(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 102905307 B
(45)授权公告日 2014.12.31

(21)申请号 201210337833.8

(22)申请日 2012.09.12

(73)专利权人 北京邮电大学
地址 100876 北京市海淀区西土城路 10 号

(72)发明人 滕颖蕾 宋梅 王景尧 秦文聪 王莉 张勇 张羽 牟善文 邢益海

(74)专利代理机构 北京路浩知识产权代理有限公司 11002
代理人 王莹

(51)Int.Cl.
H04W 24/10(2009.01)
H04W 28/08(2009.01)
H04W 36/00(2009.01)
H04W 36/30(2009.01)

(56)对比文件
CN 102395157 A,2012.03.28,
CN 102098712 A,2011.06.15,
US 2003153315 A1,2003.08.14,
CN 102256307 A,2011.11.23,

审查员 靳莉

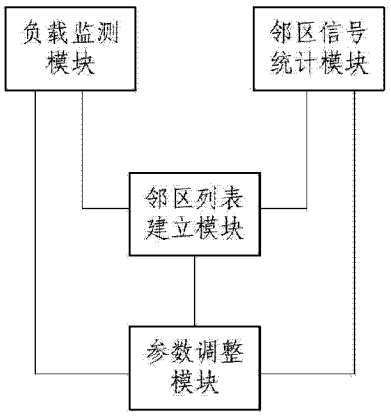
权利要求书2页 说明书7页 附图3页

(54)发明名称

实现邻区列表和负载均衡联合优化的系统

(57)摘要

本发明公开了一种实现邻区列表和负载均衡联合优化的系统，所述系统包括：负载监测模块，用于对服务小区和邻区的负载状态进行周期性监测，并交互负载状态；邻区信号统计模块，用于统计服务小区中终端测量的邻区的信号质量；邻区列表建立模块，用于依据服务小区和邻区各自的负载状态及统计的信号质量建立邻区列表，并发送邻区列表给参数调整模块；参数调整模块，用于根据所述邻区列表动态调整切换参数，以进行切换，并将调整后的参数反馈给负载检测模块和邻区信号统计模块。本发明的邻区列表兼顾信号强度和小区负载，根据该邻区列表进行切换参数调整，并按切换参数对用户进行小区切换，相对于现有的切换方式更合理。



CN 102905307 B

1. 一种实现邻区列表和负载均衡联合优化的系统，其特征在于，所述系统包括：

负载监测模块，用于对服务小区和邻区的负载状态进行周期性监测，并交互负载状态；

邻区信号统计模块，用于统计服务小区中终端测量的邻区的信号质量；

邻区列表建立模块，用于依据服务小区和邻区各自的负载状态及统计的信号质量建立邻区列表，并发送邻区列表给参数调整模块；

参数调整模块，用于根据所述邻区列表动态调整切换参数，以进行切换，并将调整后的参数反馈给负载检测模块和邻区信号统计模块；所述切换参数为指 A3 事件触发条件中的小区偏置 H：

Mn>Mp+H

其中 Mn 是邻区信号强度；

Mp 是服务小区信号强度；

H 是小区偏置；

其中，所述负载监测模块中负载状态的计算公式为：

$$l_s(t)=\frac{1}{T}\sum_{k=0}^{n-1}l(t-T+k\cdot\frac{T}{n-1})$$

其中，$l_s(t)$ 为小区负载指示器在 (t — T, t) 的时间间隔内对负载 n 次抽样并求平均值，T 为监测周期；

其中，所述邻区列表建立模块建立邻区列表的过程为：

按照预设的负载优先级的大小对邻区做优先级排序；

计算邻区信号强度的优先级，按照信号强度优先级顺序对相同负载的邻区再排序；

其中，所述邻区信号强度优先级的计算方法为：

$$p_i=\frac{m_i}{\sum_j m_j}$$

其中，m_i 是邻区 i 的 A3 事件测量报告数量，m_j 是邻区 j 的 A3 事件报告数量；

其中，所述参数调整模块参数调整的过程为：

初始化 Q(s,a)，设定折扣因子 γ 和初始学习率 α，以及动作选择算法中的初始探索概率 ε，Q(s,a) 为强化学习函数；

获取当前状态 s，s ∈ S，系统状态空间 S：{vh，h，m，l，vl} vh，h，m，l，vl 表示负载由高到低的 5 个等级；

选择执行的动作 a，a ∈ A，可选动作集 A：{-Δ，-2Δ，+Δ，+2Δ}，Δ 是参数 H 的单位调整步长，以 Δ 为基本单位，根据此状态的动作函数值 $Q_t(s,a)$，采用 ε 贪婪算法，从动作集中选择动作 a 并执行，具体以概率 (1 — ε) 选择 $Q_t(s,a)$ 值最大的动作，而以探索概率 ε 选择其它任一个动作；

获取回报 r 和下一时刻的状态 s'，根据动作执行结果按照 $r(t)=(F^*-F)+\alpha(D^*-D)+\beta(L^*-L)$ 计算当前回报 r，其中：

F^* 是 (t — T, t) 时间内服务小区统计的异常切换次数占总切换次数的比例上限；

F 是 (t — T, t) 内服务小区统计的到目标邻区的异常切换次数占总切换次数的比例；

α、β 是相关系数，由运营商设定依据相互关系设定；

D^* 是（t — T，t）内服务小区统计的掉话次数上限；

D 是（t — T，t）内服务小区统计的掉话次数；

L^* 是服务小区高负载门限；

L 是服务小区当前负载；

找到下一状态的动作值函数最大值 $\max_{a'} Q_t(s',a')$，根据 $Q_{t+1}(s,a)=(1-\alpha_t)Q_t(s,a)+\alpha_t(r_t+\gamma\max_{a'}Q_t(s',a'))$ 更新 $Q_t(s,a)$，其中：α_t 是 t 时刻的学习率，是一个变步长的参数，且 $0<\alpha_t<1$；γ 为折扣因子，且 $0\leqslant\gamma\leqslant1$，$\max_{a'} Q_t(s',a')$ 中的 a′ 是系统在 t 时刻处于 s′ 状态时，对应的所有行为中 Q 值最大的那个动作；

每轮迭代结束后更新学习率和贪婪算法中的探索概率 ε，使学习率和探索概率以负指数规律随着学习的过程逐渐减少为 0。

2. 如权利要求 1 所述的实现邻区列表和负载均衡联合优化的系统，其特征在于，所述服务小区和邻区通过 X2 接口交互负载信息。

实现邻区列表和负载均衡联合优化的系统

技术领域

[0001] 本发明涉及LTE网络下的SON(Self-Organizing Network)技术，具体的涉及一种实现邻区列表和负载均衡联合优化的系统。

背景技术

[0002] SON(Self-Organizing Network)是在LTE的网络的标准化阶段由移动运营商主导提出的概念，其主要思路是实现无线网络的一些自主功能，减少人工参与，降低运营成本。

[0003] NGMN中的移动运营商对SON的部署有强烈的需求，于是纷纷投入SON需求的研究，发布有关SON的白皮书和建议书。3GPP也在重点研究SON和当前电信管理网络的实现方案。欧盟也在进行两个相关项目，一个主要由欧洲主要运营商、设备商共同承担，从SON的技术方案、实现方法及验证平台入手，研究SON对网络运维产生的影响；另一项目是利用感知无线电和分布式感知原理进行前沿重点研究。

[0004] MLB(Mobility Load Balancing)处理的根本目的在于通过调整切换参数，以用户切换的方式把过载小区中适当比例的负载转移到不过载的小区中，也就是把两个小区中的用户终端从过载的小区切换到尚未过载的小区。要实现MLB，负载的获取必不可少，与通用移动通信系统(UMTS)相比，LTE系统的最大特点是采用了更扁平化的架构，取消了用户终端和无线移动实体/服务网关(MME/S-GW)之间的中间控制节点(即UMTS中的无线网络控制节点)，因此相比S1接口上的均衡处理而言，X2接口上的小区负载信息交换为负载获取提供了方便。

[0005] 由于MLB最终是通过切换实现，而切换又与邻区列表密切相关。在蜂窝式移动通信网络中，每个小区都有一张邻区列表，该表记录了与当前小区相关的邻区，它决定了移动终端搜索的范围和切换的方向。正确而完整的邻区关系列表非常重要，邻区关系做的太少，会出现邻区漏配的现象，这会直接导致大量的掉话；邻区关系做的太多，则不仅会导致测量报告的精确性降低而且会大大提高运营商的运营成本。传统邻区列表的建立都是基于邻区信号强度的大小，并没有考虑邻区负载的状况。

[0006] 强化学习(Reinforcement Learning,RL)可以从延迟的回报中获取最优的控制策略。一个可学习的智能体，它可以观察环境的状态并能做出一组动作改变这些状态，学习的任务是获得一个控制策略，以选择能达到目的的行为。Q-学习算法是由Watkins在1989年提出的类似于动态规划算法的一种强化学习方法，它提供智能系统在马尔科夫环境中利用经历的动作序列选择最优动作的一种学习能力，并且不需要建立环境模型。

发明内容

[0007] (一)要解决的技术问题

[0008] 本发明为了解决和实现过载小区和非过载小区之间用户切换的问题，本发明提出了一种实现邻区列表和负载均衡联合优化的系统。

[0009] (二)技术方案

[0010] 为解决上述技术问题，本发明提供了一种实现邻区列表和负载均衡联合优化的系统，所述系统包括：

[0011] 负载监测模块，用于对服务小区和邻区的负载状态进行周期性监测，并交互负载状态；

[0012] 邻区信号统计模块，用于统计服务小区中终端测量的邻区的信号质量；

[0013] 邻区列表建立模块，用于依据服务小区和邻区各自的负载状态及统计的信号质量建立邻区列表，并发送邻区列表给参数调整模块；

[0014] 参数调整模块，用于根据所述邻区列表动态调整切换参数，以进行切换，并将调整后的参数反馈给负载检测模块和邻区信号统计模块。

[0015] 其中，所述负载监测模块中负载状态的计算公式为：

[0016] $$l_s(t)=\frac{1}{T}\sum_{k=0}^{n-1}l(t-T+k\cdot\frac{T}{n-1})$$

[0017] 其中，$l_s(t)$ 为小区负载指示器在（t — T，t）的时间间隔内对负载 n 次抽样并求平均值，T 为监测周期。

[0018] 其中，所述服务小区和邻区通过 X2 接口交互负载信息。

[0019] 其中，所述邻区列表建立模块建立邻区列表的过程为：

[0020] 按照预设的负载优先级的大小对邻区做优先级排序；

[0021] 计算邻区信号强度的优先级，按照信号强度优先级顺序对相同负载的邻区再排序。

[0022] 其中，所述邻区信号强度优先级的计算方法为：

[0023] $$p_i=\frac{m_i}{\sum_j m_j}$$

[0024] 其中，m_i 是邻区 i 的 A3 事件测量报告数量，m_j 是邻区 j 的 A3 事件报告数量。

[0025] 其中，所述切换参数为指 A3 事件触发条件中的小区偏置 H：

[0026] Mn>Mp+H

[0027] 其中 Mn 是邻区信号强度；

[0028] Mp 是服务小区信号强度；

[0029] H 是小区偏置。

[0030] 其中，所述参数调整模块参数调整的过程为：

[0031] 初始化 Q(s,a)，设定折扣因子 γ 和初始学习率 α，以及动作选择算法中的初始探索概率 ε，Q(s,a) 为强化学习函数；

[0032] 获取当前状态 s，s ∈ S，系统状态空间 S：{vh，h，m，l，vl} vh，h，m，l，vl 表示负载由高到低的 5 个等级；

[0033] 选择执行的动作 a，a ∈ A，可选动作集 A：{-Δ，-2Δ，+Δ，+2Δ}，Δ 是参数 H 的单位调整步长，以 Δ 为基本单位，根据此状态的动作函数值 $Q_t(s,a)$，采用 ε 贪婪算法，从动作集中选择动作 a 并执行，具体以概率（1 — ε）选择 $Q_t(s,a)$ 值最大的动作，而以探索概率 ε 选择其它任一个动作，这也是保证了每个状态动作对都能够遍历到；

[0034] 获取回报 r 和下一时刻的状态 s′，根据动作执行结果按照 $r(t)=(F^*-F)+\alpha(D^*-D)+\beta(L^*-L)$ 计算当前回报 r，其中：

[0035] F^* 是(t — T,t)时间内服务小区统计的异常切换次数占总切换次数的比例上限；

[0036] F 是(t — T,t)内服务小区统计的到目标邻区的异常切换次数占总切换次数的比例；

[0037] α、β 是相关系数,由运营商设定依据相互关系设定；

[0038] D^* 是(t — T,t)内服务小区统计的掉话次数上限；

[0039] D 是(t — T,t)内服务小区统计的掉话次数；

[0040] L^* 是服务小区高负载门限；

[0041] L 是服务小区当前负载；

[0042] 找到下一状态的动作值函数最大值 $\max_{a'} Q_t(s',a')$，根据 $Q_{t+1}(s,a)=(1-\alpha_t)Q_t(s,a)+\alpha_t(r_t+\gamma\max_{a'}Q_t(s',a'))$ 更新 $Q_t(s,a)$,其中：

[0043] α_t 是 t 时刻的学习率,是一个变步长的参数,且 $0<\alpha_t<1$；γ 为折扣因子,且 $0\leqslant\gamma\leqslant1$,$\max_{a'} Q_t(s',a')$ 中的 a′ 是系统在 t 时刻处于 s′ 状态时,对应的所有行为中 Q 值最大的那个动作；

[0044] 每轮迭代结束后更新学习率和贪婪算法中的探索概率 ε,使学习率和探索概率以负指数规律随着学习的过程逐渐减少为 0。

[0045] (三)有益效果

[0046] 本发明的邻区列表兼顾信号强度和小区负载,根据该邻区列表进行切换参数调整,并按切换参数对用户进行小区切换,相对于现有的切换方式更合理；采用 Q 学习方法来实现对切换参数的自动控制,在不同的小区负载状态下,选择回报最大的参数调整行为,以提高系统性能。

附图说明

[0047] 图 1 本发明的实现邻区列表和负载均衡联合优化的系统结构示意图；

[0048] 图 2 切换参数示意图；

[0049] 图 3-1 负载均衡示意图,负载均衡前；

[0050] 图 3-2 负载均衡示意图,负载均衡后。

具体实施方式

[0051] 下面结合附图和实施例,对本发明的具体实施方式作进一步详细描述。以下实施例用于说明本发明,但不用来限制本发明的范围。

[0052] 本实施例的基于邻区列表的负载均衡自优化系统结构示意图如图 1 所示,包括：

[0053] 负载监测模块：

[0054] 该模块用于服务小区和邻区的负载检测,标准 ts36.300 中定义了负载指示器,因此从负载指示器中获取小区的负载状态,在整数时间周期 T 内,小区负载根据下式计算：

[0055] $$l_s(t)=\frac{1}{T}\sum_{k=0}^{n-1} l\left(t-T+k\cdot\frac{T}{n-1}\right) \qquad (1)$$

[0056] 其中,$l_s(t)$ 为小区负载指示器在(t — T,t)的时间间隔内对负载 n 次抽样并求平均值,函数 l 的具体表达形式取决于负载指示器。

[0057] 根据实际情况设置不同的门限值，将小区负载分为低、中、高、过载四个等级，小区基站开始工作后，周期性检测自身负载状态，同时，邻区也在周期性监测自身负载，当服务小区需要邻区的负载信息时，通过 X2 接口发送负载请求消息，接收到请求消息后，邻区反馈当前自身负载状态给请求小区。

[0058] 邻区信号质量统计模块：

[0059] LTE 中的切换由终端（UE）的测量报告触发，在 LTE 小区之间，通常采用 A3 事件的报告来触发切换，服务小区用户（用户手机终端）周期性测量邻区信号，如果信号质量满足 A3 事件，A3 事件是邻区的服务质量（RSRP/RSRQ）比服务小区高一个绝对门限，触发 A3 事件切换则向基站发送测量报告，基站将报告数量进行统计。

[0060] A3 事件具体公式为：

[0061] Mn+Ofn+Ocn-Hys>Mp+Ofp+Ocp+Off (2)

[0062] Mn：邻区的测量结果；Ofn：邻区频率的特定频率偏置；Ocn 邻区的特定小区偏置；Ms：服务小区的测量结果；Ofs：服务小区的特定频率偏置；Ocp：服务小区的特定小区偏置；Hys：A3 事件迟滞；Off：A3 事件偏置。

[0063] 为了便于描述，上式简化为：

[0064] Mn>Mp+H (3)

[0065] Mn 是邻区信号强度；

[0066] Mp 是服务小区信号强度；

[0067] H 是偏置。

[0068] 邻区列表建立模块：

[0069] 为了提高用户测量邻区信号的效率，基站维护了一张邻区列表，它周期性下发给用户，用户按照邻区列表中给出的信息（频点等）监测邻区的信号强度，这样用户只要监测几个列表中的小区，不需要全频段的监测，提高了测量的效率。当满足 A3 事件后，用户触发上报，当前服务基站决定是否发起切换。

[0070] 本发明中的邻区列表的建立基于优先级，主要考虑两个原则：

[0071] 1、边缘用户测量邻区信号，如果满足 A3 事件条件，则向基站发送测量报告，越多边缘用户上报的邻区，优先级越高。

[0072] 2、小区的负载状态。因此，该模块需要邻区信号质量统计模块和负载监测模块分别提供邻区信号和小区负载信息。

[0073] 首先根据以下公式计算一定时间周期内的邻区 i 的优先级：

[0074] $$p_i = \frac{m_i}{\sum_j m_j} \quad (4)$$

[0075] m_i 是邻区 i 的 A3 事件测量报告数量，m_j 是邻区 j 的 A3 事件报告数量，分母 $\sum_j m_j$ 是所有邻区的 A3 事件报告数量。m_i 越大，即 p_i 越大，意味着越多的用户测量到小区 i 的信号满足 A3 事件，那么通过调整参数能够切换到邻区 i 的用户数也越多，由于在邻区列表里排位越靠前的小区，被用户选择作为切换小区的优先级越高，因此按照 p_i 排序的邻区列表，进行负载均衡切换时的效果也越明显。

[0076] 与此同时，服务小区通过 X2 接口向列表中的邻区发送获取邻区负载的请求，邻区

把负载状态通过 X2 接口反馈给服务小区。

[0077] 邻区列表的建立依照以下步骤：

[0078] 1、按照负载的优先级由低到高的顺序对所有邻区进行排序；

[0079] 2、根据 p_i 大小对有相同负载状态的邻区重新排序，从而获得新的邻区列表。

[0080] 依照以上步骤生成最终的邻区列表，邻区列表按照以上原则，定期更新。

[0081] 参数调整模块：

[0082] 参数调整是对触发切换的 A3 事件公式中的偏置参数进行调整，本发明将经典的强化学习算法 Q 学习（Q-learning）引入控制过程，该算法可以通过跟环境的交互，学习如何控制参数。每次参数调整后，服务小区和邻区的负载状态会发生改变，变化会反馈给负载监测模块和邻区信号质量统计模块，作为更新邻区列表的依据。如果某个邻区转变为高负载，那么该邻区在邻区列表中的排名就会靠后，甚至移出列表，相应的边缘用户就很少或者不能切换到该邻区。

[0083] 参数调整开始后，UE 会根据系统配置的测量事件进行测量上报。在邻区信号质量统计模块已经提到，LTE 中的切换由 UE 的测量报告触发。在 LTE 小区之间，通常采用 A3 事件来触发切换，具体采用公式（2）来触发切换，简化式为公式（3）。

[0084] 其中公式（3）中 H 的影响范围是小区级，即 H 的改变会影响到小区内所有用户的切换判定，而且该值与邻小区无关。

[0085] 本发明是对 H 进行参数调整，由于 H 的调整是全局性的，为了更好的说明负载均衡的过程，以一个邻区为例进行说明。

[0086] 在图 2 中，H1 为小区 1 的切换偏置参数，用于触发小区 1 向小区 2 切换的 A3 事件为：

[0087] $M_2>M_1+H$ (5)

[0088] M_1，M_2 分别是终端测量到的小区 1 和小区 2 的信号质量，假设小区 1 到达高负载或者过负载状态进行负载均衡操作。如图 2 所示，小区 1 将增加 H_1 的值至 H_1'，将小区 1 向小区 2 的切换触发门限由点 A 调整至点 A′，从而减小切换触发门限值，让小区 1 的偏向小区 2 一侧的边缘用户，更容易切换至小区 2，以此来达到卸载服务小区负载的目的，当服务小区为低负载时，参数调整方向为使边缘用户更难切换到邻区，如果有高负载的邻区可以选择当前服务小区作为卸载的小区之一。附图 3-1 和 3-2 描述了负载均衡前后的变化，如图 3-1 所示负载均衡前，当前服务小区（图中位于中间的小区）处于高负载状态，边缘用户不断检测到邻区的信号强度，并上报给服务小区，图 3-2 表示负载均衡后，很多边缘用户切换到了适当的邻区，缓解了当前服务小区的负荷。一部分边缘用户切换到了邻区经过调整之后，用于触发小区 1 向小区 2 切换的 A3 事件改为：

[0089] $M_2>M_1+H_1'$ (6)

[0090] 如果调大 H_1 超过一定的限度，会引发异常切换次数增加，比如过早切换或乒乓切换，如果 H_1 调整过小，会导致小区 1 的边缘用户切换困难，可能发生过晚切换，甚至掉话。因此，参数调整需要控制在一定的范围内。参数调整过程采用强化学习的一种典型方法 Q－学习，它提供智能系统在马尔科夫环境中利用经历的动作序列选择最优动作的一种学习能力，并且不需要建立环境模型。

[0091] Q 学习是强化学习中最典型的一个算法。Q 函数 Q(s,a) 表示在状态 s 下执行动作

a，以及采取后续策略的折扣奖赏和的期望。Q 值函数的学习是通过 Q 值迭代来完成的。当 Q 值函数经过多次迭代后，所有的 Q 值都不再发生较大的变化时，即可认为 Q 值函数收敛，Q 学习结束。它在每一轮的迭代中，首先感知当前的环境状态 $s \in S$，并查找相应的所有 Q 值，根据当前的策略 λ 选择动作 $a \in A$ 作用于环境，A 是动作集合；环境状态会由此变化为 $s' \in S$，S 是状态空间集合，同时根据所执行动作的效果获得一个强化信号（称为“回报”）r(s,a)；学习者便据此按照式 (7) 更新其策略，并进入下一轮迭代。

[0092] $$Q_{t+1}(s,a)=(1-\alpha_t)Q_t(s,a)+\alpha_t(r_t+\gamma\max_{a'}Q_t(s',a')) \qquad (7)$$

[0093] 其中，α_t 是 t 时刻的学习率，是一个变步长的参数，它决定 Q 函数更新的速度，且 $0<\alpha_t<1$。当 α_t 接近 1 时，回报将随新经验值变化更加明显，即收敛越快，但是过大的 α_t 将导致不成熟收敛。折扣因子 γ 决定未来回报对当前的影响，且 $0 \leqslant \gamma \leqslant 1$。当 γ 越接近 1，未来行为预测在整个效用函数中将起更重要的作用，$\max_{a'}Q_t(s',a')$ 中的 a′ 是系统在 t 时刻处于 s′ 状态时，对应的所有行为中 Q 值最大的那个动作。随着 $t \to \infty$，若每对 (s,a) 的 Q 值能够经历无穷多次更新，且 α_t 递减至 0，则 $Q_t(s,a)$ 将以概率 1 收敛到最优值 $Q^*_t(s,a)$。如此循环下去，通过不断地“试错”学习最终目标是找到每个状态的最佳动作选择策略 $\lambda^*(s) \in A$ 以最大化期望的长期累积回报。此时，最优策略 λ^* 可以由式 (8) 得到：

[0094] $$\lambda^*(s)=\arg\max_{a}Q^*_t(s,a) \qquad (8)$$

[0095] Q 学习中式 (8) 的收敛性并不依赖于动作空间的探索方法。为了使所有的状态动作对 $Q_t(s,a)$ 都被访问到，本文采用 ε 贪婪算法来选择动作。具体地，ε 贪婪算法以概率 (1 — ε) 选择 $Q_t(s,a)$ 值最大的动作，而以探索概率 ε 选择其它任一个动作，这也保证了每个状态动作对都能够遍历到。

[0096] 问题映射如下：

[0097] (1) 状态空间

[0098] 系统状态为服务小区的负载状态，根据当前小区服务的用户数，分为四个等级 vh（很高）、h（高）、m（中）、l（低）、vl（很低），所以状态空间为：

[0099] S：{vh，h，m，l，vl}

[0100] (2) 可选动作集 A{-Δ，-2Δ，+Δ，+2Δ}

[0101] Δ 是参数 H 的单位调整步长，以 Δ 为基本单位，设定四个调整值。

[0102] (3) 回报函数

[0103] 调整 H 值要限定在一定的范围内，如果调整过大，容易引发乒乓切换和过早切换等异常切换，同时如果调整过小，服务小区边缘用户很难切换到邻区，掉话率会升高，因此设置异常切换次数和掉话次数两个量纲，

[0104] $r(t)=(F^*-F)+\alpha(D^*-D)+\beta(L^*-L)$ (9)

[0105] 其中：

[0106] F^* 是（t-T，t）时间内服务小区统计的异常切换次数占总切换次数的比例上限；

[0107] F 是（t-T，t）时间内服务小区统计的到目标邻区的异常切换次数（包括过早切换和乒乓切换）占总切换次数的比例；

[0108] α、β 是相关系数，由运营商设定依据相互关系设定；

[0109] D^* 是（t-T，t）时间内服务小区统计的掉话次数上限；

[0110] D 是(t-T,t)时间内服务小区统计的掉话次数；

[0111] L^* 是服务小区高负载门限；

[0112] L 是服务小区当前负载；

[0113] 算法实现过程：

[0114] 1) 初始化 Q(s,a),比如是随机产生的值,设定折扣因子 γ 和出示学习率 α,以及动作选择算法中的初始探索概率 ε ；

[0115] 2) 获取当前状态 s,

[0116] 3) 选择执行的动作 a,根据此状态的动作函数值 $Q_t(s,a)$,按照一定的策略 λ 选择动作 a 并执行；

[0117] 4) 获取回报(奖赏)r 和下一时刻的状态 s′ ,根据动作执行结果按照式(9)计算当前回报 r,并找到下一状态的动作值函数最大值 $\max_{a'} Q_t(s',a')$,根据式(7)更新 $Q_t(s,a)$；

[0118] 5) 参数更新,每轮迭代结束后,学习率和探索概率都要更新,为了满足 Q 学习的收敛性要求,本发明设置它们以负指数规律随着学习的过程逐渐减少为 0。

[0119] 以上实施方式仅用于说明本发明,而并非对本发明的限制,有关技术领域的普通技术人员,在不脱离本发明的精神和范围的情况下,还可以做出各种变化和变型,因此所有等同的技术方案也属于本发明的范畴,本发明的专利保护范围应由权利要求限定。

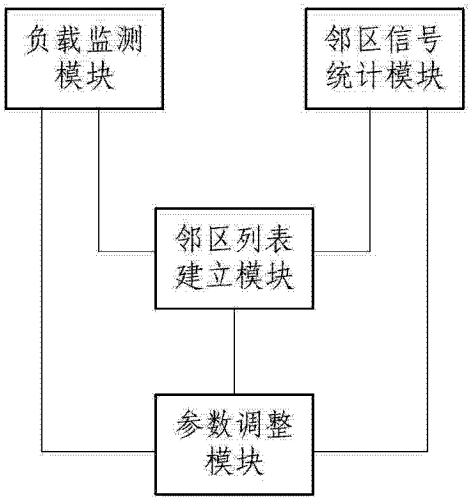

负载监测模块
邻区信号统计模块
邻区列表建立模块
参数调整模块

图 1

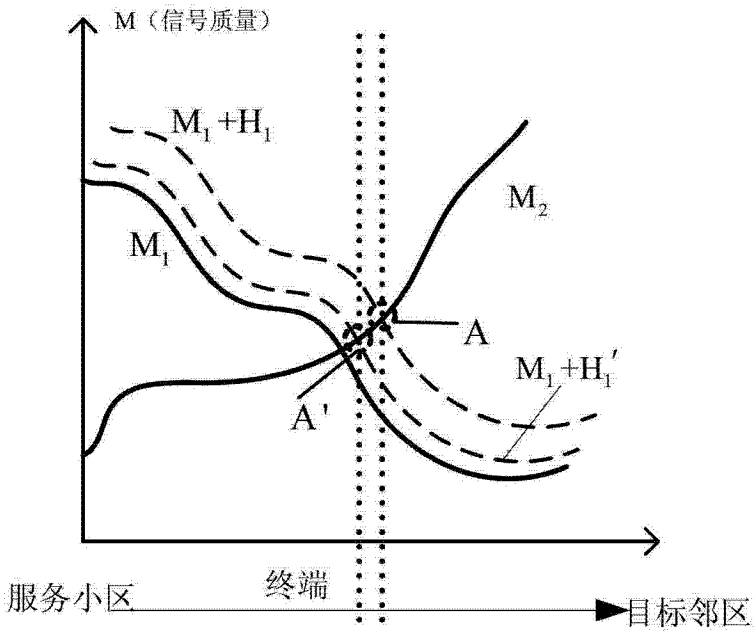

M（信号质量）
M1+H1
M1
M2
A
M1+H1′
A'
服务小区
终端
目标邻区

图 2

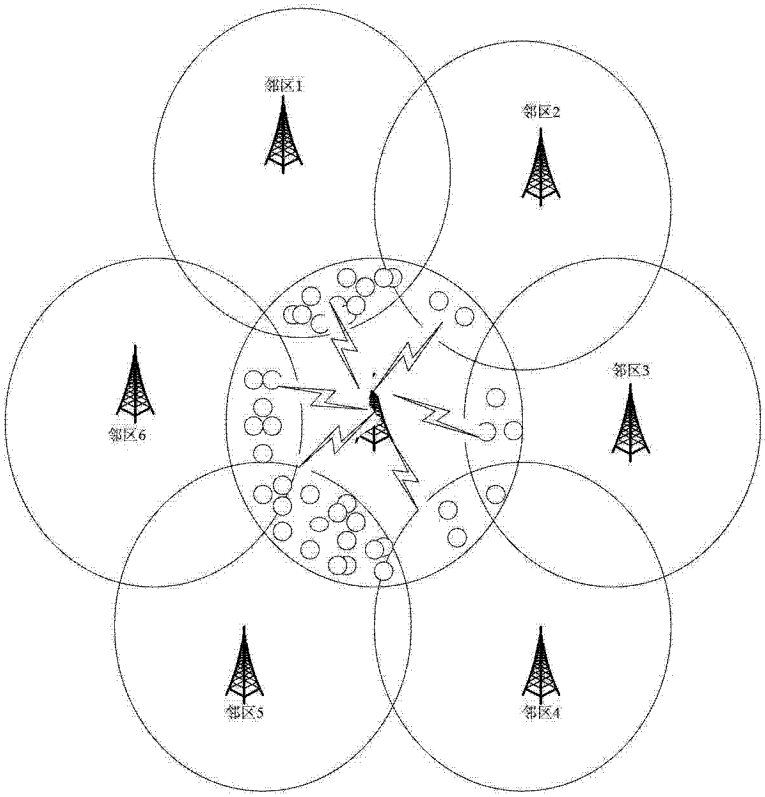


图 3-1

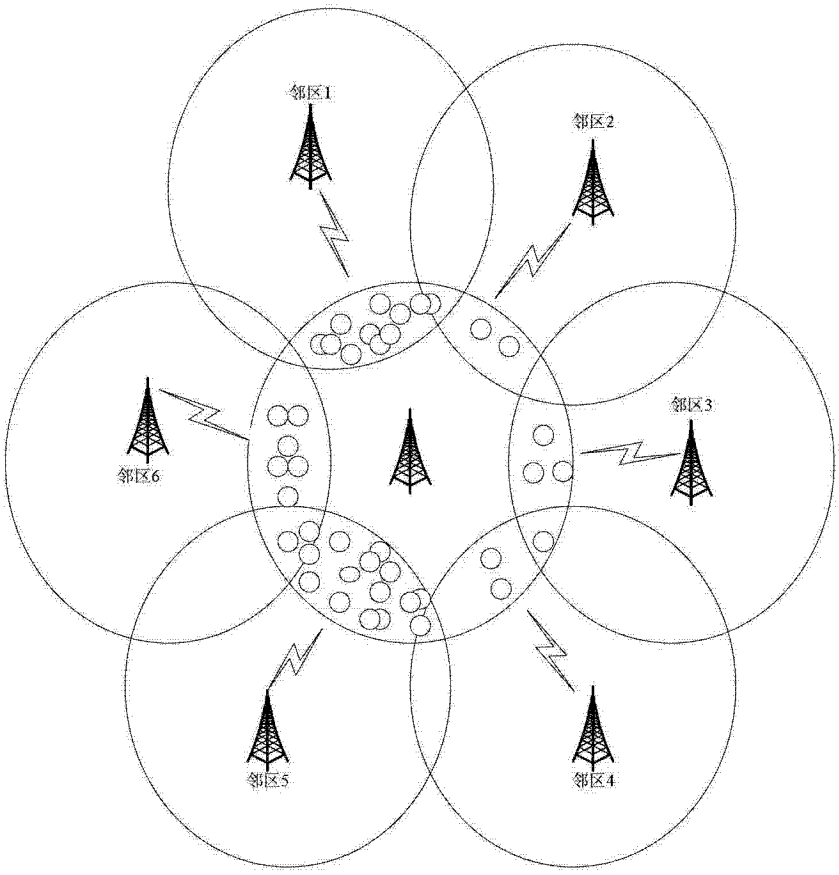


图 3-2