

doi:10.13756/j.gtxyj.2018.03.016

无线通信技术

基于 QoS 的 5G 切片间资源分配

王文琪^{1,2}, 徐 湛¹, 田志刚²

(1. 北京信息科技大学 信息与通信工程学院, 北京 100101; 2. 清华大学 信息技术研究院, 北京 100084)

摘要:新一代的无线通信实现了万物互联和多样化的业务需求,网络切片可以将网络在逻辑上进行隔离,每个网络服务于一类相似的业务,不同切片之间的业务差异性大,所需资源也不同。为了解决切片间的资源分配问题,为业务提供更好的服务,文章提出了基于服务质量的比例公平资源分配方案,权衡了各个切片之间的公平性和用户需求。将仿真结果与比例公平和最大载干比算法进行了对比,结果表明,新算法能在保证公平性的同时获得较高的用户频谱效率。

关键词:切片;资源分配;第五代移动通信技术;服务质量

中图分类号: TN929.5 **文献标志码:** A **文章编号:** 1005-8788(2018)03-0059-05

QoS - based 5G Slice Resource Allocation

WANG Wen-qi^{1,2}, XU Zhan¹, TIAN Zhi-gang²

(1. School of Information & Communication Engineering, Beijing Information Science & Technology University, Beijing 100101, China; 2. Research Institute of Information Technology, Tsinghua University, Beijing 100084, China)

Abstract: The new generation of wireless communication has to meet the requirements in the Internet of Things with diversified business. The network slices can logically isolate the network, and each network works in a similar class of business. Therefore, the business differences between different slices is large, and the required resources are also different. In order to solve the problem of resource allocation among slices and provide better service for users, this paper proposes a fair proportion of resource allocation scheme based on Quality of Service (QoS), which weighs the fairness and user's requirements in each slice. The simulation results are compared with the proportional fairness and the maximum load-keeping ratio. The results show that the new algorithm can achieve higher user spectrum efficiency while guaranteeing the fairness.

Key words: slice; resource allocation; 5G; QoS

0 引 言

未来 5G 需要支持更丰富的通信方式,不仅包括 4G 应用的增强,还有一些新兴的应用,如虚拟现实、增强现实、可穿戴设备应用、面向垂直行业和商业领域的紧急任务应用等^[1]。为了能提供更好的用户体验,势必会出现多种多样的应用场景。国际电信联盟无线电通信局(International Telecommunication Union - Radio communications Sector, ITU-R)确定未来 5G 具有以下 3 大主要的应用场景:增强型移动宽带(Enhance Mobile Broadband, eMBB)、超高可靠与低延迟的通信(ultra Reliable & Low Latency Communication, uRLLC)和大规模机器类通信(massive Machine Type Communication, mMTC)^[2]。下一代移动网络联盟提出利用虚拟化技术将网络进行切片,每个切片在逻辑上隔离,但共用相同的物理基础设施。网络切片服务的

是一类需求相似的业务。业务之间可能也有不同,所以切片可以由一个或多个子切片组成^[3]。这不仅要求 5G 网络架构能服务多种业务,还要求其能灵活地配置业务之间的资源。

切片技术可以满足 5G 网络架构的多样化和个性化的网络服务,并且可以动态地建立和运营虚拟网络,高效管理资源^[4]。例如,eMBB 场景需要为高速移动的用户提供服务;uRLLC 场景则为远程手术、车联网等可靠性要求高的业务提供保障,第三代合作伙伴计划(3rd Generation Partnership Project, 3GPP)提出,在 uRLLC 场景下,未来的蜂窝系统中的时延要低于 1 ms 以保障高可靠性^[5];智能抄表等业务则由专为大规模物联网应用而发展的 mMTC 场景提供服务,Release-15(5G 通信系统,Release-15 NSA 5G NR 规范)确定了 mMTC 场景的 3 个主要性能指标^[6],支持 ≥ 10 s 的时延,在每平方千米能够支持连接百万设备,终端设备的电池寿命要超过 10 年。

收稿日期:2017-09-11

基金项目:国家自然科学基金资助项目(61402044,61620106001);国家“八六三”计划资助项目(2015AA01A706);北京市科技新星计划资助项目(Z161100004916086);北京市优秀人才资助计划青年拔尖项目(2016000026833ZK08);北京市属高校“青年拔尖人才培养计划”资助项目(CIT&TCD201704065);北京信息科技大学科技创新服务能力建设项目资助(71F1810921)

作者简介:王文琪(1992-),女,山东烟台人。硕士,主要研究方向为宽带无线通信和移动通信等。

通信作者:徐湛,副教授。E-mail: xuzhan@bistu.edu.cn

与 4G 网络不同,5G 的网络架构在同一个基站下可有多虚拟切片,切片间的资源分配变得很重要。文献[7]提出了快速启发式的有效分配物理资源块(Physics Resource Block,PRB)的方法,但该方法没有结合不同业务的服务质量(Quality of Service,QoS)。文献[8]提出了基于机会的频谱共享无线虚拟化资源分配方案,但该方案没有充分考虑几个切片之间的均衡问题。文献[9]提出了基于比例公平的切片间资源分配的算法,该算法注重并满足切片间资源分配的公平性,但也没有考虑不同切片的不同需求。文献[10]提出了一种可以自适应调节用户优先级的调度算法,该算法优化了吞吐量和丢包率,但是该文献没有网络切片,也没有考虑到不同的业务有不同的需求。文献[11]提出了一种最大化切片的网络效用的资源分配算法,该算法根据不同切片的差异进行区分,但是没有充分考虑用户之间的公平性。本文提出一种基于 QoS 的比例公平资源分配方案,切片时根据不同切片业务的不同要求制定不同的优化目标,并且兼顾系统吞吐量和切片间的公平性。

1 系统模型

考虑一个 5G 单小区下行链路的切片网络,为了满足用户需求,将网络进行切片。这里以网络应用类型为单元划分 3 个切片,分别是 eMBB、uRLLC 和 mMTC 切片。假设 3 个切片的物理基础设施为同一个基站,每个切片逻辑上隔离,服务不同业务,如图 1 所示。为避免不必要的损失,同一业务只能接入一个切片,用户在切换业务时才可以切换切片。假设小区中共有 N 个用户, N_e 为 eMBB 切片中的用户数, N_u 为 uRLLC 切片中的用户数, N_m 为 mMTC 切片中的用户数,且用户均匀分布在小区内。

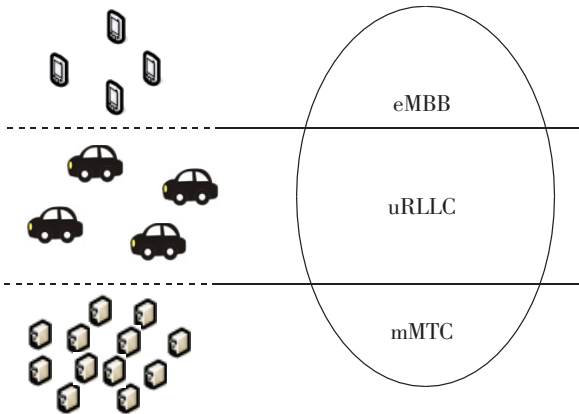


图 1 切片模型

在切片中需要使用调度算法对资源进行分配,包括颗粒度和用户优先级的计算。为了给业务提供好的 QoS 保证,这里以资源块(Resource Block, RB)为单位分配资源。最大载干比算法的基本思想是信道条件最好的用户可以得到最高的优先级^[12],使整个系统的吞吐量达到最大值。虽然最大载干比算法可以获得最优吞吐量,但是信道条件好的用户一直占用资源会造成用户之间的调度不公平。轮询(Round Robin,RR)算法的基本思想是轮流为用户提供服务^[13],不考虑信道和业务的差异,这使得系统的吞吐量很低。比例公平(Proportional Fair, PF)算法的基本思想是通过引入用户的历史吞吐量来计算优先级^[14],PF 算法可以在公平性和系统吞吐量之间有较好的权衡,因此本文选用 PF 算法。

将每个切片的所有用户看成一个整体来分配资源,假设无线资源总数为 M ,先为每个切片分配满足最低要求的资源。假设 eMBB、uRLLC 和 mMTC 切片分配到的资源总数分别为 M_1 、 M_2 和 M_3 。在每个调度时刻都要决定各个切片的资源分配,资源调度方案都要在考虑公平性和满足业务 QoS 要求的前提下将系统吞吐量最大化。

PF 算法的用户优先级 $P_k(t)$ 定义如下:

$$P_k(t) = R_k(t) / \overline{R_k(t)}, \quad (1)$$

式中, $R_k(t)$ 为 t 时刻 k 用户的瞬时速率; $\overline{R_k(t)}$ 为到 t 时刻前两个时隙 k 用户的平均传输速率,在每个传输时间间隔(Transmission Time Interval,TTI)计算一次,计算方法如下:

$$\overline{R_k(t)} = \left(1 - \frac{1}{T}\right) \overline{R_k(t-1)} + \frac{1}{T} R_k(t), \quad (2)$$

式中, T 为时间窗。将每个切片的用户作为一个整体, t 时刻 i 切片中所有用户的瞬时速率 $R_{(i,k)}(t)$ 表示为

$$R_{(i,k)}(t) = \sum_k R_k(t); \quad (3)$$

t 时刻前两个时隙 i 切片中用户的平均速率表示为 $\overline{R_{(i,k)}(t)} = \sum_k \left[\left(1 - \frac{1}{T}\right) \overline{R_k(t-1)} + \frac{1}{T} R_k(t) \right]; \quad (4)$

t 时刻 i 切片的优先级 $P_{i,k}(t)$ 表示为

$$P_{(i,k)}(t) = R_{(i,k)}(t) / \overline{R_{(i,k)}(t)}. \quad (5)$$

计算出每个切片的优先级后根据此优先级对未分配的资源进行分配。更新 eMBB、uRLLC 和 mMTC 切片分配到的资源总数为 M_1 、 M_2 、 M_3 ,如图 2 所示。

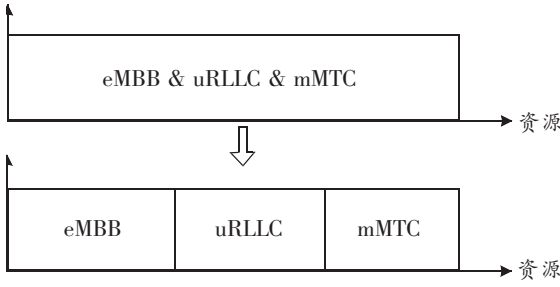


图 2 切片资源分配

2 基于 QoS 的 PF 资源分配

每个切片在逻辑上相互独立但共用物理资源，这时资源分配就包括切片间分配和切片内分配。每个切片服务的业务不同，切片间的资源分配不能只考虑公平性，也不能把吞吐量作为每个切片的衡量标准。eMBB 切片需要考虑速率要求，uRLLC 切片需要考虑时延要求，而 mMTC 切片则要考虑丢包率。资源分配根据系统能否按照需求进行资源变动而分为静态、半静态和动态，这里选择半静态，即初始分配的资源保证切片的最低开销，剩余资源可以灵活调度。切片间的资源分配选用半静态分配为基础，结合切片的服务特性，加入速率因子 α 和时延因子 β ，提出基于 QoS 的 PF 资源分配方案。

每个切片中的用户分为满意用户和不满意用户，不同的切片用不同的指标来衡量用户是否为满意用户。分别为 eMBB 切片和 uRLLC 切片设定一个 QoS 指标的阈值，用户的指标大于该阈值则为满意用户，小于该阈值则为不满意用户。

达不到界限速率的用户数 a_x 表示为

$$a_x = \{R_x < v, x \in N_e\}, \quad (6)$$

式中， v 为 eMBB 切片的速率阈值。

速率因子 α 表示为

$$\alpha = e^{\frac{a_{N_e}}{N_e}}. \quad (7)$$

式中， a_{N_e} 为 eMBB 中达不到界限的用户个数。

eMBB 切片的优先级 P_{eMBB} 表示为

$$P_{\text{eMBB}} = P_{(\text{eMBB}, N_e)} \cdot \alpha. \quad (8)$$

时延超过时延界限的用户数 b_y 表示为

$$b_y = \{t_y < \tau, y \in N_u\}, \quad (9)$$

式中， τ 为 uRLLC 切片中用户的时延阈值。

时延因子 β 表示为

$$\beta = e^{\frac{b_{N_u}}{N_u}}, \quad (10)$$

式中， b_{N_u} 为 uRLLC 中达不到界限的用户个数。

uRLLC 切片的优先级 P_{uRLLC} 表示为

$$P_{\text{uRLLC}} = P_{(\text{uRLLC}, N_u)}(t) \cdot \beta. \quad (11)$$

mMTC 切片的优先级 P_{mMTC} 表示为

$$P_{\text{mMTC}} = P_{(\text{mMTC}, N_m)}(t). \quad (12)$$

切片中满意用户越多，则 α 和 β 越低；不满意用户越多，则 α 和 β 越高。

对于切片内的资源分配，要根据切片中的业务来选择具体算法。eMBB 切片的业务传输数据量大，用户移动速度较快，选择 PF 算法；uRLLC 切片中的业务要求低时延高可靠，所以选择修正最大加权时延优先 (Modified Largest Weighted Delay First, M-LWDF) [15] 算法；mMTC 切片中的业务传输数据量小，对时延没有很高的要求，用户数量大，对公平性要求很高，因此选用 RR 算法。

3 仿真结果与分析

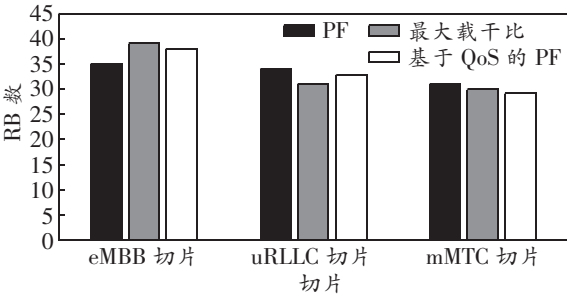
本文通过仿真实验评估所提算法的性能。利用 MATLAB 软件搭建系统级仿真平台，并与文献[9]中提出的分配方案和最大载干比资源分配方案进行比较。仿真参数如表 1 所示。

表 1 仿真参数

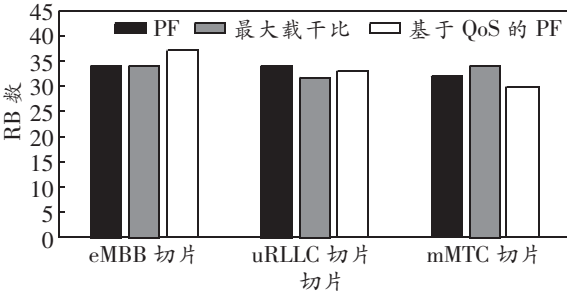
参数	数值
基站数量/个	1
基站总功率/ dBm	46
小区半径/m	500
切片数/个	3
用户数/个	240
用户分布	随机撒点
载波频率/ GHz	2
带宽/MHz	10

图 3(a)、(b) 所示分别为两个随机 TTI 中 3 种方案的切片 RB 分配比例。由图可见，每个 TTI 中 RB 分配都是不同的，PF 资源调度分配给每个切片的 RB 数差不多，只注重公平性。最大载干比算法和基于 QoS 的 PF 算法分给 eMBB 切片的 RB 最多，分配给 mMTC 切片的 RB 最少，分给 uRLLC 切片的 RB 与分配给 mMTC 切片的差不多。图 3(c) 所示为在所有 TTI 中 3 种方案 RB 分配的平均图，可见每种 RB 分配在进行平均后相差较小。3 种方案都将较多的 RB 分配给 eMBB 切片，在 uRLLC 和 mMTC 切片的分配中有一定的差异。首先是因为每个切片中的业务不同，eMBB 切片中的用户传输的数据量大，uRLLC 中的用户虽然要求实时性比较高，但数据量没有 eMBB 中的用户大，mMTC 虽然用户数量多，但是数据量很小。虽然 PF 注重公平性，但每个切片的 RB 分配还是不同的，因为 mMTC 中的业务大部分时间不需要分配资源，没有

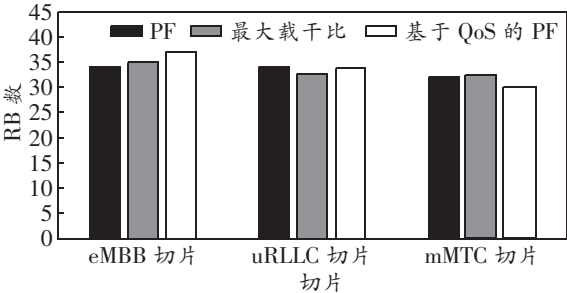
瞬时速率,却有平均速率,导致优先级低,该切片分配的资源比其他切片少。最大载干比则受用户位置影响,如果切片中的用户靠近基站的多,用户信道条件好,则优先级高,分配的 RB 多;如果切片中用户分布在小区边缘的多,信道条件差,则优先级低,分配的 RB 就少。而基于 QoS 的 PF 算法引入了速率因子和时延因子,能够很好地分配资源。



(a) 第 c 个 TTI 的 RB 分配图



(b) 第 d 个 TTI 的 RB 分配图



(c) 3 种方案 RB 分配平均图

图 3 3 种资源调度方案的比较

图 4 所示为 3 种方案中用户频谱效率的累计分布函数。从图中可以看出,PF 算法的用户频谱效率相对较差,最大载干比算法次之,基于 QoS 的 PF 算法最好。图 3 中 PF 将 1/3 的 RB 分配给 mMTC 切片,由于 mMTC 切片中业务不需要分配资源的 TTI 偏多,导致频谱利用率降低。最大载干比算法则受用户距离基站远近的影响,边缘面积大于中心面积,所以可能很多用户在边缘区域,从而影响 RB 的分配。基于 QoS 的 PF 算法分配给 eMBB 和 uRLLC 切片的 RB 多,eMBB 和 uRLLC 中的业务传输数据量大。

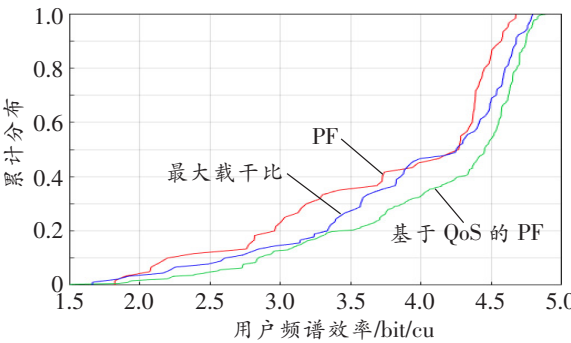


图 4 3 种方案用户频谱效率的累积分布函数

4 结束语

针对 5G 的切片架构,本文介绍了 3 种切片并且根据每种切片应用的特点提出了一种基于 QoS 的 PF 资源分配方案。根据不同切片的业务不同,在 PF 算法的基础上,eMBB 切片考虑速率并加入速率界限,uRLLC 切片考虑时延并加入时延界限。通过仿真并与 PF 和最大载干比两种资源分配方案进行对比,结果表明,本文提出的资源分配方案可以在满足切片间公平性的同时提高用户频谱效率。

参考文献:

[1] 陈亮, 杨奇. 5G 网络中无线频谱资源分配的进展分析 [J]. 光通信研究, 2016, (6): 68—71.

[2] 4G Americas'. Recommendation on 5G Requirements and Solutions [EB/OL]. (2017-09-11) [2015-10-08] http://www.5gamerica.org/files/2714/1471/2645/4G_Americas_Recommendations_on_5G_Requirements_and_Solutions_10_14_2014-FINALx.pdf.

[3] NGMN Alliance. Description of Network Slicing Concept. Version 1.0 (2016) [EB/OL]. (2017-09-11) [2016-01-13]. https://www.ngmn.org/fileadmin/user_upload/160113_Network_Slicing_v1_0.pdf.

[4] 安琪, 刘艳萍, 孙茜, 等. 基于 SDN 与 NFV 的网络切片架构 [J]. 电信科学, 2016, 32(11): 119—126.

[5] 3GPP TR 38.913-2017, Study on Scenarios and Requirements for Next Generation Access Technologies [S].

[6] 3GPP TR 38.913-2016, Technical Specification Group Radio Access Network; Study on Scenarios and Requirements for Next Generation Access Technologies [S].

[7] Kamel M I, Long B L, Girard A. LTE Wireless Network Virtualization: Dynamic Slicing via Flexible Scheduling [C]// Vehicular Technology Conference 2014. Seoul, Korea: IEEE, 2014: 1—5.

[8] Yang M, Li Y, Jin D, et al. Opportunistic Spectrum Sharing Based Resource Allocation for Wireless Virtualization [C]// International Conference on Innovative Mobile & Internet Services in Ubiquitous Computing. Taichung, Taiwan: IEEE, 2013:51—58.

[9] 栗欣, 龚金金, 曾捷. 面向 5G 网络切片无线资源分配 [J]. 电子产品世界, 2017,(4):30—32.

[10] 杨鹏, 闫俊杰, 王汝言. LTE 中队列状态感知的优化 M-LWDF 调度算法 [J]. 重庆邮电大学学报: 自然科学版, 2015, 27(4): 514—520.

[11] 唐伦, 张亚, 梁荣, 等. 基于网络切片的网络效用最大化虚拟资源分配算法 [J]. 电子与信息学报, 2017, 39(8):1812—1818.

[12] 张路. LTE 系统的下行无线分组调度研究 [D]. 北京: 北京邮电大学, 2010.

[13] 林尚静. LTE 系统跨层调度算法研究 [D]. 武汉: 武汉邮电科学研究院, 2010.

[14] Kim H, Han Y. A Proportional Fair Scheduling for Multicarrier Transmission Systems [J]. Communications Letters IEEE, 2004, 9(3):210—212.

[15] Jain R, Chiu D M, Hawe W R. A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System [M]. Hudson, MA: Eastern Research Laboratory, Digital Equipment Corporation, 1984.

(上接第 42 页)

3 结束语

本文通过对三维 MMI 的研究, 将分束与模式转换相结合, 在二维 MMI 模式转换器的基础上, 根据三维 MMI 的成像特性进行三维扩展, 设计了基于聚合物材料的三维模式转换分束器。在输入光波长为 1 550 nm 时, 通过模拟并计算器件的传输损耗和最大不均衡性, 对所设计的三维模式转换分束器的结构参数进行了优化。模拟结果显示, 优化后的器件结构总传输效率为 84.9%, 其中基模输出光场传输效率为 28.9%, 一阶模输出光场的转换效率为 56.6%; 基模输出光场的最大不均衡性为 0.010 dB, 一阶模输出光场的最大不均衡性为 0.004 dB。通过采用 ULI 技术可制备亚微米级精度的聚合物光波导。本文设计的三维模式转换分束器能够有效地解决传输效率低及扩大传输容量的问题, 从而满足下一代光通信技术发展的需求。

参考文献:

[1] Sakaguchi J, Puttnam B J, Klaus W, et al. 19-Core Fiber Transmission of $19 \times 100 \times 172$ -Gb/s SDM-

WDM-PDM-QPSK Signals at 305 Tb/s [C]// Optical Fiber Communication Conference and Exposition, Los Angeles, CA, USA: Optical Society of America, 2012:PDP5C.1.

[2] Herbert Venghaus, Grote Norbert. Fibre Optic Communication; Key Devices [M]. 2nd Ed. Berlin: Springer, 2017.

[3] Yoo S J B, Guan B, Scott R P. Heterogeneous 2D/3D Photonic Integrated Microsystems [J]. Microsystems & Nanoengineering, 2016, 2(16030): 1—9.

[4] Lindenmann N, Dottermusch S, Goedecke M L, et al. Connecting Silicon Photonic Circuits to Multicore Fibers by Photonic Wire Bonding [C]//Optical Interconnects Conference. Coronado, CA: IEEE, 2014: 131—132.

[5] Sun Y L, Jiang X Q, Yang J Y, et al. Experimental Demonstration of Two-dimensional Multimode-Interference Optical Power Splitter [J]. Chinese Physics Letters, 2003, 20(12):2182—2184.

[6] Leuthold J, Eckner J, Gamper E, et al. Multimode Interference Couplers for the Conversion and Combining of Zero- and First-order Modes [J]. Journal of Lightwave Technology, 1998, 16(7):1228—1239.