



(12) 发明专利

(10) 授权公告号 CN 102858015 B

(45) 授权公告日 2015.04.29

(21) 申请号 201110181669.1

审查员 李燕

(22) 申请日 2011.06.30

(73) 专利权人 普天信息技术研究院有限公司  
地址 100080 北京市海淀区海淀北二街 6 号

(72) 发明人 李远 宿淑艳

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018

代理人 王一斌 王琦

(51) Int. Cl.

H04W 72/10(2009.01)

(56) 对比文件

CN 101232705 A, 2008.07.30, 全文.

US 2005/0113089 A1, 2005.05.26, 全文.

CN 101227226 A, 2008.07.23, 全文.

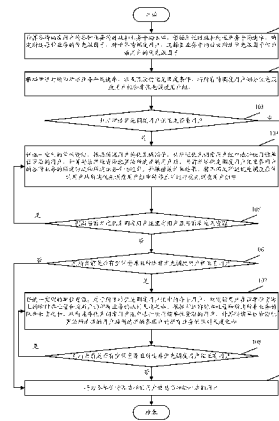
权利要求书3页 说明书9页 附图4页

(54) 发明名称

一种多业务调度方法

(57) 摘要

本发明提供了一种多业务调度方法,该方法根据业务的时延、平均速率和预设的优先调度条件,将所有待调度用户划分为优先调度用户组和非优先调度用户组,优先对优先调度用户组里的用户进行调度,调度时以满足 QoS 要求为目标,确保满足急需调度用户的服务质量需求,同时对于优先调度用户组中的用户只在优先调度用户组中选择配对用户。对于优先调度用户组中的用户,在分配完资源后,如果已经满足 QoS 要求,则放到非优先调度用户组中。另外,对于非优先调度用户组里的用户,根据瞬时吞吐量和所有业务的队列长度之和进行调度,以在保障业务需求的条件下优化系统容量。采用本发明可以提高传输速率,更好地满足业务需求和提升系统容量。



1. 一种多业务调度方法,其特征在于,该方法包括以下步骤:

a、计算各待调度用户的各种业务的时延和业务平均速率,所述时延为业务队列头的数据包进入该业务队列后的等待时间;根据所述时延和所述业务平均速率,确定每种业务的优先级因子,对于各待调度用户,选择其业务中的最大所述优先级因子作为该用户的优先级因子;

b、根据所述时延、所述业务平均速率和预设的优先调度条件,将所有待调度用户划分优先调度用户组和非优先调度用户组,如果所述优先调度用户组里有用户则执行步骤 c,否则执行步骤 d;

c、任选一空闲的单位资源,根据所述用户的优先级因子,从所述优先调度用户组中选择使用该单位资源的用户,计算将该单位资源配置给所选择的用户后,当前所述优先调度用户组里各用户的各种业务的所述时延、所述业务平均速率以及所述优先级因子,并根据该计算结果,将不满足所述优先调度条件的用户从所述优先调度用户组中转移至所述非优先调度用户组中,如果当前所述优先调度用户组里有用户且当前有空闲资源,则重复执行本步骤,否则,判断当前是否有空闲资源且所述非优先调度用户组里有用户,如果是,则执行步骤 d,否则执行步骤 e;

d、任选一空闲的单位资源,对于所述非优先调度用户组中的每个用户,确定该用户在该单位资源上的瞬时吞吐量和该用户的所有业务的队列长度之和,根据所述瞬时吞吐量和所述所有业务的队列长度之和,从所述非优先调度用户组中选择使用该单位资源的用户,计算将该单位资源配置给所选择的用户后所选择的各用户的所有业务的队列长度之和,判断当前是否有空闲资源且所述非优先调度用户组里有用户,如果是,则重复执行本步骤,否则,执行步骤 e;

e、将为各单位资源选择的用户信息通知给相应的用户;

其中,步骤 a 中所述确定所述每种业务的优先级因子为:

对于保证比特速率 GBR 业务,按照  $\omega_{k,i} = \alpha \cdot e^{d_{k,i}/D_i} + (1-\alpha)e^{(1-\bar{R}_{k,i}/GBR_i)}$ , 确定该 GBR 业务的优先级因子  $\omega_{k,i}$ ,

对于非 GBR 业务,按照  $\omega_{k,i} = \alpha \cdot e^{d_{k,i}/D_i}$ , 确定该非 GBR 业务的优先级因子  $\omega_{k,i}$ ;

其中,  $\alpha$  为加权因子,  $\alpha \in [0, 1]$ ,  $d_{k,i}$  为第 k 个用户第 i 种业务的时延,  $D_i$  为第 i 种业务的时延门限,  $GBR_i$  为第 i 种 GBR 业务的最低传输速率要求,  $\bar{R}_{k,i}$  为第 k 个用户第 i 种业

务的平均速率,  $u(\cdot)$  为阶跃函数,  $u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$ 。

2. 根据权利要求 1 所述的方法,其特征在于,所述优先调度条件为:

用户至少有一种业务满足:该业务的所述时延超过该种业务的时延门限或者所述平均速率低于该种业务的速率门限。

3. 根据权利要求 1 所述的方法,其特征在于,步骤 c 中根据所述用户的优先级因子,从所述优先调度用户组中选择使用该单位资源的用户为:

从所述优先调度用户组中,选择所述优先级因子最大的用户作为当前使用该单位资源

$n^*$ 的主用户,根据该主用户在该单位资源  $n^*$ 上的速率和该主用户的所述优先级因子,确定该主用户的配对优先值;

根据用户在该单位资源  $n^*$ 上的速率,在预设的单位资源的最大配置用户数范围内,从所述优先调度用户组中选择出与该主用户配对时能获得最大联合优先级因子且该最大联合优先级因子大于所述配对优先值的用户,并将所选择出的用户作为所述主用户的配对用户;

如果存在所述主用户的配对用户,则确定所述主用户和所述配对用户共同使用该单位资源  $n^*$ ,否则确定所述主用户使用该单位资源  $n^*$ 。

4. 根据权利要求 3 所述的方法,其特征在于,按照  $\mu_{k^*} = \omega_{k^*_{\max}} \cdot PF_{k^*}$ ,确定所述主用户的配对优先值  $\mu_{k^*}$ ;

其中,  $PF_{k^*} = \frac{R_{k^*_{SU}}}{T_{k^*}}$ ,  $R_{k^*_{SU}}$  为主用户  $k^*$ 的单用户模式下的瞬时速率,  $T_{k^*}$  为主用户  $k^*$

的平均速率;

所述联合优先级因子的确定为:

计算利用单位资源  $n^*$ 进行多用户传输时,进行传输的各用户 1 的传输速率  $R_{1\_MU}$ ;

对于进行所述多用户传输的每个用户 1,按照  $\mu_{1\_MU} = \omega_{1\_MAX} \cdot \frac{R_{1\_MU}}{T_1}$ ,确定该用户 1 的

配对优先值  $\mu_{1\_MU}$ ,其中,  $\omega_{1\_MAX}$ 为用户 1 的优先级因子;

对所有所述进行多用户传输的用户的配对优先值求和,将求和结果确定为利用单位资源  $n^*$ 进行所述多用户传输时的联合优先级因子。

5. 根据权利要求 3 所述的方法,其特征在于,按照  $\mu_{k^*} = R_{k^*_{SU}}$ ,确定所述主用户  $k^*$ 的配对优先值  $\mu_{k^*}$ ;

其中,  $R_{k^*_{SU}}$ 为主用户  $k^*$ 的单用户模式下的瞬时速率;

所述联合优先级因子的确定为:

计算利用单位资源  $n^*$ 进行多用户传输时,进行传输的各用户 1 的传输速率  $R_{1\_MU}$ ;

对于进行所述多用户传输的每个用户 1,按照  $\mu_{1\_MU} = R_{1\_MU}$ ,确定该用户 1 的配对优先值  $\mu_{1\_MU}$ ;

对所有所述进行多用户传输的用户的配对优先值求和,将求和结果确定为利用单位资源  $n^*$ 进行所述多用户传输时的联合优先级因子。

6. 根据权利要求 1 所述的方法,其特征在于,步骤 d 中根据所述瞬时吞吐量和所述所有业务的队列长度之和,从所述非优先调度用户组中选择使用该单位资源的用户为:

根据所述瞬时吞吐量  $B_1$  和所述所有业务的队列长度之和  $Q_{tot\_1}$ ,按照  $B_{eff\_1} = \min(Q_{tot\_1}, B_1)$  计算所述非优先调度用户组中各用户 1 在该单位资源  $n^*$ 上的有效速率  $B_{eff\_1}$ ,选择有效速率  $B_{eff\_1}$ 最大的用户作为当前使用该单位资源  $n^*$ 的主用户;

在预设的单位资源的最大配置用户数范围内,从所述非优先调度用户组中选择出与所述主用户配对时能获得最大联合有效速率并且该最大联合有效速率大于所述主用户在该

单位资源  $n^*$  上的有效速率的用户, 作为所述主用户的配对用户;

如果存在所述主用户的配对用户, 则确定所述主用户和所述配对用户共同使用该单位资源  $n^*$ , 否则确定所述主用户使用该单位资源  $n^*$ 。

7. 根据权利要求 6 所述的方法, 其特征在于, 按照  $B_{MU\_eff} = \sum_{m \in \Gamma} \min(B_m, Q_{tot\_m})$ , 确定所述联合有效速率  $B_{MU\_eff}$ , 其中,  $\Gamma$  为多用户传输时的用户集合,  $B_m$  为用户  $m$  在多用户传输模式下的瞬时速率,  $Q_{tot\_m}$  为用户  $m$  的所有业务的队列长度之和。

8. 根据权利要求 1 所述的方法, 其特征在于, 步骤 d 中, 按照  $Q_{tot\_m} = Q_{tot\_m} - \min(B_m, Q_{tot\_m})$ , 计算将该单位资源配置给所选择的用户后所选择的各用户的所有业务的队列长度之和  $Q_{tot\_m}$ ; 其中,  $m \in \Gamma$ ,  $\Gamma$  为所选择的用户集合,  $B_m$  为用户  $m$  在多用户传输模式下在该单位资源上进行传输的瞬时速率。

## 一种多业务调度方法

### 技术领域

[0001] 本发明涉及移动通信系统的调度技术,特别是涉及 LTE 系统的基于多输入多输出 (MU-MIMO) 的多业务调度方法。

### 背景技术

[0002] 在 3GPP 长期演进 (LTE) 系统支持多种不同类型的业务,如语音业务、视频流业务、WWW 业务和 FTP 业务等,并将这些业务按照不同的 QoS 需求,如业务时延、保证比特速率等划分为会话业务、流业务、交互业务和背景类业务四类。通常一个用户可以在同一时间并行地应用多种业务,比如可以在语音业务的同时浏览网页或传输视频流,因此为了支持这种多用户多业务的应用场景, LTE 系统的多层传输资源分配方案需要针对不同类型的业务设计合理的调度方案。一般来说,按照 QoS 需求的不同,承载可以分为两类,即保证比特速率 (GBR) 承载和非保证比特速率 (Non-GBR) 承载。前者可以用来提供语音业务和视频流业务等实时性业务,对时延有较高的要求,而且对传输比特速率有要求,需要为这类业务分配足够的带宽以保障比特速率;后者可以提供 WWW 业务和 FTP 业务等非实时性业务,不需要保障比特速率,因此对分配带宽没有要求,而且一般来说对时延要求也低于实时性业务。为了结合业务的 QoS 需求进行调度, LTE 系统的接入网利用 S1 接口传递的质量等级指示 (quality class indicator, QCI) 指示应用于 eNodeB MAC 的层资源调度。

[0003] 在 R5 以后的 UMTS 系统中,分组调度的功能被移植到基站端完成,从而可以根据实时的链路状况进行快速链路自适应和实现更为灵活的分组调度。在 LTE 系统中,分组调度算法所要完成的任务就是在每个调度周期为具有不同时延要求、不同 QoS 等级和不同速率要求的用户提供合理的资源分配,使调度结果可以满足不同业务的需求。衡量一种调度算法的性能一方面要看其是否能够满足用户的业务需求以及用户公平性,另一方面则关注其是否能够结合时变信道的瞬时质量实现频率选择性调度,实现更高的下行数据速率。常见的分组调度算法包括轮询 (Round Robin) 算法、最大载干比 (MAXC/I) 算法、比例公平 (Proportional Fair) 算法等。

[0004] 轮询算法的基本思想是:循环地调用每个用户,就被调度的概率而言,对 K 个用户,一次循环中每个用户就被调度的概率都等于  $1/K$ ,每个用户以相同的概率占有可分配的资源。该算法不考虑用户以前被调度的情况,属于没有记忆的调度算法。轮询算法认为不同用户的传输优先级是平等的,因而实现了用户间的最佳公平性,在考虑公平性时,一般都把轮询算法作为衡量的标准。轮询算法的优点在于可以保证用户间的长时间和短时间内的时间公平性,而且实现简单。但是轮询算法主要考虑了用户间的时间公平性,没有考虑到资源的使用效率和不同用户无线信道的具体情况,因此一般来说使用该算法时的系统吞吐量是很低的。

[0005] MAX C/I 算法和比例公平算法是两种基于信道衰落的调度算法。MAXC/I 算法只考虑调度信道质量最好的用户,系统资源一直给这些用户服务,使得信道质量好的用户一直得到服务,等到该用户信道质量变差后再选择信道质量更好的用户传输。由于系统永远为

信道质量最好的用户服务,因此该算法可以获得较高的频率选择性增益。最大 C/I 调度算法的吞吐量是调度算法的吞吐量极限,其优点在于当系统采用自适应调制编码技术时,采用最大载干比算法的无线系统可获得最大的系统吞吐量,而且实现简单;但是 MAXC/I 算法没有考虑到用户公平性,导致信道状况较差的边缘用户很难得到有效的服务。

[0006] 相比而言,PF 调度算法根据用户的瞬时信道质量和用户在一个时间窗内的平均速率计算得到一个优先级因子,将资源分配给优先级因子最高的用户;优先级因子公式表示为瞬时速率和平均速率的比例,如果用户的信道条件较好,比如离服务基站比较近,则该用户的瞬时吞吐量比较大,由优先级公式可知,该用户的优先级提高。反之,如果用户因为信道条件较差,特别是由于它处于小区边缘, SINR 长时间较低,得不到传输机会,则其平均传输速率就会减小,由优先级公式可知,这同样会使其优先级提高并使边缘获得传输机会。PF 算法能在吞吐量和用户公平性之间获得较好的折衷,而且通过调整滑动时间窗的长度可以调节调度的长期公平性和短期公平性。

[0007] 多天线的引入提供了更大的系统容量,也为调度算法的实现提供了更多可能性。相比于单用户 MIMO 而言,多用户 MIMO 可以通过空分多址 (SDMA) 在同一时频资源上同时调度多个用户,利用发送端的预编码消除用户间干扰,从而利用多用户分集增益提高系统容量。在 LTE R9 标准中已经支持基于闭环码本反馈的多用户 MIMO 方案以及基于波束赋形的多用户 MIMO 方案,可以通过相应的用户配对策略将资源分配给一个用户集合,并利用预编码抑制流间干扰。MU-MIMO 结合空间分集和空间复用的优点,作为一种提高频谱效率的有效方式将得到更为广泛的应用。

[0008] 上述常用的资源调度算法,如轮询算法、MAX C/I 算法和 PF 算法,均在调度时没有考虑到业务的 QoS 特性,因此不能很好地满足业务的 QoS 需求。虽然一些单业务调度算法考虑了调度时将时延、丢包率等 QoS 需求和信道衰落特性相结合,但是,由于实际系统中通常出现不同 QoS 需求的多种实时业务和非实时共存的情况,因此,这些单业务调度算法对于单用户多业务的优先级选择和调度而言,无法起到很好的效果。

## 发明内容

[0009] 有鉴于此,本发明的主要目的在于提供一种多业务调度方法,该方法能在基于 MU-MIMO 进行多业务调度时满足业务的 QoS 需求。

[0010] 为了达到上述目的,本发明提出的技术方案为:

[0011] 一种多业务调度方法,该方法包括以下步骤:

[0012] a、计算各待调度用户的各种业务的时延和业务平均速率,所述时延为业务队列头的数据包进入该业务队列后的等待时间;根据所述时延和所述业务平均速率,确定所述每种业务的优先级因子,对于各待调度用户,选择其业务中的最大所述优先级因子作为该用户的优先级因子;

[0013] b、根据所述时延、所述业务平均速率和预设的优先调度条件,将所有待调度用户划分优先调度用户组和非优先调度用户组,如果所述优先调度用户组里有用户则执行步骤 c,否则执行步骤 d;

[0014] c、任选一空闲的单位资源,根据所述用户的优先级因子,从所述优先调度用户组中选择使用该单位资源的用户,计算将该单位资源配置给所选择的用户后,当前所述优先

调度用户组里各用户的各种业务的所述时延、所述业务平均速率以及所述优先级因子,并根据该计算结果,将不满足所述优先调度条件的用户从所述优先调度用户组中转移至所述非优先调度用户组中,如果当前所述优先调度用户组里有用户且当前有空闲资源,则重复执行本步骤,否则,判断当前是否有空闲资源且所述非优先调度用户组里有用户,如果是,则执行步骤 d,否则执行步骤 e;

[0015] d、任选一空闲的单位资源,对于所述非优先调度用户组中的每个用户,确定该用户在该单位资源上的瞬时吞吐量和该用户的所有业务的队列长度之和,根据所述瞬时吞吐量和所述所有业务的队列长度之和,从所述非优先调度用户组中选择使用该单位资源的用户,计算将该单位资源配置给所选择的用户后所选择的各用户的所有业务的队列长度之和,判断当前是否有空闲资源且所述非优先调度用户组里有用户,如果是,则重复执行本步骤,否则,执行步骤 e;

[0016] e、将为各单位资源选择的用户信息通知给相应的用户。

[0017] 综上所述,本发明提出的多业务调度方法,根据业务的时延、平均速率和预设的优先调度条件,将所有待调度用户划分为两组,即优先调度用户组和非优先调度用户组,优先对优先调度用户组里的用户进行调度,调度时以满足 QoS 要求为目标,这样可以确保满足急需调度用户的服务质量需求。另外,对于非优先调度用户组里的用户即满足 QoS 要求的用户,根据瞬时吞吐量和所有业务的队列长度之和进行调度,也就是说,以最大化传输速率为目标进行调度,如此,可以在保障业务需求的条件下优化系统容量。因此,本发明在多业务调度的过程中结合 MU-MIMO 传输方案可以提高传输速率,更好地满足业务需求和提升系统容量。

## 附图说明

[0018] 图 1 为本发明的主要流程示意图;

[0019] 图 2 为本发明实施例一的流程示意图;

[0020] 图 3 为图 2 中步骤 206 的流程示意图;

[0021] 图 4 为图 2 中步骤 213 的流程示意图。

## 具体实施方式

[0022] 为使本发明的目的、技术方案和优点更加清楚,下面将结合附图及具体实施例对本发明作进一步地详细描述。

[0023] 本发明的核心思想是,在进行调度时优先针对时延和传输速率两种业务需求,将急需调度的业务分到第一优先调度集合,进行绝对优先的调度,可以使这类用户的需求得到最大的满足;另外提供了一种多业务优先级排列方法,可以结合时延和速率特性及相应的 QoS 需求对不同用户进行优先级排序;对于急需调度的用户,以满足 QoS 要求为目标,完成急需调度用户的资源分配之后,对满足 QoS 要求的用户以最大化传输速率为目标进行调度,可以在保障业务需求的条件下优化系统容量。这样,本发明在多业务调度的过程中结合 MU-MIMO 传输方案可以提高传输速率,更好地满足业务需求和提升系统容量。

[0024] 图 1 为本发明的主要流程示意图,如图 1 所示,本发明主要包括以下步骤:

[0025] 步骤 101、计算各待调度用户的各种业务的时延和业务平均速率,所述时延为业务

队列头的数据包进入该业务队列后的等待时间,所述业务平均速率为一种业务在一个时间窗内的平均传输速率;根据所述时延和所述业务平均速率,确定所述每种业务的优先级因子,对于各待调度用户,选择其业务中的最大所述优先级因子作为该用户的优先级因子。

[0026] 这里,确定所述每种业务的优先级因子的方法可以为:

[0027] 对于 GBR 业务,按照  $\omega_{k,i} = \alpha \cdot e^{d_{k,i}/D_i} + (1-\alpha)e^{(1-\bar{R}_{k,i})/GBR_i}$  确定该 GBR 业务的优先级因子  $\omega_{k,i}$ ,

[0028] 对于非 GBR 业务,按照  $\omega_{k,i} = \alpha \cdot e^{d_{k,i}/D_i}$  确定该非 GBR 业务的优先级因子  $\omega_{k,i}$ ;

[0029] 其中,  $\alpha$  为加权因子,  $\alpha \in [0, 1]$ ,  $d_{k,i}$  为第  $k$  个用户第  $i$  种业务的时延,  $D_i$  为第  $i$  种业务的时延门限,  $GBR_i$  为第  $i$  种 GBR 业务的最低传输速率要求,  $\bar{R}_{k,i}$  为第  $k$  用户第  $i$  业务

的平均速率,  $u(\cdot)$  为阶跃函数,  $u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$ 。

[0030] 步骤 102-103、根据所述时延和所述业务平均速率,以及预设的优先调度条件,将所有待调度用户划分优先调度用户组和非优先调度用户组,如果所述优先调度用户组里有用户,则执行步骤 104,否则执行步骤 107。

[0031] 这里,所述优先调度条件为:用户至少有一种业务满足:该业务的所述时延超过该种业务的时延门限或者所述平均速率低于该种业务的速率门限。也就是满足该条件的用户将被划分至优先调度用户组内。

[0032] 步骤 104-106、任选一空闲的单位资源,根据所述用户的优先级因子,从所述优先调度用户组中选择使用该单位资源的用户,计算将该单位资源配置给所选择的用户后,当前所述优先调度用户组里各用户的各种业务的所述时延和所述业务平均速率,并根据该计算结果,将不满足所述优先调度条件的用户从所述优先调度用户组中转移至所述非优先调度用户组中,如果当前所述优先调度用户组里有用户且当前有空闲资源,则重复执行步骤 104,否则,判断当前是否有空闲资源且所述非优先调度用户组里有用户,如果是,则执行步骤 107,否则执行步骤 109。

[0033] 这里,根据所述用户的优先级因子,从所述优先调度用户组中选择使用该单位资源的用户具体过程可以为:

[0034] 从所述优先调度用户组中,选择所述优先级因子最大的用户作为当前使用该单位资源  $n^*$  的主用户,根据该主用户在该单位资源  $n^*$  上的速率和该主用户的所述优先级因子,确定该主用户的配对优先值;

[0035] 根据用户在该单位资源  $n^*$  上的速率,在预设的单位资源的最大配置用户数范围内,从所述优先调度用户组中选择出与该主用户配对时能获得最大联合优先级因子且该最大联合优先级因子大于所述配对优先值的用户,并将所选择出的用户作为所述主用户的配对用户;

[0036] 如果存在所述主用户的配对用户,则确定所述主用户和所述配对用户共同使用该单位资源  $n^*$ ,否则,确定所述主用户使用该单位资源  $n^*$ 。

[0037] 需要说明的是,在实际应用中,可以采用多种方法来确定主用户的配对优先值,以及联合优先级因子。下面给出两种较佳的方法,但不限于此:



[0038] 方法一：

[0039] 按照  $\mu_{k^*} = \omega_{k^*_{\max}} \cdot PF_{k^*}$ ，确定所述主用户的配对优先值  $\mu_{k^*}$ ；

[0040] 其中， $PF_{k^*} = \frac{R_{k^*_{SU}}}{T_{k^*}}$ ， $R_{k^*_{SU}}$  为主用户  $k^*$  的单用户模式下的瞬时速率， $T_{k^*}$  为主用户  $k^*$  的平均速率；

[0041] 相应的，所述联合优先级因子的确定方法可以为：

[0042] 计算利用单位资源  $n^*$  进行多用户传输时，进行传输的各用户  $l$  的传输速率  $R_{l\_MU}$ ；

[0043] 对于进行所述多用户传输的每个用户  $l$ ，按照  $\mu_{l\_MU} = \omega_{l\_MAX} \cdot \frac{R_{l\_MU}}{T_l}$ ，确定该用户  $l$

的配对优先值  $\mu_{l\_MU}$ ，其中， $\omega_{l\_MAX}$  为用户  $l$  的优先级因子；

[0044] 对所有所述进行多用户传输的用户的配对优先值求和，将求和结果确定为利用单位资源  $n^*$  进行所述多用户传输时的联合优先级因子。

[0045] 方法二：

[0046] 按照  $\mu_{k^*} = R_{k^*_{SU}}$ ，确定所述主用户  $k^*$  的配对优先值  $\mu_{k^*}$ ；其中， $R_{k^*_{SU}}$  为主用户  $k^*$  的单用户模式下的瞬时速率。

[0047] 相应的，所述联合优先级因子的确定方法可以为：

[0048] 计算利用单位资源  $n^*$  进行多用户传输时，进行传输的各用户  $l$  的传输速率；

[0049] 对于进行所述多用户传输的每个用户  $l$ ，按照  $\mu_{l\_MU} = R_{l\_MU}$ ，确定该用户  $l$  的配对优先值  $\mu_{l\_MU}$ ；

[0050] 对所有所述进行多用户传输的用户的配对优先值求和，将求和结果确定为利用单位资源  $n^*$  进行所述多用户传输时的联合优先级因子。

[0051] 步骤 107-108、任选一空闲的单位资源，对于所述非优先调度用户组中的每个用户，确定该用户在该单位资源上的瞬时吞吐量和该用户的所有业务的队列长度之和，所述瞬时吞吐量为用户单位资源上的传输比特数目，根据所述瞬时吞吐量和所述所有业务的队列长度之和，从所述非优先调度用户组中选择使用该单位资源的用户，计算将该单位资源配置给所选择的用户后所选择的各用户的所有业务的队列长度之和，判断当前是否有空闲资源且所述非优先调度用户组里有用户，如果是，则重复执行步骤 107，否则，执行步骤 109。

[0052] 这里，本步骤中根据所述瞬时吞吐量和所述所有业务的队列长度之和，从所述非优先调度用户组中选择使用该单位资源的用户的具体方法可以为：

[0053] 根据所述瞬时吞吐量  $B_l$  和所述所有业务的队列长度之和  $Q_{tot\_l}$ ，按照  $B_{eff\_l} = \min(Q_{tot\_l}, B_l)$  计算所述非优先调度用户组中各用户  $l$  在该单位资源  $n^*$  上的有效速率  $B_{eff\_l}$ ，选择有效速率  $B_{eff\_l}$  最大的用户作为当前使用该单位资源  $n^*$  的主用户；

[0054] 在预设的单位资源的最大配置用户数范围内，从所述优先调度用户组中选择出与所述主用户配对时能获得最大联合有效速率并且该最大联合有效速率大于所述主用户在该单位资源  $n^*$  上的有效速率的用户，作为所述主用户的配对用户；

[0055] 如果存在所述主用户的配对用户，则确定所述主用户和所述配对用户共同使用该单位资源  $n^*$ ，否则确定所述主用户使用该单位资源  $n^*$ 。

[0056] 较佳地,上述方法中可以按照  $B_{MU\_eff} = \sum_{m \in \Gamma} \min(B_m, Q_{tot\_m})$ , 来确定所述联合有效速率  $B_{MU\_eff}$ , 其中,  $\Gamma$  为多用户传输时的用户集合,  $B_m$  为用户  $m$  在多用户传输模式下的瞬时速率,  $Q_{tot\_m}$  为用户  $m$  的所有业务的队列长度之和。

[0057] 较佳地,可以按照  $Q_{tot\_m} = Q_{tot\_m} - \min(B_m, Q_{tot\_m})$ , 计算将该单位资源配置给所选择的用户后所选择的各用户的所有业务的队列长度之和  $Q_{tot\_m}$ ; 其中,  $m \in \Gamma$ ,  $\Gamma$  为所选择的用户集合,  $B_m$  为用户  $m$  在多用户传输模式下在该单位资源上进行传输的瞬时速率。

[0058] 步骤 109、将为各单位资源选择的用户信息通知给相应的用户。

[0059] 下面通过本发明的实施例一对上述方案的具体实现进一步说明。

[0060] 步骤 201、初始化。

[0061] 将所有的资源、用户集合以及优先级参数进行初始化,系统所有可用带宽资源初始化为资源集合, Lv1 和 Lv2 用户集合初始化为空集。

[0062] 步骤 202、业务优先级计算。

[0063] 计算每个用户每种业务的时延和业务平均速率。业务时延表示为某种业务队列头的数据包进入队列的时间;业务平均速率表示为在一个时间窗内某种业务的平均传输速率,用于保障 GBR 需求,具体的时间窗的长度和平均速率计算方法属于公知技术,这里不再做详细说明。利用时延和平均传输速率计算得到每种业务的业务优先级因子  $\omega_{k,i}$ , 其中  $k$  表示第  $k$  个用户,  $i$  表示第  $i$  种业务;不同业务类型的优先级可以用公式 (1) 来计算,  $\omega_{k,i}$  越大则优先级越高。对于 GBR 业务,优先级因子包括时延部分和业务速率部分;对于非 GBR 业务,优先级因子包括时延部分。

$$[0064] \quad \omega_{k,i} = \begin{cases} \alpha \cdot e^{\frac{d_{k,i}}{D_i}} + (1-\alpha)e^{(1-\frac{\bar{R}_{k,i}}{GBR_i})} \cdot u(GBR_i - \bar{R}_{k,i}), & GBR \\ \alpha \cdot e^{\frac{d_{k,i}}{D_i}}, & Non-GBR \end{cases} \quad (1)$$

[0065] 其中:

[0066]  $\alpha$  是加权因子,  $\alpha \in [0, 1]$ , 调节  $\alpha$  的值可以调节调度过程中对时延和 GBR 需求的权重,  $\alpha$  越大则对时延越敏感,反之则对业务速率越敏感。

[0067]  $d_{k,i}$  表示第  $k$  个用户第  $i$  种业务的时延

[0068]  $D_i$  表示第  $i$  种业务的时延门限,不同的业务类型对时延的敏感度不同,非实时性业务的时延门限高于实时业务的门限,因而对时延的忍耐度更高。优先级因子时延部分的取值范围为  $\alpha \cdot [1, e]$ 。

[0069]  $GBR_i$  为第  $i$  种 GBR 业务的最低传输速率要求,不同业务对 GBR 的要求不同,对应不同的门限;

[0070]  $\bar{R}_{k,i}$  表示第  $k$  用户第  $i$  业务的平均速率。 $u(\cdot)$  为阶跃函数,  $u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$  当业务平均速率超过 GBR 门限速率时,优先级因子中的 GBR 部分为 0,表示满足 GBR 要求之后,业务速率不再影响优先级。当平均速率低于 GBR 门限时,优先级因子 GBR 部分的取值范围为

$(1-\alpha)[1, e]$ 。

[0071] 公式 (1) 可以用来做不同业务类型的优先级比较,所要实现的目标包括:对于 GBR 业务,影响优先级的队列因素包括时延和业务平均速率,而对于非 GBR 业务,队列因素只有

时延；业务优先级随时延的上升和平均速率的下降而提高，并且呈加速提高；当业务满足 GBR 需求时，GBR 业务的平均速率不再影响该业务的优先级，GBR 业务和非 GBR 业务之间通过时延来比较优先级；GBR 业务中，时延和速率两者对优先级的影响程度通过一个权重因子来体现，以便于基站根据结果调整两者的权重。

[0072] 在业务间优先级因子计算结束之后，可以通过公式 (2) 获取用户优先级因子  $\omega_{k_{\max}}$ ，等于该用户所有业务中优先级最高的业务的优先级因子。

$$[0073] \quad \omega_{k_{\max}} = \arg \max_i (\omega_{k,i}) \quad (2)$$

[0074] 步骤 203、用户分组。

[0075] 根据上一步中获取的每个用户每种业务的队列头时延和平均速率，判断是否有业务时延超过时延门限  $\alpha_i$ ，或平均速率低于速率门限  $\beta_i$  的用户，将这些超过门限的用户列为急需调度的用户，添加到 Lv1 用户集合  $\Omega$  中。这类用户具有最高的优先级，需要在其他用户之前优先调度，Lv1 用户调度目标是确保用户的时延和 GBR 需求，调度策略是针对用户分配资源。其余用户添加到 Lv2 用户集合  $\Psi$  中，Lv2 用户调度目标是最大化系统吞吐量，调度策略是针对资源选择用户。

[0076] 步骤 204、判断 Lv1 用户组中是否有用户，如果存在这类符合条件的急需调度的用户，则执行步骤 205，进入 Lv1 调度，如果 Lv1 用户集合  $\Omega$  为空集则执行步骤 211，进入 Lv2 资源调度。

[0077] 步骤 205、Lv1 主用户选择。

[0078] 在调度过程中，首先为 Lv1 集合中的用户进行优先级排序，排序过程只考虑用户的队列特性，不考虑信道特性，按照公式 (2) 计算得到的用户优先级因子进行排序，优先级因子越大则优先级越高。选定的用户  $k^*$  作为主用户，在可用的资源集合  $\Lambda$  中选择速率最大的资源  $n^*$  分配给用户  $k^*$ 。

[0079] 步骤 206、Lv1 用户配对。

[0080] 在 Lv1 用户配对过程中，为了确保 Lv1 集合的用户可以得到优先调度，候选配对的用户只能从 Lv1 的集合中选择。用户配对按照贪婪准则，结合吞吐量和业务需求，选择联合优先级因子最大的用户作为最佳配对用户，然后对比添加该用户后的配对用户集合和未添加该用户的配对用户集合的优先级因子从而判断是否要将该用户添加到配对用户集合中。具体步骤如下（如图 3 所示）：

[0081] 步骤 2061、初始化用户集合  $\Gamma = \{k^*\}$ ，并将步骤 205 中选择的主用户  $k^*$  添加到配对用户集合  $\Gamma$  中。

[0082] 步骤 2062、根据主用户  $k^*$  在资源  $n^*$  上的速率计算用户优先级  $F(R_{k^*,n^*})$ （即前述的主用户的配对优先值），这个优先级也是用户集合  $\Gamma$  的初始化优先级。

[0083] 步骤 2063、轮询其余用户，从剩余的用户集合  $\Omega - \Gamma$  中选择备选配对用户 1，计算将 1 添加到集合  $\Gamma$  后在资源  $n^*$  上的联合速率  $R_{\Gamma \cup \{1\}}$ ，并根据一定的联合优先级公式计算得到联合优先级因子  $F(R_{\Gamma \cup \{1\}})$ 。轮询结束后选择联合优先级因子最大的用户  $1^*$  作为最佳配对用户， $1^* = \arg \max_{l \in \Omega - \Gamma} F(R_{\Gamma \cup \{l\}})$ 。

[0084] 步骤 2064、判断添加  $1^*$  后  $\Gamma \cup \{1^*\}$  的联合优先级因子和之前  $\Gamma$  的优先级因子  $F(R_{\Gamma})$ ，若后者大于前者则停止配对过程，并执行步骤 2066，且配对用户集合中不添加该用

户；若前者大于后者则执行步骤 2064-1)，将该最佳配对用户添加到用户集合中，更新集合  $\Gamma = \Gamma \cup \{1\}$ 。

[0085] 步骤 2065、判断配对用户集合  $\Gamma$  中用户数目是否达到配对的最大用户数，如果达到最大用户数，则执行步骤 2066，否则执行步骤 2063，再选择下一个配对用户进行配对。

[0086] 步骤 2066、结束配对过程，将资源  $n^*$  分配给用户集合  $\Gamma$ 。

[0087] 步骤 207、更新资源和 Lv1 用户。

[0088] 每次为一个配对用户集合分配一个资源单位，就更新该用户集合中所有用户的队列头时延、平均速率并计算该用户的用户优先级因子  $\omega_{k \max}$ 。将资源  $n^*$  从资源集合  $\Lambda$  中删除。

[0089] 步骤 208、利用更新后的用户信息判断当前 Lv1 用户集合中的用户是否满足 Lv1 集合的时延和 GBR 要求，如果当前用户集合中某个用户更新队列信息后所有业务  $i$  的时延低于门限  $\alpha_i$  且平均速率高于门限  $\beta_i$ ，则执行步骤 209。

[0090] 步骤 209、将步骤 208 中不满足 Lv1 门限的用户从 Lv1 集合  $\Omega$  中去除，添加到 Lv2 集合中，在后续的 Lv1 调度中不再为该用户分配资源。

[0091] 步骤 210、判断 Lv1 集合  $\Omega$  中是否还有用户，如果 Lv1 集合中没有用户，则执行步骤 210-1，如果仍有用户，则执行步骤 210-2。

[0092] 步骤 210-1、判断资源集合  $\Lambda$  中是否还有资源，如果有资源，则执行步骤 211，进入 Lv2 资源调度，如果没有资源，则执行步骤 216。

[0093] 步骤 210-2、判断资源集合  $\Lambda$  中是否还有资源，如果系统还有剩余资源，则执行步骤 5，如果没有剩余资源，则执行步骤 16。

[0094] 步骤 211、Lv2 初始化。

[0095] Lv2 调度要满足的目标是最大化系统吞吐量，在资源分配过程中只考虑用户的信道特性，而不考虑用户的业务特性。由于在 Lv1 调度中，满足业务需求门限的用户被划分到 Lv2 的集合，因此 Lv2 的用户集合  $\Psi$  包含小区内的所有用户，可用资源为 Lv1 调度后的剩余资源。在剩余可用的资源中任意选择一个资源单位  $n^*$ ，获取用户集合  $\Psi$  中所有用户在资源  $n^*$  上的瞬时吞吐量  $B_k$  和用户所有业务的队列长度  $Q_{\text{tot}_k}$ 。

[0096] 步骤 212、Lv2 主用户选择。

[0097] 根据用户信道信息为当前资源上的所有用户进行排序，排序原则是最大速率原则。为了防止队列长度较短的用户分到较大的资源而造成资源浪费，在 Lv2 调度过程中结合用户的队列长度，利用有效速率表示某个资源上的用户速率。瞬时速率可以表示为用户  $k$  在资源  $n^*$  上的传输比特数目  $B_k$  (bit)，用户  $k$  所有业务的队列长度之和表示为  $Q_{\text{tot}_k}$  (bit)，那么用户  $k$  的有效速率可以表示为  $B_{\text{eff}_k} = \min(Q_{\text{tot}_k}, B_k)$ 。在资源单位  $n^*$  上选择  $B_{\text{eff}_k}$  最大的用户  $k^*$  作为主用户。

$$[0098] \quad k^* = \arg \max_{k \in \Psi} (B_{\text{eff}_k}) \quad (3)$$

[0099] 步骤 213、Lv2 用户配对。

[0100] 配对的过程仍然采用有效速率作为衡量配对是否成功的准则。考虑到遍历的算法会导致计算复杂度的提升，这里采用贪婪算法选择配对用户。具体步骤如下（如图 4 所示）：

[0101] 步骤 2131、初始化用户集合  $\Gamma = \{k^*\}$ ，并将步骤 208 中选择的主用户  $k^*$  添加到配对用户集合  $\Gamma$  中。计算用户  $k^*$  在  $n^*$  上的有效速率  $B_{\text{eff}_\Gamma}$ 。

[0102] 步骤 2132、轮询其余用户。在用户集合  $\Psi - \Gamma$  中选择备选配对用户  $l$ ，将其添加到集合  $\Gamma$  中，并获取集合  $\Gamma \cup \{l\}$  中每个用户的瞬时速率  $B_{\Gamma \cup \{l\}} = \sum_{m \in \Gamma \cup \{l\}} B_m$ ，其中  $B_m$  表示  $\Gamma \cup \{l\}$  在 MU 传输模式下的瞬时速率。利用公式 (4) 计算集合  $\Gamma \cup \{l\}$  的联合有效速率  $B_{\text{eff}_{\Gamma \cup \{l\}}}$ ，作为配对的衡量准则。

$$[0103] \quad B_{\text{eff}_{\Gamma \cup \{l\}}} = \sum_{m \in \Gamma \cup \{l\}} \min(B_m, Q_{\text{tot}_m}) \quad (4)$$

[0104] 遍历备选配对用户集合  $\Psi - \Gamma$  中的所有用户，并利用公式 (5) 选择有效速率最大的用户  $l^*$  作为最佳配对用户。

$$[0105] \quad l^* = \arg \max_{l \in \Psi - \Gamma} (B_{\text{eff}_{\Gamma \cup \{l\}}}) \quad (5)$$

[0106] 步骤 2133、判断添加最佳配对用户  $l^*$  之后的有效速率  $B_{\text{eff}_{\Gamma \cup \{l^*\}}}$  和原集合  $\Gamma$  的有效速率  $B_{\text{eff}_\Gamma}$ ，若后者大于前者，则停止配对过程，执行步骤 2135，且配对用户集合中不添加该用户；若前者大于后者，则执行 2133-1)，将  $l^*$  添加到配对用户集合  $\Gamma$  中，更新集合  $\Gamma = \Gamma \cup \{l^*\}$ 。

[0107] 步骤 2134、判断配对用户集合  $\Gamma$  中用户数目是否达到配对的最大用户数，如果达到最大用户数，则执行步骤 2135，否则执行步骤 2133，再选择下一个候选配对用户进行配对。

[0108] 步骤 2135、结束配对过程，将资源  $n^*$  分配给用户集合  $\Gamma$ 。

[0109] 步骤 214、更新资源和用户队列。将资源  $n^*$  分配给用户集合  $\Gamma$ ，更新资源  $\Lambda = \Lambda - \{n^*\}$ ，并为集合  $\Gamma$  中的所有用户更新队列长度， $Q_{\text{tot}_m} = Q_{\text{tot}_m} - \min(B_m, Q_{\text{tot}_m})$ ， $m \in \Gamma$ 。

[0110] 步骤 215、检查剩余资源，如果还有剩余资源，则进入步骤 211 继续执行资源分配，如果没有剩余资源，则执行步骤 216。

[0111] 步骤 216、结束分组资源分配过程，通过 PDCCH 信道将调度结果告知用户。

[0112] 综上所述，以上仅为本发明的较佳实施例而已，并非用于限定本发明的保护范围。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

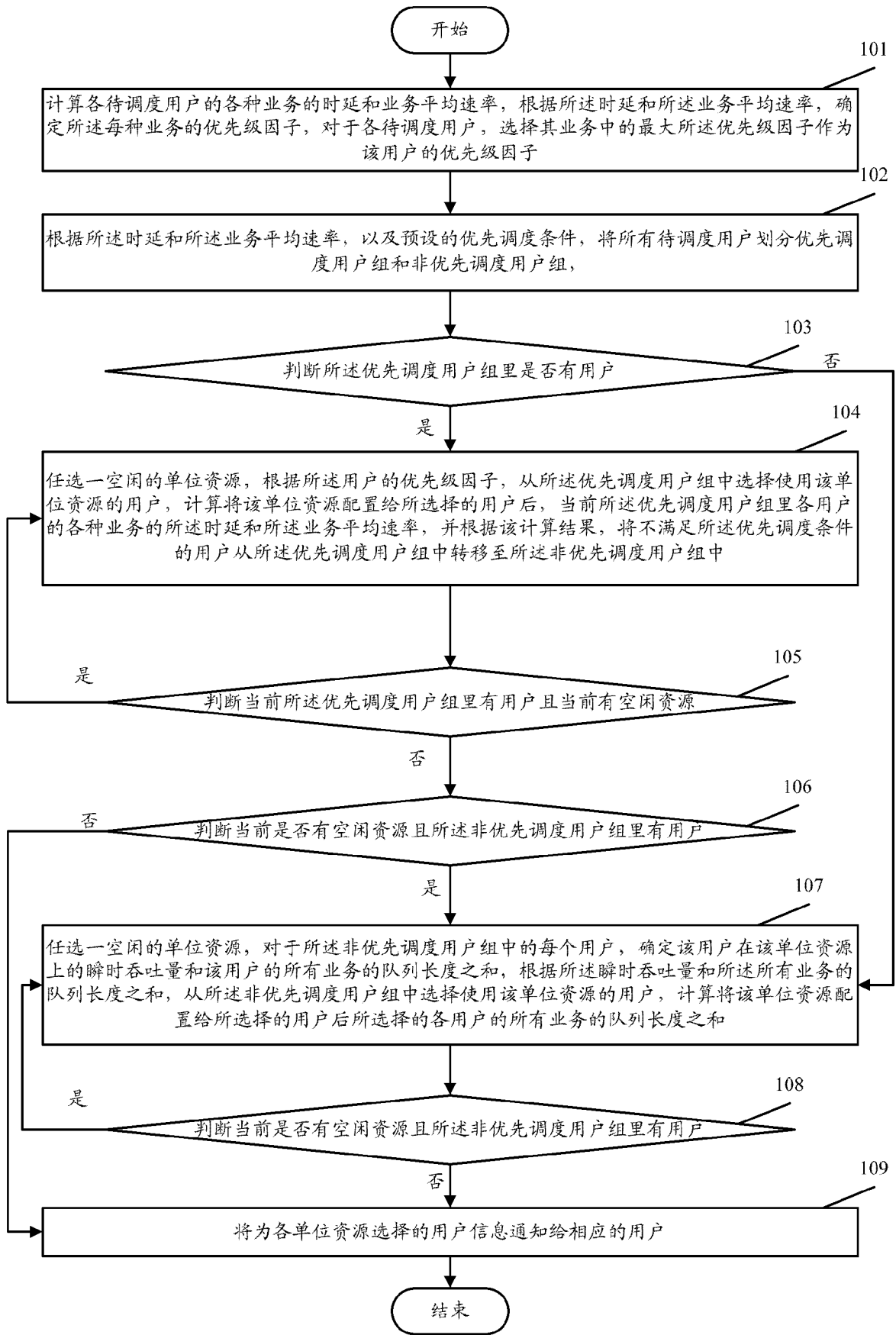


图 1

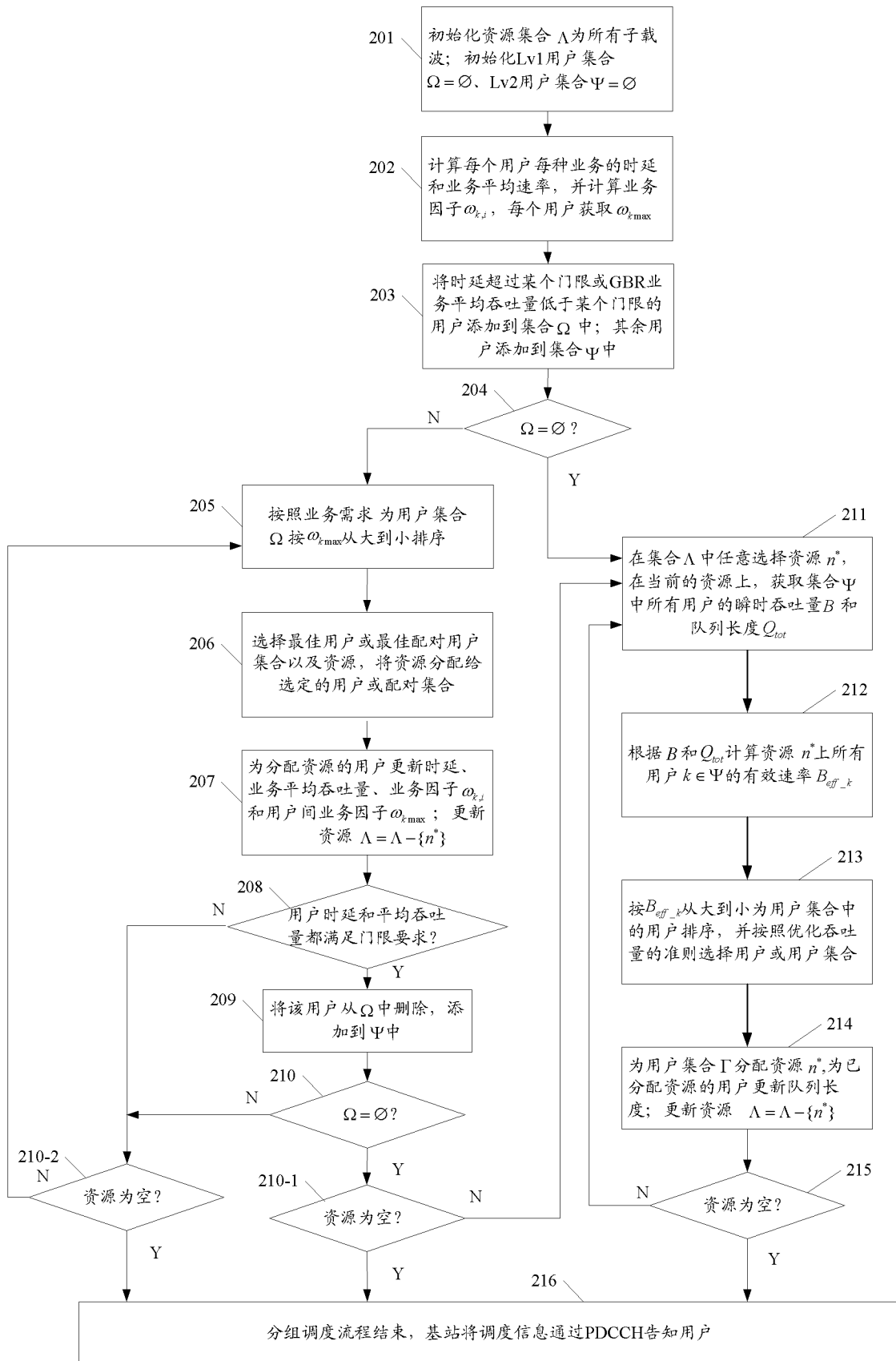


图 2

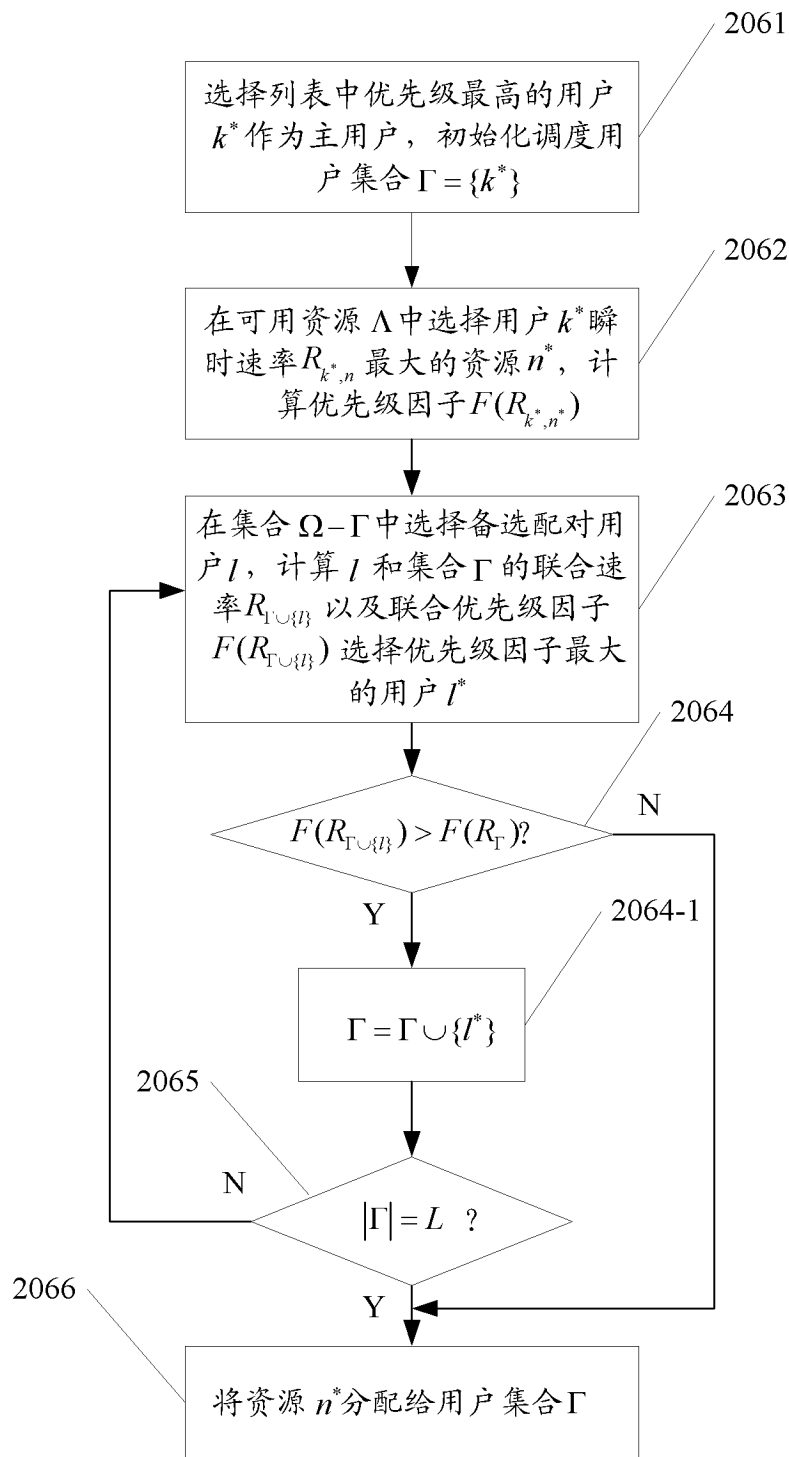


图 3



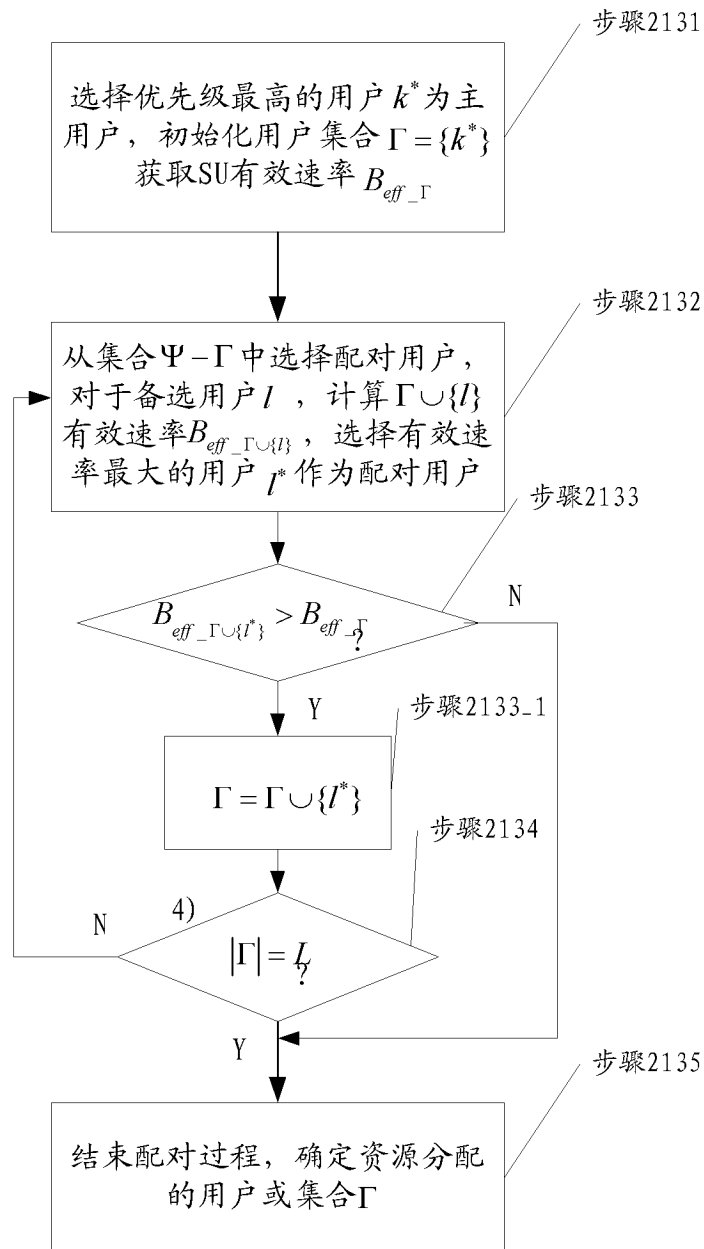


图 4