与参考标签比对,以计算待定位标签位置。

#### (6) 超宽带(UWB)定位技术

UWB是一种无载波通信技术,利用纳秒至皮秒级的非正弦波窄脉冲传输数据。通过在较宽的频谱上传送极低功率的信号,UWB能在10 m左右的范围内实现每秒数百兆至数吉比特的数据传输速率。与蓝牙、Wi-Fi等带宽相对较窄的无线技术相比,UWB具有抗干扰能力强、带宽极宽、耗能低、发送功率小、保密性好等优点,特别适合在室内定位系统中应用<sup>[22]</sup>。然而UWB受通信距离的限制,部署密度较高,在诸如首都国际机场这类大型室内场景中需要部署数百万计的节点,系统部署开销较大,这也限制了UWB的广泛应用。

#### (7) 其他室内定位解决方案

此外,针对特定环境与应用需求,还有很多其他室内定位技术,如:红外室内定位<sup>[23]</sup>、蓝牙室内定位<sup>[23]</sup>以及超声波室内定位<sup>[24]</sup>等。这些定位技术只适合某些特定场合,无法为用户提供一个适应性一致收敛的全网

解决方案。

#### (8) 现有定位技术存在的问题

现有定位技术存在的问题可以归纳为以下4个方面:

- 适应性问题:融合不同定位技术是普遍采用的方案,但是这些技术都无法做到普适于各类室内场景。例如:A-GPS基本能够覆盖室外环境,但不能应用于室内定位场景。

- 覆盖问题:无法提供大范围覆盖和无缝覆盖,如UWB和蓝牙。

- 成本问题:有些技术原理上可行,但真正建设时会涉及大量的建设成本和维护成本。

- 精度问题:精度动态范围大,很难提供一致质量的定位服务。

同时,可以看出无线定位网在室内定位方面还处于从属、补充的地位,无法充分发挥蜂窝网覆盖范围大、已有基础设施完善的特点。

### 3 室内定位产业发展及无线定位标准

导航与位置服务关键技术已成为国家安全和国民经济发展的坚实基础,广域高精度室内定位技术更是支撑导航与位置服务的关键。据估计2015年之后,中国位置服务应用行业市场规模将以大于50%速率增长,到2020年,室内位置服务市场规模将超2万亿。因此,很多著名互联网公司都重点布局了室内定位业务。其中,Google室内定位服务目前拥有1万多个场所的地图,已经进入了包括美、英、法、日等10多个国家,在机场、商城、火车站、体育馆和博物馆等公共场所提供室内导航服务;苹果公司也已经意识到地图、室内定位对未来应用和商业形态的巨大影响力,于2013年收购WiFiSLAM公司,以获取该公司的室内定位技术;Nokia、三星、Sony Mobile及Broadcom等22家手机厂商及芯片领域厂商共同成立室内定位联盟,携手建立室内定位的系统解决方案,并已经完成技术架构白皮书。

同时,为了适应高精度室内定位的需求,美国联邦通信委员会(FCC)、第3代合作伙伴项目(3GPP)、电气与电子工程师协会(IEEE)等国际组织已将广域高精度室内定位确立为下一代移动通信技术的基础功能。其中,在蜂窝移动通信网定位技术的国际标准方面,IEEE 802.11成立了NGP研究下一代高精度室内定位。目前3GPP标准中的定位方法如表1所示。从表中可以看出:相比3G时代通信网定位方式,定位方法更加丰富,要求更高的定位精度;通信网开始努力增强自身定位能力,以及融合其他定位技术。其中,3GPP LTE Release 9规范了通过测量两个或更多的基站参考信号(RS)的到达时间差(RSTD)来计算手机位置的技术观察到到达时间差法(OTDOA),在物理层的协议中增加了定位参考信号的特性,弥补了全球导航卫星系统(GNSS)在室内无法定位的缺陷<sup>[25]</sup>。在2016年6月,3GPP公布了5G标准的首个版本Release 15的计划,描述了各工作组的协调项目和检查重点,并计划于2018年6月之前完成标准发布,其中高精度定位被列为一项内容。此外,国际电信联盟(ITU)也已正式启动5G标准研究工作,于2015年6月提出IMT-2020计划,确定5G愿景、时间表等关键内容,到2020年完成5G技术标准化,5G进入到技术突破及标准研究的关键阶段。

此外,以高通为代表的传统通信技术公司,在芯片方面也已经进行了大量先期研发。由于各标准化组织对室内定位立项时间较短,绝大多数室内定位技术方案仍处于非公开状态,只能通过其专利涉及领域分析各公司提案内容,见表2。中国IMT-2020(5G)推进组2015年2月发布的《5G概念白皮书》中把“移动互联网和物联网将成为5G发展的主要驱动力”作为5G系统需求基础,而高精度室内定位技术是未来移动互联网和

▼表1 目前3GPP标准中的定位方法

方法	UE-based	UE-assisted, E-SMLC-based	eNB-assisted	LMU-assisted/ E-SMLC-based
A-GNSS	Yes	Yes	No	No
Downlink <sup>1</sup>	No	Yes	No	No
E-CID	No	Yes	Yes	No
Uplink	No	No	No	Yes
Barometric	Yes	Yes	No	No
WLAN	Yes	Yes	No	No
Bluetooth	No	Yes	No	No
TBS <sup>2</sup>	Yes	Yes	No	No

注1:包括基于PRS的TBS定位方法

注2:此处仅包含基于MBS信号的TBS定位方式

A-GNSS: 辅助定位全球导航卫星系统  
E-CID: 增强小区识别  
eNB: 演进型基站  
E-SMLC: 演进型业务移动定位中心

LUM: 定位测量单元  
MBS: 都会信标系统  
PRS: 定位参考信号  
TBS: 地面信标系统

UE: 客户端  
WLAN: 无线局域网

▼表2 各公司专利涉及领域

公司	涉及领域
爱立信	PRS 增强
高通	PRS 增强, TBS, DL-LMU
nextnav	TBS
LG	prs-Beacon, D2D
朗讯	unlicensed prs-Beacon, D2D
intel	D2D, prs-Beacon, TBS

D2D: 设备到设备  
DL-LMU: 下行链路定位测量单元  
PRS: 定位参考信号  
TBS: 地面信标系统

物联网的重要核心业务之一。未来移动通信论坛是在中国发改委、科技部、工信部的共同支持下,由移动通信运营企业、设备制造企业、科研机构、高等院校等26家单位共同发起成立的非营利性国际社团组织,其在2016年末发布了以中国联通、中兴通讯、清华大学和哈尔滨工程大学等单位撰稿的文章,专门论述5G高精度定位的关键技术。

从现有资料分析,在下一代无线通信网络标准中十分可能加入对定位功能支持,未来移动通信网络将和定位网相结合,发展出具有通信定位一体化功能的新一代网络。

#### 4 通信定位一体化网络

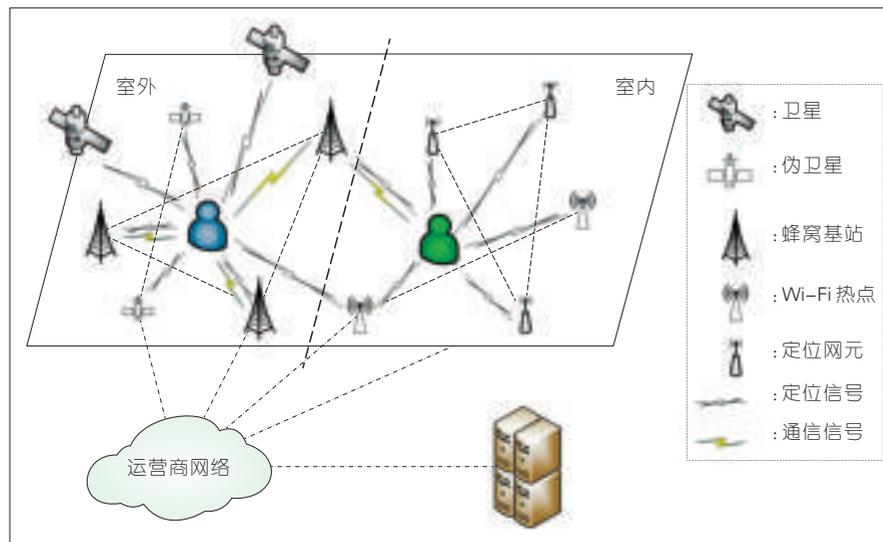
根据定位系统覆盖的范围可大

致将定位网络的发展划分为球域空间定位、地面基础定位以及局部精确定位三个维度。其中,球域空间定位技术主要针对特殊定位需求,提供能够覆盖全部近地空间的海、陆、空天一体化的立体定位服务;地面基础定位技术能够满足人们日常工作、生活中大部分需求,具有较大的地面定位覆盖范围;局部精确定位技术为迎合特定场合对定位精度的需求,只能覆盖较小的范围。移动通信网络是一种常见的地面基础定位系统,借助运营商部署的大量基站,通信网络具有室内外一体化覆盖、相对部署成本低

以及终端设备兼容性高等优点,是构建未来通信定位一体化网络的基础。

通信定位一体化网络是一种结合通信与定位功能的网络新型态,能够提供室内室外一体化的高精度定位服务,如图2所示。一般认为通信定位一体化技术的演进路线包括以下几种:(1)通信网与伪卫星系统、TC-OFDM定位系统等异构地面定位融合;(2)通信网与Wi-Fi定位系统、蓝牙定位系统等局部定位网络融合;(3)具有带内高精度定位技术的新一代通信网络。通信定位一体化网络是解决现有定位技术存在问题的有效手段,在提供室内室外一体化的定位服务覆盖的基础上,兼顾了室内高精度定位精度的要求,同时以现有移动通信系统为基础,能够最大限度地降低定位网络的部署成本,提高定位终端的兼容性。

从运营性角度分析,高精度的一体化付费定位服务能够成为新兴的业务增长点,也是摆脱管道化困境的机遇。以移动通信网络为基础发展通信定位网络需要解决定位精度的问题,可从以下两个方面着手:(1)增强移动通信网络自身的定位功能,可以通过构建地面高精度无线定位网解决通信网定位的关键瓶颈,同时引入基于设备到设备(D2D)辅助定



▲图2 通信定位一体化网络

位技术进一步增强定位精度;(2)异构定位系统的融合,引入地面伪卫星、Wi-Fi定位、UWB定位以及RFID定位等技术,为其设计相应的接口使其能够与通信网络交互信息,融合定位结果以提高定位精度。

## 5 结束语

随着移动互联网的兴起与不断发展,定位产业迎来了前所未有的发展机遇。为了满足未来智能化社会需求,解决定位精度和定位覆盖两大核心问题,从未来5G的多种演进途径可以看到未来定位技术的两个趋势:融合定位趋势和通信定位一体化趋势。现有的各种定位技术都无法同时满足覆盖范围与定位精度的需求,而现有移动通信网络具有覆盖范围大、设备兼容性高等特点,通过构建融合的地面定位网有助于提高定位精度和室内区域定位覆盖。同时,移动通信网不仅是管道,将有机形成地面定位网,并在网络结构和定位技术层面和通信网一体化发展。室内外一体化覆盖的高精度定位服务不但将在室内导航、商品索引、特殊人员照理等新兴市场有巨大的需求,而且对传统行业的改造升级具有巨大的促进作用,如:能在智能社会中的智能交通、动物管理、智能环保等方面创造巨大的社会和经济效益。现有的定位技术都无法同时满足覆盖范围与定位精度的需求,而现有移动通信网络具有覆盖范围大、设备兼容性高等特点。因此,以移动通信网络为基础,构建覆盖室内外环境的通信定位一体化网络是满足未来移动互联网中定位需求的最佳解决方案。

### 参考文献

- [1] ZICKUHR K. Location-Based Services[EB/OL]. (2013-09-12)[2017-07-15]. <http://www.pewinternet.org/2013/09/12/location-based-services/>

- [2] 翟红生,于海鹏. 在线社交网络中的位置服务研究进展与趋势[J]. 计算机应用研究, 2013, 30(11): 3221-3227. DOI: 10.3969/j.issn.1001-3695.2013.11.004
- [3] 周傲英,杨彬,金澈清,等. 基于位置的服务: 架构与进展[J]. 计算机学报, 2011, 34(7): 1155-1171. DOI: 10.3724/SP.J.1016.2011.01155
- [4] 唐科萍,许方恒,沈才樑. 基于位置服务的研究综述[J]. 计算机应用研究, 2012, 29(12): 4432-4436. DOI: 10.3969/j.issn.1001-3695.2012.12.006
- [5] BENSKY A. Wireless Positioning Technologies and Applications[M]. USA: Artech House, 2016
- [6] SAYED A H, TARIGHAT A, KHAJEHNOURI N. Network-Based Wireless Location: Challenges Faced in Developing Techniques for Accurate Wireless Location Information[J]. IEEE Signal Processing Magazine, 2005, 22(4): 24-40. DOI: 10.1109/MSP.2005.1458275
- [7] CONTI M, WILLEMSSEN J, CRISPO B. Providing Source Location Privacy in Wireless Sensor Networks: A Survey[J]. IEEE Communications Surveys & Tutorials, 2013, 15(3): 1238-1280. DOI: 10.1109/SURV.2013.011413.00118
- [8] ZENG Y, CAO J, HONG J, et al. Secure Localization and Location Verification in Wireless Sensor Networks: A Survey[J]. The Journal of Supercomputing, 2013, 64(3): 685-701. DOI: 10.1007/s11227-010-0501-4
- [9] 吴翠先,张功国,杨映红. 从网络层架构论运营商管道化趋势[J]. 数字通信, 2013 40(1): 48-51. DOI: 10.3969/j.issn.1001-3824.2013.01.011
- [10] 韦乐平. 电信业的未来与去电信化的思考[J]. 现代电信科技, 2013, (03): 1-6
- [11] HOFMANN-WELLENHOF B, LICHTENEGGER H, COLLINS J. Global Positioning System: Theory and Practice[M]. New York: Springer Science & Business Media, 2012.
- [12] NOURELDIN A, KARAMAT T B, GEORGY J. Fundamentals of Inertial Navigation, Satellite-Based Positioning and Their Integration [M]// Global Positioning System. Berlin: Springer, 2013: 65-123
- [13] GOSWAMI S. Indoor Location Technologies [M]// Global Positioning System. New York: Springer, 2013: 51-63
- [14] DUBEY A K. Understanding An Orogenic Belt[M]// Global Positioning System. New York: Springer, 2014: 215-230
- [15] TAKAHASHI S. Transmission Power Control of Terrestrial Pseudo Satellite Signal for Global Navigation Satellite Systems[C]// TENCON 2012-2012 IEEE Region 10 Conference. USA: IEEE, 2012: 1-5. DOI: 10.1109/TENCON.2012.6412250
- [16] FENG J H, DENG L. Research on Geometric Layout of Pseudo-Satellite in the Small Airport[J]. Applied Mechanics and Materials, 2014, 651: 495-498. DOI: 10.4028/www.scientific.net/AMM.651-653.495
- [17] CHAN E C L, BACIU G. Differential GPS and Assisted GPS[M]// Introduction to Wireless Localization: With iPhone SDK Examples. New Jersey: John Wiley & Sons, Ltd, 2012: 157-184
- [18] PATEL V. Are We Living In a World Full of Lies?-Can We Trust Assisted GPS

Technology to Provide Accurate Data When Using Location Based Services?[EB/OL]. (2016-10-20)[2017-07-14]. <https://computing.derbys.ac.uk/c/wp-content/uploads/2015/10/V-Patel.pdf>

- [19] 魏菲,李允俊,金华. 使用位置指纹算法的WiFi定位系统设计[J]. 单片机与嵌入式系统应用, 2014, 14(5): 29-32.
- [20] LIU H, YANG J, SIDHOM S, et al. Accurate WiFi Based Localization for Smartphones Using Peer Assistance[J]. IEEE Transactions on Mobile Computing, 2014, 13(10): 2199-2214. DOI: 10.1109/TMC.2013.140
- [21] LAI Y L, CHENG J. A Cloud-Storage RFID Location Tracking System[J]. IEEE Transactions on Magnetics, 2014, 50(7): 1-4. DOI: 10.1109/TMAG.2014.2303810
- [22] LEE S, HA K N, LEE K C. A Pyroelectric Infrared Sensor-Based Indoor Location-Aware System for the Smart Home[J]. IEEE Transactions on Consumer Electronics, 2006, 52(4): 1311-1317. DOI: 10.1109/TCE.2006.273150
- [23] 苏松,胡引翠,卢光耀,等. 低功耗蓝牙手机终端室内定位方法[J]. 测绘通报, 2015 (12): 81-84. DOI: 10.13474/j.cnki.11-2246.2015.385
- [24] 王凡,彭勇. 基于TDOA的室内超声波定位方法的改进[J]. 计算机技术与发展, 2014, 24(6): 250-252. DOI: 10.3969/j.issn.1673-629X.2014.06.063
- [25] 3GPP. Release 9[EB/OL]. (2010-03)[2017-07]. [http://www.3gpp.org/specifications/releases/71-release-9\[EB/OL\].](http://www.3gpp.org/specifications/releases/71-release-9[EB/OL].)

## 作者简介



陈诗军,中兴通讯股份有限公司高级工程师,CCSA ST7副主席、室内定位工作组组长,北斗技术与应用委员会副主席;主要研究方向为5G、室内定位技术、MIMO等;曾获深圳科技贡献奖、广东省科技进步奖二等奖、深圳科技进步一等奖、中国通信学会科技奖等10项科技奖;已发表论文30余篇。



王慧强,哈尔滨工程大学教授、博士生导师;主要研究方向为未来网络、网络安全和室内定位等;曾获省部级以上奖励11项;已发表论文400余篇,其中SCI检索20余篇,EI检索160余篇。



陈大为,中兴通讯股份有限公司高级工程师;主要研究方向为5G、室内定位等。

# 软件定义天地一体化网络: 架构、技术及挑战

## Architecture, Technology and Challenges of Software Defined Integrated Space and Terrestrial Network

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 02-0059-004

**摘要:** 提出了一种基于软件定义网络(SDN)的天地一体化网络架构,将控制平面和数据平面分离。该架构下,地面信息港和地球静止轨道卫星组成了多控制器结构的控制平面,将中轨、近地轨道卫星作为简单的转发设备,提供了更灵活的路由策略,更方便的网络配置以及更好的兼容性。同时,考虑到目前在卫星节点上部署网络设备的难度,提出了初期在地面部署控制器,之后逐步向静止轨道卫星扩展控制平面的演进思路。此外,还分析了这种架构面临的挑战。

**关键词:** 天地一体化网络; SDN; 多控制器; 演进式

**Abstract:** A space-terrestrial integrated network architecture based on software defined network(SDN) is proposed in this paper. In this architecture, (1) control plane and data plane are separated; (2) the terrestrial information station and geostationary orbit(GEO) satellites are combined as the control plane; and (3) medium earth orbit (MEO) & low earth orbit(LEO) satellites are regarded as simple forwarding devices. This architecture could provide more flexible routing strategies, more convenient network configuration and better compatibility. Meanwhile, considering the difficulty of deploying network devices on satellites, an evolutionary strategy is proposed, in which controllers are deployed in terrestrial stations at the present stage and extended to GEO satellites gradually. Moreover, the challenge of this space-terrestrial integrated network architecture is analyzed.

**Key words:** integrated space and terrestrial network; SDN; multiple controllers; evolutionary

许方敏/XU Fangmin<sup>1</sup>  
仝宗健/TONG Zongjian<sup>1</sup>  
赵成林/ZHAO Chenglin<sup>1</sup>  
秦智超/QIN Zhichao<sup>2</sup>

(1. 北京邮电大学, 北京 100876;  
2. 中国电子科学研究院, 北京 100041)  
(1. Beijing University of Posts and  
Telecommunications, Beijing 100876, China;  
2. China Academy of Electronics and  
Information Technology, Beijing 100041,  
China)

着卫星节点可以仅完成简单的转发和硬件配置功能,将路由策略和具体路由的计算交给控制器节点来实现,从而实现灵活的路由策略、方便的网络配置、更好的兼容性,以及更低的部署、升级成本。

### 1 基于SDN的天地双骨干 空间信息网络架构

#### 1.1 逻辑架构

在对地面网络SDN架构及空间信息网络分析的基础上,图1所示为基于SDN的天地一体化网络逻辑架构,分为应用层、控制平面和数据平面3层。空间的深空探测、对地观测、航天测控和导航定位等空间任务构成了逻辑架构的应用层。控制平面为整个架构的核心,由天基和地基控制器节点组成,并且利用空间网络操作系统进行抽象描述。网络操作

基于互联网协议(IP)的宽带空间网络已成为研究的主要趋势<sup>[1]</sup>,思科公司在CLEO<sup>[2]</sup>和IRIS<sup>[3]</sup>项目中分别将地面网络路由器部署到了近地轨道(LEO)卫星和地球同步轨道(GEO)卫星上,但基于IP的地面网络本身存在的问题,以及空间网络本身的特殊性,天地一体化网络还面临着

一些挑战:网络拓扑的动态性;空间网络资源限制;天地一体化网络的异构性。

软件定义网络(SDN)是一种新型的网络架构,不同于传统网络中控制平面和数据平面紧密耦合的设计,SDN架构将控制平面和数据平面分离,控制平面可以通过南向接口对数据平面中的网络设备进行集中式控制<sup>[4]</sup>,并提供灵活的可编程能力。在空间网络的情景中,SDN的结构意味

收稿时间: 2016-09-15  
网络出版时间: 2016-11-03

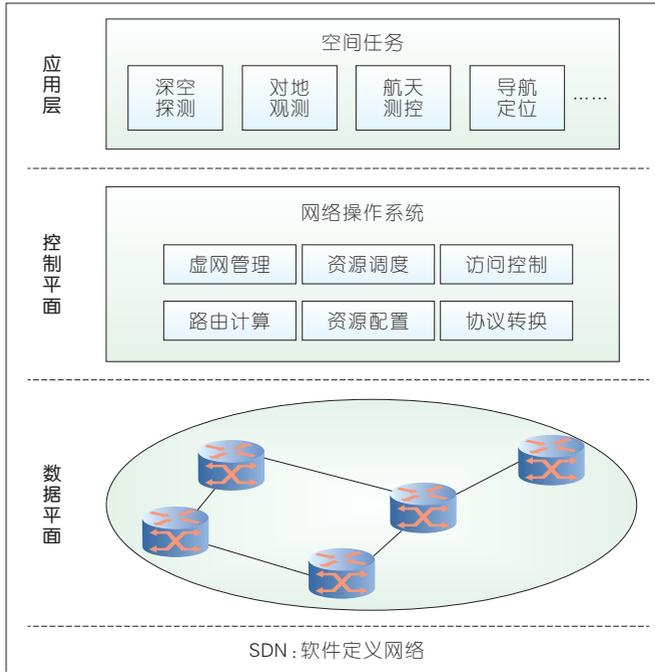


图1 基于SDN的天地一体化网络逻辑架构

系统的主要功能包括虚网管理、资源调度、路由计算、资源配置等,形成整体的网络全局视图,同时多个控制器之间还会进行网络状态信息的同步。数据平面由众多的卫星节点组成,主要负责根据控制平面下发的流表进行相应的数据处理和转发。

### 1.2 演进式的物理架构

设想基于SDN的天地一体化网络物理架构由天基骨干网、数据转发层和地面信息港3层组成。图2b)所示的天地双骨干的物理分层架构由两层卫星网络与地面网络构建而成。根据卫星网络的天然分层特性,将空间信息网络的天基网络分为两层架构,分别为高轨道卫星和中低轨道卫星。GEO层的卫星也担任控制器的职能,采用扁平式的多控制器控制方式进行控制器部署,GEO控制器组成了空间信息港。在扁平式的多控制器部署方案中,各个控制器的功能地位相同,并在逻辑上实现集中控制,每个控制器负责各自的控制区域,在获取本区域的网络状态信息后,控制器之间通过东西向接口进行网络状态信息的通信和交互并形成

网络全局视图。中地球轨道(MEO)/LEO层的卫星相对于GEO层来说,地面可视时间较短,波束之间的切换较频繁,因此不适合部署控制器,但是MEO/LEO层卫星与地面距离较近且数量较多,可以承担SDN网络中交换

机的职责,负责空间信息网络中数据的转发。地面信息港的网络节点与天基骨干网中的GEO卫星节点共同扮演控制器的角色,构成控制平面;而GEO卫星节点和中低轨道的MEO、LEO卫星节点组成数据平面。

由于目前在卫星节点上部署网络设备较为困难,而地面网络技术及设备已比较成熟,因此提出了一种演进式的方案,即在天地一体化网络发展初期,在地面部署控制器,由地面控制器实现逻辑架构中的控制平面功能。随着空间信息网络设备技术的成熟,逐步在GEO卫星上部署控制器,实现天地双骨干多控制器的控制思路。图2a)所示为地面单骨干的物理架构,图2b)所示为天地双骨干的物理架构。通过这种演进方案,兼顾空间网络的现状和未来发展的趋势。

### 1.3 典型的工作流程

由地面信息港和天基骨干网组成的控制平面负责实现路由决策和整个空间信息的管理与控制,数据转发层则根据控制平面下发的指

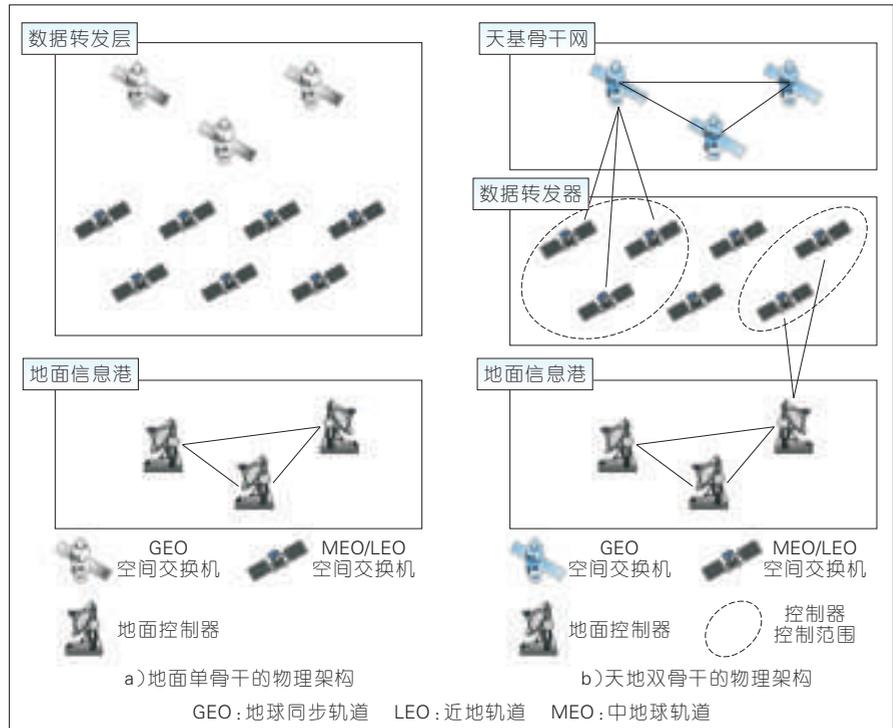


图2 天地一体化网络物理架构的演进

令进行数据转发。控制平面包括多个控制器节点,采用优化的接入控制机制、资源动态调度机制及网络虚拟化技术,完成任务到资源的映射。控制平面向上为研究人员开发各种应用层服务提供统一的北向接口,向下通过南向接口向数据平面下发各种指令,并利用南向接口的上行通道对底层交换节点上报的信息进行统一监测控制和统计,获取网络状态信息,例如:各个节点的带宽使用情况、链路状态、吞吐量、载荷情况等,然后通过东西向接口进行各个控制器间的信息同步,获取网络的全局视图。

数据平面的空间交换节点包含许多不同的流表,每个流表包含很多流表项。当数据包进入空间交换机之后,它将在流表中查找与之相匹配的流表项,若匹配成功,则根据找到的流表项中的指令前往另一个流表,或执行指令集;否则,空间交换节点将请求负责管控自己的控制器节点,为该数据包产生并下发相应的流表项。数据包将通过执行相应指令集而被发送到下一跳节点卫星。

由于MEO、LEO卫星均运行在不同高度、不同位置的轨道上,空间网络的拓扑结构是不断变化的,这也导致空间交换机中的一些流表项随之发生变化。如果空间控制器在拓扑发生变化之后才根据空间交换机的请求对相应的流表项进行更新和下发,一方面会造成短时间内的控制信息量急剧增加,另一方面也会严重影响业务传输的时效性和流畅度。然而由于卫星轨道位置和运行速度是既定的,网络拓扑的变化基本是可预知的,所以空间控制器可以预先计算出变化后的拓扑结构和相应的新流表项<sup>[5]</sup>,在与旧的流表项进行比较之后,选择发生了改变的流表项下发给空间交换机。新的流表项仍然可能与旧的流表项具有一些相同的匹配字段,流表中的优先级和超时字段将帮助交换选择网络拓扑变化前后的匹配流表项。

相比传统网络结构中基于相邻信息的分布式路由协议,利用全局视图提供的信息和多目标优化算法,控制平面对数据平面网络节点的集中式管理可以得到更合理的路由。此外,基于SDN的天地一体化网络架构中集中式的资源分配和调度将支持差异化服务机制,根据流的类型对网络流量进行分类和管理,从而更加有利于保证服务传输质量。

## 2 SDN架构带来的优势

### 2.1 灵活的路由策略

传统的卫星网络中一般使用静态的快照路由方法,以保证网络的可靠性和可控性,但卫星网络的高动态性导致负载均衡等需求得不到满足。一些基于分布式链路信息收集的动态路由方法<sup>[6]</sup>虽然可以实现上述需求中的一部分,但却使得卫星无法便捷地得到网络全局视图,只能达到局部最优的路由。

在基于SDN的架构中,控制平面通过数据平面上传的网络状态信息和控制器之间的通信,获取实时的网络状态全局视图,并以此对数据平面的卫星节点进行集中式的管理,这使得面对高动态性的网络拓扑能够提供更加灵活的全局路由计算和路由策略,如负载均衡、多播路径修正和节点失效管理等。此外,当空间信息网络需要进行扩充或更新时,拥有网络全局视图的控制平面能够有效地更新网络全局配置,让新发射的卫星无缝接入现有的空间信息网络。另一方面,有卫星节点失效或发生故障时,控制平面能够及时地对网络配置做出调整,分配相邻节点负责失效节点的覆盖区域和网络任务,以及对失效节点进行替换。

### 2.2 方便的网络配置

由于空间信息网络资源受限,内存较小,中央处理器(CPU)处理能力较低。与此同时,空间网络应用却在

不断增加,星上载荷需要处理的任务量和复杂度也大大增加。这使得与地面网络相比,空间信息网络的配置十分困难。SDN架构的核心思想——控制转发分离简化了卫星节点的处理功能,在很大程度上缓解了这一问题。数据平面的卫星节点需要做的只是接收和执行控制平面下发的各种配置信息,以及向控制平面反馈自身的网络状态信息。复杂的网络配置和控制功能,以及收集数据平面发来的信息构建网络全局视图等功能都交给天地双骨干的控制平面,由其中的控制器完成。

### 2.3 更好的兼容性

SDN架构具有统一的数据交换标准和编程接口,可在网络具有异构性的情况下对全网设备进行统一管理。SDN架构中的流表对二层转发表、三层路由表进行了抽象处理,整合了各个层次的网络配置信息,能够同时处理在空间信息网络中并存的各种协议,如延迟容忍网(DTN)等,从而很好地解决空间网络协议异构性的问题。

### 2.4 更低的硬件成本

传统卫星网络中,卫星节点要完成繁复的处理任务,因此往往是最复杂、最昂贵的部分。而采用了基于SDN的架构后,数据平面卫星节点只是简单的网络转发设备,这样简化了卫星功能的架构,能有效地减少卫星的设计和生产成本,也能够简化卫星管理的复杂度,让空间信息网络变得更加灵活和可控。另外,天地双骨干的控制平面结构和星间链路转发方式也使所需的地面站数量减少,降低了基础设施方面的投资。

## 3 存在的问题和挑战

### 3.1 控制平面接口设计

对基于SDN的空间网络架构来说,控制平面的南北向接口设计是一

个比较显著的问题。北向接口的设计需要满足卫星网络应用业务的要求,目前学术界和工业界仍在讨论。而南向接口虽然已经有比较成熟的OpenFlow协议<sup>[9]</sup>,但并不能完全提供全局的软硬件控制和策略决定功能。

### 3.2 安全问题

基于SDN的天地双骨干网络架构是一种开放的分布式无线网络结构,开放性作为SDN逻辑架构的一个重要特征,是SDN实现统一管理、配置异构网络、提供可编程特性的基础。但引入SDN架构带来的开放性,使整个网络更加容易受到安全威胁,如网络入侵、网络攻击和拒绝服务等。这是由于SDN的开放性使得控制平面的安全漏洞和策略等暴露在攻击者面前,给了攻击者足够的信息制定攻击策略。与此同时,SDN架构中控制平面提供给应用层的北向可编程接口也可能导致对网络资源的滥用和攻击变得更加频繁。

在网络结构的演进过程中,需要先将控制平面部署在地面信息港,然后逐步将控制器部署在GEO卫星上,完成从地面单骨干到天地双骨干的演进。由于存在上述的许多安全隐患,在部署控制器前需要建立一套隔离防护机制,包括控制器自身的安全防护以及控制平面南、北向接口的安全防护,以确保引入SDN架构的天地一体化网络能够安全稳定地运行。

### 3.3 多控制器协作方案

在基于SDN的天地一体化网络架构中,多控制器结构的部署和协作是非常重要的环节。空间网络具有传输距离远、覆盖面广的特点,若采用单控制器结构,控制器的管控范围过大,跨区域下发流表和通信时会产生额外的时延<sup>[10]</sup>。此外,控制平面负责把握网络的全局视图和对网络的集中控制,采用单控制器不能有效地保障服务质量和安全性。故采取了扁平分布式多控制器架构,控制器部

署在互不相交的网络区域,地位相同,逻辑上集中控制,所有控制器掌握相同的全局视图。

首先,对于扁平式多控制器架构,每个控制器管辖着自己负责的网络区域,这些地位相同的控制器组成控制平面。当涉及跨区域业务时,业务的建立与拆除均需要多控制器的协作,这使得控制器的协作方式极大程度上影响着网络业务承载能力和服务性能。其次,由于控制器储存着所管辖范围内网络的所有信息,一旦控制器失效,需要其他控制器以适当的协作方式将失去控制的部分网络重新控制起来,继续完成控制平面的功能。

## 4 结束语

基于SDN思想的演进式天地一体化网络架构将中低轨卫星节点简化为简单的转发设备,全部的控制功能由部署在地面信息港和GEO层的多控制器组成的控制平面实现,控制平面掌握实时的网络状态全局视图,对数据平面的卫星节点进行集中式的管理,实现了更灵活的路由策略、更方便的网络配置、更好的兼容性。同时,文章提出了演进式架构,该架构能兼顾空间网络亟待发展的需求和高轨卫星部署网络设备困难的现状。随着对多控制器协作方式、控制平面接口设计等问题的进一步研究,基于SDN思想的演进式天地一体化网络架构有望不断地完善。

#### 参考文献

- [1] MUKHERJEE J, RAMAMURTHY B. Communication Technologies and Architectures for Space Network and Interplanetary Internet[J]. IEEE Communication Surveys & Tutorials, 2013, 15(2): 881-897. DOI:10.1109/SURV.2012.062612.00134
- [2] IVANCIC W, STEWART D, DAN S, et al. Secure, Network-Centric Operations of a Space-Based Asset: Cisco Router in Low-Earth Orbit (CLEO) and Virtual Mission Operations Center (VMOC)[R/OL]. (2005-06-24) [2016-05-10] <https://ntrs.nasa.gov/search.jsp?R=20050180616>
- [3] CUEVAS E G, TANG Z. Preliminary Results of a Technology Demonstration of the Internet

Routing in Space Capability over a Geostationary Communications Satellite[C]// Military Communication Conference. USA: IEEE, 2010:1959-1964. DOI: 10.1109/MILCOM.2010.5680418

- [4] BERTAUX L, MEDJIAH S, BERTHOUS P, et al. Software Defined Networking and Virtualization for Broadband Satellite Networks[J]. IEEE Communications Magazine, 2015, 53(3): 54-60. DOI:10.1109/MCOM.2015.7060482
- [5] FENG J, JIANG L, SHEN Y, et al. A Scheme for Software Defined ORS Satellite Networking[C]// Big Data and Cloud Computing (BDCloud), 2014 IEEE Fourth International Conference on. USA: IEEE, 2014: 716-721. DOI:10.1109/BDCloud.2014.19
- [6] EKICI E, AKYILDIZ I F, BENDER M D. A Distributed Routing Algorithm for Datagram Traffic in LEO Satellite Networks[J]. ACM Transactions on Networking, 2001, 9(2): 137-147. DOI:10.1109/90.917071
- [7] OpenFlow Specification v1.1.0[R/OL]. <http://www.openflow.org/wp/documents/>
- [8] Casado M, Freedman M J, Pettit J, et al. Ethane: taking control of the enterprise [J]. ACM Sigcomm Computer Communication Review, 2007, 37(4):1-12.

## 作者简介



许方敏,北京邮电大学副教授、硕士生导师;主要研究方向为无线网络及天地一体化网络架构。



全宗健,北京邮电大学硕士毕业;主要研究方向为无线网络及天地一体化网络架构。



赵成林,北京邮电大学教授、博士生导师;主要研究方向为无线通信及信号处理、认知无线电、软件定义网络、物联网技术及应用。



秦智超,北京邮电大学博士;主要研究方向为物联网、天地一体化网络。

# 《中兴通讯技术》杂志(双月刊)投稿须知

## 一、杂志定位

《中兴通讯技术》杂志为通信技术类学术期刊。通过介绍、探讨通信热点技术,以展现通信技术最新发展动态,并促进产学研合作,发掘和培养优秀人才,为振兴民族通信产业做贡献。

## 二、稿件基本要求

### 1. 投稿约定

- (1)作者需登录《中兴通讯技术》投稿平台:tech.zte.com.cn/submission,并上传稿件。第一次投稿需完成新用户注册。
- (2)编辑部将按照审稿流程聘请专家审稿,并根据审稿意见,公平、公正地录用稿件。审稿过程需要1个月左右。

### 2. 内容和格式要求

- (1)稿件须具有创新性、学术性、规范性和可读性。
- (2)稿件需采用WORD文档格式。
- (3)稿件篇幅一般不超过6000字(包括文、图),内容包括:中、英文题名,作者姓名及汉语拼音,作者中、英文单位,中文摘要、关键词(3~8个),英文摘要、关键词,正文,参考文献,作者简介。
- (4)中文题名一般不超过20个汉字,中、英文题名含义应一致。
- (5)摘要尽量写成报道性摘要,包括研究的目的、方法、结果/结论,以150~200字为宜。摘要应具有独立性和自明性。中英文摘要应一致。
- (6)文稿中的量和单位应符合国家标准。外文字母的正斜体、大小写等须写清楚,上下角的字母、数据和符号的位置皆应明显区别。
- (7)图、表力求少而精(以8幅为上限),应随文出现,切忌与文字重复。图、表应保持自明性,图中缩略词和英文均要在图中加中文解释。表应采用三线表,表中缩略词和英文均要在表内加中文解释。
- (8)参考文献以20条左右为宜,不允许公开发表的资料不应列入。所有文献必须在正文中引用,文献序号按其在文中出现的先后次序编排。常用参考文献的书写格式为:
  - 期刊[序号]作者.题名[J].刊名,出版年,卷号(期号):引文页码.数字对象唯一标识符
  - 书籍[序号]作者.书名[M].出版地:出版者,出版年:引文页码.数字对象唯一标识符
  - 论文集析出文献[序号]作者.题名[C]//论文集编者.论文集名(会议名).出版地:出版者,出版年(开会年):引文页码.数字对象唯一标识符
  - 学位论文[序号]作者.题名[D].学位授予单位所在城市名:学位授予单位,授予年份.数字对象唯一标识符
  - 专利[序号]专利所有者.专利题名:专利号[P].出版日期.数字对象唯一标识符
  - 国际、国家标准[序号]标准名称:标准编号[S].出版地:出版者,出版年.数字对象唯一标识符
- (9)作者超过3人时,可以感谢形式在文中提及。作者简介包括:姓名、工作单位、职务或职称、学历、毕业于何校、现从事的工作、专业特长、科研成果、已发表的论文数量等。
- (10)提供正面、免冠、彩色标准照片一张,最好采用JPG格式(文件大小超过100kB)。
- (11)应标注出研究课题的资助基金或资助项目名称及编号。
- (12)提供联系方式,如:通信地址、电话(含手机)、Email等。

### 3. 其他事项

- (1)请勿一稿两投。凡在2个月(自来稿之日算起)以内未接到录用通知者,可致电编辑部询问。
- (2)为了促进信息传播,加强学术交流,在论文发表后,本刊享有文章的转摘权(包括英文版、电子版、网络版)。作者获得的稿费包括转摘酬金。如作者不同意转摘,请在投稿时说明。

编辑部地址:安徽省合肥市金寨路329号国轩凯旋大厦1201室,邮政编码:230061

联系电话:0551-65533356,联系邮箱:magazine@zte.com.cn

本刊只接受在线投稿,欢迎访问本刊投稿平台:tech.zte.com.cn/submission

办刊宗旨

以人为本,荟萃通信技术领域精英  
迎接挑战,把握世界通信技术动态  
立即行动,求解通信发展疑难课题  
励精图治,促进民族信息产业崛起

---

双月刊 1995年创刊 总第139期  
2018年4月 第24卷2期

主管:安徽省科学技术厅  
主办:安徽省科学技术情报研究所  
中兴通讯技术杂志社  
出版:中兴通讯技术杂志社

总编:陈杰  
常务副总编:黄新明  
责任编辑:徐烨  
编辑:卢丹,朱莉,赵陆  
排版制作:余刚  
发行:王萍萍  
编务:王坤

---

《中兴通讯技术》编辑部  
地址:合肥市金寨路329号凯旋大厦12楼  
邮编:230061  
网址:tech.zte.com.cn  
投稿平台:tech.zte.com.cn/submission  
电子信箱:magazine@zte.com.cn  
电话:(0551)65533356  
传真:(0551)65850139

编辑、发行:《中兴通讯技术》编辑部  
发行范围:公开发行  
印刷:合肥添彩包装有限公司  
出版日期:2018年4月10日  
中国标准连续出版物号:ISSN 1009-6868  
CN 34-1228/TN  
定价:每册20.00元