

5G NR (新空口) 中灵活帧结构对时延敏感业务的支持能力评估

黄忠源

中国移动通信集团广西有限公司防城港分公司 广西 防城港 538021

摘要: 随着工业互联网、车联网、远程医疗和增强现实/虚拟现实 (AR/VR) 等新兴应用场景的快速发展, 通信系统对端到端时延提出了前所未有的严苛要求。第五代移动通信系统 (5G) 通过引入新空口 (New Radio, NR) 技术, 旨在满足增强型移动宽带 (eMBB)、超可靠低时延通信 (URLLC) 和海量机器类通信 (mMTC) 三大核心场景的需求。其中, URLLC场景对时延的要求通常低于1毫秒, 这对传统固定帧结构的4G LTE系统构成了巨大挑战。为此, 3GPP在5G NR标准中引入了灵活帧结构 (Flexible Frame Structure) 机制, 通过动态调整上下行时隙配比、符号级调度粒度以及自包含子帧 (Self-Contained Subframe) 等关键技术, 显著提升了系统对时延敏感业务的支持能力。本文系统地分析了5G NR灵活帧结构的技术原理, 构建了面向URLLC业务的时延模型, 并通过仿真实验定量评估了不同帧结构配置下端到端时延性能。研究表明, 采用高比例上行或下行专用时隙、缩短传输时间间隔 (TTI) 至单个OFDM符号级别, 并结合免授权调度 (Grant-Free) 机制, 可将用户面时延有效控制控制在1毫秒以内。同时, 本文也探讨了灵活帧结构在实际部署中面临的干扰协调、资源碎片化和调度复杂度等挑战, 并对未来6G系统中的进一步优化方向进行了展望。

关键词: 5G NR; 灵活帧结构; URLLC; 时延敏感业务; 自包含子帧; 免授权调度; 端到端时延

引言

第五代移动通信 (5G) 作为数字经济发展的关键基础设施, 核心目标之一是突破传统移动通信在时延方面的瓶颈。国际电信联盟 (ITU) 定义的5G关键性能指标里, URLLC场景要求空口时延低于1毫秒, 可靠性高达99.999%, 这对自动驾驶车辆协同感知、工厂机械臂实时控制、远程手术操作等应用意义重大。但4GLTE系统因采用固定10ms无线帧和1ms子帧结构, 最小调度单元为1个子帧 (14个OFDM符号), 难以满足亚毫秒级时延需求^[1]。为应对挑战, 3GPP在Release 15及后续版本为5G NR设计全新物理层帧结构体系, 摒弃LTE僵化的上下行配置模式, 采用高度灵活的时隙和符号级资源配置机制, 不仅上下行方向可动态切换, 还能根据业务需求自适应调整传输时间间隔 (TTI) 长度, 为时延敏感业务提供前所未有的调度自由度。不过, 灵活帧结构虽理论上具备支持超低时延潜力, 实际网络环境中其性能表现受子载波间隔 (SCS)、时隙配置比例、调度策略、信道条件等多种因素影响。目前, 业界对灵活帧结构如何具体影响端到端时延, 尤其是混合业务负载下的性能边界, 缺乏系统性量化评估。鉴于此, 本文旨在深入剖析5G NR灵活帧结构技术内涵, 建立面向URLLC业务的端到端时延分析模型, 通过仿真实验定量评估不同参数配置对时

延性能的影响, 为5G网络规划者和设备制造商提供理论依据与实践指导, 优化网络配置以最大化支持时延敏感业务。

1 5G NR 灵活帧结构关键技术

1.1 基本帧结构组成

5G NR的无线帧长度仍为10ms, 由10个子帧 (Subframe) 组成, 每个子帧长度为1ms。但与LTE不同的是, NR引入了“时隙 (Slot)”作为基本的调度单元。一个时隙包含的OFDM符号数量取决于子载波间隔 (SCS)。具体关系如下:

$$N_{\text{symbol}} = 14 \quad \text{for } \mu = 0 (15\text{kHz})$$

$$N_{\text{symbol}} = 14 \quad \text{for } \mu = 1 (30\text{kHz})$$

$$N_{\text{symbol}} = 12 \text{ or } 14 \text{ for higher } \mu$$

其中, μ 是SCS配置索引。更重要的是, 一个子帧可以包含 2^μ 个时隙。这意味着, 当SCS增大时, 时隙长度缩短, 从而实现了更精细的调度粒度。例如, 在30kHzSCS ($\mu = 1$) 下, 一个时隙长度为0.5ms; 在120kHzSCS ($\mu = 3$) 下, 一个时隙长度仅为0.125ms。

1.2 自包含子帧 (Self-Contained Subframe)

这是灵活帧结构的核心创新之一。在一个自包含的传输单元 (通常是一个时隙) 内, 可以同时包含下行控制信息 (DCI)、下行数据 (PDSCH)、上行反馈

(HARQ-ACK)以及上行数据(PUSCH)^[2]。这种“发送-接收-反馈”闭环在一个极短的时间窗口内完成,极大地缩短了HARQ重传的等待时间。

典型的自包含时隙结构有两种:(1)下行自包含:包含PDCCH(下行控制)、PDSCH(下行数据)和PUCCH(上行ACK/NACK)。(2)上行自包含:包含PDCCH(上行授权)、PUSCH(上行数据)和PUCCH(下行ACK/NACK)。

这种设计使得一次完整的数据传输和确认可以在一个时隙内完成,避免了跨时隙甚至跨子帧的等待,是实现亚毫秒级时延的关键。

1.3 动态TDD与符号级调度

5G NR支持动态TDD(Time Division Duplexing),允许网络在符号级别上动态分配上下行资源。通过高层信令(RRC)或物理层信令(DCI)指示每个时隙内每个符号的用途(下行、上行或灵活)。这种灵活性使得网络可以根据瞬时业务需求,例如突发性的上行传感器数据或下行控制指令,即时调整资源分配,避免了固定配比造成的资源浪费和时延增加。

1.4 免授权调度(Grant-Free / Configured Grant)

对于周期性或半持续性的URLLC业务,传统的基于调度请求(SR)和动态授权(Dynamic Grant)的四步握手机制会引入显著时延。5G NR引入了免授权调度机制,基站预先为UE配置一组上行资源(Configured Grant),UE在有数据时可直接使用这些资源进行传输,无需等待网络授权。这消除了调度请求和授权的往返时延,是降低上行接入时延的有效手段。

2 面向URLLC业务的端到端时延建模

端到端时延(End-to-End Latency)可分解为空口时延(Air Interface Latency)和核心网处理时延。本文聚焦于空口时延,其主要组成部分包括:(1)数据包排队时延(T_{queue}):数据包在MAC层缓冲区等待被调度的时间。(2)调度授权时延(T_{grant}):从UE发出调度请求到收到基站授权的时间。对于免授权业务,此项为0。(3)传输时延(T_{tx}):数据在物理信道上传输所需的时间,等于TTI长度。(4)传播时延(T_{prop}):电磁波在空间中传播的时间,通常可忽略(微秒级)^[3]。(5)HARQ处理与反馈时延(T_{hard}):包括接收端解码、生成ACK/NACK、以及ACK/NACK在上行信道上传输的时间。

在采用自包含子帧和免授权调度的理想情况下,总空口时延可近似为:

$$T_{\text{total}} \approx T_{\text{tx}} + T_{\text{hard}}$$

假设采用120kHzSCS($\mu = 3$),一个时隙包含14个

符号,时长为0.125ms。若TTI配置为2个符号,则 $T_{\text{tx}} = 2 \times (0.125/14) \approx 18\mu\text{s}$ 。HARQ反馈在同一时隙内完成,其时延主要由处理时间和剩余符号数决定,通常在几十微秒量级。因此,理论上的单次传输时延可远低于1ms。

然而,在实际网络中,还需考虑资源竞争、信道质量导致的重传等因素,使得实际时延具有随机性和不确定性。

3 仿真实验与性能评估

3.1 仿真设置

为定量评估灵活帧结构的性能,我们搭建了一个基于MATLAB的5G NR链路级仿真平台。主要参数设置如下:

频段:2.6 GHz(n41,中国移动主流部署频段)

带宽:100MHz

子载波间隔(SCS):15,30,60,120kHz

时隙配置:采用3GPP定义的TypeA/B/C/D/E共5种半静态TDD配置,并对比全动态符号级调度。

业务模型:URLLC业务(包大小32字节,到达间隔服从泊松分布,平均到达率10kpps),混合eMBB背景流量。

调度策略:对比动态调度与免授权调度(Configured Grant Type 1)。

信道模型:3GPP UMi(Urban Micro-cell)场景。

3.2 评估指标

平均端到端时延(Mean E2E Latency)

时延超过1ms的包比例(Packet Loss due to Latency > 1ms)

系统吞吐量(Throughput)

3.3 结果分析

3.3.1 子载波间隔(SCS)的影响

仿真结果显示,随着SCS从15kHz增加到120kHz,URLLC业务的平均时延显著下降。在120kHzSCS下,采用2符号TTI和免授权调度,平均时延可降至0.35ms,且99.99%的包满足<1ms的时延要求。而15kHzSCS下,即使采用最优配置,平均时延也在1.2ms左右,无法满足URLLC需求。

3.3.2 帧结构配置的影响

在固定TDD配置中,TypeD(10:2:2,即10个下行符号,2个灵活,2个上行)在以下行为主的URLLC场景(如工业控制指令下发)中表现最佳。而在以上行为主的场景(如传感器上报),TypeB(2:2:10)更为合适。然而,全动态符号级调度在混合业务场景下展现出最强的适应性,能将时延方差降低40%以上。

3.3.3 调度策略的影响

免授权调度相比动态调度，在高负载下优势尤为明显。当URLLC业务负载达到50%时，动态调度的平均时延飙升至2.5ms，而免授权调度仍能稳定在0.8ms以内。这是因为免授权机制完全规避了调度请求队列的拥塞。

3.3.4 与eMBB业务的共存

在混合业务场景下，灵活帧结构通过为URLLC业务预留专用符号或时隙（通过半持久调度SPS），可以有效隔离其与eMBB业务的资源竞争。仿真表明，只要为URLLC预留20%的符号资源，即可保证其99.999%的可靠性。

4 挑战与未来展望

尽管5G NR灵活帧结构在理论上能有效支持超低时延业务，其实际部署仍面临多重挑战。首先，小区间干扰协调问题突出：动态TDD若在相邻小区间未同步，将引发严重的上下行交叉链路干扰（CLI），显著恶化上行接收质量，影响URLLC可靠性与时延。虽然3GPP引入了CLI指示机制，但在高密度组网下，信令开销与协调延迟仍制约性能。其次，资源碎片化降低了系统整体效率。符号级调度虽提升时延灵活性，却导致时频资源高度离散，不仅增加调度复杂度，还可能牺牲eMBB等大带宽业务的吞吐量，难以在混合业务场景中实现最优平衡^[4]。第三，基站调度器的实时性要求极高。要在亚毫秒内完成多用户、多业务的优先级判断与资源分配，对基带处理单元的算力和算法效率提出严峻挑战，尤其在高负载下易成为性能瓶颈。此外，终端功耗上升亦不可忽视。短TTI和频繁的上下行切换迫使UE射频模块频繁启停，并在极短时间内完成信号处理，显著增加能耗，对电池受限

的工业物联网设备尤为不利。

面向未来，6G有望通过深度融合人工智能与无线资源管理，实现基于业务预测的智能帧结构配置；结合通感一体化（ISAC）技术，利用环境感知信息动态优化调度；同时，太赫兹频段提供的超大带宽将支持更细粒度的时频划分，进一步释放灵活帧结构在确定性低时延通信中的潜力。这些方向将共同推动超可靠低时延通信向更高性能演进。

5 结语

本文系统评估了5G NR灵活帧结构对时延敏感业务的支持能力。研究表明，通过采用大子载波间隔（如120kHz）、自包含子帧、符号级动态TDD以及免授权调度等关键技术组合，5G NR能够有效满足URLLC场景下低于1毫秒的空口时延要求。仿真实验验证了在不同业务模型和网络负载下，灵活帧结构相较于传统固定结构的巨大优势。然而，其在实际部署中仍需解决干扰协调、资源管理和终端功耗等挑战。未来的研究应聚焦于智能化、协同化的帧结构管理机制，以充分释放5G乃至6G在超低时延通信领域的潜能。

参考文献

- [1]毛志伟.5G空口授时信号捕获设计与实现[D].中国科学院大学(中国科学院国家授时中心),2025.
- [2]张忠皓,高帅,李福昌,等.5G毫米波业务需求和灵活帧结构方案分析[J].邮电设计技术,2021,(04):1-4.
- [3]蒋伟然.面向5G NR的随机接入方案的研究及优化[D].扬州大学,2025.
- [4]任青,王森,杨东来,等.空闲模式下5G NR终端节能增强方案研究[J].邮电设计技术,2025,(02):12-18.