

文章编号:1007-5321(2018)05-0026-18

DOI:10.13190/j.jbupt.2018-195

第5代移动通信系统的设计与标准化进展

孙韶辉¹, 高秋彬¹, 杜滢², 刘晓峰², 艾明¹

(1. 电信科学技术研究院(大唐电信)无线移动通信国家重点实验室, 北京 100191;

2. 中国信息通信研究院标准研究所, 北京 100191)

摘要: 通过对第5代移动通信系统(5G)需求指标、国际标准化组织工作计划的介绍以及基于3GPP R15的5G新空口物理层关键技术、空口协议栈过程、接入网架构与接口设计、基于服务化架构的核心网关键技术,描述了一个较为完整的5G系统设计和标准化进展情况,使读者能够对5G技术整体设计思路、方案和特点有一个清晰认识,并对5G标准化内容有较为充分地了解。

关键词: 第5代移动通信系统; 接入网; 核心网; 波束赋形; 网络架构

中图分类号: TN929.53

文献标志码: A

Overview on the Progress of Design and Standardization of the Fifth Generation of Mobile Communications System

SUN Shao-hui¹, GAO Qiu-bin¹, DU Ying², LIU Xiao-feng², AI Ming¹

(1. China Academy of Telecommunications Technology, Beijing 100191, China;

2. China Academy of Information and Communications Technology, Beijing 100191, China)

Abstract: With the continuous development of information technology, the key technology and system design of the fifth generation of mobile communications system (5G) has become a hotspot in both academia and industry. A complete overview of 5G design and standardization progress, including the key performance index (KPI) of 5G was provided, the work plan of international standardization organization on 5G, and 5G design based on 3GPP R15. Specifically, the key technology of physical layer, the protocol stack and procedure of air interface, the radio access network architecture and interface design, the key technology and architecture of core network were introduced and analyzed. We try to provide a whole picture of 5G design, from radio access network to core network. Therefore, a clear understanding of the overall design ideas, solutions and features of 5G technology, as well as the 5G standardized contents can be achieved.

Key words: the fifth generation of mobile communications system; radio access network; core network; beamforming; network architecture

1 5G需求与标准规划

1.1 5G需求与技术指标

移动通信技术在过去近40年的时间里得到了

迅猛发展,给人们的生活方式、工作方式以及社会的政治、经济都带来了巨大的影响,已经实现了1G~4G的商业应用。从2012年起,第5代移动通信系统(5G, the fifth generation of mobile communications

system) 成为业界热点, 目前 5G 的标准化和产业化工作取得了长足进步, 很多国家和组织都宣布了 5G 的商用计划, 预期 5G 的最早商业化部署大约在 2019 年下半年, 而大规模商业化运营将从 2020 年开始。

移动通信大约每 10 年发展一代, 每一代都体现了强烈的技术特色和业务能力。图 1 展示了从第 1 代移动通信系统支持纯语音通信, 到 5G 大数据高速连接以及广泛指标的物联网业务的变化。

移动互联网和物联网是 5G 发展的两大驱动力。移动互联网主要面向以人为主体的通信, 注重提供更好的用户体验。移动物联网主要面向物与物、人与物的通信, 不仅涉及普通个人用户, 也涵盖了大量不同类型的行业用户。为了适应 2020 年之后移动互联网和物联网业务的快速发展, 5G 面临巨大需求挑战。

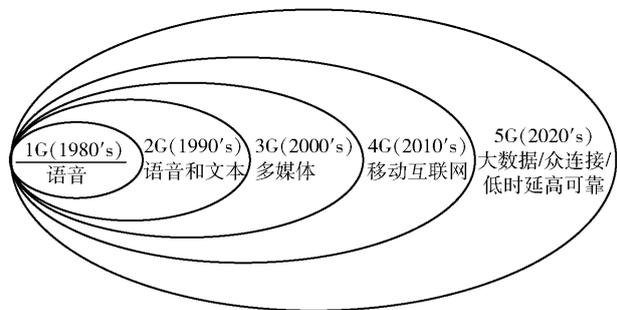


图 1 从 1G ~ 5G 支撑业务类型的变化

5G 面临的 4 个挑战主要包括: 未来 10 年移动数据业务的爆炸性增长, 预计将有千倍的容量增长和超高速率的传输需求; 物联网业务类型丰富多样, 连接数量巨大, 业务特征差异明显, 产生大连接数量和业务指标要求极端的需求; 如果沿袭原有的移动通信发展方式, 通信系统的耗电量将上百倍地增加,

通信节能的需求强烈; 为了满足业务和部署场景的多样化, 5G 需要满足多样、灵活等网络部署需求。

针对 5G 需求, 中日美韩等国都提出了对 5G 发展的趋势理解以及需求技术指标, 其中中国提出“5G 需求之花”的 9 项需求得到了广泛认可^[1]。国际电信联盟 (ITU, international telecommunication union) 经过多轮讨论, 最终确定了 5G 的三大类应用场景和技术指标^[2]。

图 2 所示为 ITU 确定的三大应用场景: 应用于移动互联网的增强移动宽带 (eMBB, enhanced mobile broadband), 它可以进一步分为广域连续覆盖和局部热点覆盖 2 个子应用场景; 应用于物联网的大连接物联网 (mMTC, massive machine type communications) 和高性能物联网 (URLLC, ultra reliable and low latency communications)。三大场景包括 8 个指标, 其中 eMBB 重点关注 6 个指标, mMTC 重点关注 2 个指标, URLLC 则强调传输时延指标。

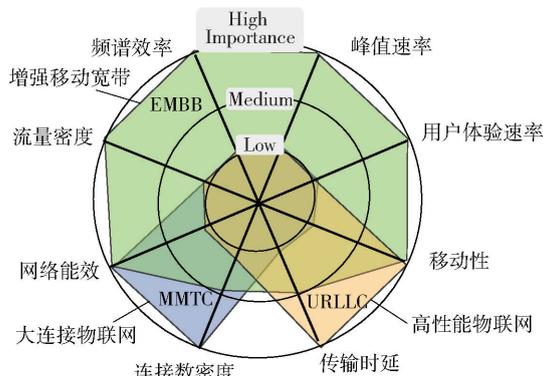


图 2 5G 系统的三大类应用场景和指标关系

图 2 所对应的具体指标要求如表 1 所示。4G 与 5G 的性能指标对比如图 3 所示。

表 1 ITU 制定的 5G 系统性能指标

指标名称	流量密度	连接数密度	传输时延	移动性	网络能效	用户体验速率	频谱效率	峰值速率
性能指标	10 Tbit · s ⁻¹ · (km ²) ⁻¹	1 × 10 ⁶ /km ²	空口 1 ms	500 km/h	100 倍提升 (相对 4G)	0.1 ~ 1 Gbit/s	3 倍提升 (相对 4G)	10 Gbit/s

1.2 5G 关键技术与标准规划

为了满足 ITU 制定的 3 大应用场景技术指标, 需要采用先进的无线传输与网络传输关键技术。

在无线传输技术方面, 关键技术包括大规模波束赋形、超密集组网、新型编码调制、毫米波高频段通信、终端直通 (D2D, device to device) 等。其中大规模波束赋形通过数十到数百根天线实现多用户的

空分复用, 可以有效提升频谱效率 5 ~ 10 倍。设计和研究超大规模的天线结构和高精度的波束赋形算法、传输方案成为提升大规模波束赋形性能的难点。超密集组网研究在热点地区提供超高速传输速率和容量的实现技术, 可以提供百倍的超高流量密度, 其研究关键在于新型接入网架构、干扰管理、承载管理等; 新型低密度校验 (LDPC, low density parity

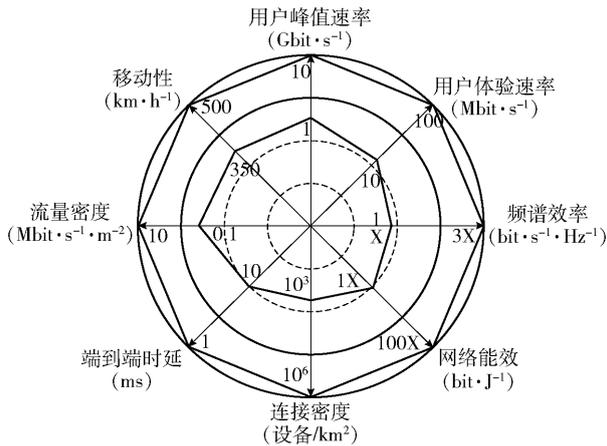


图 3 4G 和 5G 在关键性能指标方面的对比

check) 码和极化 (Polar) 码可以为 5G 提供不同性能需求的多样化业务和部署场景下的可靠传输, 并提升频谱效率. 在 LDPC 中如何设计满足高可靠性需求的信道编码以及在 Polar 码中如何实现低时延和低复杂度的高性能编码方案成为设计中的难点. 高频段提供超高带宽的频谱资源, 通过波束的跟踪和管理, 可以实现超高速传输和支持大容量业务. 如何设计低时延、高可靠性的波束配对和恢复机制成为研究的重点. D2D 是实现不同应用场景和需求的物联网业务关键技术, 通过终端间直接通信, 可使传输形式更为灵活, 降低传输时延, 并能够与蜂窝网络形成异构层叠结构, 提升系统容量.

在核心网传输方面, 新业务的出现使得核心网面临高速传输、垂直行业的高可靠性、低时延和大连接等挑战; 同时虚拟化、软件定义网络、云计算、软件化等 IT 新兴技术的日渐成熟给通信行业的变革带来了新机遇. 5G 核心网的服务化架构、软件定义网络、网络切片、网络功能虚拟化、边缘计算等核心技

术成为 5G 核心网关键技术. 在这些关键技术与新型服务化网络架构、新设计理念, 如控制面与用户面分离、网络能力开发、计算与数据分离等交织和推动下, 形成了 5G 核心网革命性地变化, 能够支持多种 5G 创新性业务发展需求.

5G 研究和标准化制定大致经历 4 个阶段. 第 1 阶段是 2012 年 5G 基本概念提出; 第 2 阶段是 2013—2014 年, 这个阶段主要关注 5G 愿景与需求、应用场景和关键能力; 第 3 阶段是 2015—2016 年, 主要关注 5G 定义, 开展关键技术研究 and 验证工作; 第 4 阶段是 2017—2020 年, 主要开展 5G 标准方案的制定和系统实验验证.

在标准化方面, 制定 5G 国际标准主要在 ITU-R WP5D 和 3GPP 两大标准化组织中进行. 其中 ITU 重点制定 5G 系统需求、指标以及性能评价体系, 在全球征集 5G 技术方案, 开展技术评估, 确认和批准 5G 标准, 不做具体的技术和标准化规范制定工作. 3GPP 作为全球各通信主要产业组织的联合组织, 从事具体的标准化技术讨论和规范制定, 并将制定好的标准规范提交到 ITU 进行评估, 满足 ITU 的 5G 指标后将被批准为全球 5G 标准.

ITU-R WP5D 是 ITU 中专门负责地面移动通信业务的工作组. 2010 年, 当 4G 标准之争刚落下帷幕, WP5D 就启动了面向 2020 年的业务发展预测报告起草工作, 以支撑未来 IMT 频率分配和后续技术发展需求, 5G 工作开始启动. WP5D 除了完成频率相关工作外, 还启动了面向 5G 的愿景与需求建议书撰写, 面向后 IMT-Advanced 的技术趋势研究工作以及 6 GHz 以上频段用于 IMT 的可行性研究. 面向未来 5G 的频率、需求、潜在技术等前期工作在 ITU 全面启动. 2014 年, WP5D 初步制定了

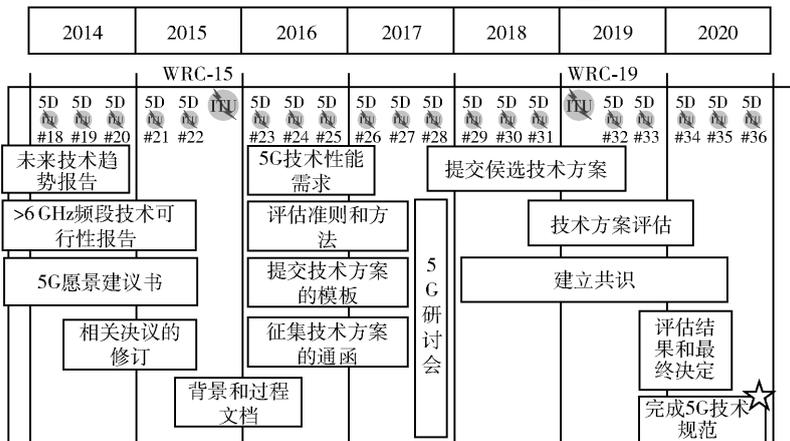


图 4 ITU 定义的 5G 标准时间计划

5G 标准化工作的整体计划^[3], 如图 4 所示, 并向各外部标准化组织发送了联络函。在 2015 年中, ITU-R 完成了对 5G 的命名, 决定 5G 在 ITU 正式命名为 IMT-2020。

ITU 的 5G 标准化整体分为 3 个阶段: 第 1 阶段为前期需求分析阶段, 开展 5G 的技术发展趋势、愿景、需求等方面的研究工作; 第 2 阶段为准备阶段 (2016—2017 年), 完成需求制定、技术评估方案以及提交模板和流程等, 并发出技术征集通函; 第 3 阶段为提交和评估阶段 (2018—2020 年), 完成技术方案的提交、性能评估以及多个提交方案融合等工作, 并最终完成详细标准协议的制定和发布。

3GPP 作为全球 3G 和 4G 主流移动通信标准制定组织, 其标准化工作主要由 RAN (radio access network)、SA (services & systems aspects) 和 CT (core network & terminals) 3 个工作组开展。其中, 负责接入网与空口标准化的工作组 RAN 在 2015 年 9 月召开了 5G 研究的研讨会, 制订了具体的工作计划。3GPP 确定从 2016 年 3 月启动 5G 新空口 (NR, new radio) 研究和标准化的制定工作, 到 2019 年 12 月完成所有标准规范, 持续 R14、R15 和 R16 等 3 个版本。3GPP 于 2018 年第 3 季度向 ITU-R 提交第 1 个基于 R15 版本的初步技术方案, 该版本主要包括基本的 eMBB 和 URLLC 2 个应用场景的技术协议以及核心网架构和协议, 支持基于 5G 新空口的独立组网以及 LTE 和 NR 联合组网的方式, 并于 2019 年提交基于 R16 的全面满足 5G 需求的增强型版本, 包括 3 个应用场景和性能增强。目前 3GPP 已完成 5G 第 1 个版本规范 R15 的标准化工作。

在网络架构和核心网方面, 2016 年, 负责系统需求定义的 SA1 工作组完成了 5G 及业务需求的定义, 负责系统架构设计的 SA2 工作组完成了 5G 架构的研究。截至 2017 年年底, SA2 完成了 5G 第 1 阶段的标准制定工作, 发布了 5G 架构、5G 流程、策略和计费控制框架这 3 个方面的标准。3GPP CT 各工作组同期积极展开了协议的设计工作, 并于 2018 年 6 月发布了 32 个接口及协议的规范, 这标志着 R15 版本规范工作全部完成。目前, R16 的工作已经全面展开, 预计 5G 第 2 阶段的标准制定工作将于 2019 年 6 月完成。

下面介绍基于核心网和接入网标准规范对 3GPP R15 的 5G 标准。

2 NR 接入网架构与接口

为了满足 5G 高吞吐量、低延迟的需求, 密集部署成为 5G 网络的重要形式。针对密集组网特点, NR 采用了中心控制节点 (CU, central unit) 和分布节点 (DU, distributed unit) 相结合的部分分离架构形式, 实现网络的密集部署, 同时尽可能地减少资本支出, 降低维护成本。另外, 为了支持 4G 和 5G 网络的融合以及满足早期部署需求, 还引入了 LTE 和 NR 的双连接。本节主要介绍 NR 接入网整体架构设计、接入网内部分离架构以及 LTE 和 NR 的双连接解决方案。

2.1 接入网整体架构设计

NR 的接入网由 gNB 和 ng-eNB 两类节点组成, 其中 gNB 是提供 NR 接入技术的基站, 而 ng-eNB 是提供 E-UTRA 接入技术并连接到 5G 核心网的基站, 这两类节点统称 NG-RAN 节点。NG-RAN 节点和 5G 核心网通过 NG 接口进行连接, NG-RAN 节点之间通过 Xn 接口进行连接, 如图 5 所示^[4]。

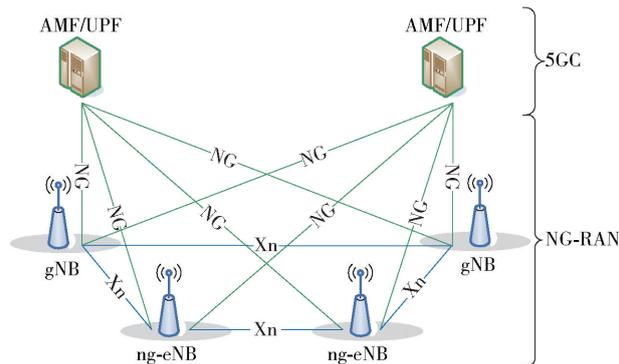


图 5 5G 系统整体架构^[4]

2.1.1 NG 接口

NG 接口分为用户面和控制面。类似于 LTE 的 S1 接口, NG 接口用户平面的传输网络层基于 IP 传输, UDP/IP 协议之上采用 GTP-U 来传输核心网和接入网之间的用户平面 PDU。NG 接口控制平面与用户平面类似, 也是基于 IP 传输的, 不同的是控制平面在 IP 层的上面采用流控制传输协议 (SCTP, stream control transmission protocol), 为无线网络控制层信令消息提供可靠的传输。NG 接口的控制面和用户面协议栈如图 6 所示^[4]。

NG 控制面的主要功能包括 NG 接口管理、UE 上下文管理、UE 移动性管理、NAS 消息传输、寻呼以及分组数据单元 (PDU, packet data unit) 会话管

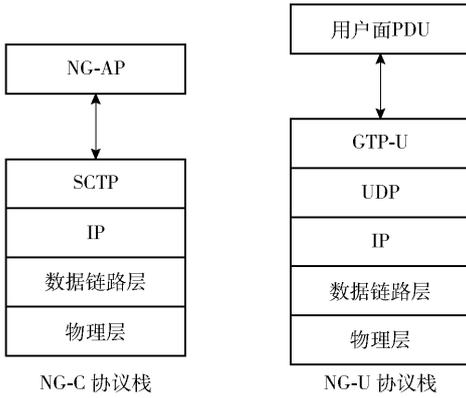


图 6 NG 接口的控制面和用户面协议栈

理等. 用户面提供 NG-RAN 节点和 5G 核心网之间的用户数据传输功能.

2.1.2 Xn 接口

Xn 接口也分为用户面和控制面,其用户面协议结构和控制面协议结构与 NG 接口类似.

Xn 的用户面提供 NG-RAN 节点之间的用户数据传输功能. 控制面主要提供 Xn 接口管理、UE 移动性管理(包括切换和寻呼)以及双连接等功能.

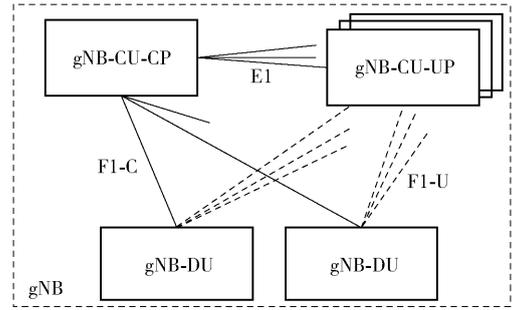
2.2 接入网内部分离架构

为了支持灵活的网络部署方式及接入网络的虚拟化, NR 引入了 gNB-CU/gNB-DU 分离的架构. 其中, gNB-CU 是中心控制节点, 包括 RRC 和 PDCP 功能, gNB-DU 是分布节点, 包括 RLC、MAC 和物理层. 通过 CU/DU 的架构, 可以提升各节点间资源协调和传输协作能力, 并通过将 CU 进行云化和虚拟化提升网络资源的处理效率. 另外, 为了进一步增强部署的灵活性和实现的便利性, 可将 gNB-CU 的控制面和用户面部署在不同的位置, 引入 gNB-CU-CP/gNB-CU-UP 分离(用户面和控制面分离)的架构, 如图 7 所示^[5]. 可以看出, 一个 gNB 可由一个 gNB-CU-CP 和多个 gNB-CU-UP 以及多个 gNB-DU 组成. gNB-CU-CP 通过 E1 接口和 gNB-CU-UP 连接, gNB-DU 通过 F1 接口和 gNB-CU 连接, 其中 F1-C 终止在 gNB-CU-CP, F1-U 终止在 gNB-CU-UP.

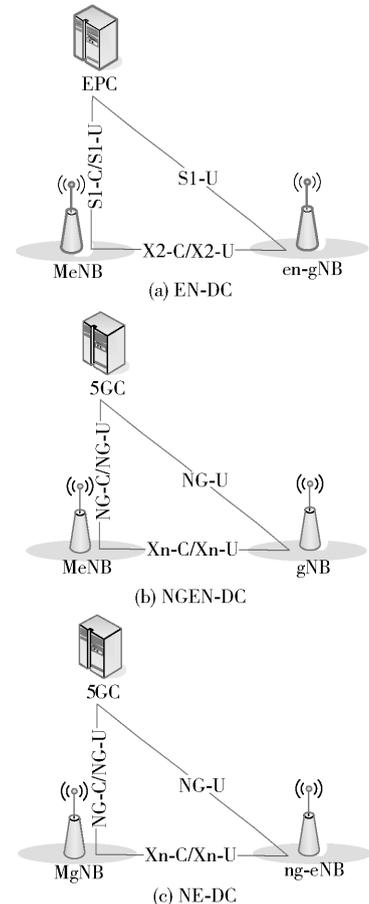
E1 接口主要支持接口管理和承载管理的功能. F1 接口分为控制面和用户面. F1 接口控制面提供接口管理、系统信息管理、UE 上下文管理、寻呼及 RRC 消息传递等功能; F1 接口用户面主要在 gNB-DU 和 gNB-CU-UP 之间提供数据传输, 同时, F1 用户面还支持数据传输的流控机制, 以进行拥塞控制.

2.3 LTE 和 NR 双连接

考虑到 5G 未来应用频谱资源包含了更高频

图 7 RAN 内部分离的架构^[5]

段, 如 6 GHz 以上, 同时考虑到 4G 与 5G 网络融合及更早推动 5G 网络部署等方面, 3GPP 标准支持 LTE 与 NR 联合组网的方式, 即由 LTE 提供基本的覆盖和移动性支持, NR RAN 节点辅助提供更高的容量, 这种组网方式称为 MR-DC (multi-RAT dual connectivity). MR-DC 根据接入网连接的核心网类型以及主辅节点的类型进一步分为如下 3 种方式, 如图 8 所示^[6].

图 8 LTE 和 NR 双连接^[6]

1) EN-DC: EN-DC 指 E-UTRAN/NR 双连接, 其中主节点是 eNB, 辅节点是 en-gNB. 接入网和 EPC

的控制面信令通过 eNB 和 EPC 交互, en-gNB 和 EPC 之间只有用户面连接,没有控制面连接。

2) NGEN-DC: NGEN-DC 指 NG E-UTRAN/NR 双连接,其中 ng-eNB 是主节点, gNB 是辅节点, ng-eNB 和核心网通过 NG 控制面连接。

3) NE-DC: NE-DC 指 NR/E-UTRAN 双连接,其中 gNB 是主节点,提供和 5GC 的控制面信令连接, ng-eNB 是辅节点。

在 5G 部署初期,连接到 EPC,同时又能提供 5G 高速率服务的 EN-DC 的方式对现有网络影响小,是很多运营商在部署初期的选择。后续可演进为 NE-DC 和 NGEN-DC 方式,以提供完整的 5G 服务。

3 NR 空口设计与关键技术

3.1 空口物理层关键技术与传输方案

NR 物理层信号传输框架以 4G LTE 为基础设计,支持 FDD 和 TDD 双工方式,下行采用 CP-OFDM,上行支持 CP-OFDM 和 DFT-s-OFDM 两种波形调制。但为支持多样的 5G 应用场景和业务以及更大的候选频谱范围(100 GHz 以内),其物理层设计体现出 3 个特点: 1) 支持灵活动态 TDD 双工方式为系统主要部署形式; 2) 围绕波束赋形的信号传输体系设计; 3) 支持多样性应用场景和业务的传输信号参数灵活配置。

下面对 NR 的物理层传输设计方案、大规模波束赋形、高频段传输以及信道编码的设计方案和特点进行介绍^[7-13]。

3.1.1 物理层信号传输方案

1) 帧结构与 OFDM 参数设计

OFDM 符号的时域和频域信号基本参数设计决定了系统传输的基本性能和设计体系。在时域上通过设计信号传输帧结构,决定了 TDD/FDD 双工信号的传输时序关系;频域的 OFDM 子载波间隔决定了信号传输性能。

LTE 支持频域 15 kHz 子载波间隔,单一的子载波间隔无法满足 5G 系统的需求。这是由于 NR 需要支持 100 GHz 以内的频谱范围,并且根据部署场景和支持的业务多样性,需要支持更多的子载波间隔,以对系统进行优化设计。为此, NR 支持 15、30、60、120 和 240 kHz 的多种子载波间隔,子载波间隔之间呈 2 的幂次方倍数关系。

时域内,一个无线帧的长度为 10 ms,分为 10 个子帧。与 LTE 不同的是, NR 的子帧只作为记时单

位,不作为调度单位,目的是支持更灵活的资源调度方式。一个子帧进一步分割为若干个时隙,具体时隙个数取决于子载波间隔。无论子载波间隔多大,一个时隙包括 14 个 OFDM 符号。FDD 系统下行频谱所有的时隙均为下行时隙,上行亦然。TDD 系统的一个时隙内的 OFDM 符号可能包括下行符号、上行符号和灵活符号 3 种类型。其中灵活符号既可以充当传统 TDD 系统中的 GP 的角色,也可以根据其他的控制信令用于进行上行传输或者下行传输,最大化了 TDD 系统的灵活性。

NR 支持基于时隙的资源调度。为了支持 URLLC 等对时延敏感的业务, NR 也支持基于微时隙的调度。下行微时隙的长度可以是 2、4 或者 7 个 OFDM 符号,上行微时隙的长度则可以是 14 个 OFDM 符号以内的任意长度。

2) 灵活分级的带宽设计

NR 支持分级的多系统带宽设计,最小为 5 MHz,最大为 100 MHz (6 GHz 以下频段) 和 400 MHz (6 GHz 以上频段) 系统带宽,目的是提供高速率的数据传输。同时,从终端节电和成本等方面考虑, NR 允许终端只工作在系统带宽的一部分,例如 20 MHz。终端通过初始接入过程接入网络之后,网络可以通过专用信令为终端配置其工作的带宽部分 (BWP, bandwidth part), 每个终端最多可以配置 4 个 BWP,但是在任意时刻仅有 1 个 BWP 是激活的。除 RRM 测量之外,终端仅在 1 个激活的 BWP 上收发数据。终端支持的 BWP 个数以及 BWP 的带宽作为终端能力上报给网络,网络根据终端的能力进行配置。

3) 下行同步信道设计

NR 采用大规模天线技术有效地克服了高频段 (特别是毫米波频段) 带来的覆盖受限难题: 通过大规模天线的波束赋形增益,提升覆盖范围。在 4G LTE 系统中同步、广播等需要进行小区范围内广播传输的信号,一般都采用全向 (或宽波束) 传输,而业务信道通常采用窄波束宽度的赋形来提升信号传输速率和频谱效率。在高频段传输中,为了保证同步、广播等广播信道的信号能与业务信道的覆盖范围相匹配,同步等信号传输也需要采用波束赋形方式。由于波束赋形的波束宽度较窄,为了使小区内所有终端都能够接收到信号,采用一种称为波束扫描的过程来实现同步等广播信道发送。

波束扫描的具体做法是用多个波束在不同方向

重复发送下行同步信道,这些非重叠的波束集可以覆盖整个小区的范围。通常多个波束在不同的时间轮流发送,因此被称为波束扫描。

下行同步信道包括主同步信号、辅同步信号和物理广播信道。因为主同步信号、辅同步信号和物理广播信道需要按照相同的方式进行波束扫描, NR 将它们组合起来定义为一个同步块(SSB),波束扫描以 SSB 为单位进行。

图9所示为SSB波束扫描的配置方案。一个扫描周期内的SSB构成一个SSB突发集。一个SSB突发集内不同编号顺序的SSB可以用不同方向的波束发送,形成扫描波束。同时,SSB突发集以一定时间周期进行重复发送,每个发送周期内相同编号顺序的SSB所采用的波束方向相同。终端可以接收多个后续波束,并选择最为合适的下行波束用于后续操作。

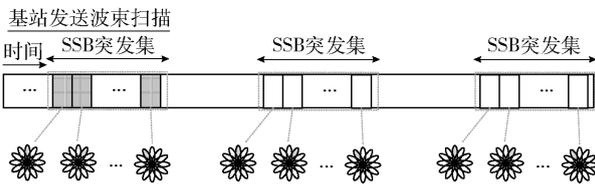


图9 SSB波束扫描

一个SSB突发集内的SSB个数主要取决于2个因素:系统开销和覆盖的要求。SSB的个数越多,基站可以用更多更窄的波束发送SSB,获得更大的波束赋形增益和更好的覆盖效果(越窄的波束需要的天线数量越多)。但是相应系统的开销也等比例地增加。NR频段定义了允许最大的SSB个数:①3GHz频点以下最多允许4个SSB;②3~6GHz频点最多允许8个SSB;③6GHz以上频点最多允许64个SSB。随着频点的升高,对于提升SSB覆盖的需求增强,因此允许的SSB数目也增大。

在实际的网络中,基站可以根据系统频点、用户分布等因素选择实际发送的SSB个数和SSB的波束方向。

4) 上行随机接入过程设计

在高频段传输中,随机接入过程也需要通过波束赋形来满足覆盖范围需求,因此终端也应采用波束赋形方式发送上行前导序列和接入信息,才能接入网络。随机接入过程中,由于终端无法通过预先的信令交互完成波束的配置,随机接入过程中的4个消息所用发送和接收波束是标准化过程中重点讨论的议题。

首先介绍基本概念:波束互易性。波束互易性指一个终端/基站可以通过接收波束确定其发射波束,或者反之。波束互易性是比TDD系统中的信道互易性弱化的概念,波束互易性对TDD和FDD系统都可以应用。波束互易的存在也是因为无线信号在空间传播的路径是互易的,这时只要一个设备的收发链路是校准好的,就可以实现波束互易。

随机接入过程需要确定的波束包括终端发送Msg. 1和Msg. 3的发送波束,终端接收Msg. 2和Msg. 4的接收波束,基站则是相反的情况。下面对随机接入过程4个步骤中消息的发送和接收波束进行简要描述。

① Msg. 1

终端在发送Msg. 1之前已经实现了下行的同步,检测到了一个最适合己方接收的SSB,获取接收波束信息,并通过该SSB指示的系统信息获得了物理随机接入信道(PRACH)的资源配置信息。

如果终端侧具有波束互易性的话,那么终端就可以通过接收波束确定发送波束。如果终端不具备波束互易性,则只能采用上行波束扫描方式,发送多个上行前导序列Preamble码。

基站通过接收Preamble码,确定所用的接收波束。如果基站侧的波束互易性不成立,基站需要对多个可能的接收波束进行尝试,即采用波束扫描的方式接收。

如果基站侧波束互易性成立,基站可以由某一个发送波束确定对应接收波束。因为终端在发送Preamble之前,已经确定了一个SSB,那么该SSB对应的发送波束是确定接收波束的最佳候选。NR通过建立PRACH资源以及Preamble和SSB的映射关系,终端检测到一个SSB之后从该SSB映射到的PRACH资源以及Preamble中选择资源和序列。也就是说终端选择的PRACH资源和Preamble指示了终端检测到的SSB编号。那么基站在某一特定的PRACH资源上检测特定的Preamble时就可以用其映射到的SSB的发送波束确定接收波束。

② Msg. 2

终端所发送的Preamble码和PRACH资源与特定的SSB下行波束对应,因此基站发送的Msg. 2可以采用与终端选择的SSB相同的波束发送;而终端接收Msg. 2时,用终端选择的SSB的接收波束进行接收。

③ Msg. 3

Msg. 3 是上行消息, 可以采用和 Msg. 1 一样的发送和接收波束。

④ Msg. 4

Msg. 4 是下行消息, 可以采用和 Msg. 2 一样的发送和接收波束。

5) 下行控制信道

下行控制信道(PDCCH)用于承载下行控制信令, 主要包括上下行调度信令、上行功率控制命令等。NR PDCCH 信道的设计特点包括以下 3 点。

① 灵活的资源配置。为避免类似于 LTE 的 PDCCH 在小区之间产生的恒定干扰, NR PDCCH 资源配置更加灵活, 在频域上 PDCCH 仅占用部分带宽, 不必占用所有带宽, 从而可以支持小区间干扰协调, 以及终端支持和使用不同带宽的情况。

② 支持低时延业务。LTE 的 PDCCH 固定在每个子帧的前几个符号上, 这意味着在紧急业务(如 URLLC)待发送的情况下, 基站也只能等到下一个子帧才能传输 PDCCH, 无法满足低时延业务的需求。NR 允许一个时隙的多个 OFDM 符号位置具有发送 PDCCH 机会, 灵活的 PDCCH 发送位置与 Mini-slot 调度结合, 可以有效地支持低时延业务。

③ 大规模天线传输。为与业务信道的覆盖范围相匹配, PDCCH 可利用大规模天线技术进行波束赋形传输。

6) 上行控制信道

上行控制信道(PUCCH)承载的信息包括 HARQ-ACK、宽带 CSI 和 SR 请求等。LTE PUCCH 时域长度固定为 14 个 OFDM 符号, 占用的 PRB 个数随 PUCCH 格式而不同。NR PUCCH 设计的一个重要改变是引入了两种 PUCCH 结构: 长 PUCCH 和短 PUCCH。长 PUCCH 的长度可以占用 4 ~ 14 个 OFDM 符号, 短 PUCCH 的长度可以是 1 或者 2 个 OFDM 符号, 由基站配置确定。长 PUCCH 的设计与 LTE 类似, 短 PUCCH 的设计主要用于降低 HARQ-ACK 反馈的时延, 支持 URLLC 等低时延业务以及支持上行波束扫描。

短 PUCCH 可以设置为在一个称为“自包含”的时隙中, 如图 10 所示。图中前面部分 OFDM 符号用于下行数据传输, 中间是下行到上行的保护间隙, 后面是一个 1 ~ 2 个符号的 PUCCH 信道, 确保下行传输数据在一个时隙内获得应答消息。

同时, 由于高频段模拟波束赋形在一个时间点

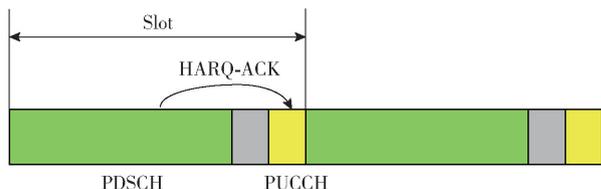


图 10 短 PUCCH 在自包含时隙中传输配置

上只能发送一个波束, 如果上行传输采用长 PUCCH 进行波束扫描, 会占用更多时域资源, 导致系统开销增加, 同时也增加了 PUCCH 传输的时延。而用短 PUCCH 可以节省频谱资源, 降低扫描时延。

7) 上下行业务信道

NR 的上下行业务信道(PDSCH 和 PUSCH)均以专用导频(DMRS)为解调导频, DMRS 和数据的预编码相同。为满足上下行的峰值速率要求, PDSCH 和 PUSCH 单用户分别最多支持 8 流和 4 流传输。多用户 MIMO 调度时, 每用户最多支持 4 流传输; 最多可配置 12 个正交 DMRS 端口。

DMRS 采用了 Front-loaded 设计, 即 DMRS 尽量前置, 目的是降低终端解调、译码的时延。在此基础上, 可以通过配置 Additional DMRS 来支持高速移动的终端。NR 为了降低高频段相位噪声对传输性能的影响, 引入相位跟踪补偿导频(PT-RS)用于相位噪声的跟踪和补偿。

3.1.2 大规模波束赋形技术

随着有源天线技术商业成熟度的提升, 垂直维数字端口的开放与天线规模的进一步扩大逐渐成为可能。在这一背景下, 3GPP 从 R12 阶段开始了针对 3D 信道与场景模型问题的研究, 并在 R13、R14 及后续版本中对 FD-MIMO 技术进行了研究与标准化。至此, 开启了大规模天线技术进入标准化发展的新篇章。随着 5G 时代的来临, 面对诸多更加严苛的技术指标需求, 大规模天线技术仍然被认为是 5G 系统中最重要的一项物理层技术。

天线规模的增大及高频段模拟波束的使用使得大规模天线技术在发展过程中将面临一些新的挑战, 下面重点介绍大规模波束赋形在标准化中信道测量机制、码本设计、信息反馈机制、波束管理流程以及导频的相关设计。

1) 灵活的 CSI 反馈框架。NR 系统引入了一套统一的反馈框架, 能够同时支持 CSI 反馈和波束测量上报。该反馈框架内, 所有和反馈相关的参数都是可以配置的, 如测量信道和干扰的参考信号、反馈

的 CSI 的类型、所使用的码本、反馈所占用的上行信道资源、反馈的时域特性(周期、非周期、半持续等)、反馈的频域特性(CSI 的带宽)等. 网络设备可以根据实际的配置参数. 相比之下, LTE 需要使用多种反馈模式, 并且将反馈和传输模式绑定, 灵活度欠佳. NR CSI 反馈框架的应用实例如图 11 所示. CSI 反馈框架中包括 Reporting setting 配置和 Resource setting 配置, 其中 Reporting setting 用于配置 CSI 上报相关的参数, 例如所使用的码本、时频域颗粒度等; Resource setting 用于配置 CSI 测量所需的参数, 即信道测量参考信号以及干扰测量资源. Reporting setting 和 Resource setting 之间有一定的链接关系. 一个 Reporting setting 可以链接到多个 Resource setting, 分别用于不同的目的, 如图 11 中 Reporting setting 0 链接到 3 个 Resource setting. 同时, 一个 Resource setting 也可以链接到多个 Reporting setting, 如图 11 中的 Resource setting 0.

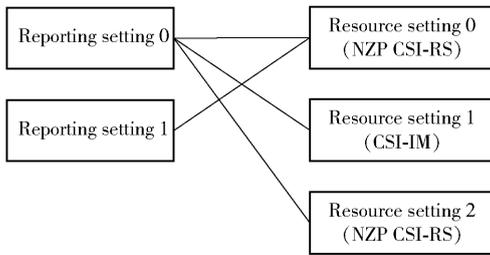


图 11 CSI 反馈框架

2) 高精度码本设计. MU-MIMO 系统的性能依赖于 CSI 的获取精度以及后续的预编码与调度算法的优化程度. CSI 的获取是大规模天线系统设计及标准化的一个关键议题. 针对这一问题, NR 系统中定义了两种类型的码本, 即常规精度 (Type I) 码本与高精度 (Type II) 码本. 其中 Type I 主要针对 SU-MIMO 或 MU-MIMO, 而 Type II 则主要针对 MU-MIMO 传输的增强. Type I 码本延续了 LTE 的双级码本方案, 即预编码矩阵分解为 2 个矩

阵的乘积:

$$\mathbf{W} = \mathbf{W}_1 \mathbf{W}_2$$

其中: \mathbf{W}_1 为宽带信息, 具体为由 DFT 向量组成的波束集合, \mathbf{W}_2 为子带信息, 从 \mathbf{W}_1 的波束集合中进行波束选择(每一列上只有 2 个元素取值非零, 其他都为零)以及实现极化方向之间的同相位合并. R15 的 Type II 码本采用了线性合并方式构造预编码矩阵, 能够显著地提升 CSI 精度进而极大地改善 MU-MIMO 传输的性能. Type II 码本在结构上和 Type I 码本一致, 即也表示成 2 个矩阵的乘积形式. 区别有两点: 一是 Type II 码本中 \mathbf{W}_1 由彼此正交的 DFT 波束组成; 二是 Type II 码本中 \mathbf{W}_2 的作用是对 \mathbf{W}_1 中的波束进行线性合并, 各合并系数为非横模的复数. Type II 码本的构造方式如图 12 所示. 图 12 中所示的配置有 8 个正交的 DFT 波束, \mathbf{W}_1 从中选择了 4 个 DFT 波束 ($\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$). \mathbf{W}_2 中的 4 个合并系数对这 4 个 DFT 波束进行线性合并. 合并系数分为幅度 ($\mathbf{a}_0 = 1, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$) 和相位 ($\mathbf{p}_0 = 1, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$) 两个部分分别反馈.

3) 波束管理和恢复过程. 毫米波频段由于天线尺寸以及传播条件限制, 需应用大规模波束赋形技术来补偿路径传播损耗和遮挡. 且出于成本和功耗等因素考虑, 模拟或者混合 (模拟 + 数字) 波束赋形是主要的技术手段. 为此 NR 设计了波束管理机制使基站和终端可以对齐发射和接收波束, 包括波束的测量和上报机制、波束指示机制等. 如图 13 所示, 基站选择了一个发射波束之后, 信号沿着特定的方向传播, 终端需要使用与基站的发射波束对应的接收波束进行接收, 否则终端接收信号的质量会下降. 因此, 发射波束和接收波束之间有一定的对应关系, 称之为一个波束对. 为实现收发波束的对齐, 基站以波束扫描的方式传输参考信号 (如 CSI-RS). 如果一个基站能够发送 M 个模拟波束, 可以为每个波束配置一个参考信号集合用于波束的测量, 每个

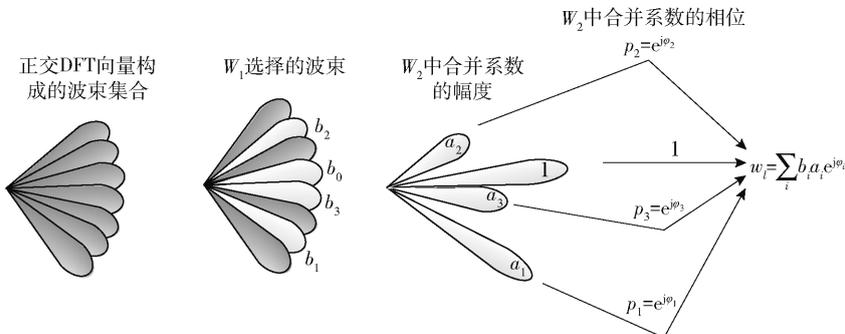


图 12 Type II 码本构造说明

参考信号用所对应的模拟波束赋形. 这 M 个参考信号在不同时域或频域资源上传输, 以便于基站能够针对每个波束方向调整移相器的配置来实现模拟波束赋形. 同时, 终端通过对 N 个接收波束分别对 M 个参考信号进行测量, 选择合适的接收波束, 如图 14 所示. 因此, 基站与终端间一共需要测量 MN 个波束对, 寻找到最佳的收发配对波束. 终端将选择出的 K (具体个数由网络配置) 个最佳的发射波束的标识信息上报给网络, 发射波束对应的接收波束不需要上报, 终端自行记录下来即可. 在进行数据传输时, 因为收发波束之间有对应关系, 基站需要将数据传输所使用的发射波束通知给终端, 终端则根据记录的收发波束之间的对应关系和基站的通知设置合适的接收波束.

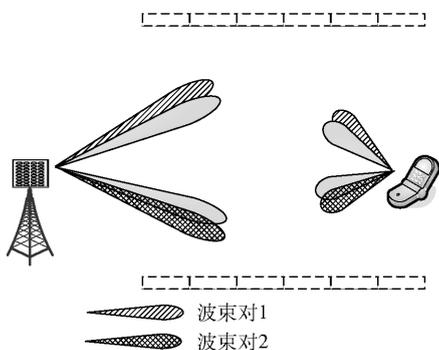


图 13 收发波束对齐

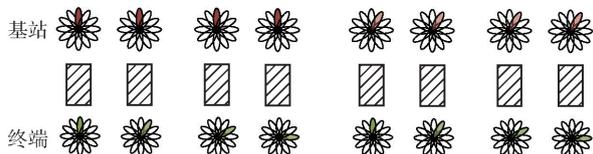


图 14 波束扫描与测量

同时由于人体遮挡、旋转等因素会导致已经对齐的波束失效, 为了能快速地从失效状态恢复, NR 设计了波束失败恢复机制, 包括波束失败检测、波束失败请求发送和响应接收等.

4) 灵活可配置的 CSI-RS 导频设计. 为了保证前向兼容性和降低功耗, NR 尽量减少了持续发送的参考信号, 基本上所有的参考信号的具体功能、发送的时频位置、带宽等都是可以配置的. NR 对 LTE 已经存在的 CSI-RS 进行了进一步的扩展, 除了支持 CSI 测量外, 还支持波束测量, RRM/RLM 测量, 时频跟踪等. CSI-RS 支持的端口数包括 1、2、4、8、12、16、24 和 32. CSI-RS 的图样由基本图样聚合得到, 并且支持多种基本图样和 CDM 类型.

3.1.3 信道编码技术与设计

信道编码是现代通信系统用于对传输过程中错误进行前向纠错的有效手段. 无线移动通信系统中, 信号传输过程中受到慢衰落和快衰落的影响, 更容易出现错误. 采用先进的信道编码技术能够满足不同业务对可靠传输的需要, 并提升频谱效率.

NR 标准制定过程中, 主要针对上下行控制信道及业务信道的信道编码方案进行研究和评估. 在标准化讨论伊始, 主要有 LDPC 码、Turbo 码 (LTE Turbo 与增强方案)、Polar 码、卷积码 (LTE TBCC 与增强方案) 等候选方案. 其中在业务信道编码方案的研究和选择过程中, Turbo 码由于其内在的串行编译码特性, 虽然做了一些提升并行化处理和降低译码复杂度的改进, 无法实现 5G 超高速率、低时延的大容量数据传输, 并在误码平层 (error floor) 性能方面存在短板, 被 NR 放弃. NR 最终采用了 LDPC 这种具有天然并行化译码和良好的错误平层的信道编码方案. 在控制信道方案的评估和分析中, Polar 码在低码率上相比于 Turbo 码和卷积码具有性能优势, 因此被选为控制信道的编码方案.

NR 所采用的 LDPC 码方案是一种被称为二进制准循环低密度校验码 (QC-LDPC, quasi-cyclic LDPC) 的结构化 LDPC 码设计方案.

NR LDPC 码校验矩阵 H 由基矩阵 B 扩展得到, 其中基矩阵 B 是一个维数为 $mb \times nb$ 的二元矩阵, 即 $b_{i,j}$ 取值为 0 或 1. 基矩阵中每个元素被扩展为一个循环子矩阵或者 0 矩阵, 则 H 可表达为

$$H = \begin{bmatrix} Q(p_{0,0}) & Q(p_{0,1}) & \cdots & Q(p_{0,nb-1}) \\ Q(p_{1,0}) & Q(p_{1,1}) & \cdots & Q(p_{1,nb-1}) \\ \vdots & \vdots & \ddots & \vdots \\ Q(p_{mb-1,0}) & Q(p_{mb-1,1}) & \cdots & Q(p_{mb-1,nb-1}) \end{bmatrix}$$

其中: $p_{i,j}$ 是循环移位值, 取值于整数集合 $\{-1, 0, 1, 2, \dots, Z-1\}$; 当 $b_{i,j} = 0$ 时, $p_{i,j} = -1$, $Q(-1)$ 为 $Z \times Z$ 零矩阵; 当 $b_{i,j} = 1$, $p_{i,j}$ 取值于整数集合 $\{0, 1, 2, \dots, Z-1\}$, $Q(p_{i,j})$ 为 $Z \times Z$ 循环置换阵, Z 是扩展因子. NR LDPC 设计的基矩阵个数为 2, 大基矩阵 (BG1, base graph#1) 对应的信息位列数为 22, 最低码率为 1/3; 小基矩阵 (BG2, base graph#2) 对应的信息位列数为 10 时, 最低码率为 1/5. 基矩阵支持大量的扩展因子 $Z = a \cdot 2^t$, $a = \{2, 3, 5, 7, 9, 13, 15\}$, $t = \{0, 1, 2, 3, 4, 5, 6, 7\}$, 所以 NR LDPC 码可以灵活地支持不同长度、不同码率的信息比特编码, 从而支持不同场景和业务传输需求.

NR LDPC 码的校验矩阵由 5 个子矩阵构成,具体如图 15 所示,其中子矩阵 A 和子矩阵 E 是由循环置换矩阵和全零矩阵组成的矩阵阵列,子矩阵 O 为全零矩阵,子矩阵 I 为单位矩阵. 其中矩阵 A 对应系统比特,矩阵 D 对应校验比特,矩阵 $[A D]$ 对应一个高码率的 LDPC 码;矩阵 $[E I]$ 对应支持 IR-HARQ 的扩展冗余比特,其中单位阵 I 实际上对应一个度为 1 的单校验比特. 该结构等价于一个高码率的 LDPC 码与许多个单校验码串行级联,而且随着扩展矩阵行数与列数的增加可以得到码率任意低的 LDPC 码校验矩阵.

Polar 码基于信道极化,可以将 N 个比特信道中一部分变成纯噪声信道,另一部分变成无噪声信道,只需将所要传输的数据加载在无噪声信道,就可以实现数据的可靠传输^[14].

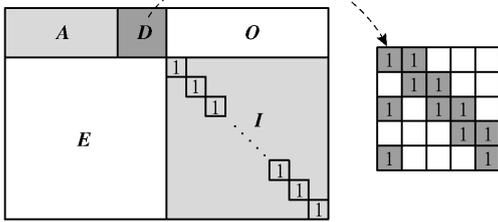


图 15 NR LDPC 码校验矩阵结构

NR 基本 Polar 码编码过程如下:信源比特为 u_0^{N-1} ,包含需要传输的 M 比特信息, $M \leq N$, Polar 编码器输出的码字为 $x_0^{N-1} = u_0^{N-1} G_N$. 其中, $G_N = F^{\otimes n}$ 为 $N \times N$ 的生成矩阵, $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $F^{\otimes n}$ 表示矩阵 F 的 n 次克罗内克积.

为了提高基本 Polar 码的性能,CRC-Aided Polar (CA-Polar) 码被提出^[15];信源序列 d 经过 CRC 编码器产生的 CRC 序列附加在序列 d 后得到序列 u ,再将序列 u 送入 Polar 编码器得到码字序列 c . 在 SCL (successive cancellation list) 译码结束时,CRC 译码器将对 L 个候选译码路径进行校验,将可靠度最大的且通过 CRC 校验路径上的译码序列作为最终译码结果 \hat{d} 输出,如图 16 所示.

CA-Polar 码被 NR 采纳为上行控制信道编码方案(信息比特 $K_{UCC} > 19$). 在信息比特很短时,在 CRC 编码之前引入几个校验(PC, parity check)比特,能进一步提高 CA-Polar 的性能,即 PC-CA Polar^[16],该方案被用于信息比特 $12 \leq K_{UCC} \leq 19$ 时的编码方案. 当 K_{UCC} 小于 12 bit 时, NR 重用了 LTE 的编

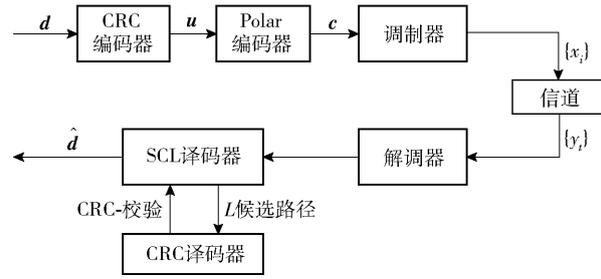


图 16 CA-Polar 编译码框图

码方案,即 1 bit 采用重复码,2 bit 采用 simplex 码,3 ~ 11 bit 采用 LTE RM 码.

与上行控制信息编码不同,下行控制信息进行 Polar 编码时,通过交织器重新排列 u 中的信息比特和 CRC 比特的位,使得部分 CRC 比特位置提前,且每一个 CRC 比特都位于它校验的所有信息比特之后. 这样 Polar 码译码就具备了早停功能,即 Distributed CRC Polar 码^[17],该方案利于降低终端检测下行控制信道的功耗.

3.1.4 高频段传输

NR 设计的目标是用统一的空口支持 100 GHz 以内的频段,因此前述 NR 物理层设计的各个方面均考虑了对高频段(毫米波频段)的支持. 对毫米波频段的特殊考虑简要总结如下.

1) 帧结构方面. 60 kHz 和 120 kHz 的子载波间隔的重要应用场景是毫米波频段. 此外,微时隙调度也可以支持在一个时隙内用户间的 TDM 复用,克服模拟波束赋形带来的调度灵活度的问题.

2) 波束管理和波束恢复机制. 专门针对毫米波频段的模拟波束赋形和混合波束赋形而设计.

3) PT-RS 设计. 设计目的是跟踪和补偿相位噪声,因为相位噪声在毫米波频段有显著的影响.

4) 基于波束的下行同步信道和上行初始接入过程设计. 为补偿毫米波频段的路径损耗,上下行的初始接入信道均设计了波束扫描的发送/接收机制,扩展覆盖范围.

3.2 空口高层协议设计与过程

空口高层协议通过与物理层信号传输相配合,确保空口无线数据按序可靠传输,无线资源高效管理,实现网络控制与切换管理等重要功能的系统传输过程. NR 的空口协议栈和过程基于 LTE,并根据 5G 传输特点,有针对性地增加了相应功能. 下面首先介绍 NR 的空口协议栈架构,然后针对空口传输与控制过程的新增功能进行介绍.

3.2.1 空口协议栈

NR 的空口控制面协议栈采用了和 LTE 一样的架构,如图 17 所示^[4]。

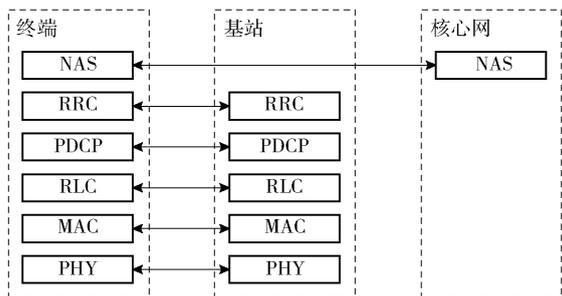


图 17 NR 空口控制面协议栈

控制面协议栈的核心协议层为 RRC 层^[18],其功能大部分与 LTE 类似,比如支持系统信息、寻呼、接入控制、连接控制、安全、移动性管理、测量和 NAS 消息传输等。但在一些具体特性上, NR 相比 LTE 有了增强,在后续章节中将介绍具体的增强技术。

NR 的用户面协议栈,相比 LTE 做了一定的增强,如图 18 所示^[4]。

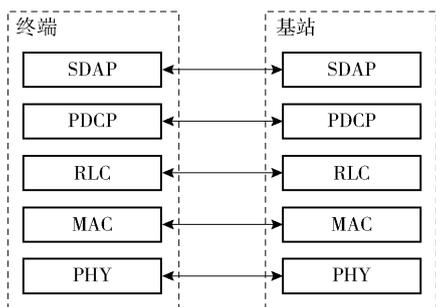


图 18 NR 空口用户面协议栈

用户面协议栈自上至下包括 SDAP (Service Data Adaptation Protocol)、PDCP、RLC 和 MAC 4 个协议层。

SDAP 层是 NR 新引入的协议层^[19],该协议层的引入是因为 5G 核心网相比于 4G 核心网在服务质量 (QoS, quality of service) 管理和承载粒度方面有重大的变化,需要一个新的协议层来进行 5G 的核心网到 NR 空口之间的承载映射。在 4G 阶段,核心网的数据以演进分组系统 (EPS, evolved packet system) 承载的粒度发送到基站,基站将 EPS 承载和空口的无线承载之间维持简单的一对一关系;而 5G 核心网发送到基站的数据粒度为 QoS 流,基站需要将相同或者相近 QoS 需求的 QoS 流映射到一个无线承载上完成空口的传输。SDAP 的主要功能是管

理核心网 QoS 流与空口无线承载之间的映射关系,支持 NAS 层和 AS 层的反向映射功能,在数据包中携带 QoS 流标识信息,支持无损 QoS 流与空口无线承载之间的映射关系变更。

NR PDCP 层^[20]继承了 LTE PDCP 层的基本功能,主要为数据传输提供头压缩和安全方面的操作。相比于 4G LTE, NR PDCP 新增对数据业务进行完整性保护的功能,支持数据重复传输和重复数据删除,支持重排序和乱序递交功能。

NR RLC 层^[21]也是在 LTE RLC 的基础上发展而来的,同样支持 TM、UM 和 AM 三种传输模式,为业务提供不同的传输可靠性保障。相比于 4G LTE, NR RLC 去掉了数据级联功能,RLC SDU 和 RLC PDU 之间是一一对应的关系,主要为了满足在大数据处理时进行预处理的需求,并且将接收端的数据包重排序功能从 RLC 移到了 PDCP,便于提升接收端处理效率。同时针对 5G 不同需求,对业务数据包和控制数据包格式进行了重新设计。

NR MAC 层^[22]沿袭了 LTE 的基本特性,如对双链接 (DC, dual connectivity)、载波聚合 (CA, carrier aggregation) 的支持,MAC 层基本过程等。同时, NR MAC 层针对 5G 引入的一些新特性以及高层和物理层的增强进行了针对性设计。MAC 层支持的新特性包括: BWP、BFR (beam failure recovery)、SUL (supplementary uplink) 和 PDCP duplication 等。MAC 层过程仍包括随机接入、上行 TA 维护、调度传输、资源预分配、BSR (buffer status report) 上报、PHR (power headroom report) 上报和 DRX (discontinuous reception) 等,但每个过程针对 NR 特性各有增强。基于数据预处理的需求,MAC 层对 MAC PDU 结构做了根本性变更,MAC PDU 格式改为由多个 MAC subPDU 组成,每个 MAC subPDU 是一个独立的 MAC 子头和必要的负载的组合。

3.2.2 系统信息

NR 的系统信息设计以 LTE 系统信息为基础,对系统信息进行了更明确的分类,并引入新的传输机制。其主要目的是为了保证终端可以正常地驻留到可以为其提供服务的小区,提供空闲态和非激活态终端接入网络的必要的公共信道的信息以及保证其在空闲态和非激活态下的移动性。

按照系统信息内容的重要程度及传输方式的不同,系统信息分为最小系统信息和其他系统信息,其中最小系统信息仅包含终端驻留和接入一个小区所

必须的系统信息,剩下的都是其他系统信息.并且为了更合理地利用网络资源,其他系统信息的传输机制又分为不间断广播类型以及按需广播类型.按需广播是 NR 引入的新机制,终端如果需要某个系统信息,那么就向网络发起请求,网络基于终端的请求,广播相应的系统信息.这样可以降低网络一直广播系统信息的开销.

3.2.3 非激活态

非激活态是 NR 引入的一个新的 RRC 状态.当终端处于该状态时,终端的 NAS 层仍然保持在连接态(终端与核心网的连接仍然保持),终端的空口连接是断开的,基站侧保留终端的上下文信息,以及保留该终端与核心网的 NG 连接.终端可以在基站配置的一个区域范围内移动而无须通知网络,以节省信令开销.终端进入非激活态时,最后一个服务基站存储着该终端的上下文以及与服务核心网的 NG 连接,终端的 AS 层也保存相应的上下文信息,包括承载、非激活态的标识、归属区域等.通过该方式,基站能在所配置的区域范围内通过无线接入网寻呼机制寻呼到处于非激活态的终端,终端可以基于终端侧和基站侧所存储的上下文信息,快速恢复数据传输,实现低时延传输.

3.2.4 测量

在 NR 系统中,基站在高频段会使用波束扫描方式发送信号.终端侧针对同一个小区会测量到多个波束上对应的测量参考信号结果,此时为了及时、稳定地评估某个小区信号质量,小区信号质量通过平均 $N(N \geq 1)$ 个波束的信号质量获得.因此其相应的测量模型也与 LTE 系统有所不同,增加了对多波束测量结果的滤波和平均.另外,由于高频和低频的信号特性不同,测量间隙的配置也有所不同,相对于 LTE 系统,引入了多套测量间隙的配置.

3.2.5 接入控制

接入控制是系统约束终端接入网络的一种控制技术,主要用于在网络过载时,通过减少用户接入,以降低网络负荷,对各小区业务进行均衡.接入控制参数适用于所有 RRC 状态的终端,包括空闲态、非激活态和连接态.终端通过系统信息获取接入控制参数.终端每发起一种类型的业务,都会触发相应的接入控制过程.网络侧通过应用接入控制技术,可以很好地管理整个网络的资源使用情况.在 LTE 系统中,对应不同类型的业务、不同状态的终端,有多种接入控制机制,实现起来比较复杂. NR

中对接入控制机制做了简化,引入统一的接入控制机制,不同状态、不同业务的终端统一基于接入等级进行接入控制的判断,网络通过在系统信息中广播不同的接入等级参数,实现接入控制.

3.2.6 寻呼

网络可以向空闲态、非激活态和连接态的终端发送寻呼消息.寻呼过程可以由核心网侧触发,用于通知某些终端接收寻呼;或者由接入网侧触发,用于对非激活态用户进行寻呼;或者通知终端进行系统信息更新.因为在 NR 系统中,基站在高频段会使用波束赋形技术扩大覆盖范围,所以寻呼会采用波束扫描方式发送,通过波束扫描,对于不同波束方向的终端分别发送一份寻呼消息,不同于 LTE 全向发送的方式. NR 中继承了 LTE 中的寻呼无线帧计算公式,但是对于具体的寻呼位置,重新定义了公式中部分参数的含义,以更好地适配波束扫描方式的寻呼发送.

3.2.7 重复传输

重复传输是 NR 新引入的一个重要特性,主要是应对 URLLC 业务高可靠性低时延的业务传输需求,将相同的数据在不同的载波上进行重复传输.

重复传输包括 DC 重复传输和 CA 重复传输两种类型.其中 DC 重复传输复用了 LTE DC 分离承载的架构,在分别位于 MCG (master cell group) 和 SCG (secondary cell group) 的两条逻辑信道中传输重复的数据. CA 重复传输采取的是 1 个承载对应 1 个 PDCP 实体和 2 个 RLC 实体的架构,在 2 个 RLC 逻辑信道上传输重复的数据,同时为了保证这种重复传输的增益,为这 2 个 RLC 逻辑信道分别配置不同的载波集合,保证重复的数据是在不同载波上传输,以达到一定的重复增益.

重复传输功能在 RRC 层进行功能开启和初始状态配置,并通过 MAC CE 进行动态激活和去激活操作.当重复传输被激活,则 PDCP 层将所有待发的数据 PDU 均复制之后向两个对应的 RLC 实体进行发送,控制 PDU 仅在主 RLC 实体发送.当重复传输功能被关闭,则 PDCP 层仅向主 RLC 实体发送单发数据, PDCP 指示辅 RLC 实体删除没有传输的重复数据包,辅 RLC 实体不复位,但不再有新数据传输.为了解决两个 RLC 实体传输速率不匹配的情况,已经在一个 RLC 实体传输成功的数据,可以通过 PDCP 层指示删除另一个 RLC 实体未传输的重复包.

4 5G 核心网设计与标准化

对于 5G 核心网, 标准化主要涉及服务化架构设计、核心网的服务功能单元及与移动性管理、会话管理、网络切片、QoS、语音业务等相关的内容, 下面将对此进行介绍。

4.1 服务化架构

区别于 4G 等传统的网络采用网元(或网络实体)来描述系统架构, 5G 系统中引入了网络功能(NF, network function)和服务的概念。不同的 NF 可以作为服务提供者与其他 NF 提供不同的服务, 此时其他 NF 被称为服务消费者。NF 服务提供和消费之间的关系灵活: 一个 NF 既可使用一个或多个 NF 提供的服务, 也可以为一个或多个 NF 提供服务。服务化架构基于模块化、可重用、自包含的思想, 充分利用了软件化和虚拟化技术。每一个服务为软件实现的一个基本网络功能模块, 系统可以根据需要对网络功能进行编排, 就像一块积木, 需要时就可以添加到系统架构中, 不需要时就移除, 这使得网络的部署和演进非常方便灵活, 也有利于引入对新业务的支持。

5G 系统的非漫游场景的架构如图 19 所示, 包含核心网中主要的 NF 和 NF 之间的连接关系。由于标准制定会议的时间有限, R15 中仅有控制面的 NF 实现了服务化, 用户面 NF 之间, 控制面 NF 和用户面 NF 之间, 仍采用点对点方式。R16 将会对服务化架构进一步增强。

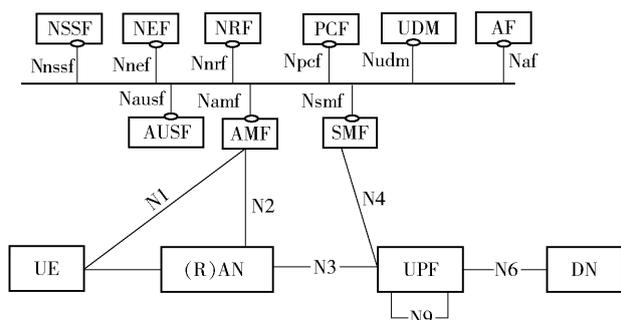


图 19 5G 系统架构^[23]

5G 核心网中主要的 NF 名称和主要功能如下。

1) 接入和移动性管理功能 (AMF, access and mobility management function)。终结来自 UE 的非接入层消息、实现对 UE 的接入控制和移动性管理功能; 终结接入网的控制面接口 (N2) 等。

2) 用户面功能 (UPF, user plane function)。PDU

会话用户面相关功能, 即连接接入网和外部数据网络 (DN, data network) 之间采用特定的封装传递用户数据报文, 实现 QoS、监听、计费等方面的功能; UPF 不但实现 4G 网络中服务网关 (SGW, serving gateway)、分组数据网关 (PGW, PDN gateway) 中的用户面的各项功能外, 还支持边缘计算等新特性所需的用户面功能。

3) 会话管理功能 (SMF, session management function)。PDU 会话管理 (建立、删除、修改等)、UPF 选择、终端 IP 地址分配等; SMF 实现了 4G 网络中 SGW、PGW 中的控制面的各项功能。

4) 网络存储功能 (NRF, NF repository function)。实现服务的管理功能。NF 启动时将自己提供的服务注册到 NRF。当 NF 需要使用服务时, 先查询 NRF, 即可发现提供该服务的 NF 信息。

5) 统一数据管理功能 (UDM, unified data management)。用户签约数据和鉴权数据的管理。

6) 鉴权服务器功能 (AUSF, authentication server function)。实现对用户的鉴权的相关功能, 与安全锚点功能 (SEAF, security anchor function) 配合完成密钥相关的操作。

7) 策略控制功能 (PCF, policy control function)。实现统一的策略和计费控制的节点, 制定并下发策略给控制面 NF、UE。

8) 网络开放功能 (NEF, network exposure function)。实现将网络能够提供的业务和能力“暴露”给外部如第三方实体。

9) 网络切片选择功能 (NSSF, network slice selection function)。为 UE 选择为其服务的网络切片、AMF 等。

10) AF (application function)。与核心网交互, 以提供业务 (如 IMS 的 AF 提供 IMS 语音呼叫服务)。

利用计算和存储相互分离的思想, 5G 核心网还引入了可选的网络功能: UDSF (unstructured data storage function), 实现非结构化数据的存储, 并为任意控制面的 NF 提供检索功能。例如, 将 AMF 中 UE 上下文数据交由 UDSF 存储, 其他的 AMF 也可以访问, 并在必要时比如某 AMF 死机时接管这些用户数据。这种分离不但提升了网络的鲁棒性, 还天然地支持 NF 的虚拟化部署, 如运行在虚拟环境中的 NF 可以按需调增或调减计算能力。

3GPP CT3 和 CT4 工作组经过研究决定^[24], 服

务化架构的 NF 之间的通信协议采用情况是:

- 1) 应用层:HTTP/2;
- 2) 传输层:TCP;
- 3) 序列化协议:JSON;
- 4) 应用 RESTful framework 进行接口设计,并采用 OpenAPI 3.0.0 作为接口定义语言。

4.2 网络切片

网络切片是 5G 系统架构的关键特性之一,指为服务一个行业或者一类终端或者某些特定的场景,从一个公共陆地移动网络 (PLMN, public land mobile network) 中选取特定的特性和功能,定制出的一个逻辑上独立的网络。网络切片使得运营商可以部署功能、特性服务各不相同的多个逻辑网络,分

别为各自的目标用户服务。目前定义了 3 种网络切片类型:eMBB、URLLC、MIoT,如图 20 所示:

- 1) 终端同时从切片 1、切片 2 接收服务,切片 1 和 2 共享 AMF、PCF、NRF,由一个 AMF 对终端的所有业务进行控制;
- 2) 终端的用户面服务可以从多个切片获得,如分别通过切片 1 和 2,可以实现与数据网络 1 和 2 分别进行通信;
- 3) 切片 3 是由所有的网络功能组成的单独的切片。

图 20 中的切片 1 和 2 虽然与公共的 AMF 等 NF 交互,但可根据业务需求,分别提供完全不同 QoS 的数据传输服务。

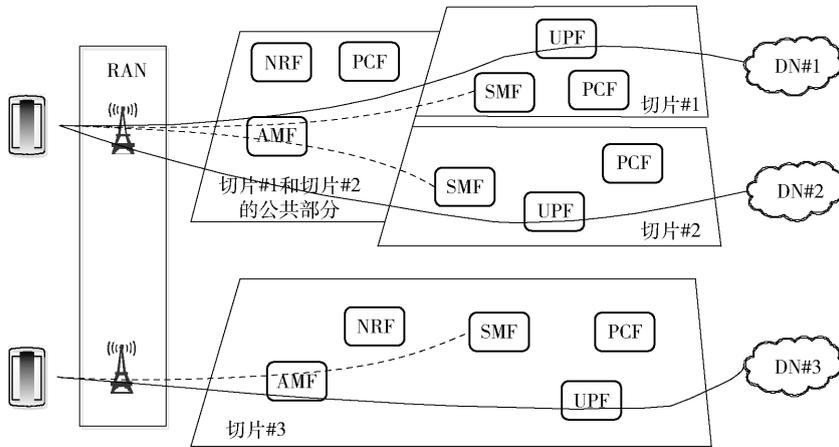


图 20 5G 网络切片部署示例^[25]

4.3 定制化移动性管理

5G 核心网对移动性管理的功能进行了增强,主要包括定制化的移动性管理和统一的非接入层协议等。

定制化的移动性管理能力可以根据不同用户的特点和部署场景,对用户进行定制化移动性管理和功能定制,如 AMF 可以根据终端的移动特点定制以终端为颗粒度的移动性管理功能,如确定注册区域和相关定时器的时长等;移动性功能与应用层功能紧密结合,通过 PCF 实现移动性限制区域、UE 移动性相关策略的制定和下发等增强。

通过统一的非接入层协议,无论是 3GPP 接入还是非 3GPP 接入,都可以实现统一的移动性管理过程,简化了网络的移动性管理。

4.4 会话和业务连续性

为了满足多种数据业务,如支持 IP 报文、非结构化数据、Ethernet 等的数据传输需求,5G 网络需要

实现灵活高效的移动性管理。

如图 21 所示,针对终端移动时的业务连续性需求,定义了 3 种会话和业务连续性模式 (SSC, session and service continuity mode)。

- 1) 模式 1. 终端移动时,其 IP 锚点不变。适合 IP 多媒体子系统 (IMS, IP multimedia subsystem) 话音等低中断时延要求的业务,也是 EPS 系统的唯一默认模式;
- 2) 模式 2. 也称“先断后建”,终端移动后会采用新的 IP 锚点,适合网页浏览等业务;
- 3) 模式 3. 也称“先建后断”,终端移动时保持旧的锚点,移动后获得新的 IP 锚点,此后根据需要断开旧的锚点;这种模式适合优化用户面路径,避免业务中断的业务,如某些视频播放业务等。

为了支持移动边缘计算等应用,SMF 可以为 PDU 会话选择位置合适的 UPF。同时具有在用户面路径中合适的 UPF 处插入上行分类器 (uplink clas-

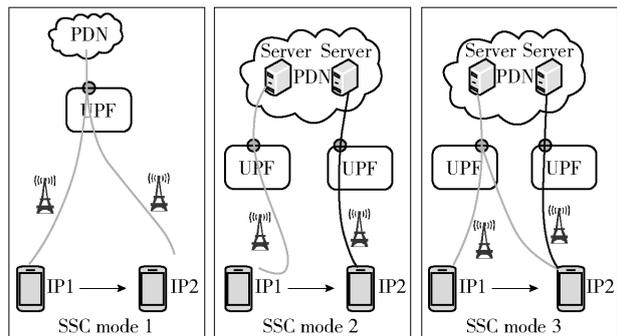


图 21 3 种会话和业务连续性模式^[25]

sifier) 或者分支点 (branching point) 功能. 通过该功能可将用户数据直接路由到相应的数据网.

在网络策略许可的情况下, 通过与网络协调, 可为 PDU 会话选取适合的模式, 位置恰当的 UPF, 以及设置上行分类器或者分支点, 实现业务数据路径的优化.

支持上行分类器的用户面网络架构如图 22 所示.

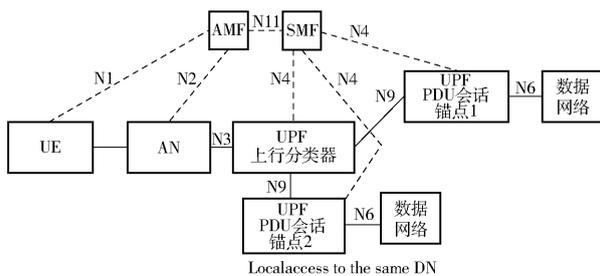


图 22 支持上行分类器的用户面架构^[23]

4.5 QoS

5G 网络定义了面向多种接入网的新的 QoS 框架, 在满足多样化应用需求的同时, 可以有效地利用接入网资源.

这个框架具有以下特点.

- 1) 更细颗粒度的 QoS 控制, 核心网基于 IP 流实现, 接入网支持 IP 与无线承载的灵活映射;
- 2) 在标准化的 5QI 参数基础上, 支持核心网根据需求灵活定义非标准化的 5QI 和细粒度的 QoS 参数对应关系, 通过信令发送到接入网;
- 3) 引入 Reflective QoS 机制, 实现上行数据和下行数据的对称式 QoS 处理, 避免了显式控制信令对系统的影响;
- 4) 针对 non-3GPP 接入网, 引入了 In-band QoS 控制, 即接入网根据数据包头的 QFI 获知 QoS 参数, 无需核心网与接入网之间的信令.

4.6 PCC

5G 的策略和计费控制 (PCC, policies and charging control) 架构主要实现了如下新功能:

- 1) 在支持 PDU 会话策略控制的基础上, 增加了对终端相关策略、移动性相关策略的支持;
- 2) 支持根据网络数据分析功能 (network data analytics function) 得出的分析结果, 制定网络控制相关策略 (由于标准化时间限制, R15 版本仅仅实现了网络切片负荷相关).

4.7 网络的演进与互通

5G 系统需求明确了 5G 网络不与 3G、2G 等网络之间进行互操作^[26]. 同时, 为了避免将 4G 网络中原有设计的不足或者说“历史包袱”引入和影响 5G 核心网性能, 采取了尽可能少地在 5G 和 4G 系统之间建立接口的设计原则. 基于这个原则, 4G 网络和 5G 网络的互通方法采用将 4G 接入网直接连接到 5G 核心网, 而不像以往的新 4G 网络通过 3G 核心网节点连接 3G 的接入网的处理方式. 4G 接入网同时也可以连接到 4G 核心网, 为 4G 终端服务.

在 4G 和 5G 系统并存期间, 为了实现数据业务的连续性, 设计了“双注册”模式, 即终端同时在 4G 和 5G 系统完成注册、鉴权等过程, 这样做可以减少系统间切换的时延. 虽然这种方法在某种程度上实现了数据业务的连续性, 也不需要两个系统间引入接口, 但是无法满足 IMS 话音业务对中断时间延迟的要求. 为此, 系统设计时不得不在 4G 和 5G 系统间引入 N26 接口, 即移动性管理实体 (MME, mobility management entity) 和 AMF 之间的接口, 用于在切换时传递终端的上下文信息, 特别是 PDU 会话的信息, 在终端接入到目标系统前, 目标网络为终端准备好用户面资源, 以满足有较高连续性需求的业务的需要.

4.8 5G 网络和话音业务

5G 系统可以采用基于 IMS 系统的 VoNR 技术或者话音回退到 4G 系统, 通过 VoLTE 方式来支持语音业务.

对于 VoNR 功能, 需要终端、接入网、核心网三者都支持相关功能时, VoNR 才能正常工作. 接入网需要在无线信道环境中建立承载 IMS 话音服务质量流 (QoS Flow); 核心网需要部署支持 IMS 系统的相关功能, 如支持系统间业务连续性, 核心网与 IMS 系统间设立接口等. 当接入网、核心网都支持 IMS 时, 网络将指示终端当前系统支持 IMS 语音.

终端通过 NR 网络进行话音业务传输时,如 5G 接入网的信道质量能够满足终端发起建立 IMS 特定的 QoS Flow 的要求,则可以直接通过接入网进行呼叫,建立 VoNR 的业务连接;如果无法满足,则接入网指示终端“回退”到 EPS 系统,或者“回退”到连接 5G 核心网的演进 E-UTRAN 接入网,通过 VoLTE 方式实现 IMS 话音呼叫。

当 5G 和 4G 网络质量都无法支持 IMS 业务时,5G 系统需要将话音呼叫转移到 3G 网络中,具体的实现方案,由于 R15 的标准化实现限制,未开展讨论,相关解决方案将在 R16 研究和标准化。

5 结束语

随着支持独立组网和非独立组网的 3GPP R15 标准正式发布,基于灵活动态 TDD 信号传输设计,新型的接入网和核心网络架构,以及大规模波束赋形、新型 LDPC 和 Polar 信道编码、网络切片、网络功能虚拟化等关键技术应用,5G 系统能够初步满足未来无线移动通信业务对更为严苛的技术指标、更为灵活多样的部署环境、更为丰富的频谱资源的需求。

同时,3GPP 已经开始针对 R16 的标准化工作,其主要标准化内容可以分为:

1) 基于 R15 的基础能力提升,如非授权频段的标准制定,定位技术,终端节能等;

2) 增强移动宽带能力提升,如大规模波束赋形增强,移动性管理增强,支持在更高频段传输优化,非正交多址接入;

3) 物联网业务扩展,如基于 NR 的车联网技术,工业互联网等。

R16 的标准化将进一步提升 5G 系统性能,并能在更多场景中得到应用,满足移动通信业务不断发展的需求。

致谢 在本文撰写过程中,电信科学技术研究院的梁靖、刘爱娟、湛丽和刘佳敏参与空口高层协议及接入网架构设计等内容的撰写,中国信息通信研究院的朱浩对核心网设计部分亦有贡献,电信科学技术研究院的王加庆对信道编码部分撰写提供了帮助,在此一并表示感谢!

参考文献:

[1] IMT-2020(5G)推进组. 5G 愿景与需求白皮书[EB/OL]. 2014(2014-05-09). <http://www.imt-2020.org.cn/zh/documents/1>.

[2] ITU-R. M. 2410-2017, Minimum requirements related to technical performance for IMT-2020 radio interface (s) [EB/OL]. 2017(2017-11-07). <https://www.itu.int/pub/R-REP-M.2410>.

[3] ITU-R. WP5D Document 5D/TEMP/455, Liaison statement to external organizations on the work plan, timeline, process and deliverables for the future development of International mobile telecommunications (IMT) [EB/OL]. 2015(2015-06-06). <https://www.itu.int/md/R12-WP5D-C-1042>.

[4] 3GPP. TS38. 300. 2018, NR and NG-RAN overall description stage 2 [EB/OL]. 2018(2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191>.

[5] 3GPP. TS38. 401-2018, NG-RAN architecture description [EB/OL]. 2018(2018-06-22). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3219>.

[6] 3GPP. TS37. 340-2018, E-UTRAN and NR multi-connectivity stage 2 [EB/OL]. 2018(2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3198>.

[7] 3GPP. TS38. 201-2017, NR physical layer general description [EB/OL]. 2018(2018-01-03). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3211>.

[8] 3GPP. TS38. 202-2018, NR services provided by the physical layer [EB/OL]. 2018(2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3212>.

[9] 3GPP. TS 38 211-2018, NR physical channels and modulation [EB/OL]. 2018(2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213>.

[10] 3GPP. TS 38. 212-2018, NR multiplexing and channel coding [EB/OL]. 2018(2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3214>.

[11] 3GPP. TS 38. 213-2018, NR physical layer procedures for control [EB/OL]. 2018(2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3215>.

[12] 3GPP. TS 38. 214-2018, NR physical layer procedures for data [EB/OL]. 2018(2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3216>.

[13] 3GPP. TS 38. 215-2018, NR physical layer measure-

- ments [EB/OL]. 2018 (2018-06-29). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3217>.
- [14] EArikan. Channel polarization; A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels [J]. *IEEE Trans Inform Theory*, 2009, 55(7): 3051-3073.
- [15] NiuK and Chen K. CRC-Aided decoding of polar codes [J]. *IEEE Comm Lett*, 2012, 16(10): 1668-1671.
- [16] 3GPP. R1-1706193, Huawei, Polar coding design for control channel[S].
- [17] 3GPP. R1-1708833, Nokia, Design details of distributed CRC[S].
- [18] 3GPP. TS 38. 331-2018, NR radio resource control (RRC) protocol specification [EB/OL]. 2018 (2018-06-21). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3197>.
- [19] 3GPP. TS 37. 324-2018, E-UTRA and NR service data adaptation protocol (SDAP) specification [EB/OL]. 2018 (2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3282>.
- [20] 3GPP. TS 38. 323-2018, NR packet data convergence protocol (PDCP) specification [EB/OL]. 2018 (2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3196>.
- [21] 3GPP. TS 38. 322-2018, NR radio link control (RLC) protocol specification [EB/OL]. 2018 (2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3195>.
- [22] 3GPP. TS 38. 321-2018, NR medium access control (MAC) protocol specification [EB/OL]. 2018 (2018-06-20). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3194>.
- [23] 3GPP. TS 23. 501-2018, System Architecture for the 5G System [EB/OL]. 2018 (2018-06-19). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>.
- [24] Yue Song. CP-172074, LS on conclusion on service based architecture protocol selection [EB/OL]. 2017 (2017-09-11). ftp://www.3gpp.org/TSG_CT/TSGC_77_Sapporo/Docs.
- [25] Frank Mademann. System architecture milestone of 5G Phase 1 is achieved [EB/OL]. 2017 (2017-12-21). http://www.3gpp.org/NEWS-EVENTS/3GPP-NEWS/1930-SYS_ARCHITECTURE.
- [26] 3GPP. TS 22. 261-2017, Service requirements for next generation new services and markets [EB/OL]. 2017 (2017-09-22). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107>.