



研究与开发

5G 轻量化方法及实现路径研究

仇剑书, 朱子园, 师瑜

(中国联合网络通信有限公司研究院, 北京 100176)

摘要: 5G 轻量化是通过对性能指标的适当裁剪降低终端成本和功耗, 以满足物联网等应用场景的需求。结合第三代合作伙伴计划 (3rd Generation Partnership Project, 3GPP) 的 RedCap 技术规范, 研究了 5G 轻量化方法、收益以及其对网络的影响, 并对端网协同精简、终端自主精简两种轻量化实现方式进行对比分析。在此基础上提出一种融合化的 5G 轻量化终端实现方案, 使终端产品可以更好地适应不同运营商差异化的 5G 轻量化部署策略。

关键词: NR 轻量化; 能力裁剪; 端网协同精简; 终端自主精简; 融合化轻量终端

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2024221

Research on 5G lite approaches and implementation strategies

QIU Jianshu, ZHU Ziyuan, SHI Yu

Research Institute of China United Network Communications Co., Ltd., Beijing 100176, China

Abstract: 5G lite is achieved by appropriately reducing performance metrics to trade off for lower terminal costs and power consumption, thereby meeting the needs of application scenarios such as the IoT. In conjunction with the 3GPP's RedCap technology specifications, 5G light-weighting methods, their benefits, and impacts on the network was presented. A comparative analysis of two implementation approaches was also conducted: terminal and network coordinated reduction, and terminal self-reduction. Based on this analysis, a converged 5G lite terminal implementation scheme was proposed, enabling terminal products to better adapt to varied 5G lite deployment strategies of different operators.

Key words: NR lite, RedCap, terminal and network coordinated reduction, terminal self-reduction, converged lite terminal

0 引言

国际电信联盟 (ITU) 为 5G 移动通信定义了三大典型应用场景, 即增强型移动宽带 (eMBB)、海量机器类通信 (mMTC) 和超可靠低时延通信

(URLLC) [1], 第三代合作伙伴计划 (3rd Generation Partnership Project, 3GPP) 为此制定了新一代无线 (new radio, NR) 技术规范。NR 技术在 eMBB 场景可实现千兆级传输速率, 目前已经被大规模应用。但自 5G 商用以来, 以降低成本为目

的,终端轻量化方案时常被提及。例如,5G手机对于天线配置的要求,已从最初的2T4R降低为1T4R^[2],个别厂商甚至推出过1T2R产品。

对大多数物联网场景而言,NR技术所定义的高性能严重过剩。移动物联网按传输速率可大概分为4类场景,即超高速场景(>100 Mbit/s)、中高速场景(10~100 Mbit/s)、中低速场景(100 kbit/s~10 Mbit/s)、低速场景(<100 kbit/s)。目前,低速场景一般使用NB IoT技术,中低速场景一般使用LTE Cat1技术,中高速场景一般使用LTE Cat4技术。NR技术仅适用于超高速场景,针对其他场景可以适当降低性能,以达到降低终端成本、功耗和尺寸的目的。5G轻量化是在NR终端的基础上裁剪硬件能力,当前目标是使传输速率仅维持在与LTE Cat4相当的水平。配合5G所特有的切片、大上行、URLLC、5G LAN等能力,可以更好地满足工业传感器、视频监控和可穿戴设备等中端物联网应用的需求。为此3GPP在Release17版本中也提出一种5G轻量化实现方案,即5G RedCap^[3]。

为了便于描述,本文将支持全部性能指标的5G技术规格称为“完整版NR”,将性能裁剪之后的技术规格称为“轻量化NR”。同时,本文研究主要针对FR1频段,与FR2频段略有区别,但基本原则一致。

1 5G轻量化基本原理

1.1 降低传输速率

(1) 裁剪参数

单用户峰值速率是5G最关键的性能指标之一^[1]。决定峰值速率的参数包括:单位时间内的资源元素(resource element, RE)数量、调制阶数、MIMO流数、控制面开销等^[4]。其中,单位时间内的资源元素数量由载波带宽决定,MIMO流数与天线数量密切相关。对于工作在TDD模式的载波,上下行峰值速率还受上下行时隙配比的影响。

一般而言,5G网络的核心载波带宽为100 MHz,使用TDD模式;完整版NR终端的天线配置是2T4R,上下行均支持256QAM调制。在采用2.5 ms双周期帧结构、典型时隙配比的情况下,根据NR峰值速率计算式^[4]可计算得到理论峰值速率为下行1.4 Gbit/s/上行430 Mbit/s。

在NR标准中,载波带宽、调制阶数、MIMO流数/天线数等属于可调整参数。由于NR与LTE均采用OFDM调制技术^[5]、多天线MIMO技术^[6-7],且帧结构类似,要获得与LTE相当的单用户性能,只需将相关参数裁剪到与LTE相当。例如,可以将终端支持的载波带宽收窄到20 MHz,支持的调制阶数可下调为6(64QAM),MIMO流数/天线数降低为1T2R或1T1R。仍以采用2.5 ms双周期帧结构为例,对裁剪前后的单用户速率进行概算,TDD模式性能裁剪后的峰值速率变化见表1。

(2) 成本收益分析

终端成本一般是指5G通信接口模块的成本,其核心功能包括基带、射频收发两个部分^[8]。其中,收窄载波带宽和减少MIMO流数/天线数对成本影响较大。收窄载波带宽主要影响基带成本,尤其是接收处理模块、HARQ缓冲区、FFT数据缓存、AD/DA转换器等模块。减少MIMO流数/天线数则可以降低射频部分的收发机以及基带部分的接收处理等模块成本。根据估算^[9],轻量化后终端成本有显著下降。性能裁剪带来成本降低(TDD模式)见表2,可以看出对于TDD模式,终端支持带宽从100 MHz收窄为20 MHz、下行MIMO层数/接收天线数从4减少为2所带来的成本收益。此外,最大调制阶数降为6(64QAM),也可以带来约6%的成本降低。

1.2 降低移动性能

物联网是轻量化NR的主要应用场景。有的物联网终端虽然接入移动网络,但其实是在静止状态下使用;有的物联网终端数据收发频率较



表1 TDD模式性能裁剪后的峰值速率变化

对比项	完整版NR		轻量化NR	
载波带宽/MHz	100		20	
子载波间隔/kHz	30		30	
资源块(RB)数	273		54	
RE个数/时隙	45 864		9 072	
时隙数/s	2 000		2 000	
最大有效信道码率	92.5%		92.5%	
传输方向	下行	上行	下行	上行
调制阶数	8(256QAM)	8(256QAM)	6(64QAM)	6(64QAM)
MIMO流数	4	2	2	1
信道开销	14%	8%	14%	8%
符号占比	0.64	0.33	0.64	0.33
峰值速率	1.4 Gbit/s	390 Mbit/s	104 Mbit/s	29 Mbit/s

表2 性能裁剪带来成本降低(TDD模式)

功能模块		收窄终端支持的带宽	减少MIMO流数/天线数
射频	功率放大器	-1.20%	0
	滤波器	-0.30%	-7.40%
	收发机	-2.00%	-24.60%
	双工器/转换器	0	-0.10%
	射频总成本降低	-3.50%	-32.10%
基带	数模转换器/模数转换器	-7.00%	-3.80%
	快速傅里叶变换/反变换	-2.90%	-1.80%
	快速傅里叶变换后数据缓存	-7.90%	-4.70%
	接收机处理模块	-19.10%	-13.30%
	LDPC解码	-5.50%	-0.30%
	HARQ缓冲区	-8.70%	-0.40%
	下行控制处理和解码器	-0.30%	0
	同步/小区搜索模块	0	-4.20%
	上行处理模块	-1.30%	0
	MIMO专用处理模块	-0.60%	-1.10%
	基带总成本降低	-53.30%	-29.60%
射频与基带综合成本降低		-33.40%	-30.60%

低,对连续性通信要求不高。基于这些特点,可以有针对性地降低终端移动性能,使终端更多时间处于休眠状态,从而实现终端节能的效果。例如,可以增加扩展非连续性接收(extended discontinuous reception, eDRX)周期,可以放松对邻区的测量要求、减少单位时间内测量频次。评估结果表明,在RRC空闲态或非激活态,终端增

大DRX周期至10.24 s以上有明显的节能增益,测量周期扩大4倍时,可以获得3.6%~13.4%的节电增益;在连接态,测量周期扩大4倍时,可以获得11.1%~26.6%的节电增益^[9-10]。

1.3 其他轻量化方法

(1) 支持半双工FDD模式

在3GPP Release15/Release16版本中,NR在

频分双工（FDD）频段只可以使用全双工（FD-FDD）模式。如果以半双工方式工作，可以使用开关器件和低通滤波器来代替双工器，预计可以使成本降低约 7%^[10]。例如，在 RedCap 中就新引入了半双工 FDD（HD-FDD）模式。但 HD-FDD 功能在实际应用中的价值仍有待观察。

（2）降低缓存需求

由于轻量化之后传输速率降低，对缓存的需求也相应减少。例如，在 RedCap 规范中通过降低部分协议要求来减少缓存，包括：PDCP 序列号长度和 RLC AM 序列号长度都只需支持 12 bit、18 bit 为可选^[4]。

（3）裁剪部分增强特性

轻量化 5G 终端可以不支持 CA、MR-DC、DAPS、CPAC 以及 IAB 等功能^[4]。

2 轻量化对网络性能的影响

现有 5G 网络是针对完整版 NR 终端进行规划和优化的。在引入轻量化终端之后，会对 5G 网络性能造成一定不利影响。这主要体现在网络覆盖和网络容量两个方面能力下降。

（1）天线数量影响网络覆盖能力

轻量化终端的天线数量从 2T4R 减少为 1T2R，使终端在小区边缘的性能显著降低。以 n78 频段（3.5 GHz）为例，在采用通常建议的 353 m 站间距规划的情况下，完整版 NR 终端边缘速率为上行 5 Mbit/s、下行 100 Mbit/s，轻量化 NR 终端边缘速率会降低为上行 1.6 Mbit/s、下行 71.8 Mbit/s，降幅比较明显^[11]。其中上行方向的性能受影响程度更大。物联网应用往往对上行性能的要求比较高。对于视频监控、车联网等场景，一般要求速率不低于 2 Mbit/s^[12]，否则会影响正常使用。因此网络运营商在引入轻量化终端时，需要重点考虑自身网络的覆盖情况，必要时需使用补充上行、超级上行等技术增强上行覆盖能力^[13]。另一方面，轻量化终端在功耗方面有较

好的表现。基于某款实际终端的测试表明，轻量化 NR 终端在同等情况下的功耗约为完整版 NR 终端的 55%~60%。

（2）天线数量和调制阶数影响网络容量

5G RedCap 规范允许在 100 MHz 载波带宽上配置 5 个互不重叠的带宽分段（BWP）。但即使在满配情况下，RedCap 容量还是会低于完整版 NR。其中，当业务负荷较高时，对网络容量的影响较为明显。根据仿真评估，如果使用 2 根接收天线、64 QAM 调制方式，可能导致 TDD 频谱效率最高降低 30%~50%^[14]。在某 TDD 网络的实际测试中，测得 RedCap 的下行容量是完整版 NR 的 54%，上行容量则为 60%。不过，物联网应用通常不是高负荷场景。因此，在实际使用中轻量化对网络容量的影响会优于仿真和测试结果。

3 轻量化实现方式

5G 轻量化可以有两种不同的实现方式。一种由网络与终端协同精简，即在终端设备进行功能裁剪的同时，网络侧也进行相应调整和优化，使两者之间高效匹配。3GPP 的 5G RedCap 是这种方式的代表。另一种实现方式是终端自主精简，即仅在终端侧独立进行能力裁剪，网络设施保持不变。第一种方式已由标准定义，本文重点讨论后一种方式。

3.1 终端自主精简的可行性

终端自主精简方式在功能实现上相对简单。3GPP NR 技术规范已为终端赋予了较高灵活性，并不要求终端必须支持最高性能指标。终端接入网络时需上报自身的能力参数，网络据此进行资源调度^[15]。例如，终端可自主简化为 20 MHz 带宽、1T2R 天线配置、64QAM 调制。对于带宽的调整，如果载波带宽较大，为了支持简化后的终端接入，可在网络侧配置相应的窄带宽下行 BWP。对于天线配置和调制阶数的调整，现有空口协议、信令以及网络侧能力都可以直接支持。



3.2 终端自主精简面临的问题

尽管终端自主精简具有可行性，但其能力裁剪也会带来一些性能上的问题。除了前文所述的减少天线数量、降低调制阶数对网络性能带来的不利影响之外，收窄载波带宽也会给网络配置方面带来一些问题。在端网协同精简方式中，对这些问题有针对性的解决方案。对于终端自主精简方式而言需要接受性能降低的结果。

(1) 初始下行BWP问题

如果终端支持的带宽收窄到20 MHz，则要求载波上配置的初始下行BWP也必须小于或等于20 MHz，终端才能完成初始接入。在实际网络配置中，为了提高接入性能，有时候初始下行BWP会大于20 MHz。例如，有的小区会配置为100 MHz。为了解决这个问题，在RedCap规范中引入了RedCap独立初始下行BWP的概念，即在已有配置之外，可以为轻量化终端另外配置一个不大于20 MHz的专用初始下行BWP^[3]。如果采用终端自主精简方式，则必须修改小区的初始下行BWP配置。

(2) BWP包含SSB问题

从理论上来说，下行BWP可以不包含SSB所在的带宽。但是，在终端处于连接态、工作于激活BWP中时，也需要通过SSB来进行小区测量。因此，为了降低终端实现的复杂性、避免频繁进行RF调谐，一般建议将下行BWP配置为涵盖SSB^[16]。为了用满100 MHz带宽，在RedCap规范中引入了非小区定义SSB（non-cell defining SSB，NCD-SSB）的概念，原SSB则称为小区定义SSB（cell defined SSB，CD-SSB）。BWP与SSB的关系如图1所示，当轻量化终端的BWP不包含SSB所在带宽时，可以为其配置NCD-SSB^[3]。与CD-SSB相比，NCD-SSB较为简单，不携带完整的系统信息，终端使用时的复杂性和功耗要求较低。如果采用终端自主精简方式，其使用的20 MHz的BWP一般会配置在SSB

频点附近。由此，当接入小区的轻量化终端数量较多时，会造成这部分频带拥塞，而其他频带则得不到充分利用。

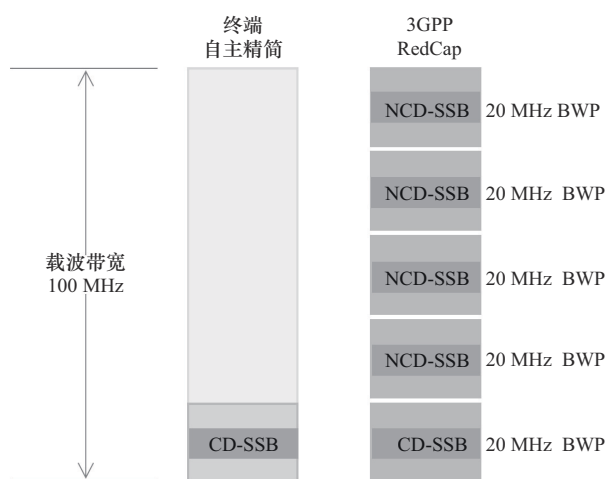


图1 BWP与SSB的关系

3.3 带宽较小的载波

在实际网络中，5G载波并不总是100 MHz。例如，国内运营商的部分5G频段，包括n1、n5、n8等，从2G/3G/4G重耕而来。这些频段带宽都不超过20 MHz（单向）^[12]。在这些频段上，更适合终端自主精简的实现方式。使用端网协同的轻量化方式并不能带来太多收益，反而会增加网络设备的复杂度。其原因，一是在小带宽频段上，终端无须特意收窄带宽，也就不存在由此带来的配置问题；二是小带宽频段的频率一般较低，通常使用FDD模式，所需天线数量也较少，减少天线数量对这些频段的性能影响也较小。

4 融合化方案

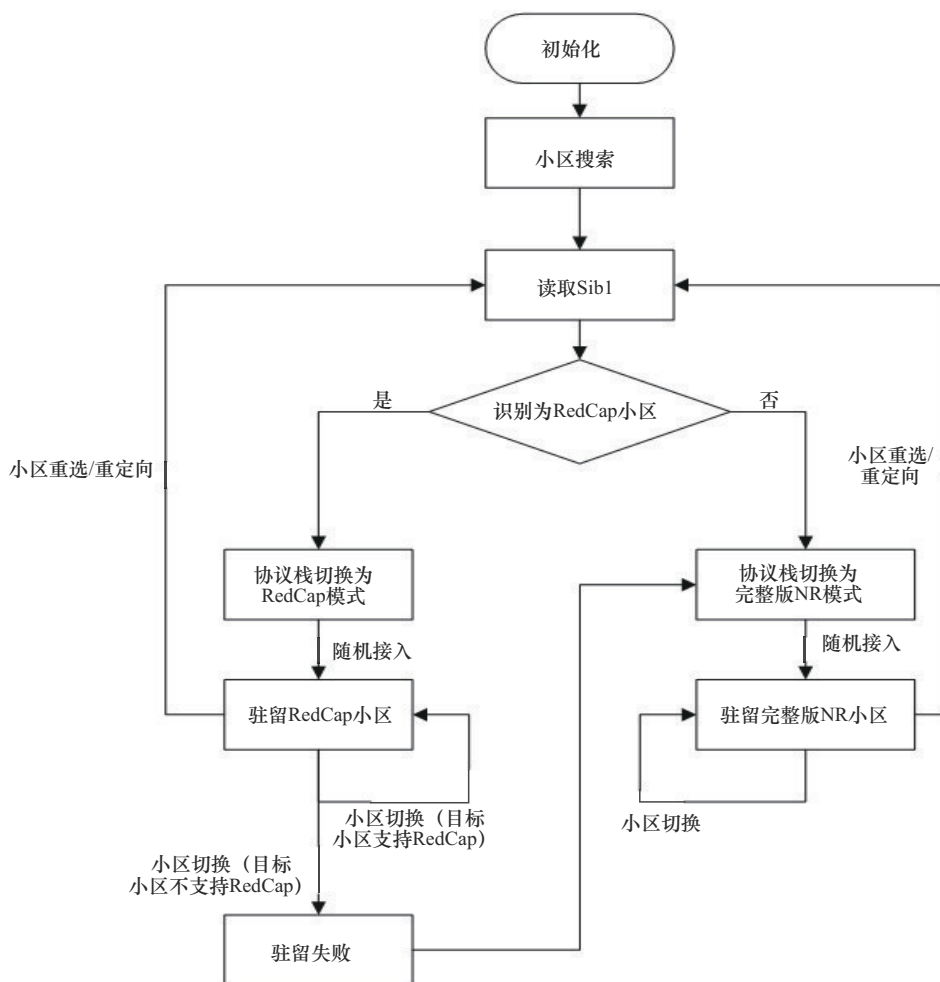
5G RedCap所提供的端网协同精简方案是实现5G轻量化的一条稳妥路径。如果运营商的5G频谱带宽远超20 MHz，且物联网业务量比较大，则应优先考虑采用RedCap方案。但是，如果运营商的5G频谱带宽不超过20 MHz，则RedCap并非必需。同时还有部分运营商既有大带宽频段，也有小带宽频段，可以根据业务需求选择

用正确的协议。通过使用NR协议，在特定情况下可以避免5G基站升级，从而降低运营商的建网成本。

4.1 RedCap 小区识别

5G 小区可以在 Sib1 系统广播消息中下发 IFRI (IntraFreqReselectionRedCap) 标识, 以此指示该小区是否支持 RedCap。终端在开机并完成小区搜索后, 需读取小区系统消息, 如果在 Sib1 中读取到 IFRI, 则认为该小区支持 RedCap, 终端可以执行 RedCap 的随机接入过程^[3,16]。

融合化轻量终端基本处理流程如图2所示。在首次接入、小区重选、小区重定向等过程中，终端需识别出待接入小区是否支持RedCap，如





果支持,则将协议栈切换为RedCap模式,在驻留该小区的过程中使用RedCap协议;如果不支持,则将协议栈切换为完整版NR模式,在驻留该小区的过程中使用完整版NR协议。

小区切换情况有所不同。当驻留在RedCap小区的终端发生小区切换时,网络会在信令消息中携带目标小区信息指示终端直接发起切换流程,不能提前通过Sib1消息进行小区识别。因此,如果目标小区不支持RedCap,则终端切换过程会发生失败。此时,需要终端将协议栈切换为完整版NR模式,并驻留到完整版NR小区。当驻留在完整版NR小区的终端发生小区切换时,由于不能提前通过Sib1消息进行小区识别,即使目标小区支持RedCap,终端也将保持完整版NR模式完成驻留。

5 结束语

3GPP 5G RedCap被视为推动5G物联网规模化发展的关键技术,目前已经完成产品化研发,进入商用发展阶段^[17]。但5G轻量化目的是简化终端,其实现路径并非只有一条。只有根据业务发展实际需要制定合适的部署策略,对网络与终端进行综合考量,选择最恰当的技术方案,才能真正实现成本与性能的平衡。

参考文献:

- [1] ITU. Minimum requirements related to technical performance[R]. 2017.
- [2] 工业和信息化部. 5G数字蜂窝移动通信网 增强移动宽带终端设备技术要求(第一阶段): YD/T 3627—2024[S]. 北京: 人民邮电出版社, 2024.
MIIT. 5G digital cellular mobile telecommunication network technical specification of eMBB user equipment (Phase 1): YD/T 3627—2024[S]. Beijing: Posts & Telecom Press, 2024.
- [3] 中国通信标准化协会. NR RedCap 关键技术研究: SR 448-2023[R]. 2023.
CCSA. NR RedCap key technology research: SR 448-2023[R]. 2023.
- [4] 3GPP. User Equipment (UE) radio access capabilities (Release 17): TS 38.306[S]. 2022.
- [5] 黄陈横. 3GPP 5G NR 物理层关键技术综述[J]. 移动通信, 2018, 42(10): 1-8.
HUANG C H. Review on physical layer key technology of 3GPP 5G NR[J]. Mobile Communications, 2018, 42(10): 1-8.
- [6] 姜东玉. LTE网络4×4 MIMO理论分析及现阶段应用探索[J]. 电信技术, 2018, 8(10): 83-84.
JIANG D Y. Theoretical analysis and application of 4×4 MIMO in LTE networks[J]. Telecommunications Technology, 2018, 8(10): 83-84.
- [7] 李艳芬, 朱雪田. 面向5G Massive MIMO部署的挑战研究[J]. 电信技术, 2019(1): 65-67.
LI Y F, ZHU X T. Research on challenges of 5G massive MIMO deployment[J]. Telecommunications Technology, 2019(1): 65-67.
- [8] 5G应用产业方阵. 精简化5G芯片能力和技术要求: 011-2024[R]. 2024.
5G AIA. The specification of ability and technical requirements of simplified 5G chips: 011-2024[R]. 2024.
- [9] 3GPP. Study on support of reduced capability NR devices (Release 17): TR38. 875[R]. 2021.
- [10] 李晗阳, 翁玮文, 李男, 等. 5G NR RedCap关键技术研究[J]. 电信科学, 2022, 38(3): 93-101.
LI H Y, WENG W W, LI N, et al. Research of 5G NR RedCap technology[J]. Telecommunications Science, 2022, 38(3): 93-101.
- [11] 李宏辉, 关成哲. NR RedCap覆盖能力分析与增强技术[J]. 通信世界, 2022(23): 43-45.
LI H H, GUAN C Z. Ability analysis and enhancement technology of NR RedCap coverage [J]. Communications World, 2022(23): 43-45.
- [12] 中国联通. 中国联通5G RedCap白皮书[R]. 2022.
CUCC. China Unicom 5G redcap white paper[R]. 2022.
- [13] 胡煜华, 王鑫炎, 李贝. 5G网络上行覆盖增强研究[J]. 电信科学, 2021, 37(7): 134-141.
HU Y H, WANG X Y, LI B. Research on uplink coverage enhancement of 5G network[J]. Telecommunications Science, 2021, 37(7): 134-141.
- [14] IMT2020. 适配RedCap的网络升级方案研究报告[R]. 2022.
IMT2020. Network upgrade solution report for RedCap[R]. 2022.
- [15] 3GPP. Radio resource control (RRC) protocol specification (Release 17): TS 38.331[S]. 2022.
- [16] 黄陈横. 5G BWP关键技术及应用研究[J]. 邮电设计技术, 2021(12): 44-49.
HUANG C H. Research on key technology and application of 5G BWP[J]. Designing Techniques of Posts and Telecommuni-

cations, 2021(12): 44-49.

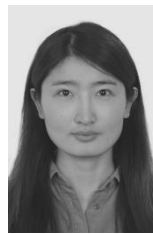
- [17] 黎卓芳. RedCap 标准及产业进展[J]. 通信世界, 2023(15): 7-8.

LI Z F. RedCap standard and industrial progress[J]. Communications World, 2023(15): 7-8.

[作者简介]



仇剑书 (1974-), 男, 中国联合网络通信有限公司研究院首席研究员、正高级工程师, 主要研究方向为移动通信、智能终端、电信智能卡、信息安全、端网业协同等。



朱子园 (1994-), 女, 中国联合网络通信有限公司研究院研究员, 主要研究方向为 5G/6G 终端关键技术标准及新技术验证。



师瑜 (1988-), 女, 中国联合网络通信有限公司研究院首席研究员、高级工程师, 主要研究方向为 5G 终端演进、RedCap、端网协同技术。