



LTE-Advanced 异构网络中负载均衡研究进展及展望 *

许娟雄, 刘伊莎, 唐 伦, 陈前斌

(重庆邮电大学移动通信技术重点实验室 重庆 400065)

摘 要: 新一代移动通信网络不断演进为高密度、无定型的异构网络, 业务需求与资源供给的匹配问题, 即负载均衡问题是异构网络(heterogeneous network, HetNet)资源管理的挑战问题之一。阐述了负载均衡的一般性问题, 根据负载均衡算法的主要特征给出了综合分类, 按照分类对近年来提出的主要负载均衡策略进行了细致和深入的对比分析; 同时也对比分析了几种减少负载均衡过程中信令开销的方案。在总结现有研究成果的基础上, 进一步分析负载均衡技术的最新发展趋势, 为下一步的研究提出了新的问题和思路。

关键词: LTE-Advanced; 异构网络; 负载均衡

doi: 10.3969/j.issn.1000-0801.2014.10.018

Survey on Research Progress of Load Balancing in LTE-Advanced Heterogeneous Network

Xu Juanxiong, Liu Yisha, Tang Lun, Chen Qianbin

(Key Lab of Mobile Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: With next-generation mobile communication networks evolving to high dense, amorphous and heterogeneous network, matching the demand for business with the supply of resources has been one of the challenging problems of resource management in heterogeneous network (HetNet), namely load balancing problem. The general load balancing problems are formulated. Following that, a comprehensive taxonomic approach to classifying algorithms of load balancing was proposed, based on which detailed analyses and thorough comparisons for the main strategies in recent years were made; and several schemes to reduce signaling overhead were compared and analyzed in the load balancing process. Finally, research results were summarized, and some new issues in load balancing strategies conforming to the trends to be further studied were pointed out.

Key words: LTE-Advanced, heterogeneous network, load balancing

1 引言

在 LTE-Advanced (以下简称 LTE-A) HetNet^[1-3] 中, 为了增强网络的灵活性与自适应性, 提升网络室内以及热点场景的容量与覆盖, 引入了大量的低功率节点的多层蜂窝网络, 在 LTE-A 系统中主要支持的低功率节点包括远程

射频头 (remote radio head, RRH)、微微基站 (pico eNB, Pico)、毫微微基站 (femto eNB, Femto) 以及中继站 (relay node, RN), LTE-A 系统中异构蜂窝网络的拓扑结构如图 1 所示。由于用户终端接入网络的随机性、时变性, 业务需求类型的差异性, 使得网络中负载分配不均匀, 即某些小区中接入了大量的用户业务, 出现过载的情况, 造成节点拥

* 国家科技重大专项基金资助项目 (No.2014ZX03003010-004), 国家高技术研究发展计划 (“863”计划) 基金资助项目 (No.2014AA01A701)

塞;而其他一些小区中业务需求较少甚至处于空闲状态,使得无线资源未得到充分利用,资源利用率极低。为了解决资源与业务需求的匹配问题,引入了负载均衡(load balancing, LB)技术^[4],它是提高系统的资源利用率和系统容量的关键技术,是一种网络的自优化技术,是无线资源管理的关键技术之一。

由于 LTE-A HetNet 的节点类型的多样化和节点数目的增多,业务需求增大,负载均衡信令复杂度明显提高。如何将超载节点的用户积极“推”到低负载的低功率节点上进行服务,充分利用网络中各个节点的资源;如何感知各个节点的负载信息变化;如何提升用户的切换成功率,自适应调整网络节点的参数;提升网络节点的智能性与协作性,将有限的网络资源与用户的业务需求进行优化匹配,是当前研究的热点。

2 负载均衡概述

负载均衡,又称为负载分担,是指将负载在宏蜂窝、微微基站、毫微微基站、中继站及无线局域网络间进行平衡,以提升整个网络的性能。

随着日益增大的网络业务量的需求,有限的网络资源仍然是限制系统性能的主要因素,即使有针对性地将小基站部署在高流量的区域,大多数用户仍然能够接收来自宏蜂窝的较强的下行信号,因而异构蜂窝网络中存在的跨层干扰^[5]同样是影响系统性能的又一个重要因素。

基于负载均衡的基本定义以及上述分析的两大限制网络性能的因素,LTE-A 异构网络中的负载均衡技术应运而生,其主体思想是将网络中高负载节点下的业务流量分摊到低负载节点上进行服务,缓减网络拥塞效应,将有限的网络资源与分配不均的用户业务需求进行优化匹配,从而提升网络的资源利用率,降低用户的呼叫阻塞率,提高网络的吞吐量,充分利用网络中部署的低功耗基础设备,最终实现全网的负载均衡。

3 负载均衡算法分类

负载均衡技术是网络性能优化的关键技术之一,核心是负载均衡算法,本节主要是对目前在 LTE-A 异构网络系统中研究较热的负载均衡算法进行分类研究,大体分为 3 类算法:一类是基于触发负载均衡操作因子算法;另一类是从全网的综合性能考虑的基于效用函数的负载均衡算法;还有就是在传统的负载均衡算法基础上、目前研究最热的、基于小区范围扩张的负载均衡算法,该类算法的实用性与稳定性都较高。上述的 3 类算法之间相互关联,相互依赖,如图 2 所示。下文主要介绍这 3 类负载均衡算法。

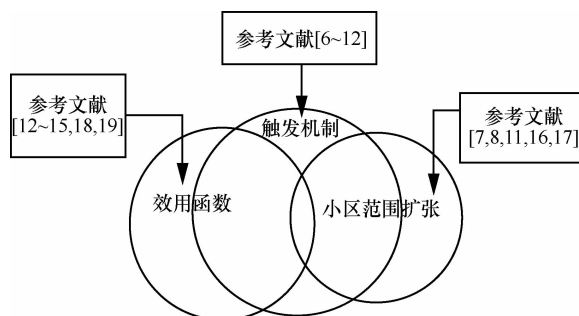


图2 负载均衡算法分类

3.1 基于触发机制的 LB 算法

如果网络的负载分配比较均匀或者网络中接入节点的负载值较低,负载均衡机制将不会被触发,当网络流量出现拥塞状况、接入某些节点的用户信道质量差(即需要网络状态达到某种条件时)或者执行网络运营策略时,就会执行负载均衡操作,因而负载均衡操作需要触发事件。根据现有的负载均衡策略触发因子来看,可将负载均衡算法分为以下 3 类。

3.1.1 基于负载值触发的 LB 算法^[6,7,9,10]

该类算法直接使用各个节点的负载信息作为触发负载均衡操作的因子。但是在衡量各个节点的负载信息时,采用不同的衡量指标,对于全业务模型的用户,即只要该

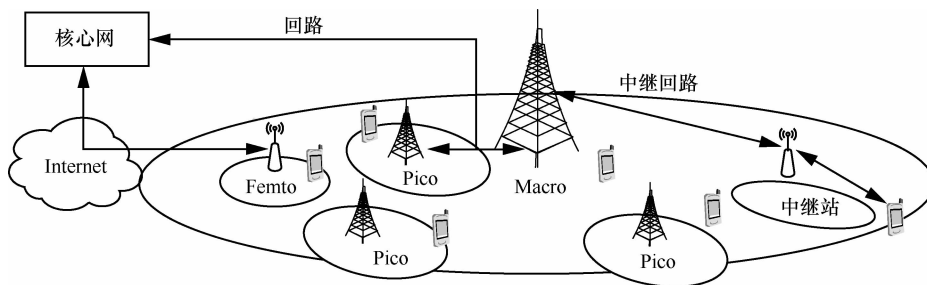


图1 3GPP HetNet 拓扑结构示意图



用户与其服务基站连接,那么它就会一直传输数据,直到它与服务基站断开为止,此时,可以直接用连接到各个基站的 用户数量作为衡量负载信息的指标,即:

$$load_i(t) = N_i^{\text{user}}(t) \quad (1)$$

其中, i 表示基站 ID, $load_i(t)$ 表示当前时刻 t 基站 i 的负载值, $N_i^{\text{user}}(t)$ 表示当前时刻连接到基站 i 的用户的数目。上述考虑负载信息是最简单的一种情况,参考文献[9]中就是采用上述方法。但是基于现实工程中,由于各个用户的业务模式不同,不可能全部都是全业务模式,因而在衡量负载信息时,不能采用式(1)的方法。根据参考文献[10]可知,将网络中各个节点的物理资源块(physical resource block, PRB)的占用率作为衡量负载信息指标,即:

$$load_i(t) = \frac{\sum_{u=1|X(u)=i}^{U_i(t)} N_{\text{PRB}}^u}{N_{\text{PRB}}^{\text{total}}} \quad (2)$$

其中, $U_i(t)$ 表示当前时刻连接到小区 i 的总的用户数目, $X(u)$ 表示用户 u 所述小区 ID, $N_{\text{PRB}}^{\text{total}}$ 表示小区 i 总的资源块数目。

当且仅当上述 $load_i(t) > Load_i^0$ 时,则执行负载均衡策略, $Load_i^0$ 表示预先设定的负载门限值,超过它则表示该基站处于高负载状态。

3.1.2 基于负载差值触发的 LB 算法^[8,11,12]

该算法考虑的是各个节点与相邻节点之间的负载差值作为触发负责均衡策略的事件。它也是基于负载值触发的 LB 算法的改进算法,相比于第 3.1.1 节所述的 LB 算法,该算法既考虑了各个节点当前的负载值,同时还考虑了与之相邻各个节点的负载情况,通过与其相比较,可以很清楚地看出各个节点的负载情况,即不仅仅考虑单个节点的负载值,要更加具有全局性地优化网络的无线资源,能够使有限的资源与网络中用户的需求进行匹配,达到更好的负载均分的效果。同理可得该算法的核心计算式为:

$$load_{\text{diff}}^i(t) > \Delta load_i \quad (3)$$

其中, $load_{\text{diff}}^i(t)$ 表示当前时刻 t 节点 i 与其相邻节点的负载差值, $\Delta load_i$ 表示预先设定的差值门限值,只有满足式(3)条件时,才触发负载均衡算法。

基于差值触发的负载均衡得到的负载均分效果较好,因为该类算法不仅仅考虑当前本身的信息,还要与其他相邻节点的信息相比较,这样可以避免由于网络环境变化或者一些其他因素导致某个节点当前的信息不正确使得负

载均衡策略失效,也正是因为考虑的参考信息较多,使得各个节点之间需要进行信息交互,带来的信令开销较大,会产生一定的时延。

3.2 基于效用函数的 LB 算法

一般而言,基于效用函数的负载均衡算法比上述单纯基于触发因子的负载均衡算法的性能要好,因为该算法不仅要考虑网络中各个节点的负载值是实时变化的,同时还需要考虑网络中的各种限制条件,比如小区选择与切换以及网络的各种性能指标,该算法会对所选择的指标的增益进行权衡,使得算法中选择的效用函数的值最大化。但是该类算法同样存在着不可避免的缺点,即算法的复杂度较大,而且负载均衡过程中产生的开销也大。下面主要介绍几类比较常见的基于效用函数的 LB 算法。

3.2.1 基于用户关联的 LB 算法^[18,19]

基于用户关联控制的 LB 算法是一种次优化方法,它能主动将用户“推”到低功率基站上进行服务,在这种算法中,超载节点上的用户可以通过使用关联因子将其关联到其他小蜂窝上。在参考文献[18]中主要采用的就是比例公平的用户关联算法实现负载均衡,其目标就是找到用户最优的长期关联,在该文献所述的算法中,每个用户只能与网络中的一个节点相关联,当最初与其相关联的节点的负载较大时,用户就会采用相应的关联算法重新选择目标节点接入,从而实现负载均衡。

基于上述的原理分析,该类算法的核心就是寻找用户与基站的关联式,一般采用以下两种关联方法。

(1) 最大 SINR(max-SINR)关联

该算法的基本思想就是在预测出没有超负载的基站集合之后,计算各个基站可以给需要进行切换的用户提供的信干噪比(signal to interference plus noise ratio, SINR)值,选出提供的 SINR 值最大的基站作为新用户的最佳接入基站。具体计算式为:

$$j^* = \max_{j \in S_i} SINR_j^i \quad (4)$$

其中, i 表示切换用户的 ID, j 表示低负载基站的 ID, j^* 表示选择的目标基站的 ID。

(2) 最大速率关联

该算法的基本思想就是在预测出没有超负载的基站集合之后,计算各个基站可以给需要进行切换的用户提供的数据传输速率的值,选出提供的数据传输速率值最大的基站作为新用户的最佳接入基站。具体计算式为:

$$j^* = \max_{j \in S_L} c_i^j \tag{5}$$

变量表示的含义与式(4)相同。以上两种关联算法是目前最常用的关联算法,尤其是最大 SINR 关联算法,因为用户接入网络的性能指标直接与其相关,同样采用式(5)的关联算法,不仅考虑了目标节点可以为用户提供的 SINR 值,还考虑了目标节点可以提供的资源量,因而获得的网络性能和关联效果更佳。

在选择了合适的关联算法之后,需要进一步考虑优化算法,因而在所选的关联因子的前提下,在最终优化的网络性能指标指导下,建立优化模型。

3.2.2 基于动态 ICIC 的 LB 算法^[13]

在异构蜂窝网络中,不仅存在小区间干扰,而且在不同类型的蜂窝之间还存在较强的跨层干扰。在通信系统中干扰是主要影响系统性能的参数,结合干扰协调技术 ICIC (inter-cell interference coordination)将能在 LB 自优化基础上提升网络性能。

该算法引入中继节点来提升网络的容量和覆盖,而且将中继节点主要部署在宏蜂窝边缘,因为处于宏蜂窝边缘的用户受到最强的小区间干扰。当位于小区边缘的中继节点的负载值较大时,可以将其切换到其他相邻负载较低的中继节点上进行服务,从而提升小区边缘用户的性能。考虑到每个用户的业务类型不相同,完成通信所需要的 PRB 数目也是不一样的,这里主要将全网用户的平均传输速率作为考虑的效用函数,即:

$$U = \prod_{k,i,m \in \text{relay UE}} \frac{\delta_{k,i,m} \cdot f_{ki}}{n_{eki}} \prod_{k,i,m \in \text{macro UE}} \frac{\delta_{k,i,m} \cdot f_3}{n_{cki}} \tag{6}$$

其中, $\delta_{k,i,m} \cdot f_{ki}/n_{eki}$ 表示用户 m 的数据率, $\delta_{k,i,m}$ 表示单个 PRB 提供的速率,对于每一个 PRB, $\delta_{k,i,m}$ 是相同的。在获取式(6)的最大值之后,再进行负载均衡策略,提升小区边缘用户的容量。

结合干扰协调技术的 LB 算法,能够更灵活地适应用户的不均匀分布或者用户的移动情况,提升小区边缘用户的性能指标,同时提升整个系统的容量,但仍旧不可避免网络资源利用率不高的缺点。

3.2.3 基于节能机制的 LB 算法^[12,14,15]

该算法是从能耗的角度出发,实现系统的负载均衡,降低能耗是实现绿色无线通信的关键技术之一,利用小区休眠技术能够使主要耗能的设备通过开关机制达到节能的效果。该类算法的基本思想是:当网络中某些节点的资源利用率非常低甚至处于空闲状态下,就可以将该节点进入休眠模式,此时,该节点的覆盖范围为零,并且将这些节点下链接的用户转交到与之相邻的节点下进行服务,以保证网络的正常覆盖。

小区休眠主要有两种触发方式:基于基站的小区休眠和基于运营管理维护 (operation administration and maintenance, OAM)的小区休眠,表 1 是两种触发方式下的小区休眠与唤醒对比。

根据表 1,可以看出此类算法的触发过程的信令流程,当然在为休眠小区的接入用户选择目标基站接入时,同样与超载基站下的用户选择目标基站接入的标准一样,即目标节点的负载值低,而且能够给切换用户提供最大吞吐量值。该算法的流程如图 3 所示。

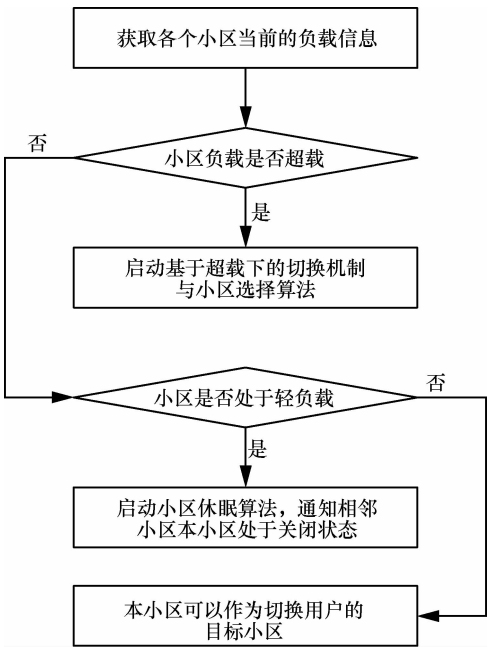


图 3 基于小区休眠的 LB 算法流程

表 1 基于基站和 OAM 的小区休眠和唤醒对比

	基站	OAM
信息	当前基站所在小区的实时流量负载	一定范围内多个小区负载的历史信息
休眠	通过小区内部触发机制进入休眠状态,完全进入休眠状态前通过 X2 接口通知相邻小区	通过北向接口 (Itf-N) 传递小区休眠的 OAM 信息给需要休眠的基站,小区状态信息由 X2 接口传递给相邻基站
唤醒	相邻基站通过 X2 接口传递小区的唤醒命令将休眠小区唤醒	通过 Itf-N 传递小区唤醒的 OAM 信息给需要休眠的基站,小区状态信息由 X2 接口传递给相邻基站



该算法与其他负载均衡算法不同的是,不仅在小区超载的情况下触发负载均衡机制,而且在小区的负载很低甚至处于空闲状态的情况下也能进行负载均衡策略,因而该类算法的复杂程度比其他的算法更高,开销更大,但是它在实现负载均衡的同时,达到了节能的目的,因而是一种重要的 LB 算法。但是在该算法中需要考虑到能耗的降低与网络吞吐量的折中,因为在判断各个小区的负载情况时,需要设置小区的过载阈值与休眠阈值,而采用休眠式的 LB 主要是从节能的角度出发,如果休眠阈值设置过大,使得过多的小区进入休眠状态,影响整个网络的吞吐量,休眠阈值太小,又达不到节能的效果,因而如何选取合适的阈值是本算法的一个难点。

3.3 基于小区范围扩张的 LB 算法

异构分层网络结构中,大部分用户都是直接与宏基站相连接,因而如何将宏基站下的用户推到小基站下进行服务,缓解宏基站的流量阻塞问题,充分利用其覆盖范围内小基站的资源则是该类 LB 算法的核心思想。

3.3.1 Pico 节点范围扩展算法^[16]

这里主要介绍异构蜂窝网络中 Pico 节点的范围扩张(range extension, RE)技术,其基本原理是基于 A3 事件触发的切换条件:

$$\overline{RSRP}_j > \overline{RSRP}_i + HOM_{i \rightarrow j} \quad (7)$$

其中, $\overline{RSRP}_i$ 与 $\overline{RSRP}_j$ 分别表示源基站 i 与目标基站 j 的参考信号接收功率(reference signal received power, RSRP)的平均值, $HOM_{i \rightarrow j}$ 表示用户由源基站 i 与目标基站 j 的切换偏移量,在满足式(7)的条件下,即可以使用户由源基站切换到目标基站。其基本原理如图 4 所示。

基本原理主要是参照 A3 事件触发用户切换,通过调整切换偏移量即图 4 中所示的补偿(offset)改变网络中 Pico 节点的覆盖范围,即将原本属于宏基站覆盖下的用户(图 4 中为深灰色环形范围内)切换到 Pico 节点下进行服务。值得注意的是,offset 的值不应该设置得太大,因为 Pico 等低功率节点的发射功率比宏基站的发射功率远小得多,当其值较大时,即图 4 中深灰色环型部分将变大,那么 Pico 节点的覆盖范围变大,此时位于深灰色型形边界的用户接收到的 SINR 值将偏低,会导致切换失败。表 2 为对 Pico 节点采用 RE 技术(offset 的值为 3 dB)前后系统吞吐量增益对比。

表 2 很明显地表明,在异构蜂窝网络中,对于低功率节点,引入 RE 技术很好地提高了系统吞吐量,当宏基站处于超载的情况下时,可以将其覆盖范围内的用户通过 RE 技术均衡到其他低功率节点上,以使用户获得更好的服务质量。

该算法能够很好地适应网络中各个节点的负载变化信息,并能够自适应调整各个小区的切换偏移量,从而调整各个小区的覆盖范围,最终将超载节点的用户切换到负

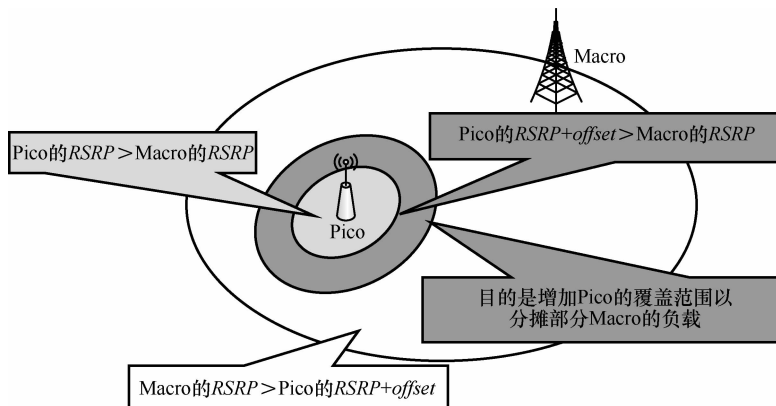


图 4 Pico 节点 RE 技术原理示意

表 2 采用 RE 技术前后的系统吞吐量增益对比

		宏蜂窝下 Pico 节点的数目			
		offset=0 dB		offset=3 dB	
		4	10	4	10
下行	5% 吞吐量	×1.5	×1.6	×2.1 (1.40)	×2.6 (1.59)
	50%吞吐量	×4	×7.6	×4.8 (1.21)	×9.7 (1.28)
上行	5% 吞吐量	×4.6	×4.7	×13.7 (2.98)	×6.2 (1.33)
	50% 吞吐量	×8.3	×10.2	×8.3 (1.00)	×11.6 (1.14)

载值较低的节点上。从上述分析中可以很明显看出该类算法的优点：

- 动态调整各个小区的切换偏移量；
- 全网负载均分的效果较高；
- 网络具有一定的感知能力。

同样缺点也较明显：

- 算法复杂度明显增大；
- 系统开销增大。

3.3.2 结合 SON 的 LB 算法^[7,8,17]

自组织网络 (self-organized network, SON)^[3,20-22] 技术概念最早是基于 Ad Hoc 通信机制提出的，它是一种无固定拓扑，无中心节点，可以随时随地动态组建的无线通信网络形式，具有自发现、自配置、自组织和自治愈等核心特征。在 LTE-A 异构网络中引入 SON 技术，旨在网络自动完成异构蜂窝移动通信网络设备的参数配置与优化工作，减少网络操作管理中的人工参与，以使得超载节点的用户成功切换到目标节点下进行服务，并获得较好的数据传输速率，进一步将网络的有限资源与用户的业务需求进行优化匹配。提高网络的可靠性和智能性，降低运营成本。目前主要的 SON 优化算法有 Q-Learning 算法^[17]、模糊逻辑 (fuzzy logic)^[7,8] 算法、小波神经算法等。

在参考文献[8]中，主要介绍的是一种结合模糊逻辑的加强 Q-Learning 的负载均衡算法，主要感知网络中各个节点的阻塞率变化，通过上述两种智能算法自适应调整小区

切换偏置参数以及发射功率，系统设计框图如图 5 所示。

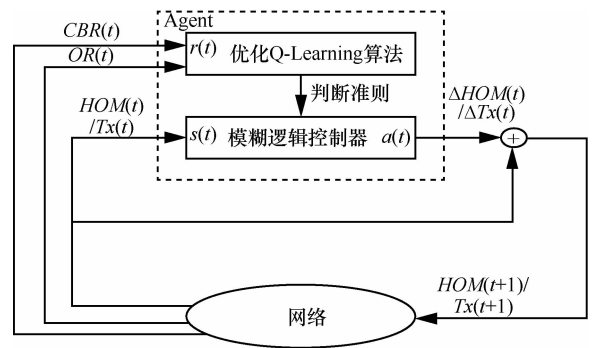


图 5 结合模糊逻辑与 Q-Learning 的负载均衡算法示意

图 5 中，根据自组织网络的基本特性，感知网络环境的变化，并对网络的环境变化进行学习，做出相关决策，在图中主要是调整网络中相关节点的切换偏移量 HOM 或发射功率 T_x ，感知的网络变化主要是呼叫阻塞率 (CBR) 与信道质量太差引起的呼叫失败 OR ，具体的计算式这里将不再详述。

该算法的主要优点就是能够感知当前网络环境的变化，自适应调整网络的相关参数，提高网络的容量，降低用户的切换失败率和呼叫阻塞率，对于以后的高密集异构蜂窝网络，是一种十分重要的负载均衡算法。缺点就是相关的智能算法比较复杂，会带来一定的时延。

上述的负载均衡算法，按照不同的侧重点进行分类讨论，各自有各自的优缺点，表 3 是对上述各种类型算法进行的总结。

表 3 不同类型负载均衡算法性能评估及对比

	触发机制		效用函数			小区范围扩张	
	负载值	负载差值	关联	ICIC	节能	Pico	SON
相关技术	LB, RRM		LB, FFR, RRM、节能			LB, RRM, SON	
核心参数	负载值(资源利用率/呼叫阻塞率/接入用户数)		关联因子	全网效用函数的选取		各个节点之间的切换偏移量	
折中因子	不考虑		资源利用率与阻塞率	用户干扰	降低能耗	不考虑	
干扰与频谱效率				干扰与频谱效率	吞吐量与能耗		
信令开销	小	大	大	大	大	中	大
复杂度	小，只需要与预设的门限比较	中，比较次数多	较高，用户与基站进行匹配	中，分布式比集中式开销大	高，休眠与超载都要切换	较高，依据求解切换量的算法而定	高，取决于 SON 智能算法和调整切换偏移量算法
切换时延	较低	中	中	中	中	中	高
切换失败率	较高	较低	低	较低	较高	较低	低
呼叫阻塞率	较高	较低	低	较低	较高	较低	低
资源利用率	较低	较高	高	较低	较低	较高	高
全网吞吐量	较高	较高	高	较低	较低	较高	高
选择目标基站端	用户/基站	用户/基站	基站	基站	基站	用户	用户
难点	无难点	获取相邻节点的信息并循环比较	获取相邻节点的信息并设置偏置因子	获取相邻并同频干扰节点的信息	休眠小区阈值设置	Pico 等低功率节点的切换偏移值设置	引入智能算法的求解难度高



4 负载均衡过程中减少信令开销方案

第3节主要介绍目前研究的负载均衡相关算法,该部分毫无疑问是负载均衡策略的核心部分,但是负载均衡过程中负载报告的交付流程也是非常重要的,网络中各个节点在进行负载信息交换过程中会产生很大的信令开销,如何减少系统的信令开销,提升网络的时效性也是目前研究负载均衡策略的重点之一。下面主要介绍目前工程上减少不同系统之间进行负载均衡策略信令开销的几个方案^[12,23-28]。

首先介绍在场景1^[28]下5种增强型负载均衡报告方案,场景1如图6所示。

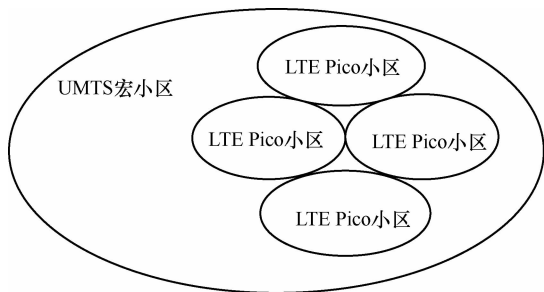


图6 网络部署场景1

场景1中以通用移动通信系统(universal mobile telecommunications system, UMTS)提供全部的覆盖范围, LTE Pico eNB 为了改善覆盖范围容量而提供部分覆盖。

4.1 主基站将 UTRAN 小区负载响应分配给从基站

该解决方案(即方案1)的主要目标是减少在接口 Iu、S3 和 S1 经过重复事件触发 UMTS 的无线接入网(UMTS terrestrial radio access network, UTRAN)小区的负载报告。如图7所示,在步骤(1)~步骤(7),主基站通过利用现有负载交换信令(RIM)流程过程请求并获取相邻无线网络控制器(radio network controller, RNC)的 UTRAN 的小区负载;在第步骤(8)中,主基站通过基于 UTRAN 小区负载请求的 X2 接口给从基站 eNB1、eNB2 和 eNB3 发送 UTRAN 小区负载。

该解决方案不会影响到主基站、管理实体(management entity, MME)、服务 GPRS 支持节点(serving GPRS support node, SGSN)和 RNC 间的 RIM 过程,其中 RIM 流程是负载均衡报告交换具体过程。当主基站从 RNC 中接收到 UMTS 小区负载时,有3种方式可以使主基站发送小区负载情况给从基站。

- 主基站通过一些修改将接收到的包含 UTRAN 小区负载请求/更新的 RIM 协议数据单元(protocol data unit, PDU)转交给从基站,在转交之前,主基站接收到的 RIM PDU 中的目标小区 ID 将由主基站 ID 改为从基站 ID。
- 主基站不需要经过任何修改,直接将接收到的包含 UTRAN 小区负载请求/更新的 RIM PDU 转交给从基站,接收到 RIM PDU 的从基站应该忽略在 RIM PDU 中的目的小区 ID。

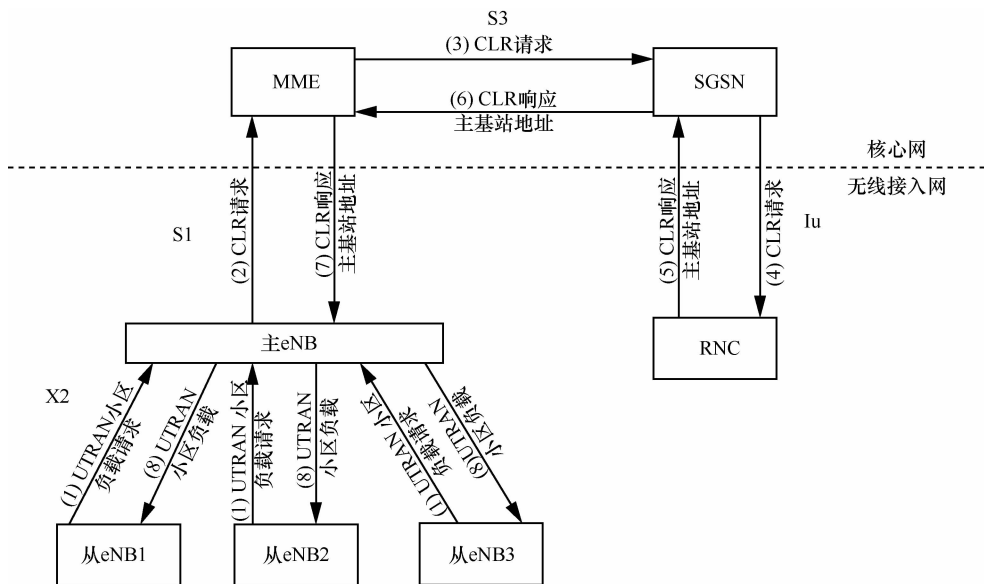


图7 主基站将 UTRAN 小区负载响应分配给从基站流程

- 主基站从 RIM PDU 中导出 UTRAN 小区负载请求/更新,并通过 X2AP 利用新的信元(IE)或者新的程序将 UMTS 小区负载发送给基站。

但是从参考文献[29]看,该方案也有相应的缺点,即:

- 相同负载水平的报告要求应该适应于所有的从 eNB,与当前的标准一样不灵活;
- 在这个解决方案中,需要事先配置从 eNB 的分组;
- 在现在的标准中还不存在主/从 eNB 的结构,引入一个主基站可能会导致瓶颈问题,如当主基站需要在从基站之间广播负载信息时。

4.2 MME 将 UTRAN 小区负载响应分配给 eNB

该解决方案(即方案 2)的目标是在 Iu、S3 接口处减少重复事件触发 UMTS 小区负载报告。如图 8 所示,每个基站利用已存在的 RIM 过程从 RNC 中获得小区负载信息,如果所有的 eNB 都要求相同的事件门限,那么 RNC 会发送一个组合 RIM 负载报告。RNC 通过 RAN 将 UTRAN 的小区负载报告给 MME,而且这个 MME 会把小区负载报告发送给不同的 eNB,MME 从 RIM 的信息头中得到路由信息。

在该方案中,如果 eNB1、eNB2 和 eNB3 共享相同的 TAI (tracking area identity, 跟踪区域标识),RNC 增加 eNB1、eNB2 和 eNB3 的 ID 或者 TAI 在 RIM 的头部中;MME 根据路由信息在 S1 接口上分布消息。从参考文献[29]中知道该解决方案需要注意以下几点。

- 为了将 UTRAN 小区负载消息发送给多个 eNB,MME 将会受到影响,以便能够理解包含在 RIM PDU 中的新目标地址即多个目标全局 eNB ID 或者 TAI,并与在当前目的地址中的一个目标全局 eNB 的 ID 进行比较。

- 为了减少负载响应消息,当请求报告事件发生时,RNC 应该对请求相同报告级的 eNB 聚合负载响应,信令减少程度取决于具有相同请求报告水平 eNB 的数目。

- 在这个解决方案中,MME 将不会聚合小区负载报告请求,因为在大多数方案中,eNB1、eNB2 和 eNB3 在不同的时间向 MME 发送请求。

4.3 eNB 将 UTRAN 小区负载分配给其他 eNB

该解决方案(即方案 3)的主要目标是在 Iu、S3 和 S1 接口处减少重复事件触发的 UTRAN 小区负载报告。如图 9 所示,每个 eNB 利用已存在的 RIM 程序从 RNC 中获得小区负载信息,如果所有的 eNB 要求的是相同的事件阈值,那么 RNC 会发送一个组 RIM 负载报告。RNC 通过 RAN 信息将 UTRAN 的小区负载信息报告给某个 eNB,而这个 eNB 会通过 X2 将小区负载信息转发给其他 eNB,该基站会在 SON 转发容器中获得路由信息,假设基站与基站之间的接口是 X2。

在该方案下,RNC 会将 eNB2、eNB3 的全部 eNB ID 增加到 RIM PDU 中,并发送给 eNB1,eNB1 通过 X2 将 UTRAN 小区负载分配给 eNB2 和 eNB3,类似于第 4.1 节方案,UTRAN 小区的相邻 eNB 应该分组,eNB 组应该由 OAM 配置或者通过新定义的 RIM 程序自己建立。在该方案下,从参考文献[29]中知道需要注意以下几个方面。

- 需要知道一种配置和组 eNB1~eNB3,需要从相同的 UMTS RNC 获得相同的负载信息。
- 为了减少负载响应消息,当请求报告事件发生时,RNC 应该对请求相同报告级别的 eNB 聚合负载响应,信令减少程度取决于具有相同请求报告水平 eNB 的数目。

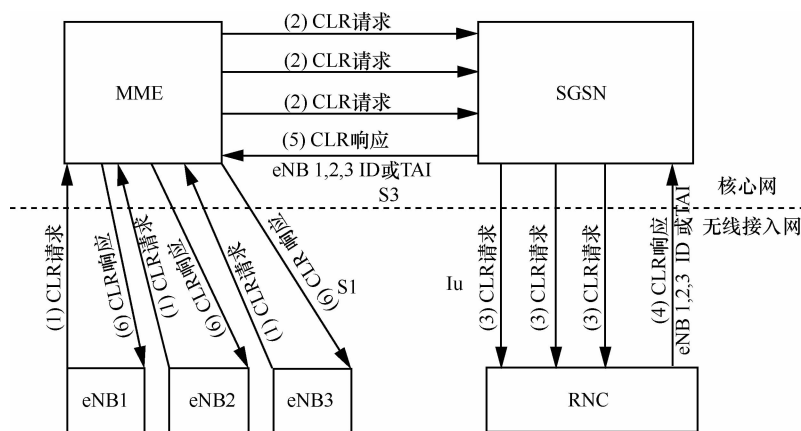


图 8 MME 将 UTRAN 小区负载响应分配给 eNB 流程

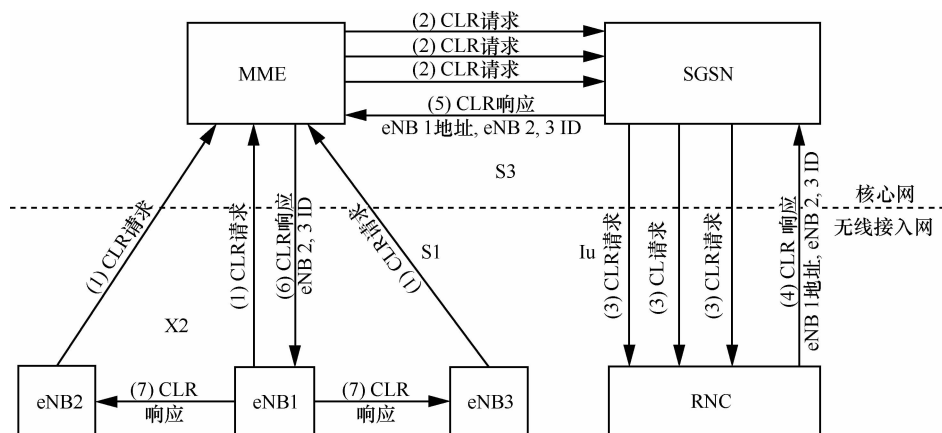


图9 eNB将UTRAN小区负载分配给其他eNB

- 这种解决方案相比解决方案1,主eNB更加灵活,而且不需要预先配置。
- 在现在的标准中还不存在主/从eNB的结构,引入一个主基站可能会导致瓶颈问题,当主基站需要在从基站之间广播负载信息时。
- 应该将UTRAN小区的相邻eNB分组,RNC将会根据OAM配置以及请求相同负载报告级别的eNB进行分组。

4.4 主基站聚集从基站的E-UTRAN小区负载响应给RNC

该解决方案(即方案4)的主要目标是减少在接口Iu、S3和S1中演进型UTRAN(evolved universal terrestrial radio

access network, E-UTRAN)系统小区负载报告请求/响应的信令负载。如图10所示,RNC要求主基站获得从基站eNB1、eNB2和eNB3的E-UTRAN的小区负载信息,将主基站的ID设置为目标小区ID,与此同时将从基站ID设置为报告小区ID。一旦主基站接收到小区负载请求的RIM PDU,它将会检查在报告小区标识中是否有可用的小区负载。如果不可用,主基站会识别那些基站和小区是否属于报告小区识别列表,并通过X2请求小区负载。

在E-UTRAN小区负载请求之前,UTRAN小区相邻的eNB应该被分组。基站分组可以由OAM配置或者通过新定义的X2AP程序来自己建立。每个eNB组包含几个从基

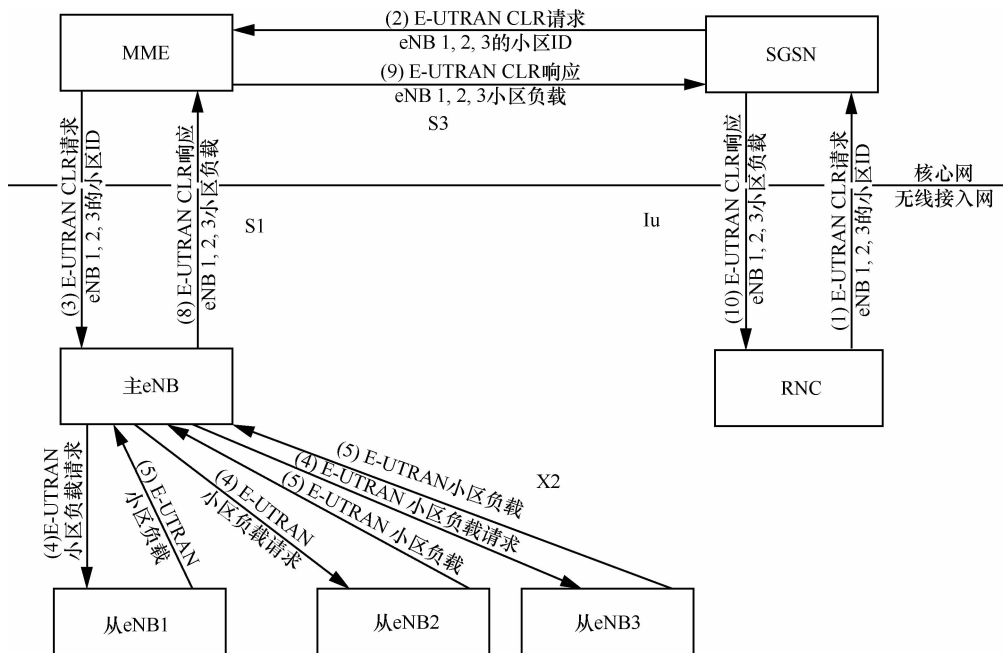


图10 主基站聚集从基站的E-UTRAN小区负载响应给RNC流程

站和一个主基站。eNB 组信息也能由 OAM 配置在 RNC 中或者可以通过新定义的 RIM 程序从主基站中接收。然后, RNC 可以请求主基站为组中一个以上的其他基站报告负载信息。基于这个请求, 主基站通过 X2 获得从基站的小区负载然后将负载信息报告给 RNC。为了简化, 在图 10 中只有 3 个从基站 ($N=4$), 通过利用加强 RIM 传输, 在 MME、SGSN 和 RNC 传输的 RIM PDU 数量将减少为 $1/N$ 。从参考文献[29]中知道该方案需要注意的是:

- 相同负载水平的报告要求应该适应于所有的从 eNB, 和当前的标准一样不灵活;
- 在现在的标准中还不存在主/从 eNB 的结构, 引入一个主基站可能会导致瓶颈问题, 如当主基站需要在从基站之间广播负载信息时;
- 主 eNB 应该记录从 eNB 的报告;
- 信令减少程度取决于从 eNB 请求事件发生的次数, 从这些基站触发的小区负载报告在同一时间到达主 eNB。

4.5 MME 聚集 eNB 的 EUTRAN 的小区负载响应给 RNC

该方法(即方案 5)的目标是减少场景 1 下的信令开销和 RNC 需要的位于其区域范围内的 eNB 的负载信息。这种解决方案如图 11 所示, 可能会采用: 当 UMTS 宏节点过载时, RNC 会寻找在其覆盖范围内的所有 LPN eNB 卸载的 UE。此时, RNC 将会从 eNB 的列表中请求获取负载信息。该解决方案归纳如下。

- RNC 向不同的 eNB 发送负载信息请求, 这些负载信息包含在属于不同 eNB 的小区列表中。
- MME 解复用由 RNC 发送的请求, 并发送给若干个 eNB。
- MME 向 eNB 发送请求, 启动一个定时器, 收集它们

的响应, 直到超过定时器的期限。定时器应该反映主干网的拓扑结构: eNB 需要一个较大的值来连接几跳。

- 一旦 MME 将从 eNB 的响应捆绑起来, 它会以一个单一的消息形式发送给 RNC。
- 如果 MME 收到迟的 eNB 的响应, 它们将会转为单响应不会捆绑到 RNC, 以保持低的响应时延。

从参考文献[29]中知道该方案中需要注意的是:

- RNC 和 MME 需要进行修改, 以能够生成/理解包含在 RIM 的 PDU, 即多目标全局基站 ID 或者 TAI;
- RNC 需要通过 OAM 配置 eNB ID 和 TAI;
- MME 需要记录和存储 eNB 的负载响应;
- 对于 eNB 的快速响应可能会带来额外的时延。

表 4 是对上述 5 种方案的对比。

从表 4 中对上述 5 种方案的评估与对比可知, 得到如下 3 个结论^[29]。

- 方案 2 和方案 5 低效。因为这两种方法对 CN 的影响处于中等或者较高的, 但是对于 S1/Iu 接口的信令减少程度来说是较低的, 而且对 CN 的资源利用率来说也是处于中间的, 即方案 2 和方案 5 不是十分有效的解决方法。
- 方案 1 和方案 3, 这两种方法在减少开关时延、核心网的资源利用、UE、eNB、CN 和 OAM 等方面上达到的效果是相似的, 但是从 S1/Iu 接口的信令开销上来说, 方案 1 比方案 3 要高得多, 而且, 方案 1 对 RNC/Node B 没有影响, 但是方案 3 则对 RNC/Node B 的影响较高。基于以上分析, 推论出方法 1 在场景 1 下加强 UTRAN 到 E-UTRAN 的小区负载报告比方案 3 更有效, 可以进一步研究。

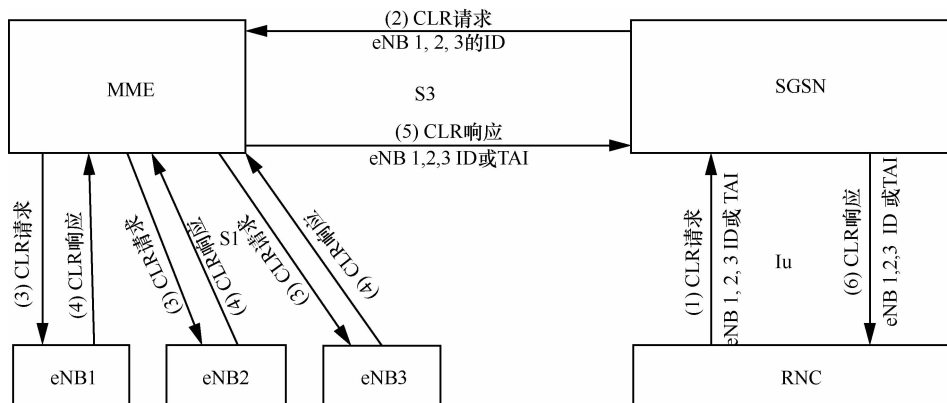


图 11 RNC 请求多个 eNB 的负载信息且 MME 以组的方式响应示意图



表 4 场景 1 下 5 种方案评估对比

	方案 1	方案 2	方案 3	方案 4	方案 5
应用场景	场景 1 由 UTRAN 到 E-UTRAN 小区负载报告	场景 1 由 UTRAN 到 E-UTRAN 小区负载报告	场景 1 由 E-UTRAN 到 UTRAN 小区负载报告	场景 1 由 E-UTRAN 到 UTRAN 小区负载报告	场景 1 由 E-UTRAN 到 UTRAN 小区负载报告
减少的信令开销	高; 减少了 S1 和 Iu 接口上的请求和响应消息的数量。根据每个组的从基站的数量, S1 与 Iu 接口上的信令开销减少到 1/N	低; 减少了 Iu 接口上的响应消息的数量, S1 接口不变。根据请求相同级别的报告的基站数目, Iu 接口上的信令减少到 1/N	中; 减少了 S1 和 Iu 接口上的响应消息的数量。根据请求相同级别的报告的基站数目, S1 和 Iu 接口上的信令减少到 1/N	中; 减少了 S1 和 Iu 接口上的请求消息的数量。如果主基站在短时间内得到多个从基站小区负载信息, 就会减少 S1 和 Iu 上的响应消息。S1 与 Iu 接口上的信令开销减少到 1/N	低; 减少了 Iu 接口上请求和响应消息的数量, S1 接口不变。在 Iu 接口上减少的请求消息的数量为 1/N, N 列表中基站的数目 Iu 接口上减少的响应消息的数目最多为 1/N (如果在计时器规定的时间内多个基站的负载信息到达 MME)
减少切换时延	没有影响	没有影响	没有影响	没有影响	没有影响
接入网资源利用率	中; 由 RNC 发送的请求和响应消息更少; 主基站需要更多的内存将负载信息响应分发给从基站	中; 由 RNC 发送的响应消息更少; RNC 在负载信息交互过程中需要更多的内存处理组基站的信息	低; 由 RNC 发送的响应消息更少; RNC 在负载信息交互过程中需要更多的内存处理组基站的信息	中; 由 RNC 发送的请求和响应消息更少; 主基站需要更多的内存来收集从基站响应的负载信息	低; 由 RNC 发送更少的请求消息与接收更少的响应消息; RNC 需要更多的内存处理目标基站的负载信息
核心网资源利用率	高; 通过 MME 与 SGSN 转发的请求和响应消息更少	中; 通过 MME 与 SGSN 转发的请求和响应消息更少; MME 需要更多的内存来处理负载流程, 将响应消息发送给基站	高; 通过 MME 与 SGSN 转发的请求和响应消息更少	中; 通过 MME 与 SGSN 转发的请求和响应消息更少	中; 通过 MME 与 SGSN 转发的请求和响应消息更少; MME 需要更多的处理流程
对用户的影响	没有影响	没有影响	没有影响	没有影响	没有影响
对基站的影响	高; 主基站将会终止从基站的 UTRAN 小区的负载报告请求, 并将接收到的 UTRAN 小区的负载信息分发给从基站; 主基站与从基站之间需要进行请求响应消息交互, X2 接口将会有更多的信令交互	低; 当接收到一个 RIM PDU 时, 基站需要检测是否包含在其目标小区的识别列表中	高; 当接收到 RIM PDU 时, 如果一个基站通过 S1 接口在 SON 容器中检测到其他基站的 ID, 它将会通过 X2 接口将 UTRAN 的小区负载信息发送给其他的基站; 其他的基站通过 X2 接口接收响应消息, 通过 S1 接口发送请求消息, X2 接口将会有更多的信令交互	中; 当接收到 EUTRAN 小区负载请求时, 主基站应该检测所请求小区的额定负载信息是否可用, 否则主基站通过 X2 接口向小区请求负载状态消息; 主基站应该将所有的 EUTRAN 小区负载请求消息封装在一个 RIM PDU 中, 并将它们发送给 RNC, X2 接口将会有更多的信令交互	没有影响; 基站的行为没有变化
对 RNC/Node B 的影响	没有影响	高; RNC 将包括一个新的目标地址, 在响应消息中应该包括多个基站的全局 ID 或 TAI; RNC 需要根据请求级别将基站分组	高; 在响应消息中, RNC 应该在 SON 转发容器中包含其他基站的 ID; RNC 根据请求级别和基站的 X2 接口的关系将基站分组	中; 在请求消息中, RNC 应该设置报告小区的 IE 识别号给请求小区的 ID	高; 在请求消息中, RNC 应该设置报告小区的 IE 识别号给请求小区的 ID
对核心网 (CN) 的影响	没影响	中	没影响	没影响	高
对 OAM 的影响	中; 配置主/从基站信息; 配置组内相同报告级别基站信息	没影响	中; 配置基站到 RNC 与组内基站的 X2 接口关系	中; 配置 RNC 的组信息; 将 UTRAN 小区的相邻基站划分为一组, 每个组由一个主基站和几个从基站组成	中; RNC 需要被配置 eNB 和 TAI, 以使得它知道哪个基站需要被联系

- 在场景 1 中加强 E-UTRAN 到 UTRAN 的小区负载报告,方案 4 是一种可行方案。

5 结束语与展望

异构网络的发展为无线资源管理技术提出了更多的挑战和需求,与此同时,各种新技术、新业务、新场景的不断涌现,需要现有无线资源管理技术不断增强适应性、扩展性、准确性、通用性。负载均衡技术作为无线资源管理部分的关键技术之一,它对优化整个网络的性能具有十分重要的意义。但是在 LTE-A 异构网络中,负载均衡技术的研究仍然处于初级阶段,具有以下几方面的挑战。

(1) 如何对各个小区的负载进行感知预测

在未来移动通信系统中,网络的智能化是必要的。这是因为网络规模越来越大,网络结构的多样化,多种控制接入方式并存,多种系统架构之间需要协同合作,如果全部采用人为操作,将会使运营成本急剧加大,而且在操作过程中也会使效率极低。负载均衡技术是随着无线网络的发展应运而生的,是一种网络性能自优化技术,它的发展必须与移动通信网络的发展相适应,实时感知网络负载变化。

(2) 如何减少负载交换过程中的信令开销

系统采取负责均衡策略时,网络中的节点需要交换各自的负载信息,这个过程中会带来大量的信令开销,势必会导致网络性能的下降,特别是时效性,因而减少负载均衡过程中的信令开销对提升网络的性能具有重要的意义。而在 LTE-A 异构网络中,采用的仍然是传统负载均衡的控制模式,并没有提出具体的标准化的负载交互流程。

(3) 如何为超载小区的用户选择目标基站接入

这部分实质是小区重选,即为超载节点下的用户选择最佳目标节点接入,使用户的服务质量最优化。因而该部分需要研究的是小区重选的优化算法,在未来移动通信技术中,同样需要与 SON 技术相结合,能够依据网络的历史信息和当前用户的状态信息更好地为切换用户选择目标节点。

(4) 如何自适应调整目标基站的相关参数

上述介绍的负载均衡算法中,有调整各个小区切换偏置量的,有调整目标小区的发射功率的,还有调整不同类型小区的偏置因子的,但是未来的网络会发展为高密集异

构蜂窝网络,系统中节点数目的不断增多会使调整参数的难度加大,因而需要引入 SON 自适应调整算法,实时感知网络环境变化,提升网络的有效性。

(5) 如何建立接入点间有效协作、协调机制

未来异构网络是一种融合网络,各个节点存在大量物理资源块(时域和频域),还有空域、码域等资源,需要建立接入点间有效协作、协调机制,将有限的网络资源与分配不均的用户业务需求进行优化匹配。

本文对近年来的负载均衡算法的分类描述、比较以及负载均衡过程中减少信令开销的方案的介绍,希望起到抛砖引玉的作用,随着未来异构网络的快速发展,负载均衡技术的重要性也日益突出,如何在网络节点间将有限的网络资源与用户的业务需求匹配,以动态适应网络负载的变化,降低用户的呼叫阻塞率,提升网络的资源利用率,仍将成为未来异构网络研究中的一个挑战性问题。

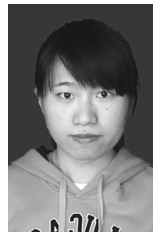
参考文献

- Demestichas P, Georgakopoulos A, Karvounas D, *et al.* 5G on the horizon: key challenges for the radio-access network. *IEEE Vehicular Technology Magazine*, 2013,8(3)
- Damnjanovic A, Montojo J, Wei Y B, *et al.* A survey on 3GPP heterogeneous networks. *IEEE Wireless Communications*, 2011,18(3)
- Peng M G, Liang D, Wei Y, *et al.* Self-configuration and self-optimization in LTE-advanced heterogeneous networks. *IEEE Communications Magazine*, 2013,51(5)
- Andrews J, Singh S, Ye Q Y, *et al.* An overview of load balancing in HetNets: old myths and open problems. *IEEE Wireless Communications*, 2014,21(2)
- Androne C, Palade T, Puschita E, *et al.* Study of co-channel cross-layer interference for the downlink communication in femtocell networks. *Proceedings of ISSCS, Lasi*, 2011
- Li Z H, Wang H, Pan Z W, *et al.* Joint optimization on load balancing and network load in 3GPP LTE multi-cell networks. *Proceedings of 2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China, 2011
- Sheng J, Yang Z, Tang L R. A novel load balancing algorithm based on utility functions and fuzzy logic in heterogeneous wireless networks. *Proceedings of FSKD, Sichuan, China*, 2012
- Muñoz P, Barco R, Ruiz-Avilé J M, *et al.* Fuzzy rule-based reinforcement learning for load balancing techniques in enterprise LTE femtocells. *IEEE Transactions on Vehicular Technology*, 2013,62(5)



- 9 Kyuho Son, Song C. Dynamic association for load balancing and interference avoidance in multi-cell Networks. IEEE Transactions on Wireless Communications, 2009,8(7)
- 10 R3-082698. Load Balancing. Vodafone, 2008
- 11 Yang Y, Li P F, Chen X H, *et al.* A high-efficient algorithm of mobile load balancing in LTE system. Proceedings of IEEE Vehicular Technology Conference (VTC Fall), Quebec City, 2012
- 12 R3-130695. Enhancements for Cell Load Reporting. Kyocera Corp, 2013
- 13 Wang J, Liu J G, Wang D Y, *et al.* Load balance based dynamic inter-cell interference coordination for relay enhanced cellular network. Proceedings of IEEE The 75th Vehicular Technology Conference (VTC Spring), Yokohama, 2012
- 14 Sung-rae Cho, Wan Choi. Energy-efficient repulsive cell activation for heterogeneous cellular networks. IEEE Journal on Selected Areas In Communications, 2013,31(5)
- 15 Peng J L, Hong P L, Xue K P. Performance analysis of switching strategy in LTE-A heterogeneous networks. Journal of Communications and Networks, 2013,15(3)
- 16 R1-103824. Aspects of Pico Node Range Extension, Nokia Siemens Networks. Nokia, 2010
- 17 Faganello L R, Kunst R, Both C B, *et al.* Improving reinforcement learning algorithms for dynamic spectrum allocation in cognitive sensor networks. Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 2013
- 18 Ye Q Y, Rong B Y, Chen Y D, *et al.* User association for load balancing in heterogeneous cellular networks. IEEE Transactions on Wireless Communications, 2013,12(6)
- 19 Yu Y, Hu Q Y, Bontu C S, *et al.* Mobile association and load balancing in a cooperative relay cellular network. IEEE Communications Magazine, 2011,49(5)
- 20 3GPP TR 37.822. Study on Next Generation Self-Optimizing Network(SON) for UTRAN and E-UTRAN, (Release 12), 2014
- 21 Aliu O G, Imran A. A survey of self organisation in future cellular networks. IEEE Communications Surveys & Tutorials, 2013, 15(1)
- 22 3GPP TS32.541. Telecommunication Management; Self-Organizing Networks (SON); Self-Healing Concepts and Requirements, 2011
- 23 R3-130873. Discussion on Load Balancing Enhancements for Scenario 1. Huawei, 2013
- 24 R3-131270. Cell Load Reporting Enhancement Scenario 1a. Huawei, China Unicom, LG Electronics, Samsung, 2013
- 25 R3-131403. Evaluation Method for Current Load Balancing Mechanisms in the Context of UMTS-LTE Interworking. NSN, 2013
- 26 R3-131366. Discussion on Load Balancing Enhancements. ZTE, 2013
- 27 R3-130559. Load Balancing Based on Access Mode Change. ZTE, 2013
- 28 3GPP TR 37.852. RAN Enhancements for UMTS/HSPA and LTE Interworking (Release 12), 2013
- 29 R3-131593. Cell Load Reporting Enhancement Scenario 1a. Huawei, China Unicom, LG Electronics, Samsung, 2013

[作者简介]



许娟雄,女,重庆邮电大学硕士生,主要研究方向为 LTE-Advanced 异构网络中负载均衡研究、SON。



刘伊莎,女,重庆邮电大学硕士生,主要研究方向为 LTE-Advanced 异构网络中负载均衡研究、SON。



唐伦,男,博士,重庆邮电大学副教授,主要研究方向为新一代无线通信网络资源管理、车载组织网络。



陈前斌,男,博士,重庆邮电大学教授、博士生导师,重庆邮电大学通信与信息工程学院院长,主要研究方向为新一代移动通信系统、未来网络、LTE-Advanced 异构小蜂窝网络等。

(收稿日期:2014-07-12)