

LTE 语音解决方案--VOLTE 调度机制的研究

摘 要

随着 LTE 无线接入标准的日益成熟，LTE 研究热点转向了应用业务的性能提升和优化上，语音业务作为运营商最主要的业务之一需要重点研究。

本文首先对 LTE 网络下的目标语音解方案--VoLTE 技术做了详细的分析，涉及 VoLTE 技术实现的各个层面，包括 VoLTE 技术的基本架构、基本业务流程、设备功能要求和互通技术等，并对几种 LTE 语音解决方案的应用进行了比较和总结。接着分析了 LTE MAC 层资源调度机制，对现有调度方案进行了总结和对比；最后提出一种基于优先级的半持续调度方案，并对其性能进行分析。

关键词：VOLTE；调度；算法；语音技术

目 录

摘 要.....	1
第一章 引言.....	3
1.1 研究背景及意义.....	3
1.2 LTE 网络演进过程.....	4
1.3 研究内容.....	5
第二章 LTE 语音方案分析.....	6
2.1 基本架构.....	6
2.2 基本流程.....	7
2.3 设备功能要求.....	13
2.4 互通技术.....	14
第三章 LTE 语音业务性能改进的研究.....	16
3.1 LTE MAC 层概述.....	16
3.2 MAC 信道特点.....	17
3.3 MAC 层调度技术研究.....	19
3.3.1 调度概述.....	19
3.3.2 资源调度的基本单位.....	19
3.3.3 现有调度算法研究.....	21
3.3.3.1 动态调度算法.....	21
3.3.3.2 半持续调度算法.....	25
3.3.3.3 组调度算法.....	27
3.3.3.4 现有调度方案对比.....	27
3.4 本章小结.....	28
第四章 VOLTE 调度机制及性能的优化策略.....	29
4.1 VOLTE 调度机制及性能的优化方案设计.....	29
4.2 优先级算法.....	错误！未定义书签。
4.3 基于优先级的半持续调度方案性能的分析.....	34
结 语.....	36
参考文献.....	37

第一章 引言

1.1 研究背景及意义

虽然演进的分组系统（EPS：Evolved Packet System）最终可以为用户提供 LTE 语音（VoLTE：Voice over LongTerm Evolution）业务，但是，由于各个运营商现有的网络运行模式不同，语音业务从电路域往全 IP 网络的业务模式转变路线和时间点也将会各不相同。对于国外有些小规模移动运营商以及新建网络的运营商来说，有可能直接进入 EPS 网络运营阶段，采用 IP 多媒体子系统（IMS：IP Multimedia Subsystem）业务控制系统来提供各种业务，包括实时语音业务，可以完全实现电路域现有的全部业务。但是，对于已经部署了成熟的 2G/3G 网络的运营商，考虑到目前的电路域网络已经能够提供高质量的语音业务、补充业务和增值业务，而且其网络管理能力也非常完善，出于对建设和网络成本的考虑，最大限度地重用运营商在传统网络的投资的目的，这些运营商的网络演进过程会是阶段性的。业界普遍认可的一种演进步骤，是在 EPS 网络的初期建设只做热点覆盖，并只给 EPS 用户提供数据卡业务；随着 EPS 网络部署规模的扩大，逐渐提供各种实时语音业务及消息类业务等。

基于以上考虑，业界对如何在 EPS 网络发展的过渡阶段实现语音业务连续性，进行了全面的分析和技术评估，提出了电路域回落（CSFB：CircuitSwitched Fallback）和单无线频率语音呼叫连续性（SRVCC：Single Radio Voice Call Continuity）两种过渡期方案。

CSFB 就是在 LTE 和 2G/3G 的双覆盖区域，用户可以从 LTE 系统回落到 CS 域来使用电路域现有的业务，包括话音、定位业务和补充业务等。

支持 CSFB 的终端需要在 EPS 和 CS 域进行双注册。当终端有语音业务的时候，由网络根据策略决定将该终端切换到 2G/3G 电路域完成语音业务。CSFB 方案要求 LTE 与 2G/3G 区域重叠覆盖，这样终端才能同时注册到 LTE 和 2G/3G 网络，并且随时可以切换到 2G/3G 网络执行电路域业务。

此外，在运营商的 EPS 网络发展过程中，可能会出现 EPS 网络在部分地区能够提供 VoIP 业务，但不能实现 VoIP 业务全覆盖的场景，此时可以考虑借用

2G/3G 网络电路域的覆盖来保证业务本身的全覆盖。在 EPS 网络覆盖好的地区使用 IMS 的 VoIP 业务,当用户到达 EPS 网络覆盖边缘却具备良好 2G/3G 网络覆盖时,可在 EPS 网络控制下,将用户正在进行的 VoIP 通话接续到 2G/3G 网络电路域。为了实现两个接入系统间语音业务切换的连续性,3GPP 提出 SRVCC 方案,通过 IMS 域网元对语音业务进行信令和媒体的接续、转接。SRVCC 实际上是个切换过程,通过 IMS SRVCC 的应用服务器 (AS: Application Server) 和承载网络实体 (MME: Mobility Management Entity) / 移动交换服务器 (MSC Server) 的配合,实现语音业务的连续性。语音业务的信令从 IMS 向 MSC Server 转接,承载信令从 MME 向 MSC Server 转接,用户媒体数据流从 IMS- 媒体网关 (MGW: Media Gateway) 转接到 MGW。

与 CSFB 方式不同的是,CSFB 是在 EPS 与 2G/3G 重叠覆盖区域内发生回落切换,而 SRVCC 则是在 LTE 网络失去覆盖的时候,才发生到 2G/3G 网络电路域间的切换。

1.2 LTE 网络演进过程

下面给出了 LTE 网络可能演进的一个过程举例。需要注意的是,各个运营商的 LTE 网络演进方式各不相同,有可能跳过以下提到的某一步,直接到达 VoLTE 目标网架构。

阶段 1: 最初的 LTE 网络规模小,终端模式简单,双模双待机,同时驻留在 LTE 网络和 2G/3G 网络,LTE 网络只提供数据业务,语音业务由 2G/3G 网络提供。

阶段 2: 随着网络的改造,终端功能完善,系统提供 CSFB 功能,此时终端是双模单待模式,可以接入 LTE 网络和 2G/3G 网络,但只能驻留在一个系统下。一般终端都是驻留在 LTE 网络下,可以执行数据业务,一旦有语音呼叫,终端将无线接入切换到 2G/3G 系统下,即驻留到 2G/3G 网络中,执行普通的 2G/3G 语音呼叫。呼叫完成后将再切换到 LTE 系统下。

阶段 3: 如果运营商 IMS 系统成熟,终端能力允许,VoLTE 语音逐渐开展,在未覆盖 LTE 区域,系统通过 SRVCC 技术,实现语音业务的切换,将用户的语音业务从 LTE 切换到 2G/3G 系统中。

阶段 4: LTE 全面覆盖, 技术和设备完全成熟, 将全网实现 VoLTE 业务, SRVCC 和 CSFB 功能将不再使用。

1.3 研究内容

本文主要对 LTE 网络下的目标语音解决方案--VoLTE 技术做了详细的分析, 涉及 VoLTE 技术实现的各个层面, 包括 VoLTE 技术的基本架构、基本业务流程、设备功能要求和互通技术等, 并对几种 LTE 语音解决方案的应用进行了比较和总结。接着对 LTE 语音业务性能从 MAC 层研究出发, 对 VoLTE 调度算法性能改进进行研究。最后提出一种基于优先级的半持续调度模型, 并对其性能分析。

第二章 LTE 语音方案--VOLTE 技术分析

2.1 基本架构

EPS 系统在部署初期有可能只为用户提供数据业务，在这个阶段核心网部分仅有演进的分组核心网（EPC：Evolved Packet Core）网络就能满足对运营商的业务需求；但随着网络的发展和演进，终将要为 EPS 系统为用户提供 VoIP 业务，此时仅有 EPC 网络就无法满足需求了，需要同时部署 IMS 和 PCC 系统提供业务控制和 QoS 策略控制。

尽管从理论上说，用户通过 EPS 分组域就可以实现所有的业务，但如果没有独立于 EPS 分组域之上的业务层来提供业务控制，用户在业务的服务质量（QoS）、计费等方面都将会存在问题。用分组域提供实时多媒体业务的主要问题是网络只能尽力而为地提供业务而不能保证业务的服务质量。例如，对于 VoIP 业务，用户在通话过程中的语音质量就可能变化很大，可能会出现声音时好时坏的现象。跟现有的电路域语音质量比，这样没有 QoS 保证的通话是无法让人接受的。但如果采用 IMS 系统作为业务控制层，就可以提供具有 QoS 保证的实时多媒体会话控制机制，很好地解决这个问题。因此，鉴于 IMS 系统的特点和优势，业界已倾向采用选择 IMS 系统为 EPS 系统的用户提供 VoIP 语音业务控制。

采用 IMS 作为业务控制层系统的另一个原因是能够合理、灵活地对多媒体会话进行计费。运营商可以基于用户的 QoS，针对用户业务的不同内容，提供不同的资费标准。比如，需要区别是 VoIP 会话还是一次网页浏览或者是一条即时消息等不同业务进行计费。

采用 IMS 的第三个原因是运营商可以对不同的业务进行整合。IMS 定义了为业务开发商使用的标准接口，运营商能够在多厂商环境下提供业务，避免绑定在单一厂商来获取新业务。

如果采用 IMS 作为业务控制层系统，那么，从整个通信系统的网络架构来看，EPS 网络其实是位于承载层，提供通信所需的信息的传递。在实现 VoIP 业务的过程中，除了由 EPS 系统提供承载，由 IMS 系统提供业务控制外，还要由

策略和计费控制 (PCC: Policy and Charging Control) 架构实现用户业务 QoS 控制以及计费策略的控制。EPS 系统提供 VoIP 业务的系统架构如图 1 所示

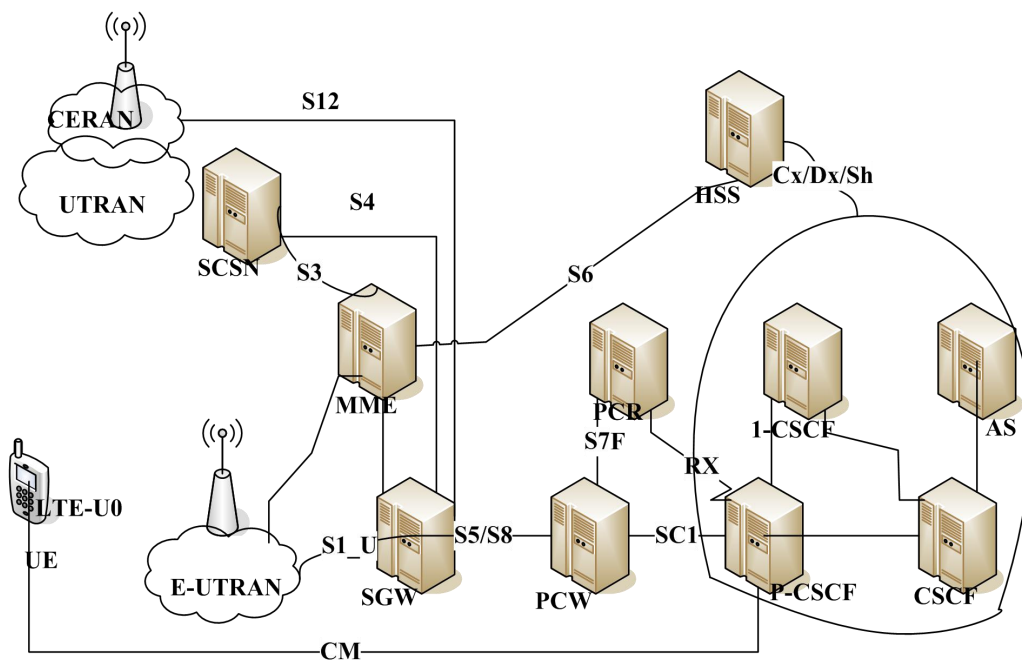


图 1 VoLTE 系统架构示意

2.2 基本流程

从业务实现流程来看，一个初次签约到 EPS 系统的用户，如果要实现端到端的 VoIP 业务，要经过 EPS 附着、IMS 注册、业务发起和会话控制过程（包括专有承载和 IMS 层信令交互）、资源释放过程等几个阶段。

一般来说，用户开机就会附着到 EPS 网络中，归属用户服务器（HSS：Home Subscriber Server）里有了用户的位置信息，MME 里也有了用户的签约信息，同时系统还会为用户建立默认承载。但是，用户在附着完成之后，未必马上就打电话，而只是简单地附着在网络上。如果用户长时间没有执行业务，根据运营商的策略，是有可能将空口和 S1-U 接口的默认承载资源释放掉的，而仅保留服务网关（S-GW：Serving GateWay）与分组数据网网管（P-GW：Packet Data Network-GateWay）之间的默认承载。

如果用户同时也是 IMS 用户，签约了 IMS 业务，那么，在完成承载层的附着之后，通常就会进行 IMS 层的注册。此时，IMS 层的注册消息就要在附着过程中建立好的默认承载上传递了。VoIP 的用户要打电话的时候，IMS 层最初的 INVITE 消息同样属于呼叫控制信令，也是在默认承载上传递的。在话音建立的

过程中，一般会在主、被叫两段的资源协商好的情况下，才建立满足话音业务 QoS 需求的专用承载资源。

下面以默认承载来传递 IMS 信令为例，描述新开机用户执行一个端到端的 VoIP 业务的过程。

(1) 附着过程

在附着过程中，用户设备（UE: User Equipment）的相关位置信息、能力等将会登记到相关的网络实体中。附着过程中还为 UE 建立“永远在线”的连接，同时为 UE 分配 IP 地址。

附着过程完成之后，系统中用户的状态就会从“未注册”变为“已注册”，核心网中记录了用户的位置信息，核心网相关节点中建立了用户的上下文，为用户提供业务做好准备。附着过程中执行的主要操作如图 2 所示，

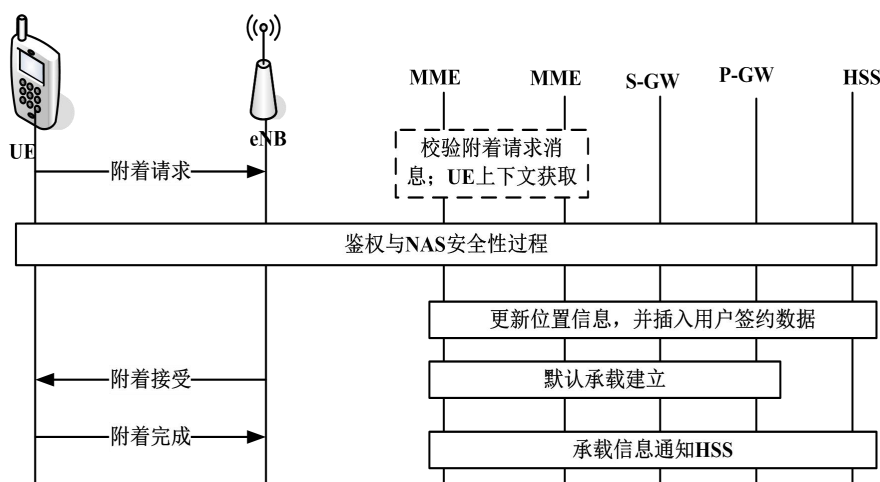


图 2 附着过程

包括：用户发起附着过程，发送非接入层消息“附着请求”给 MME。MME 根据消息中所带的用户标识信息，可判断是否要求用户先前注册的 MME 对附着请求消息进行完整性校验，并在成功后传递安全相关参数。如果在原先的 MME 中没有找到用户的上下文，MME 还可要求用户提供其国际移动用户识别码（IMSI: International Mobile Subscriber Identity）。MME 可根据运营商配置发起安全过程，其中包括鉴权和非接入层（NAS: Non-Access Stratum）安全过程的建立。之后，MME 要求删除可能存在于其他节点的、该用户未正确释放的资源，并与 HSS 交互更新位置信息。以上过程完成后，MME 将发起默认承载建立的过程。默认承载建立的过程中或过程后可为用户分配 IP 地址。默认承载建

立成功后，网络向用户返回“附着接受”消息，该消息触发空中接口为用户配置相应的承载。用户接着返回“附着完成”消息，指示完成建立的 EPS 承载以及 NAS 相关信息。

在用户第一次附着到网络的过程中，网络记录用户的位置信息，相关节点为用户建立上下文；附着到网络后，即建立了从 S-GW 到 P-GW 之间的默认连接，为用户提供默认承载；附着过程为用户分配了 IP 地址、用户驻留的跟踪区标识（TAI: Tracking Area Identity）列表、临时标识（GUTI: Globally Unique Temporary Identity）等会话必要的参数。

（2）IMS 注册

EPS 用户要使用 IMS VoIP 业务，除了要完成上述 EPC 网络的附着过程，通过 EPC 网络的安全鉴权认证以外，在建立好默认承载之后，还要在 IMS 业务层中进行认证、鉴权和授权，以确认该用户具有使用 IMS 业务的权限。一旦用户被认证授权，就会使用 IPSec 安全关联来保护自己的会话发起协议（SIP: Session Initiation Protocol）消息。用户的认证、授权以及建立 IPSec 安全关联都是通过 IMS 层的注册过程来实现的。IMS 层注册过程中的关键操作如图 3 所示。

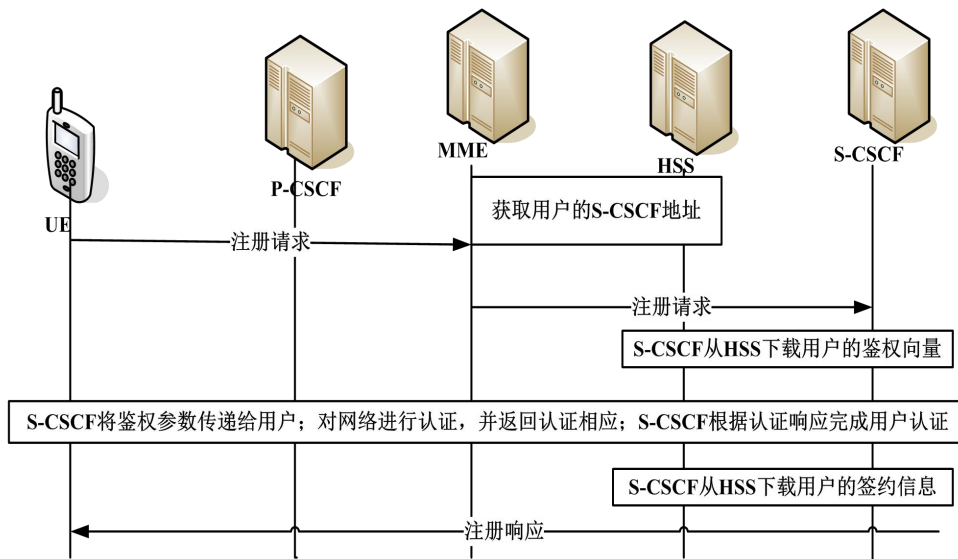


图 3 EPS 用户的 IMS 层注册过程图

在 IMS 层的注册过程中，HSS 会把用户签约信息传递给服务呼叫会话控制功能（S-CSCF: Serving Call Session Control Function），由 S-CSCF 根据用户签约的业务信息，完成对用户的业务控制。在注册过程中，用户和网络还会完成彼

此间的双向鉴权过程。一方面，S-CSCF 根据从 HSS 下载的认证向量对用户进行认证和授权；另一方面，用户也会对网络进行认证，以确定自己不是在和非法网络进行通信。双向鉴权机制提高了 IMS 网络的安全性。在安全方面，用户和代理呼叫会话控制功能（P-CSCF: Proxy CallSession Control Function）之间建立安全关联也是注册过程的一部分。

此外，由于 IMS 系统是由归属网络来进行业务控制，并为用户提供业务的，而用户接入 IMS 系统的第 1 个网元是 P-CSCF，因此，在用户作被叫的时候，用户归属的 S-CSCF 需要知道用户接入到哪个 P-CSCF 了，P-CSCF 的地址就是在注册过程中通过网元间的信息交互来传递给归属 S-CSCF 的。

IMS 注册包括 2 段注册过程：第 1 段注册过程，网络将鉴权参数传递给用户，同时把 P-CSCF 地址传递给 S-CSCF。用户在收到鉴权参数后，实现对网络的认证。在第 2 段注册过程中，用户会将计算得出的鉴权响应返回给 S-CSCF，由 S-CSCF 根据该鉴权响应，完成网络对用户的鉴权，该过程中还会将 HSS 中的用户信息下载到 S-CSCF 上。

（3）业务发起和会话控制过程

用户发起业务的时候，有可能是处于连接状态，也有可能是处于空闲状态。这是因为用户在附着以后，其 EPS 连接管理（ECM: EPS Connection Management）状态模型虽然会从空闲状态变成连接状态，但是，很多用户并不是在开机附着以后，就会马上执行业务，而基本上处于待机状态。在这种情况下，如果还为用户保持着空口资源，就会对系统资源造成极大的浪费。因此，在 EPS 中，也定义了跟 2G/3G 系统类似的机制，在用户附着完成以后，仅保留 S-GW 和 P-GW 之间的连接，而释放空口资源和 S1 接口的资源，只在用户有业务需求的时候，才会进行空口资源和 S1 接口资源的重建。

用户的状态不同，其信令交互过程也有所不同。如果用户是在连接状态发起业务的，那么就能直接在默认承载上传递 IMS 的 SIP INVITE 消息，开始 IMS 会话建立的交互过程。在这个过程中，会伴随有用于用户语音业务的专有承载过程的建立，专有承载的服务质量等级标识（QCI: QoS Class Identifier）可以为 1 或者 3。用户在连接状态下建立 VoIP 过程的 IMS 信令与 EPC 信令交互如图 4 所示。

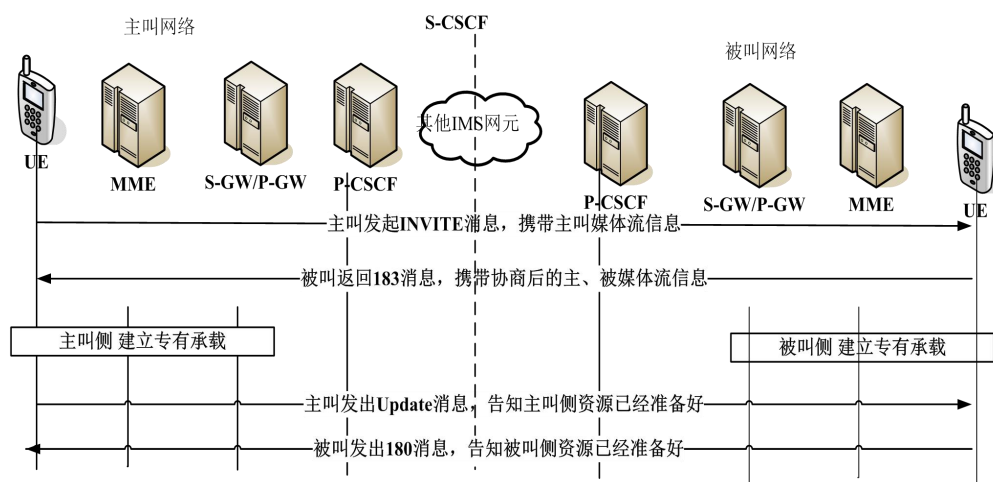


图 4 VoIP 业务 IMS 信令与 EPS 信令交互过程（用户在连接状态）

需要注意的是，图 4 中只包括了 IMS 层的请求消息，确认消息或响应消息限于篇幅原因，没有在图中体现。另外，该示意图是以 IMS 系统启动了资源预留机制为例的。从图中可以看出，在被叫发出 183 消息之后，就说明主、被叫两侧已经协商好资源了，因此 EPS 系统可以为主、被叫用户建立专有承载了。准确的时机应是主叫用户在收到 183 的 200 OK 之后建立专有承载的，而被叫用户在发出 183 的 200 OK 之后建立的。

如果系统不启动资源预留机制，那么就没有图中的 183 和 Update 过程，而是在 INVITE 消息之后以及 180 消息之前，建立好主、被叫侧的 EPS 层专有承载。

上面所讲的是用户在连接状态下发起的业务过程，如果用户是在空闲状态下发起业务的，那么在传递 IMS SIP INVITE 消息之前，必须要恢复空口资源和 S1 接口资源，这要通过 EPS 的业务请求过程来实现的。在这个过程中，除了重建无线资源和 S1 资源外，还需要将重新建立的 S1 资源信息告诉核心网节点，以保证用户数据的正确传递。空闲状态的用户，在完成无线资源和 S1 接口资源重建之后，才会继续图 4 中的消息流程。一旦完成上述 IMS 信令交互以及 EPS 层的专有承载建立，用户的语音业务就能够在所建好的专有承载上进行传递了。

（4）资源释放过程

VoIP 用户在结束通话以后，系统需要释放相应的资源。由于资源建立涉及到 IMS 层和 EPS 层，因此资源释放也包括业务层和承载层的资源释放。

业务层的资源释放过程如图 5 所示。用户结束通话后，会产生 BYE 消息发送给 P-CSCF，随后 BYE 消息会在会话中涉及到的 IMS 层的节点之间依次传递，业务层所占用的资源也就一一释放掉了。

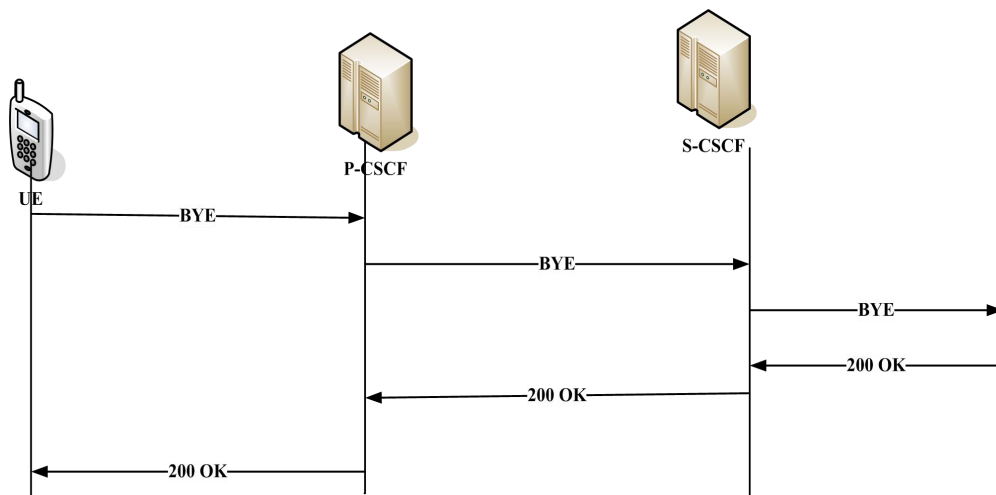


图 5 VoIP 业务 IMS 层资源释放过程

承载层的资源释放，首先是由 P-GW 发起的。因为 P-GW 是关联业务层和承载层的节点，它通过业务层知道用户此次与会话相关的专有承载资源需要释放。随后，MME/S-GW/P-GW 以及无线侧设备中的相应的专用资源都需要进行释放。如同前述建立默认承载时的原则一样，如果在结束 VoIP 业务之后，用户在一段时间内没有进一步的业务动作，空口与 S1 接口的资源也会释放，但是用户的 S5 接口的默认承载还是会继续保留的（图 6）。

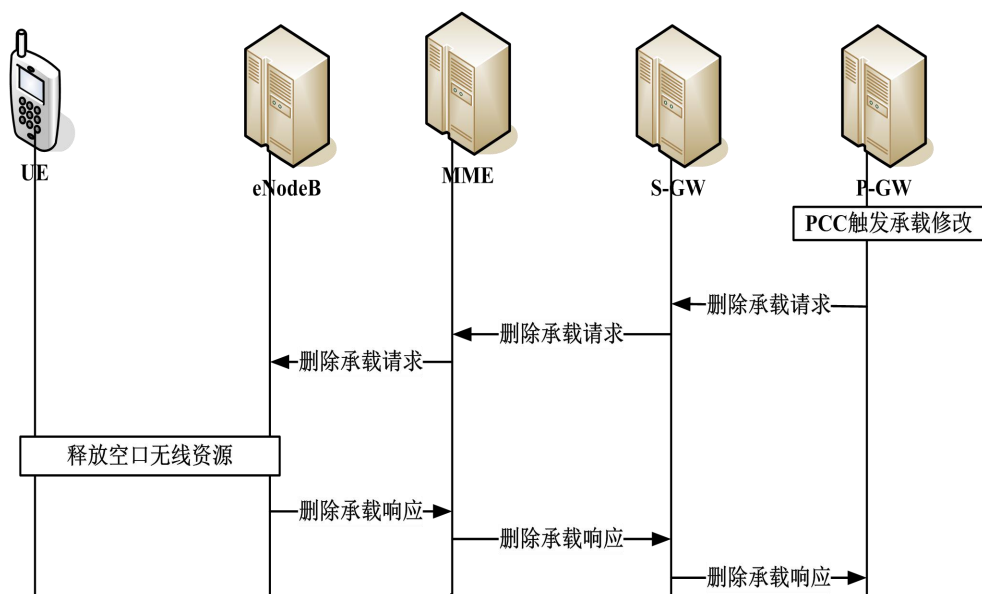


图 6 VoIP 业务 EPS 专有资源释放过程

2.3 设备功能要求

3GPP 国际标准没有对 VoLTE 业务的实现给出详细的流程和具体的要求，运营商为了简化业务实现流程，保证不同厂家设备间互通以及不同运营商间的业务互通，在全球移动通信协会（GSMA: Global System for Mobile Communications assembly）制定了一系列与 VoLTE 业务相关的标准文件，如：GSMA IR.92（IMS 要求）、GSMA IR.94（视频呼叫）、GSMA IR.88（LTE 漫游）等，其中对 VoLTE 设备（包括：无线设备、EPC 核心网设备、IMS 设备和 PCC 设备）要求做出了明确的定义。

（1）无线设备侧设备要求

头压缩（ROHC: Robust Head Compression）：这是因为 VoIP 是由实时传送协议（RTP）/ 用户数据包协议（UDP）/IP 来承载的，话音的数据包很小，但是 RTP、UDP 和 IP 的包头控制信息很长。因此，UE 和网络需要使用头压缩以提高传输效率。

TTI 绑定：用以扩展 VoLTE 的无线覆盖范围以弥补因为 LTE 时间非连续传输造成的链路预算损失。

半静态调度（SPS: Semi Persistent Scheduling）：话音数据包的产生比较规则。因此可以用 SPS 来减少调度的开销。

无线承载：UE 需要支持 SRB1 + SRB2 + 4 x AM DRB + 1 x UM DRB 的无线承载组合。网络需要支持 SRB1 + SRB2 + 2 x AM DRB + 1 x UM DRB 的无线承载组合。需要支持 QCI=1（用于话音包传输）的 GBR 承载, QCI=5（用于 SIP 信令传输）和 QCI=8/9 的 non-GBR 承载。

非连续接收（DRX: Discontinuous Reception）：为减少 UE 的耗电，基站应支持 LTE 非连续接收。分组域切换（PSHO: PS Handover）：LTE 网内的分组域的切换。如果网络也支持 VoHSPA（Voice over HSPA），最好支持 HSPA（High Speed Packet Access）和 LTE 间的分组域切换。

（2）核心网设备功能要求

由于非保证比特率（Non-GBR: Guaranteed Bit Rate）承载不能提供无线链路上可保证的比特速率，所以不适用于 IMS 语音业务。此外，虽然跟 LTE 的

峰值速率相比，语音业务只是一个低速率的数据，但是它却要求为语音业务建立 GBR 承载的业务。

在 UE 创建用于 IMS 语音的分组数据网络(PDN: Packet Data Network) 的连接时，需要建立一个缺省承载来承载 IMS SIP 信令，该缺省承载的 QCI 值为 5。

网络必须发起专有承载创建用于传递 VoLTE 的语音媒体，用于 VoLTE 的专用承载的标准 QCI 为 1。此外，专有承载的建立必须要由 PCC 来参与创建。

对于 VoLTE 业务，为防止路由迂回，接续时延过长，必须采用本地疏导方式，即用户的 VoLTE 业务由用户所在地的 P-GW 进行接续，并通过拜访地 P-CSCF 访问 IMS 系统。

为了支持国际漫游，VoLTE 业务通过全球统一的 IMS APN 接入网络，同时 VoLTE 采用 IMS TAS (Telephony Application Server) 作为业务服务器。

UE 要支持 SIP 资源预留，并且被叫侧 UE 的资源预留实现不依赖于主叫侧的 UE。

UE 和 IMS 核心网中的用户面网元必须支持自适应多码率 (AMR: Adaptive multi-rate) 语音编码，包括所有的 8 种模式，并且能够支持 8 种编码模式的所有子集。

2.4 互通技术

根据以上介绍，从 LTE 用户的角度看，在 LTE 网络发展的不同时期，要实现语音业务可能会采用 CSFB 技术，也可能会采用 SRVCC 技术，或者采用目标网 VoLTE 技术；同时，从网络部署角度看，还会出现各个运营商网络发展的不一致性。这两个因素都导致 LTE 用户的语音业务互通会出现多种互通场景，比如：VoLTE 用户间互通、VoLTE 与 CSFB/SRVCC 用户间互通等，不同的互通场景其涉及到的互通模式也是不同的。

(1) VoLTE 与 VoLTE 业务互通如果主、被叫用户均采用 VoLTE 业务，通过以上分析可以看出，EPS 网络对于 VoLTE 业务来说相当于接入网，PCC 系统则提供资源分配的策略控制，都不参与业务互通。真正实现运营商间的 VoLTE 业务互通的是 IMS 系统，其中，信令面的互通网元是互联边界控制功能 (IBCF: Interconnection Border Control Function)，互通协议采用 SIP 协议；媒体面的互

通网元是过渡网关（TrGW：Transition Gateway），互通协议采用 RTP 协议。因此，也可以说，VoLTE 业务的互通要求其实就是 IMS 网间的互通要求。

（2）VoLTE 与 SRVCC 业务互通

如果一侧用户网络采用 VoLTE 业务，而另一侧用户的网络采用 SRVCC 技术，那么，SRVCC 用户语音发生切换前，两侧均为 VoLTE 业务，只是 SRVCC 侧的网络架构中多了访问转移控制功能（ATCF）/ 访问转移网管（ATGW）功能实体，但这两个功能实体不是网间互通实体。因此，这种场景下的网络互通模式和互通要求与上述 VoLTE 业务间的互通场景一致。

切换后，SRVCC 侧用户的接入从 LTE 切换到 2G/3G 系统，UE 通过 BSC/RNC 接入到 MSCServer/MGW，信令面，随后接入到 ATCF，媒体面接入到 ATGW，依然会通过统一的业务集中和连续性（SCC：Service Centralization and Continuity）应用服务器（AS）控制呼叫的接续。因此，可以说用户切换后只有接入网络部分发生了改变，业务控制和互通模式、互通要求也未变化。

（3）VoLTE 与 CSFB 业务互通

如果一侧用户采用 VoLTE 业务，而另一侧用户采用 CSFB 业务，那么 CSFB 的语音业务将是在 2G/3G 系统完成的，因此两者的业务互通将是 IMS 网络与 2G/3G 电路域间网元的互通。其中，信令面的互通网元是 IMS 侧的媒体网关控制功能（MGCF：Media Gateway Control Function）和 2G/3G 侧的网关移动交换中心（GMSC：Gateway Mobile Switching Center）服务器（Server），互通协议可为 ISUP/BICC/SIP-I 协议；媒体面的互通网元是 IMS 侧的 IMS-MGW 和 2G/3G 侧的 MGW，互通协议是 RTP 协议。

第三章 LTE 语音业务性能改进的研究

在对 LTE 及 VoLTE 规范进行研究后，我们发现 LTE 网络主要在无线侧上做了变革，对于无线资源合理利用的规范尚不完全，也确定了我们研究的重点在于无线侧 MAC 层对于无线资源的管理，所以本章针对 LTE 的 MAC 层以及其上无线资源分配方案进行深入研究。

3.1 LTE MAC 层概述

MAC 子层在 LTE 协议栈中的位置如图所示，主要负责完成信道映射、无线资源调度 Scheduling、混合冗余重传 HARQ 过程等工作。

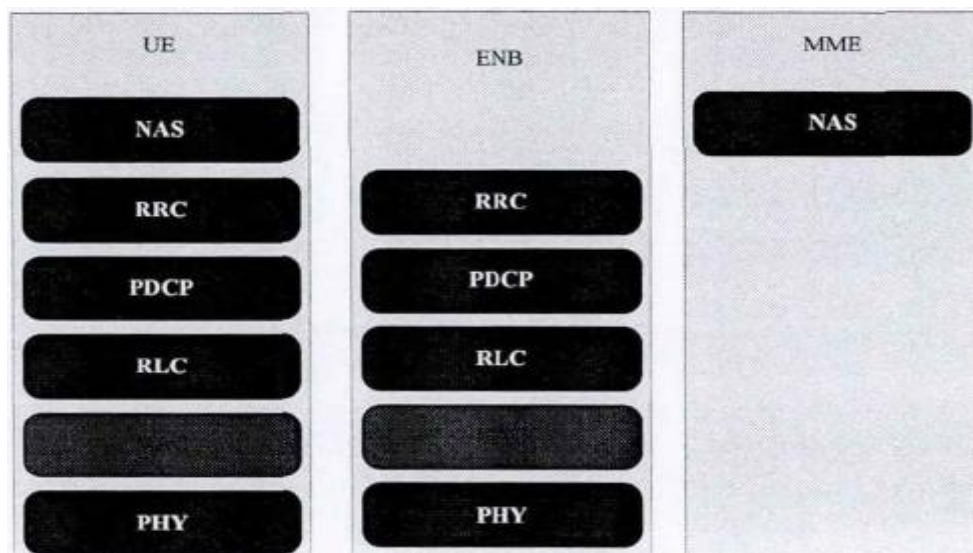


图 3-1 M;1C 层位置示意图

在 UE 和 eNB 中的 MAC 层功能不完全相同，信道映射、复用解复用、HARQ 纠错都是二者共有的功能，无线资源管理则是由 eNB 负责，UE 只需要了解信道状况如路损、信噪比等，并将之通过报告通知 eNB。具体的功能如下表所示^[5]：

表 3-1 MAC 层功能

MAC function	UE	eNB	Downlink	Uplink
逻辑信道和传输信道之间的映射	×		×	×
		×	×	×
复用	×			×
		×	×	
解复用	×		×	
		×		×
通过HARQ的纠错	×		×	×
		×	×	×
传输格式选择		×	×	×
UE间的优先级管理及无线资源分配		×	×	×
同一个UE不同逻辑信道间的优先级管理		×	×	×
逻辑信道优先级排序	×			×
UE信息报告	×			×

通过对 MAC 层功能的了解，我们发现对于无线资源的利用有两点需要深入研究，一是 LTE 的 MAC 层信道映射方式，是否仍和 3G 时代的空中接口相同，二是 LTE 的基站 eNB 是如何对无线资源进行管理和分配的，这其中用到了什么样的策略，方案是否能够完善。

3.2 MAC 信道特点

LTE 沿用了 UMTS (Universal Mobile Telecommunications System, 通用移动通信系统)里面的信道划分：逻辑信道，传输信道，物理信道。从无线接口的协议栈来看，物理信道是物理层所包含的，传输信道是物理层和 MAC 层之间抽象出的信道，逻辑信道则是位于 MAC 层和 RLC 层之间的抽象信道概念^[4]：

(1)逻辑信道主要负责具体传输什么内容，是信道中最高层的，也是最接近我们普通逻辑思维分类的，例如 BCCH C Broadcast Control Channel，广播信道)，那么就是用来收发广播控制消息的，对应的下层传输信道会负责如何传输这个消息。

(2)传输信道主要的任务是说明怎样传，例如 DL-SCH (Downlink Shared Channel，下行共享信道)，业务数据或控制消息使用此信道的时候，说明它们通过共享的下行空中资源来传输。传输信道都会有对应的物理信道，传输信道会指定空间复用等方式，告诉物理层如何去传消息。

(3)物理信道是最底层的实际传输信道，是信号在空中传输的承载，比如 PBCH (Physical Broadcast Channel, 物理广播信道)，也就是在实际的物理位置上采用特定的编码调制方式来真正传输广播消息。

跟 MAC 层相关的信道是逻辑信道和传输信道。MAC 实体负责上行逻辑信道到上行传输信道的映射，如下图 3-2, 3-3 所示。相对于 UMTS 的传输信道，LTE 最大的变化在于取消了信令专用通道和数据专用通道，比如针对业务数据，不再设置专用传输信道与专用控制信道，而使用了共享信道这一概念。上下行主要的无线资源都被设置为共享的信道 DL-SCH ， UL-SCH，除了特定的一小部分资源用于系统广播消息和上下行信道控制信息外，其他的资源对所有用户来说都是共享的，进行统一调度^[5]。如果对 UMTS 与 LTE 的传输信道进行一下对比，就会发现 LTE 的传输信道要少，功能都并入共享信道，这样的传输信道设置，和 WiMAX 对资源管理的方式接近，这样做的目的是整合无线资源，减少信令和数

据信道的空闲概率，使无线资源达到更高的利用率。

由于无线资源都变为共享，那么 MAC 层在调度方面的工作就变得更加重要起来，既要做到考虑业务优先级，又要从公平性的角度保证大部分用户的利益；既要保证信令的及时传输，又要考虑数据的及时调度。所以如何是无线资源高效使用对 MAC 的无线资源调度设计提出了比较高的要求。不同设备商的基站性能跟其在 MAC 层选择采用何种调度息息相关。

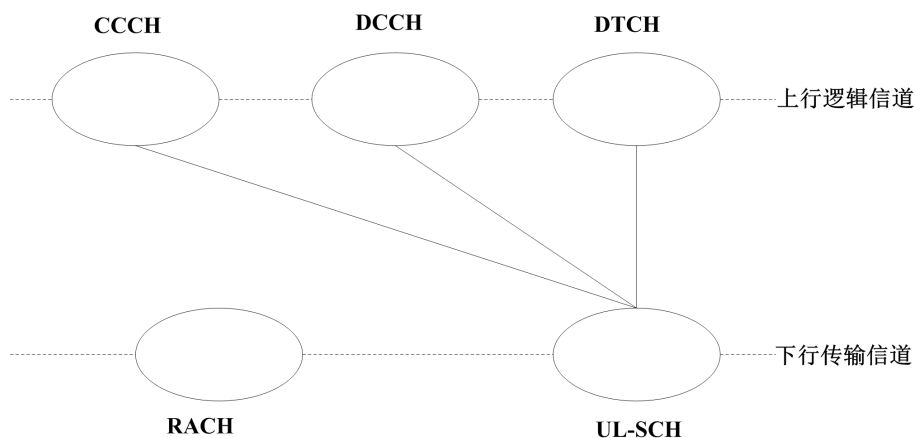


图 3-2 上行逻辑信道映射图

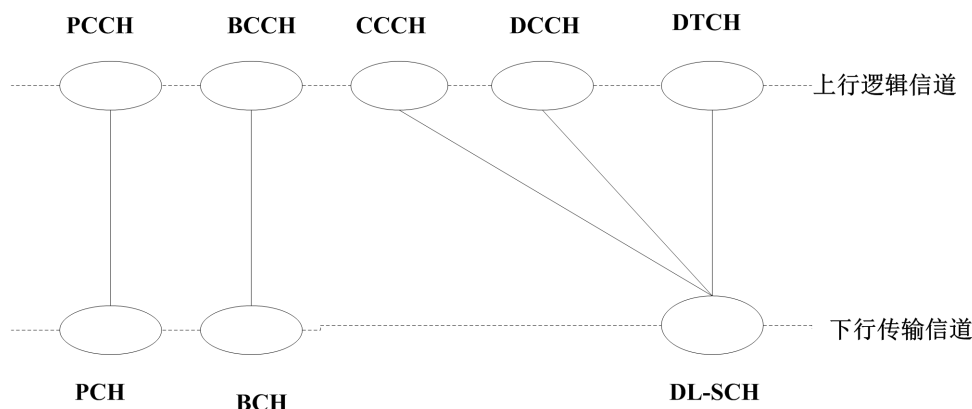


图 3-3 下行逻辑信道映射图

3.3 MAC 层调度技术研究

3.3.1 调度概述

在前面提到 LTE 系统的共享信道，无线资源是以共享的方式存在，那么对无线资源块进行分配的调度方案的设计的要求自然就有更高的要求。

调度一般表示的是对事物先后次序的安排，这里针对 LTE 所提到的无线资源调度，是由于 LTE 无线资源被数据和信令共享的特性，需要将无线资源划分为较小的单位并合理分配给所有正在竞争资源的用户。在 3GPP 中，资源调度的定义为：基站调度器动态地控制时/频资源的分配，完成在一定时间内对某一用户特定资源的分配。对 LTE 系统而言，时间和频率是系统调度的两类主要资源^[6]。

调度方案的好坏对于系统的性能影响很大，对于 LTE 同时能够支持多少用户并发十分重要。最好的分配时频资源给所有的用户终端，除了满足业务的服务质量外，还必须保证系统的容量，否则一个系统只能支持少数用户的高质量服务也没有意义，要想设计良好的调度方法，首先需要了解一下可以被 LTE 调度的资源是什么。

3.3.2 资源调度的基本单位

LTE 的资源基本单位有两种划分方式 FDD (Frequency Division Duplexing, 频分双工) 与 TDD (Time Division Duplexing, 时分双工)。在中国主要的 LTE 相关研究都是基于 TDD 方式的，因为虽然 TDD 系统的覆盖范围相对半径要小，但 TDD 技术有一个明显的优势就是灵活的带宽配比，使用 TDD 技术可以灵活的设置规范中各种不同的上行和下行配比以及转换时刻，对于实现上下行不对称的互联网业务有着明显的优势。而且因为其上下行由时间来区别，TDD 还可以使用

零碎的频段，不必要求带宽对称的频段，对于资源利用更充分，我们的研究也是基于 TDD 方式进行的。

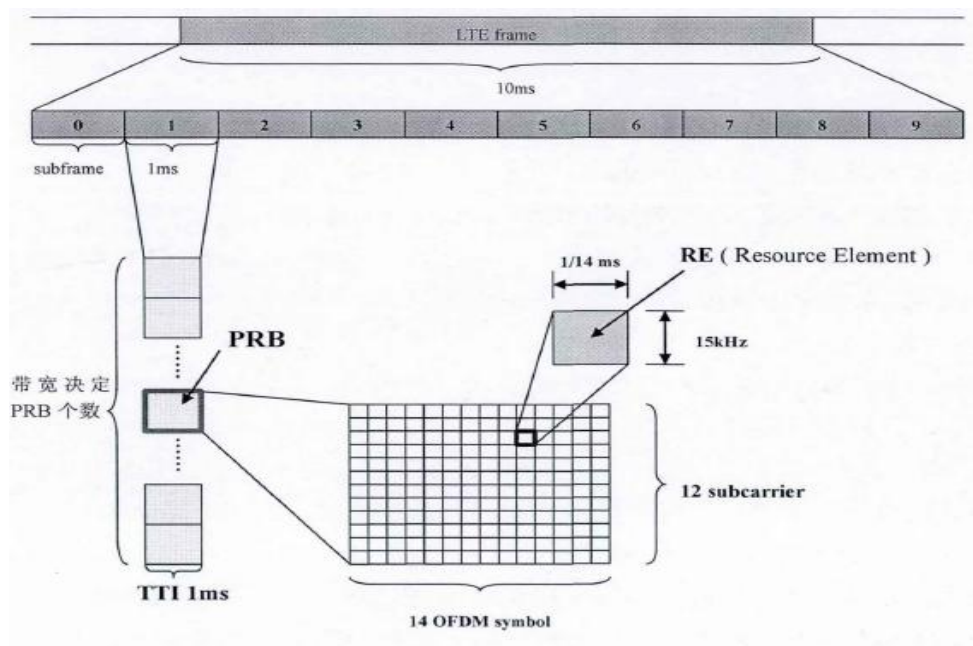


图 3-4 TDD 模式无线资源示意图

如上图所示，调度的基本单位称作 PRB (Physical Resource Block，物理资源块)，也叫做时频块，如图中蓝色方块所示。顾名思义由时域和频域两部分构成，无线资源的最小单位称作一个 RE (Resource Element)，如图中红色方块所示，在时域上包括一个 OFDM 符号(1/14ms)，在频域上是 1 SkHz 的子载波^[7]。

一个 PRB 在频域上是 1ms，包括 14 个 OFDM 符号，在频域上包括 12 个子载波，所以频域上为 180kHz，一个 PRB 总共包括 14*12 个 RE^[7]。

那么无线资源中到底有多少可用 PRB 呢?首先从时域上看，1 个 TDD 帧为 10ms ，1 个半帧 5ms，这个时间划分和 3G 是相同的。1 个帧包括 10 个子帧，1 个子帧又再细分为 2 个 slots，但是在调度的角度上，时域都是以一个子帧为单位进行的，调度的时域单位由此定义为 TTI (Transmission Time Interval，传输时间间隔)等于 1ms。

而频域上，实际的带宽决定了资源块的个数，LTE 最大带宽为 20MHz，系统可配置为 20MHz, 15MHz, 10MHz, 5MHz, 3MHz, 1.4MHz。VoIP 作为重要业务之一，应该被分配足够的带宽，由于调研发现 LTE 调度研究的大部分文献^{[9][11][12]}都采用了 5MHz 带宽来研究 VoIP 的用户容量，所以为了结果的可参考性，同时

这个分配也确实比较合理，本论文也使用 5MHz 带宽来进行分析。一个 PRB 的带宽是 $12 \times 15\text{KHz} = 0.18\text{MHz}$ ，加上保护带宽为 0.2MHz，所以 5MHz 带宽在每个子帧也就是时域 TTI 为 1ms 上的 PRB 个数为 $5\text{MHz}/0.2\text{MHz} = 25$ 个。

LTE 理论最大速率 100 Mb/s 也是根据 PRB 计算出来的，一个 OFDM 中最多可以包含 6 个 bits。PRB 中的 OFDM 个数为 $14 \times 12 = 168$ 个，1ms 上 20M 带宽包括 100 个 PRB， $6 \times 168 \times 100 \times 1000 = 100800000\text{Bits/s} = 100\text{Mb/s}$ 。

3.3.3 现有调度算法研究

MAC 层通过一个包含了调度算法的调度器对申请 PRB 的分配的请求进行一一调度。下行传输数据时效率比较高，对无线资源的调度并不是瓶颈问题，因为 eNB 是数据发送方，而调度器也处在 eNB 的 MAC 层中，因此很容易了解到各个等待无线资源的用户都缓存了多少数据，让每个 PRB 都可以尽可能的装满数据，吞吐量比较高，PRB 利用也很充分。

而在上行信道上，调度器需要通过用户端发送来的信令才能获取用户缓存、信道质量等信息，由于信令的传输也需要被调度器分配无线资源，所以 LTE 则只能采用基于用户的调度。每个 UE 需要通过信令来告诉 eNB 自己缓冲的数据、信道质量信息等 eNB 用来做判断的参量，eNB 基站调度器通过这些参量来确定为每个 UE 分配哪一个 PRB 上行资源。这种方式中信令会占据很大一部分无线资源，而且每个 PRB 中很难保证填满数据。加之上行速率 50Mb/s 本身就低于下行速率 100Mb/s，所以无线资源调度的瓶颈在于上行资源的调度，所以目前在 LTE 系统的相关研究中，主要针对的都是上行调度。

当前的主流调度算法主要有两类：分组调度和半持续调度。一前者属于一般的调度算法，面向所有的数据业务都可通用；后者则是针对 VoIP 语音业务等实时业务而设计的方案^[6]。这类实时业务的数据包有一个特点，包的尺寸小且发送规律间隔，发送频繁，所以需要减小相应的控制信令的开销来增加共享信道上数据包可利用的资源。

3.3.3.1 动态调度算法

动态调度是 LTE 规定的最基本的调度方式，适用于所有业务，我们研究的语音业务当然也可以使用动态调度进行资源分配，下图是上行动态调度示意图^[5]：

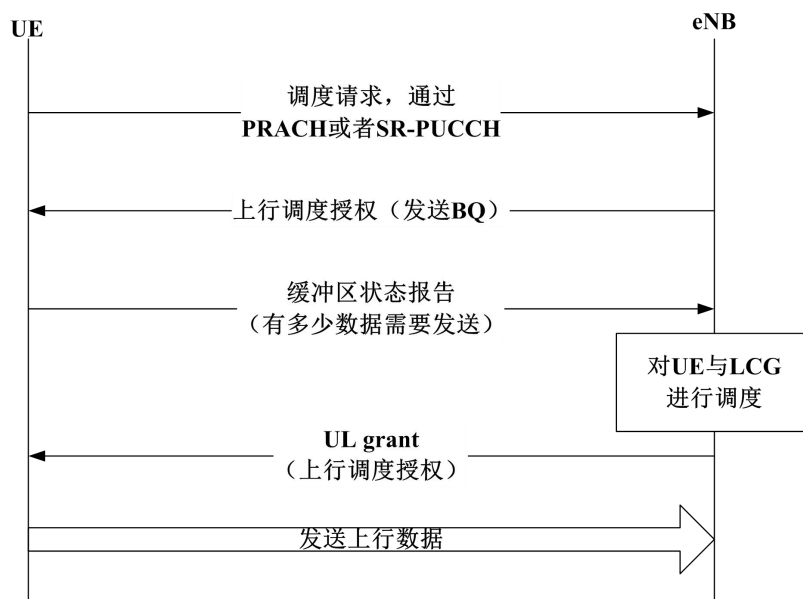


图 3-5 上行动态调度示意图

上图说明简单描述了一个上行动态调度的示意图，

(1)首先在 UE 端产生上行数据发送事件，将数据放入缓冲区后需要为这些数据申请上行资源用于发送。

(2) eNB 如有空闲资源就会通过上行调度授权(UL grant)告诉 UE 可以发送缓存报告。

(3) UE 发送报告通知 eNB 有多少数据要发送。

(4)然后 eNB 收到用户的请求，根据资源情况，按照预先设定的调度算法原则来给用户分配适当的 PRB 资源，然后通过上行调度授权 UL grant 通知 UE。

调度过程是比较清晰明确的，其中需要考虑的就是第(4)部分中调度算法如何实现，最开始出现的调度算法包括最大载干比算法、轮询算法、比例公平算法^[8]，此后又有学者提出了改进的最大权重时延优先算法^[8]、正比公平调度算法的改进算法等增强型算法，下面分别介绍并分析其优缺点。

(1)最大载干比调度算法

调度算法 Max C/I (Maximum Carrier to Interference)又称为最大载干比算法，其中 C 表示有用载频功率，I 表示干扰信号功率。这个算法的主要实现思想是让信道质量好的用户优先传输，以增加吞吐量减少重传所浪费的资源：所有等待调度的用户，依据 eNB 基站端接收到的用户载干比预测值进行由高到低的排序，

载干比高的优先级高，载干比低的优先级低，并按此顺序进行调度^[6]。算法数学表达式如下：

$$j=\max(C/I_i) \quad \text{式(3-1)}$$

如前所述， C/I_i 是载干比的比值，表征信道质量好坏。 J 是用户的编号，基站通过对比所有等待调度用户的载干比，选取其中载干比最大的用户分配 PRBo

由于这种载干比算法保证的是信道质量最好的用户优先使用资源，而信道条件不好的用户很有可能会始终无法得到发送机会而一直处于饥饿的等待状态，而且这种状况很可能出现在很大一部分用户身上，导致用户满意度不可能被接受。所以由于这种算法没有对公平性做任何考虑，Max C/I 算法存在的意义只是在理论研究时充当对比参考量或者作为系统吞吐量的极限值。

(2)轮询调度算法

RR (Round Robin, 轮询)调度算法是一种以公平性为主的算法，主要的调度实现方式是让所有等待中的用户全部按照到达的顺序依次享受 PRB 资源。在这种规则的调度方式下，所有被调度的用户都拥有同样的优先级，能够很好的体现和维护公平性，特别是体现在每个用户分得的资源所需等待的时间长度是公平的。

在各种动态调度算法的对比中，轮询调度算法是一种最公平的算法，但是过分的把重心倾斜在公平性上也限制了轮询调度算法对其它可参考情况的利用，即使信道质量很差的用户也会被公平的分配资源，造成一些可以通过用户优先级放弃掉的多次重传占用了一些无线资源，使得一些信道质量好，一次传输即可成功的数据包没有资源可用，资源的利用率也因此并不是十分合理。

(3)比例公平调度算法

在前面两种算法的基础上，PF (Proportional Fairness, 比例公平)调度算法被提出，这个算法的主要实现思想是牺牲在信道质量和公平性上取一个平衡，兼顾了最大化系统吞吐量的需求和用户间无饥饿的公平性需求，在轮询调度的基础上牺牲部分的公平性，换取更高的资源利用率以提高性能，这个算法也因此成为是无线移动网络中被普遍推荐使用的一种调度算法。

eNB 在每次进行调度时，都需要计算每个用户的优先级，在有些地方称作例因子，这个比例因子的计算公式为：

信道状况 $((\log(1 + \text{sinr})) / \text{平均吞吐量})$ 式(3-2)

其中的 sinr 是 Signal to Interference plus Noise Ratio(信号与干扰加噪声比)，它在 LTE 中用来指示信道状况的好坏。每次调度 eNB 的调度器都会选择优先级最高的用户，也就是比例因子最大的用户。所以被调度的用户是相对来说信道质量比较好的用户。比例公平算法的思想类似于最大载干比算法，但是与之不同的是，比例公平算法还将平均吞吐量作为反比纳入考虑，如果一个用户一直在占用资源，那么它的吞吐量就会上涨，比例因子会随之下降，在资源竞争中会逐渐输给其他没有分配到资源的用户，最终会使所有用户的比例因子趋于和信道质量有关的一种平均。

由于比例公平算法融合了之前两种算法的优点，于是成为了应用的最为广泛的动态调度算法之一。该算法也有一些缺点，就是没有考虑区分不同的 QoS，把所有业务和用户一视同仁，这在多种业务共同支持时可能会造成实时业务的不能及时传送。

(4)改进的最大权重时延优先算法

基于上面的三种优先级方案，还有一个被提及的较多的动态调度方案，改进的最大时延权重优先算法 M-LWDF (Modified Largest Weighted Delay First)是一种针对实时业务提出的新方案，其主要思想是丰富优先级的计算参量，不仅包括了用户信道质量，还和时延、可等待时间长度、吞吐量等参数有关系。如何将分组包的时延、信道信息、QoS 等进行平衡的考虑是这个算法考虑的问题，M-LWDF 算法优先级定义如下：

$$P_i[n] = \log(\delta_i) \times \frac{R_i[n]}{\lambda_i[n]} \times \frac{D_i[n]}{T_i} \quad \text{式(3-3)}$$

其中 δ_i 指示的是 QoS 参数值，主要是对不同业务类型进行区分对待； R_i 为用户当前的数据传输速率，信道质量越好，传输速率也就越高，由于作为优先级计算的分子，于是这类用户的优先级也相对较高； λ_i 是用户的平均吞吐量，这是对于用户服务的公平性的考量，对信道质量好一直占用 PRB 资源传数据的用户通过吞吐量的反比来一定程度的限制其优先级，使信道条件差的用户的优先级获得一定程度提高； D_i 为用户的 HOL (Head of Line，头阻塞)部分分组延时，它作为优先级计算分子的意义是如果分组在用户侧缓冲队列中等待的时间越长，优先级就越高，这同样也是为了公平性所设的参量； T_i 为用户可以接受的最长

等待时延限制,如果待调度用等待的时间超过阈值,则此分组会被判定调度失败。可等待时间长的用户,调度失败的可能性相对较小,可以稍微降低其优先级^[9]最终计算出一个优先级 P1,在每个 TTI 中调度器都选取优先级最高的用户来分配资源。

其余关于动态调度的改进算法还包括正比公平调度算法的改进方案、M-LWDF 算法优化优化等,都一定程度上增强了动态调度的性能,但是对于 VoIP 等实时业务来说,动态调度算法的改进对于业务的优化还是比较有限的,而下一节介绍的半持续调度算法大大提高了 VoIP 业务的用户容量,是一种被广泛认可的算法。

3.3.3.2 半持续调度算法

由第二章我们知道 VoLTE 系统通过 VoIP 承载语音业务。要在 LTE 系统中有效地支持 VoIP 业务,保证其实时性,就必须解决一个重要的问题:LTE 系统中存在控制信令开销的限制。因为 VoIP 的语音业务数据包比较小,如果如上一节中所述采用动态的分组调度,大量的小数据包会带来大量的信令,信令开销会占据一大部分共享信道资源。如果因为控制信令过多而导致数据包反而无法被调度器分配资源而延迟发送,就会进而导致数据包的超时丢弃,用户语音业务无法保障。

VoIP 业务的特点是数据包的编码发送间隔固定,每个数据包包含的数据量较小。既然是周期性需要的,那么如果采用事先一次配置长期占用就可以减少信令对于资源的浪费,节省出空间传输更多的数据。因此,针对 VoIP 业务的这些小数据包设计好对应的资源预留周期,就可以采用一种持续性的调度算法,不需信令地来连续一个时间段分配同样的资源给用户,达到提高节省资源提高系统容量的目的。

针对 VoIP 的这种特点,半持续调度方案应运而生,半持续调度 (SPS}semi-persisting scheduling)是一种针对 VoIP 等实时业务的减小调度过程中信令开销的调度算法。半持续调度的中心思想是用户只需要在新的会话开始建立连接时向 eNB 申请使用调度资源,此后以 20ms 为周期一直占用同一位置的 PRB 时频块,直到连续语音结束,此后对于静音数据包采用动态调度方案,周而复始直到会话结束。该调度方案的具体定义如下:

a)对于初次传输的数据分组，采用持续调度，以 20ms 为周期持续占用资源直到 UE 端检测到激活期结束，静音期开始。

b)对于重传的分组，采用动态调度方案。

c)对于静默描述符 SID (Silence Desriptor)数据包，也采用动态调度方案。静默描述符的作用是在用户的静默期按一定的间隔(一般为 160ms)传传递背景噪声信息给 eNB 的接收端，直到用户再次激活。

这三者的传输优先级相比较来说：初次传输的数据分组 new transmissions of voice packets>重传分组 retransmissions>静默分组 SID packets。

半持续调度的示意图如下，每种颜色代表了不同的用户所被分配的资源，时频块根据用户数据分组的大小，以 20ms 为周期固定分配给用户。重传的分组和静音分组使用白色的空闲部分进行动态调度。

当有新用户进入激活态时，eNB 首先根据该用户的数据量来计算出用户需要使用的资源块数量，然后从当前子帧的 TTI 开始寻找，直到某个子帧上的空闲 PRB 资源足够提供该用户所需要的 PRB 个数为止，将这几个连续资源块以 20ms 为周期分配给此用户。然后 eNB 通过信令通知用户设备分配资源的位置和数量，以及半持续调度的周期 20ms。

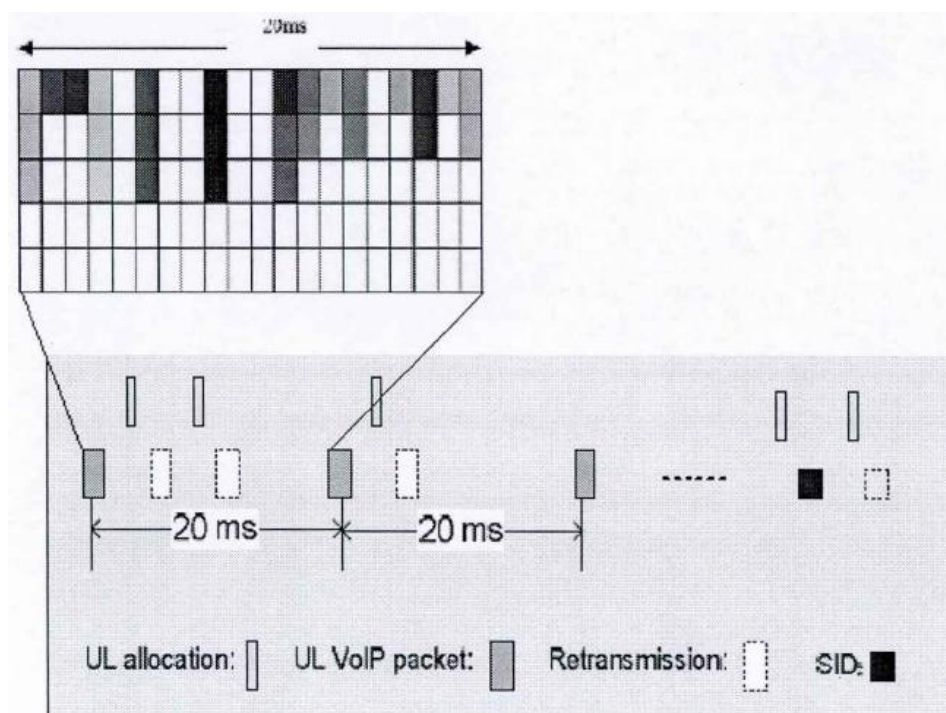


图 3-6 半持续调度算法示意图

3.3.3.3 组调度算法

组调度(Group Scheduling)是由摩托罗拉在 2006 年提出来的调度技术，基本思想是把用户划分成几个用户组，eNB 针对组来进行调度以减少资源竞争的冲突。如下图所示，每个 UE 会根据 group 分配状况和自身条件被划分到一个组中，它只能使用这个组内的资源。在每一个 group 的区间到达时，eNB 会统一处理与此组相关的 UE 资源申请并向 UE 发送资源分配指示，这样减少了扎堆的用户同时争抢有限 TTI 资源的情况，每个用户的所有数据、信令都被局限在一部分 PRB 中，减少了空闲 PRB 的资源浪费。但这样做会造成有时某一个 group 有空闲资源，但是其他 group 因为权限的关系无法使用这部分资源。论文中对这种调度方式进行了仿真，效果是略差于半持续调度的。

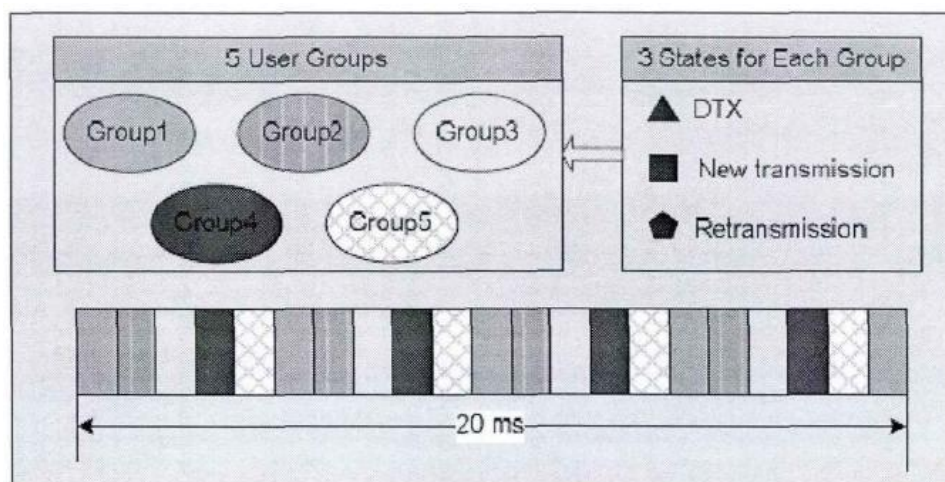


图 3-7 组调度算法示意图

3.3.3.4 现有调度方案对比

总的来说，动态调度的优点是：适用于全部的数据业务，资源分配灵活，调度算法可研究的范围比较广；缺点是调度信令占据了太多的资源导致无线资源利用率不高。半持续调度的优点是：通过周期性持续分配 PRB 资源，减少了信令的损耗，使 PRB 资源更充分的用于数据传输；缺点是：只针对实时语音业务，对于发包间隔不短且没有规律的其他业务不适用。组调度的优点是：减少了抢占资源时的冲突，用户管理比较清晰；缺点是：在实时语音业务上性能不及半持续调度，一些空闲组的资源不能被其他比较忙的组利用。

VoIP 最直观的系统性能衡量标准是 VoIP 的容量，我们就通过这个用户容量来对比几种算法的性能。VoIP 的容量定义为一个小区内 95% 的用户为满意用户

时的用户数，而满意用户是指一个用户在 60s 内不能被正确调度的数据包个数占此用户发送数据包的比例小于 2%^[11]。有研究者对这三种解决方案的 VoIP 容量进行了仿真和对比^{[11][12]}，综合起来结果如下图所示：半持续调度的容量最大，组调度低于半持续调度，高于动态调度。所以综合来看，目前最好的调度方案依旧是半持续调度。

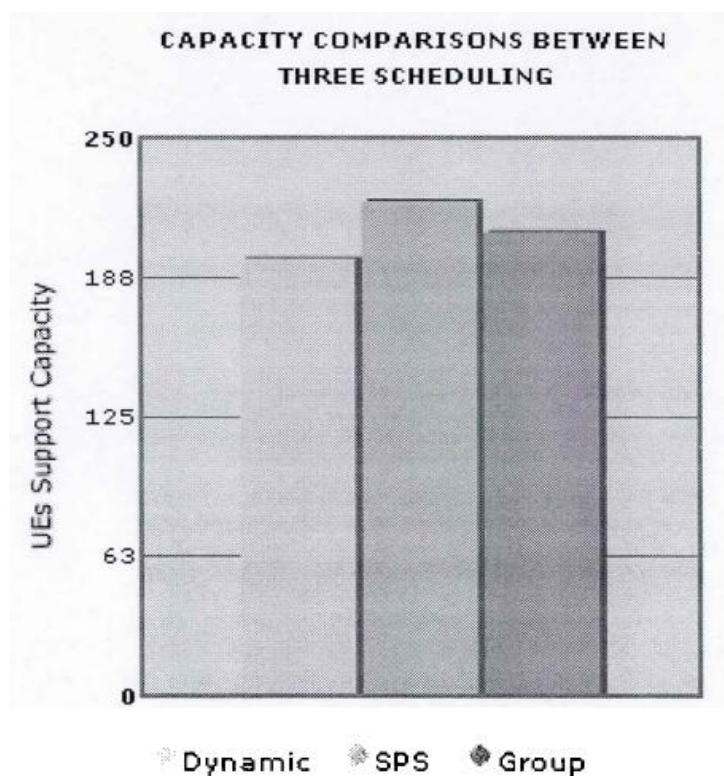


图 3-8 三种调度算法用户容量对比示意图

但在半持续调度中，用户的请求都是先到先服务的，从动态调度的方案中我们可以看到根据信道质量等参数确定优先级后的调度方案，比轮询的调度方案性能更好，因此，如果将动态调度的优先级机制引入到半持续调度中，可能能够进一步提高 VoLTE 的用户容量。

3.4 本章小结

本章首先研究了 LTE 采用共享信道模式的特点，了解无线资源分配形式是在 MAC 层进行调度。随后现有调度算法进行了总结，分析了各方案的优缺点，并对比了各方案的用户容量差异，最终确定了对用户容量最大的半持续调度方案通过引入优先级进行改进的方向。

第四章 VOLTE 调度机制及性能的优化策略

在进行 VOLTE 调度机制与调度方案的分析之后，基于动态调度算法、半持续调度算法、组调度算法等进行对比，最终发现半持续调度方案对用户的容量最大，通过引入优先级能够实现性能的优化改进。

4.1 VOLTE 调度机制及性能的优化方案设计

4.1.1 PrSPS 模型设计

文中提出的基于优先级的半持续调度(Priority Semi—Persisting Scheduling, PrSPS)方案的设计主旨是，针对半持续调度中无线资源利用率有待提高的问题，在其基础上引入了用户优先级评定机制，并根据优先级分配 PRB 资源。PrSPS 的目标是减少信令开销，提高系统的用户容量，降低用户因始终无法获得资源而造成饿死的可能性，并在尽可能提高资源利用率的情况下保证公平性。PrSPS 模型如图 4.1 所示。图 4.1 中，PRB 是物理资源的抽象，MAC 层通过分配 PRB 来实现对物理层的时频资源进行调度。调度器是 eNB MAC 层的调度核心，通过不同的调度算法对所有连接到此 eNB 的 UE 进行统一调度，包括两个功能模块：(1) 优先级计算模块。该模块接收来自 UE 的 PRB 请求，根据 UE 发来的实时用户信息和 UE 信息存储模块存储的历史用户信息，通过适当的优先级算法计算出这个 UE 的优先级，并将该 PRB 请求根据优先级插入待调度优先级队列中适当的位置。(2) PRB 分配模块。该模块存储物理层时频资源 PRB 的分配信息，并根据优先级队列和 PRB 请求类型依次对无线资源进行半持续调度或者动态调度。用户信息存储模块保存用于计算 UE 优先级所需的各种参数和属性值，如用户的服务等级、用户的平均传输速率、用户的平均吞吐量、用户的实时信道质量等。优先级的半持续调度方案的设计主要思想就是基于半持续调度方式，评定用户的优先级，然后根据优先级实施半持续调度的 PRB 分配。优先级的半持续调度的最终目标是将用户容量提升、减少信令的小号以及用户由于资源缺乏出现饿死的可能性，尽量实现资源利用的公平性。该模型的工作机制如下：(1) 调度器接收 UE 的调度请求，其中包含有用户设备的惟一标识无线网络临时标识(Radio Network Temporary Identity, RNTI)以及用户的信道质量标识(Channel Quality Indicator, CQI)。

(2) 优先级计算模块根据用户信息存储模块保存的信息和用户的 PRB 请求中的 CQI 信息计算用户优先级，并根据优先级由高到低将请求插入 PRB 分配模块中的等待调度队列。(3) PRB 分配模块根据空闲资源情况依次调度等待调度队列的队头，为其分配 PRB 资源。(4) 在用户信息存储模块中，记录或更新当前被调度用户的最新服务等级、平均传输速率、实时信道质量等信息。

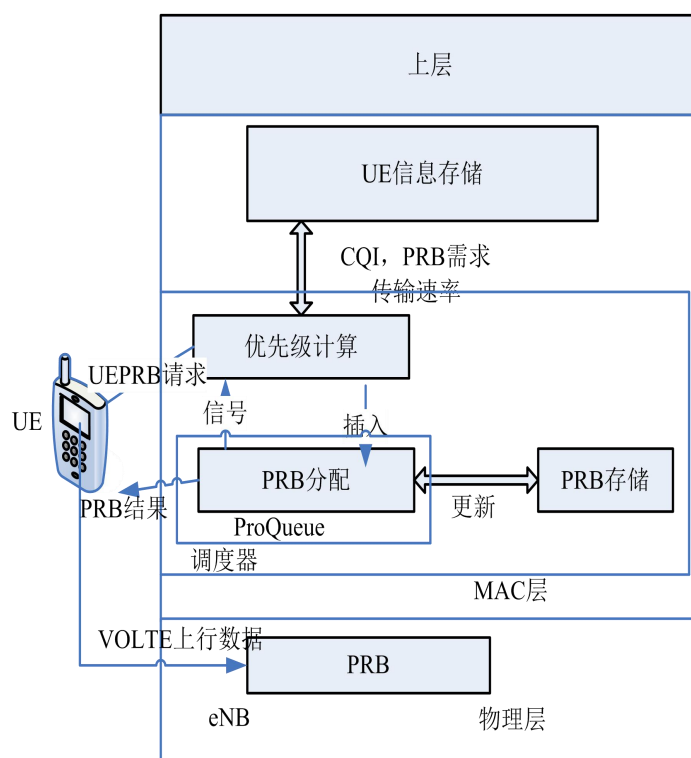


图 4.1 PrSPS 模型

4.1.2 用户优先级计算方法

上述调度过程的核心是用户优先级的计算。计算用户优先级需要考虑两方面：(1) 提高用户的吞吐量和传输效率，减少用户数据的重传和碎片及空闲资源的浪费。(2) 需要兼顾用户间的公平性，减少不满意用户的数量。基于这个思想，在 PrSPS 方案中提出如下的优先级计算方法：

$$P(i) = Pri(i) \lg \frac{\overline{PrbN(i)CQI(i)} + 6}{\lg R_{ave}(i)} \quad \text{式(4-1)}$$

式(4-1)包括两部分，第 1 部分是预设优先级 Pri，描述了用户的重要程度，越重要的用户优先级越高。第 2 部分为修正优先级，由几个参数共同组成，用于

根据用户信道质量 CQI、平均传输速率 R_{ave} 、用户 PRB 资源需求数 PrbN 等实时调整用户的优先级，这个值和预设优先级应该具有对等的效力，共同决定用户的优先级。

4.3 基于优先级的半持续调度方案性能的分析

4.3.1 仿真场景

利用 NS-3 软件进行仿真，基于 CTTC 组开发的 LTE 扩展模块，可以实现 PrSPS 的仿真和对比分析。仿真场景是由一个 LTE 接入基站 eNB 以及若干均匀分布于基站周围 5KM 范围内的用户终端组成，如图 4.2 所示，选择 5KM 是因为 3GPP TR25.913[11]规范中规定了 LTE 在 5KM 范围内保证用户的服务性能指标。所有用户均处于通话状态，即处于讲话和收听状态的平均概率是 50%。

仿真场景参数配置为：调度算法为 SPS 和 PrSPS，传播频谱模型为 Friis Spectrum Propagation Loss Model，上行带宽为 5MHz，下行带宽为 5MHz，上行发射功率为 10dB，下行发射功率为 30 dB，上行环境噪音为 9 dB，下行环境为 5dB。

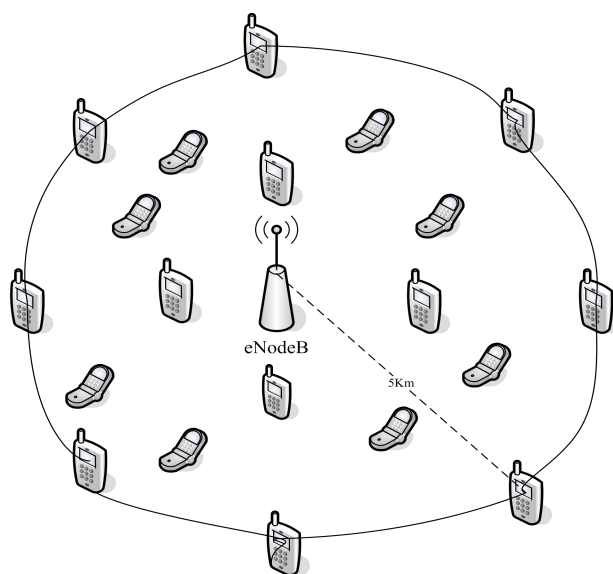


图 4.2VOLTE 仿真现场示意图

4.3.2 仿真结果及分析

用户容量定义为一个小区内满意用户数为 95%时的用户总量。其中满意用户的定义为：在 60s 的时间内调度失败的数据包个数占此用户发送数据包的比例小于 2%。数据包调度失败的原因有两种可能：（1）因为信道质量较差，经过两

次重传都未达到能被成功调度。（2）在调度队列中等待时间过长而导致调度超时。

4.3.2.1 最大用户量对比

两种调度算法的用户满意度对比如图 4.3 所示。图 4.3 中横轴为基站 eNB 接入的用户总数，纵轴为用户的满意度。左侧曲线是半持续调度算法的满意度，右侧曲线是基于优先级的半持续调度的满意度，水平线为用户容量。可以看出，SPS 的用户容量为 266，PrSPS 的用户容量为 276，PrSPS 的用户容量增加了 10 个左右，约提升了 4%。

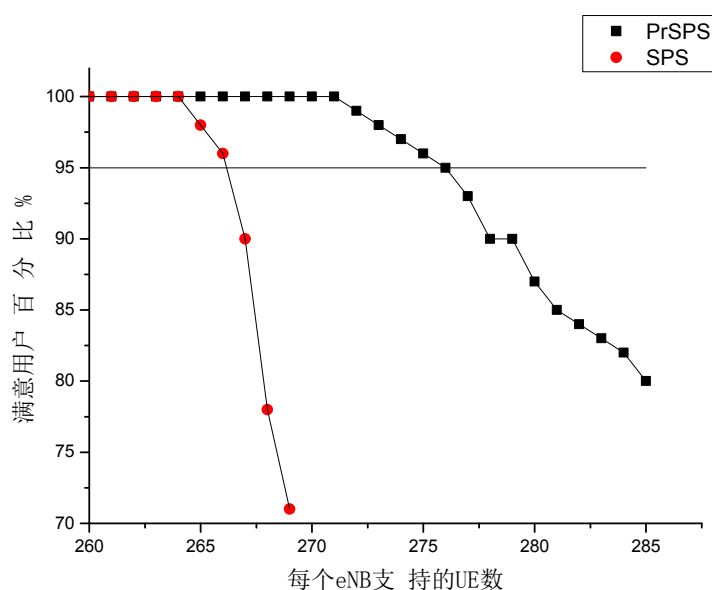


图 4.3 两种调度算法的用户满意度对比图

4.3.2.2 无线资源使用率对比

通过仿真结果可统计出两种算法的 PRB 使用率，如图 4.4 所示，SPS 的使用率如左侧柱状所示为 92.08%，PrSPS 的使用率如右侧柱状所示为 97.96%，PRB 的使用率提升了 6%左右。

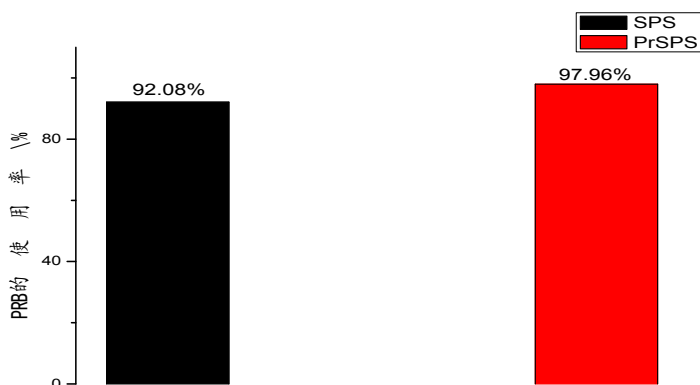


图 4.4 两种调度算法的 PRB 使用率对比图

4.3.3.3 调度失败数分布对比

从图 4.3 可以看到, PrSPS 的用户满意度随终端总数增加而下降的速率大大降低, 这对基站超负荷运行时尽可能保证用户服务质量有重要意义。产生这一现象的原因可以从用户调度失败率得到解释。满意度下降的原因是每增加一个用户终端就意味着它会在一段时间半持续地占用 PRB 资源, 进而造成其他用户等候时间增加, 每个用户调度失败的次数都可能增多。由于 60 S 的仿真中每个用户调度的数据包个数约为 1 500, 调度失败率数量级比较小, 统计调度失败的次数如图 4.5 所示。

从图 4.5(a)可以看出, 使用 SPS 方案时, 每个用户调度失败次数分布比较集中, 266 个用户中的大部分用户失败次数集中在 15~20 次, 也就是所有用户的调度失败率都在 2% 左右, 这样一旦增加了新用户, 很容易出现大量用户调度失败率同时超过 2% 的情况, 进而导致用户满意度迅速下降。

PrSPS 方案中, 由于信道质量好的用户优先使用了资源, 减少了它们的数据因为调度超时导致的调度失败, 使得调度失败次数的统计出现了两级分化。如图 4.5(b)所示, 276 个用户中的大部分用户调度失败次数集中在 10 次以下。这样的方案同样也牺牲了部分的公平性, 相对于半持续调度的分布, 有少量用户的调度失败数会大于 25, 这样在增加用户终端后, 约 50% 的用户不会因此成为不满意用户, 只有调度失败次数大于等于 16 的这 22% 的用户有可能成为不满意用户, 而这类用户在 SPS 方案中的比例接近 80%。所以 PrSPS 方案大大降低了用户满意度随终端数增加而下降的速度。

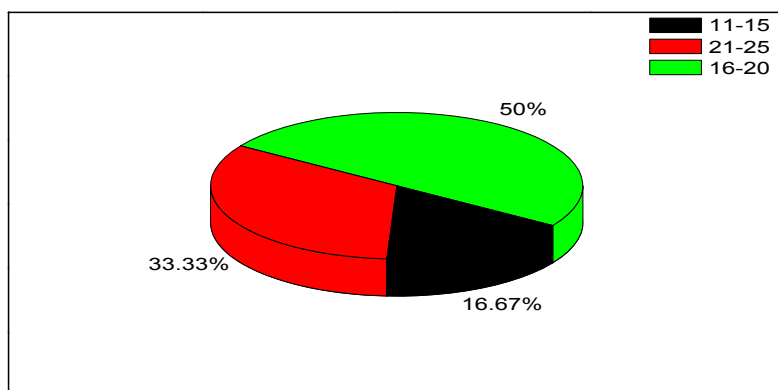


图 4.5 (a) 使用 SPS 时 UE 调度失败次数分布

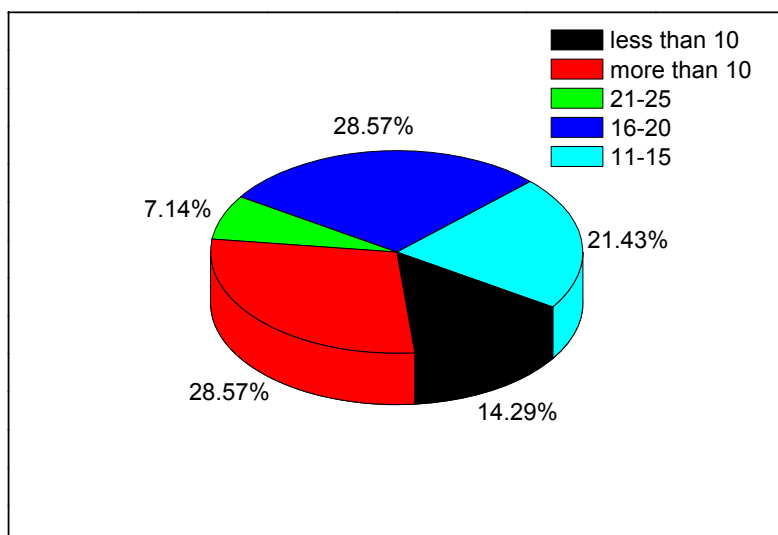


图 4.6 (b) 使用 PrSPS 时 uE 调度失败次数分布

4.4 基于优先级的半持续调度方案性能的分析

VOIP 系统性能的衡量标准也就是 VOIP 的容量, 将其定义为一个小区当中 95% 的用户为满意用户, 则需要满足一个用户在 60s 的时间中期发送的数据包中的难以正确调度的数据包不能超过总量的 2% 以上。LTE 的调度特点为基于调度分配 PRB 的个数, 而决定用户量的主要条件就是资源的使用率, 每一位用户的 PRB 需

求、重传概率以及公平性因素都会影响到使用率，所以需要对基于优先级的半持续调度方案用户容量进行估算。

首先是资源总量，应当将可用资源块的个数确定下来，在 3GPP 规范中基于 TD-LTE 帧的上下行配置进行了相应的规定，默认为用户的进行语音通话的过程中，上行与下行手法数据的吞吐量接近，每一位用户既发送也接受，基于这样的条件计算出上行资源能够支持的容量。此外本文中选择较大的宽带容量分析，以 5MHz 作为参数研究用户的容量，提高结果的可参考性。一个 PRB 宽带为 $12 \times 15\text{KHz} = 0.18\text{MHz}$ ，加上保护宽带等得到在时域 TTI 为 1ms 的 PRB 个数为 25 个，进而计算出 VOIP 可用的资源块总数量为 179.2 个。

其次是资源的使用量，基于上文的计算继续探 N 个用户需要的 PRB，然后估算出用户的容量，基于 PRB 资源的总量与用户对 PRB 资源总量的使用量之间的关系，得到 PRB(Used) 应当小于 PRB(Total)，得到用户容量：

$$N_{UE} \leq \frac{PRB_{total}}{0.95(1 + R_{ave_re})[(PRB_{need} + (1 - R_{active})PRB_{need} \cdot 0.125)]}$$

R(active) 为活跃用户的比例，在但持续调度方案中取 50%，也就是有一般的用户处于活跃的状态下，voip 静音帧发送周期为 160ms，因而能够进一步得到每个 PRB 的平均承载能够为 840 比特；R(ave_re) 为用户平均重传比例，取 20%。在优先级半持续调度算法中得到调整的优化，具有更多的资源被有效的传输，丢包的比率得到很大的下降。同时优先级半持续调度相对于普通调度算法合理优化用户的使用次序，进而等待的时间更少，重传的比例减少。基于公式就能够针对不一样的编码方式得出优先级半持续调度方案基于普通方案于用户容量上的提高效果。

结 论

随着 4G 牌照发放时间点的临近，LTE 技术商用步伐也逐渐加快，国内运营商也加大了给 LTE 用户提供语音业务的研究和验证的力度。其中，CSFB 技术作为一种过渡技术，能够充分利用现有 2G/3G 网络资源，同时也是不少国际运营商选择使用的方案，预计未来商用的可能性比较大；SRVCC 技术相对比较复杂，而且，仍有不少涉及到与 2G/3G 网络间互通的技术尚未有明确方案，国外已商用的 LTE 网络都未采用 SRVCC 技术，国内部分运营商由于网络规模和需求的不同性，有可能会采用 SRVCC 技术；VoLTE 技术作为 LTE 语音的目标解决方案，最终将替代 2G/3G 传统语音业务，被各个运营商广泛采纳和部署。

本文首先重点对 LTE 网络下的目标语音解决方案--VoLTE 技术做了详细的分析研究，涉及 VoLTE 技术实现的各个层面，包括 VoLTE 技术的基本架构、基本业务流程、设备功能要求和互通技术等，并对几种 LTE 语音解决方案的应用进行了比较和总结。

第二，通过对负责分配 LTE 无线资源的 MAC 进行了解，希望基于 LTE 共享无线信道资源这一特点进行研究。首先对当前的无线资源调度算法进行了调研和对比，分析了这些方案的优点和存在的不足。随后在现有效果最好的半持续调度方案基础上提出了改进方案，这个方案的主要思想是通过优先级的引入，牺牲部分公平性，从而使更多用户获得高质量服务，减少重传数据对于 PRB 资源的浪费，从而增加用户容量，最后从理论上分析了对语音业务性能的提升。

最后提出了一种基于优先级的半持续调度模型方案，设计了方案，分析其算法，并对方案的性能进行分析。

微信扫描以下二维码，免费加入【5G 俱乐部】，还赠送整套：5G 前沿、NB-IoT、4G+ (VoLTE) 资料。



参考文献

- [1] GSMA IR.92 IMS Profile for Voice and SMS[S], 2013.
- [2] 辛伟, 杨红梅. 演进分组系统 (EPS) 业务应用技术[M]. 北京: 人民邮电出版社, 2010.
- [3] 姜怡华. 3GPP 系统架构演进(SAE)原理与设计(第2版)[M]. 北京: 人民邮电出版社, 2013.
- [4] 姚斌. 3GPP LTE 无线接口协议及体系结构[J]. 移动通信, 2007, 31(12):54-57.
- [5] 3GPP, TS 36.321-2011, Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification, 2011.
- [6] 王喆. 3GPP LTE 系统中基于 QoS 保证的资源调度研究[D]. 华中科技大学, 2009.
- [7] 3GPP. TS 36.211-2011, Evolved Universal Terrestrial Radio Access (E-UTRA): Physical Channels and Modulation, 2011.
- [8] 沈嘉, 索世强, 赵训威等. 3GPP 长期演进技术原理与设计[M], 人民邮电出版社, 2008, p352-p356.

- [9] 仇娟娟, 杨丰瑞. HSDPA 中 M-LWDF 实时业务分组调度算法的改进[J]. 无线电工程, 2009, 39(4):11-13.
- [10] Kim Y. An Efficient Scheduling Scheme to Enhance the Capacity of VoIP Services in Evolved UTRA Uplink[J]. Eurasip Journal on Wireless Communications & Networking, 2008, 2(2):723-728.
- [11] Jiang D, Wang H, Malkamaki E, et al. Principle and Performance of Semi-Persistent Scheduling for VoIP in LTE System[C]. //Wireless Communications Networking & Mobile Computing .wicom .international Conference. IEEE, 2007:2861 - 2864.
- [12] Wang H, Jiang D. Performance Comparison of Control-Less Scheduling Policies for VoIP in LTE UL[C]. //IEEE Wireless Communications & Networking Conference. IEEE, 2008:2497 - 2501.
- [13] 3GPP, TS 36.213-2011 , Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures, 2011.
- [14] Fan Y, Lunden P, Kuusela M, et al. Efficient Semi-Persistent Scheduling for VoIP on EUTRA Downlink[J]. Vehicular Technology Conference IEEE, 2008:1 - 5.