

异构无线网络终端服务感知的动态负载均衡机制

罗俊辉¹ 白光伟^{1,2} 沈航² 曹磊³

(南京工业大学计算机科学与技术系 南京 210009)¹

(南京理工大学高维信息智能感知与系统教育部重点实验室 南京 210094)²

(南京大学计算机软件新技术国家重点实验室 南京 210093)³

摘要 异构无线网络负载均衡是提高网络服务质量的关键技术之一。现有负载均衡机制大多没有考虑用户的个性化业务在请求网络资源时存在的差异,同时缺乏一种用户满意度保障机制。针对这一问题,提出一种终端服务感知的动态负载预先均衡机制。首先,采用模糊数学的推理方法,计算出终端对各个候选网络的接入期望值。在此基础上,充分考虑各候选网络的资源拥塞状况,有选择性地将该终端用户业务链接切换到低资源拥塞率、高接入期望值网络中,从而缓解网络中的局部热点,提高网络资源有效利用率与用户满意度。仿真结果表明,该机制能够显著均衡网络中各个无线接入点的工作负荷,有效减少拥塞发生的概率,改善用户的网络体验。

关键词 移动互联网,异构网络切换,动态负载均衡,松耦合模式,模糊数学

中图分类号 TP393.03 **文献标识码** A

Dynamic Load Balance Mechanism Based on Mobile Terminal Service-aware in Integrated Heterogeneous Wireless Networks

LUO Jun-hui¹ BAI Guang-wei^{1,2} SHEN Hang² CAO Lei³

(Department of Computer Science and Technology, Nanjing University of Technology, Nanjing 210009, China)¹

(Key Laboratory of Intelligent Perception and System for High-dimensional Information of Ministry of Education of China, Nanjing University of Science and Technology, Nanjing 210094, China)²

(State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, China)³

Abstract Load balance in heterogeneous wireless networks is one of the key technologies to enhance service quality of networks. Most existing load balance mechanisms do not take the difference of personalized business demand of users for network resources into account, and moreover, lack a satisfaction guarantees mechanism. To address this problem, a novel dynamic load balance in advance mechanism based on end-user service awareness was proposed. At first, a fuzzy mathematics model was introduced to achieve access expectations with respect to each candidate network of terminal. On this basis, the services of mobile end-user in current network could be transferred to low resources congestion rate and high exceptions access network selectively, in full consideration of the congestion status of these candidate network resources, such that local hot-spots could be alleviated, as a result to enhance the resource utilization efficiency and improve user satisfactions. Simulations demonstrate that the proposed mechanism achieves significant performance improvement, in terms of access point load, as well as congestion avoidance and user network experiences.

Keywords Mobile network, Heterogeneous network handoff, Dynamic load balancing, Loose coupling, Fuzzy mathematics

1 引言

近年来,随着移动互联网中网络泛在性的日渐普及、移动计算技术的逐步推进,产生了各种各样的无线接入技术与多

模终端设备^[1],造成了网络接入技术冗余与业务扩展多样化、多标准化^[2],如各种接入网络在空间上重叠覆盖,各种移动多媒体业务在服务内容上耦合。因此,根据网络及用户的固有属性,建立不同决策标准进行网络切换,为异构网络中的移动

到稿日期:2013-08-13 返修日期:2013-11-14 本文受国家自然科学基金项目(60673185,61073197),江苏省自然科学基金项目(BK2010548),江苏省科技支撑计划(工业)项目(BE2011186),江苏省普通高校研究生科研创新计划项目(CXLX11_0262,CXZZ12_0425),江苏省六大高峰人才基金资助项目,南京大学计算机软件新技术国家重点实验室开放课题(KFKT2010B08)资助。

罗俊辉(1985—),男,硕士生,主要研究方向为移动互联网,E-mail:luojunhui@163.com;白光伟(1961—),男,博士,教授,博士生导师,CCF高级会员,主要研究方向为移动互联网、无线传感器网络、网络体系结构和协议、网络系统性能分析和评价、多媒体网络服务质量等;沈航(1984—),男,博士生,CCF学生会员,主要研究方向为无线网络编码、移动互联网、无线多媒体通信协议等;曹磊(1980—),男,博士生,CCF学生会员,主要研究方向为无线网络编码、移动互联网、无线传感器网络等。

用户提供无缝个性化服务,是移动互联网络中值得研究的一个关键技术。然而,不同接入技术的移动网络,其服务能力存在很大差异。3G网络提供大范围的连接服务,同时支持终端用户的高速移动,但是需要支付高服务费用,可用带宽较低;WIFI网络覆盖范围有限,却能够向用户提供高带宽、低费用的数据服务^[3];WiMAX弥补了3G与WIFI的缺陷,其数据传输速率高,覆盖范围较大,具有一定的移动性。因此,在通过切换决策来为用户提供个性化服务时,考虑各个网络的业务服务能力,特别是局部热点网络中用户负载的均衡,将能极大提高网络的资源利用率,延长网络服务周期。

目前,关于上述异构网络切换及局部负载均衡问题,已有大量文献资料做出了深入研究。在文献[4]中,通过对接收信号强度、服务费用和可用带宽的动态评估与计算,为用户选择最佳的切换网络。文献[5]通过访问移动终端、网络与用户的环境信息,提出了一种基于用户感知的QoS触发切换决策算法。文献[6]考虑网络的接收信号强度、数据率、服务费用、时延、比特错误率与用户偏好等参数,提出了异构网络中基于模糊逻辑控制的垂直切换算法。文献[7]提出了一种智能切换机制,即通过网络与终端的协同为用户选择最佳目标网络。然而,这些研究工作着重于为用户选择高满意度的网络,没有考虑网络的资源使用状况。文献[8]提出了一种在3G/WLAN网络中终端移动与业务认知的动态负载均衡机制DLB-MSA,以为进入3G/WLAN重叠覆盖区域内的用户提前部署无线通信资源,避免新业务链接建立时引入的实时计算时延,实现热点地区AP间的负载动态均衡。文献[9]提出的优化的联合业务接入控制机制JSAC,通过3G与WLAN之间的密切协作,完成网络负载的动态均衡。然而,上述研究在执行网络负载均衡分配时,大多独立追求网络自身资源使用效益最大化,没有考虑用户对网络的接入期望,虽然提高了网络通信资源的利用率,但是无法保证所选的接入网络能够使用户获得较高的网络服务体验。

因此,在网络融合过程中,结合终端与网络的利益完成网络切换与负载均衡,不但能够高效利用网络资源,而且可以更好地满足终端用户的个性化需求。本文在移动互联网融合框架下,基于资源预留与负载预先均衡的思想,结合网络资源的使用状况与移动终端的自身需求,在各网络的协助下完成终端对接入网的个性化自适应选择,实现松耦合模式下各融合网络中用户负载的动态均衡,从而一方面可以很好地满足用户日益增长的业务需求,另一方面能够提高整个网络的资源利用率。

本文第2节分析现有的垂直切换与负载均衡机制存在的问题;第3节提出基于终端服务感知的动态负载均衡机制;第4节仿真验证上述机制的有效性;最后对全文进行总结。

2 问题描述

现有的异构无线网络中的垂直切换机制,大多是以网络为核心的实时业务切换,通过网络通信资源的高效使用来完成实时切换任务,容易造成网络负载过重而致使网络瘫痪;另外,大部分的切换机制未引入面向用户的资源预留机制,当有大量业务到达时,繁重的实时计算不仅增加了连接建立时延开销,还严重降低了用户的用户体验。

DLB-MSA^[8]机制由网络接入路由器根据用户的移动、通信状态,为其预先选择一个合适的接入点,提前部署无线通信资源,实现网络负载预先均衡。然而,该机制主要从网络角度

考虑了网络覆盖以及网络负载,并结合终端的移动状态进行网络负载均衡,忽略了用户业务的个性化需求,缺乏对用户QoS需求及服务费用等因素的考虑,造成移动用户体验降低,打击了用户追求个性服务的积极性,严重影响了用户的网络服务体验。终端用户对目标AP的选择,通常基于如下期望:

- 1)重叠覆盖区域内的信号强度越强,用户接入的期望值越高。
- 2)终端电池剩余能量越高,获得个性化服务的期望越高。
- 3)网络服务开销越低,获得网络服务的期望越高。
- 4)用户有业务到达时,网络提供的QoS保证越高,获得网络服务的期望越高。
- 5)网络可用资源越充裕,获得网络服务的期望越高。

基于上述情形,若采用以用户为核心的垂直切换机制,缺乏面向用户服务的资源预留机制,容易导致网络负载过重;另外,采用DLB-MSA机制实现异构网络的切换,由于没有考虑终端用户个性化多样化需求而带来了用户满意度过低等问题。因此,本文运用资源预留与负载预先均衡的思想,采取以用户满意度为核心、网络协助的方式实现负载动态均衡。在选择切换网络时,充分考虑终端的电池剩余能量、QoS需求、网络服务费用以及终端接受信号强度等环境信息,以此为输入并通过模糊数学算法评估各个候选接入网络,计算每个候选AP的APEI(AP Expectations Index)值,基于各个AP的APEI值及网络资源拥塞状况,执行切换决策,实现重叠覆盖区域内的业务负载动态预先均衡。通过实验发现,本文提出的终端服务感知动态负载均衡机制节省了网络通信资源,提高了终端用户对网络服务的满意度。

3 终端服务感知的动态负载均衡机制

本节首先引出网络模型和相关定义,用于机制描述;然后在此基础上,给出终端接入点APEI的模糊化推算过程;最后提出终端服务感知的动态负载均衡机制。

3.1 网络模型和相关定义

目前的移动互联网,通过不同接入技术的无线网络相互交叠覆盖形成。保持现有无线网络结构的完整性,增加某些中间节点实现无线网络融合,是一种为移动用户提供无缝服务的主要技术方案,它可以分为紧耦合和松耦合^[3]两种模式。考虑到网络无缝服务性能及资源的优化利用,本文采用图1所示的松耦合网络拓扑结构。

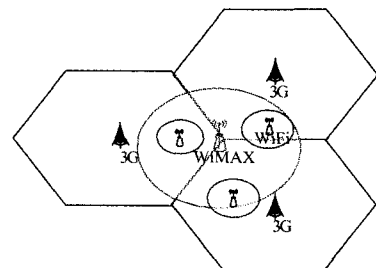


图1 异构无线网络融合模型

在该拓扑结构中,3G蜂窝网基站、WiMAX无线城域网接入点、WLAN无线局域网WiFi热点相互独立,使用各自的人网机制并通过各自的网关与因特网相连;各网络的无线信号相互交叠覆盖,形成了一个新的无缝服务区域。各个异构网络的接入点记为 AP_i^T (T 为网络类型标识, $T=0,1,2$,即3G网络、WiMAX网络及WLAN网络分别用0、1、2标识; $i=1,2,\dots,I$, I 表示重叠覆盖域内的无线接入点个数,即终端用户候选列表内的可选接入点个数)。

移动终端选择合适的切换网络时,会综合考虑自身的条件与需求,如终端自身的能量、成本、接收信号强度及 QoS 需求等信息。考虑到决策参考因子的复杂性及模糊性,本文采用图 2 所示的模糊推理模型,对用户的模糊输入参数进行推理评价。根据该模型,首先基于经验拟合构建 IF-THEN 模糊规则集;然后将需要考虑的因素作为模糊推理系统的输入,进行模糊化处理;最后根据规则库中定义的对应不同输入组合的规则,得到模糊化的决策结果。在此,利用模糊逻辑控制模型,对终端用户输入的模糊决策因子执行推理判决,计算出各个用户对各个 AP 的满意隶属度,完成模糊量到精确量的转化,计算出各候选网络的 APEI 值。

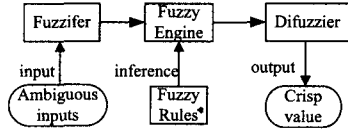


图 2 模糊逻辑控制架构

移动互联网中,从网络与用户的角度出发,可以划分出以网络或终端为核心的若干切换执行策略。本文采用图 3 所示的以终端为核心的 NAHO(network assisted handover^[3],网络协助切换)切换执行模式。终端经过切换发现、切换决策与切换执行 3 个过程,在网络负载预先均衡的前提下,自主选择具有最佳用户满意度的接入网络,实现网络通信资源预先均衡分配,完成网络超前切换。

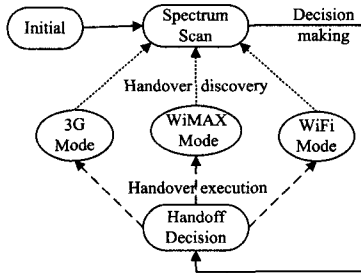


图 3 终端自主决策模型

另外,为了便于机制描述,将本文涉及的部分符号标记列于表 1 中。

表 1 本文使用的符号标记

参数符号	参数意义
MU	主要潜在用户, Underlying Main User
Best _{AP}	终端业务到达时提供业务链接的首选 AP
$P_{block,i}(t)$	AP _i 在 t 时刻的业务呼叫阻塞概率
APAV	网络候选接入点 AP 的接入向量
APEI	网络候选接入点的 AP 的接入满意值
δ	用户满意度的最小门限值
$\mu_i^k(t)$	第 i 个网络中第 k 个评价因子的隶属函数
ω_j	接入网络第 j 个指标的加权因子
λ_0	呼叫服务率
$\rho_i(t)$	t 时刻 AP _i 中业务的呼叫服务比, $\rho_i(t) = \lambda_0 / \theta_i(t)$
$\theta_i(t)$	t 时刻 AP _i 中业务的业务服务率
$f(\lambda_0)$	业务到达概率门限值, λ_0 为标准呼叫率

3.2 终端期望指数模糊化运算

为了选择最佳期望指数值 APEI 的 AP 来建立连接,本文参照文献[10,11]中的方法,对接收信号强度 RSSI、电池剩余能量、网络计费、网络负载、QoS 等参数进行模糊化评估,计算出终端接入各个 AP 的 APEI。详细的模糊化计算过程描述如下。

为了衡量重叠覆盖区域内终端用户接收到的网络信号的强弱,定义截下型接收信号强度隶属函数为:

$$\mu_i^1(t) = \begin{cases} 0, & R_i(t) < R_i^{\text{th}} \\ \frac{R_i(t) - R_i^{\text{th}}}{R_i^{\text{max}} - R_i^{\text{th}}}, & R_i^{\text{th}} \leq R_i(t) < R_i^{\text{max}} \\ 1, & R_i(t) \geq R_i^{\text{max}} \end{cases} \quad (1)$$

式中, $i=1,2,\dots,n$; $R_i(t)$ 为终端用户在 t 时刻接收到 AP_i 的信号强度; R_i^{th} 为 AP_i 发出信号强度的网络门限值; R_i^{max} 为 AP_i 的最大信号发射强度值。用户根据当前接收到的信号强度值,计算自身接收信号强度的隶属度。

终端电池剩余能量越低,用户获得网络服务的期望值就越低。因此,终端剩余电量多少的隶属度函数可定义为:

$$\mu_i^2(t) = \begin{cases} 0, & e_i(t) < \epsilon E_i \\ \frac{e_i(t) - \epsilon E_i}{E_i(1 - \epsilon)}, & \epsilon E_i < e_i(t) < E_i \end{cases} \quad (2)$$

式中, $e_i(t)$ 为终端在 t 时刻的电池可用电量; E_i 表示终端电池的额定电量; ϵ 表示终端的最低工作门限值指数,一般取值 5%~10%。终端能量越充裕,期望获得网络服务的愿望就越强烈。

网络服务费用是终端用户选择切换网络时的一个重要参考因素。对于用户,服务花费越低,入网期望就越大。因此,用户资费高低的截上型隶属度函数表示为:

$$\mu_i^3(t) = \begin{cases} 0, & m_i(t) \geq M_i^{\text{max}} \\ \frac{M_i^{\text{max}} - m_i(t)}{M_i^{\text{max}} - M_i^{\text{min}}}, & M_i^{\text{min}} \leq m_i(t) \leq M_i^{\text{max}} \\ 1, & m_i(t) \leq M_i^{\text{min}} \end{cases} \quad (3)$$

式中, $m_i(t)$ 为用户位于网络 AP_i 内的预期网络费用,即 MU 用户业务到达时的预估开销; M_i^{min} 、 M_i^{max} 表示网络 AP_i 提供的最低、最高服务资费。对于免费的 AP,相应的隶属度取最大值。

终端用户选择最佳接入网络时,相应的 QoS 保障是一个重要的决策参考指标。考虑到计算复杂性,对于网络提供的资源消耗性业务,仅从各 AP 的可用带宽和连接时延两个方面来考察该网络提供的 QoS 期望值。采用指派法来确定各 AP 的 QoS 模糊隶属函数 $\mu_i^4(t)$,定义 AP 的 QoS 论域为:

$$QoS \in \{bw, dl\} \quad (4)$$

算符“ ∞ ”定义为 QoS 论域由可用带宽(available bandwidth)和连接时延(link delay)构成,其中带宽与时延输入参数的模糊程度描述因子定义为: L-A-H(Low, Average, High) 3 个状态,如图 4 所示。

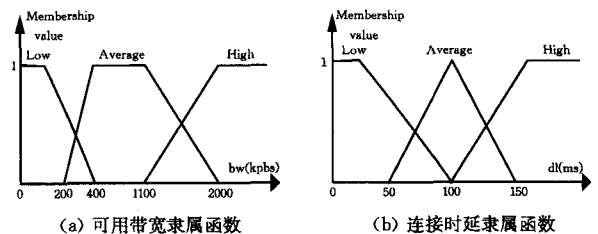


图 4 输入参数模糊化隶属函数

QoS 模糊程度的评价因子设置为 VL(Very Low)、LL(Low Level)、AL(Average Level)、HL(High Level)、VH(Very High) 5 个等级:

$$QoS = \{VL, LL, AL, HL, VH\}$$

采用重心法进行反模糊化处理,对应的量化数字分别为

{0.1, 0.3, 0.5, 0.7, 0.9}。通过量化数字确定 QoS 的隶属度。根据经验,针对资源消耗性业务,建立 Mamdani 模糊推理系统的 if-then 准则(见表 2),完成模糊输入因子的模糊化输出。

表 2 模糊准则

If-Then Rules				
RULE	bw	dl	QoS	数字量化
RULE1	L	L	LL	0.3
RULE2	L	A	LL	0.3
RULE3	L	H	VL	0.1
RULE4	A	L	HL	0.7
RULE5	A	AND	AL	0.5
RULE6	A	H	LL	0.3
RULE7	H	L	VH	0.9
RULE8	H	A	HL	0.7
RULE9	H	H	AL	0.5

根据模糊准则表和用户输入的 QoS 模糊评价因子,推导出对应的 QoS 服务等级,生成 QoS 数字量化值,确定各个 AP 的 QoS 隶属度。

基于上述隶属函数,计算出各个 AP_i 对应的不同隶属度矩阵,如表 3 所列。

表 3 各个接入点隶属度

	AP ₁ ^T	AP ₂ ^T	...	AP _n ^T
接收信号强度	$\mu_1^1(t)$	$\mu_2^1(t)$...	$\mu_n^1(t)$
剩余能量	$\mu_1^2(t)$	$\mu_2^2(t)$...	$\mu_n^2(t)$
接入成本	$\mu_1^3(t)$	$\mu_2^3(t)$...	$\mu_n^3(t)$
QoS 需求	$\mu_1^4(t)$	$\mu_2^4(t)$...	$\mu_n^4(t)$

对于网络 i ,定义其接入向量 $APAV_i$,如式(5)所示:
 $APAV_i = [\mu_1^i(t), \mu_2^i(t), \mu_3^i(t), \mu_4^i(t)]^T$ (5)

用户的候选 AP 的候选接入矩阵 $APAV$ 表示为:
 $APAV = [APAV_1, APAV_2, \dots, APAV_n]$ (6)

终端在执行切换运算时,网络环境、用户的需求及偏好不断变化。为了反映决策因子之间的相关性权重差异,从用户角度出发,定义一个用户偏好加权向量 W ,如式(7)所示:

$$W = [\omega_1, \omega_2, \omega_3, \omega_4] \quad (7)$$

其中

$$\omega_j = \frac{\sigma_j}{\sum_{k=1}^4 \sigma_k}, j \in \{1, 2, 3, 4\} \quad (8)$$

式中, σ_j 为网络评价指标 j 的隶属度的标准差。

$$\sigma_j = \sqrt{\frac{1}{n-1} \sum_{k=1}^n [\mu_k^j(t) - \frac{1}{n} \sum_{k=1}^n \mu_k^j(t)]^2} \quad (9)$$

通过式(6)和式(7),用户可以计算出每个接入网络 AP 的用户满意向量 $APEI_i$,如式(10a)所示:

$$APEI_i = W \cdot APAV_i = \sum_{k=1}^4 (\omega_k \cdot \mu_k^i(t)) \quad (10a)$$

因此,对于用户候选列表中的所有候选 AP,可以得到一个用户满意度向量 $APEI$,如式(10b)所示:

$$APEI = \{APEI_1, APEI_2, \dots, APEI_n\} \quad (10b)$$

在确保重叠覆盖区域内用户负载动态均衡条件下,终端从用户满意度向量 $APEI$ 中选取具有最大满意度值的 AP 作为该用户的接入 AP,为用户预留网络通信资源,实现基于用户满意度的网络负载均衡。

3.3 终端自主决策

针对终端业务在请求资源时存在的差异性,本文在 DLB-MSA 的基础上提出了移动终端服务感知动态负载均衡机制

SA-DLB,由终端完成对高 APEI 值、低资源拥塞 AP 的自主选择,详细过程如下。图 5 给出了该机制的算法流程。

算法 1 终端服务感知的 SA-DLB 算法

```

1 MU roaming a new Overlap Hot Area
2 Exchange Network Signalling and Load the context information
3 Estimating the call arrive probability  $f(t)$  in the future
4 if  $f(t) < F\_th$  then
5   connected to the old AP
6 else
7   label MU as MUU
8   if  $V(t) > Velocity\_th$  then
9     connected to the old AP
10  else
11    calculating APEI for current AP
12    if  $APEI > APEI\_th$ 
13      connected to the old AP
14  else
15    load candidate AP from Network Interface Manager
16    calculating APEI for each AP
17    descending sort the candidate AP_list based on APEI
18    calculating Pblocki(t) for each candidate AP
19    while  $i \leftarrow 1$  to  $I_{his}[apei]$ 
20      temp1  $\leftarrow$  Pblocki(t)
21      temp2  $\leftarrow$  Pblockmin(t)
22      if (MIN(|temp1-temp2|) is true) then
23        selecting APi as the Best_AP
24      endif
25    endwhile
26  endif
27 endif
28 endif

```

图 5 终端服务感知的 DLB 算法

切换发现过程,分为两个阶段:

1)推算终端用户业务阻塞概率阶段。当终端用户 MU (Mobile User)漫游到一个新的局部热点网络的重叠覆盖区域 AP_i 内,通过与各网络的信令交换,获取网络服务性并将该用户标记为 MU_j^i (即位于网络 i 内的第 j 个用户)。本文采用了以用户为中心的切换模型,在执行网络选择性切换之前,先采用文献[8]中的方法推断出用户自身在将来某时段 $T_0 = 1/\lambda_0$ 内某资源消耗型业务到达的概率值 $f_j(\lambda_0)$,并将其与各网络广播的最大门限值 F_th 相比较,确定该 MU_j^i 用户是否为 MUU:若满足

$$f_j(\lambda_0) > F_th$$

则确定该用户为 MU_j^i ,然后使用自主决策的服务感知机制选择合适的接入 AP;反之,对于非 MU_j^i 用户,由于其业务到达可能性很小,则直接使用 MU_j^i 原来的 3G 或 WiMAX 的 AP 建立业务链接,避免冗余切换带来的终端资源消耗。

2)评估当前 AP 的 APEI 是否满足用户的需求,避免发生不必要的切换。对于高业务到达概率 MU_j^i 用户,根据终端用户的移动速度,滤除重叠覆盖区域内驻留时间很短的高速用户 MU_j^i ,即由高速 MU_j^i 用户原来的 AP 为其服务。对于低速 MU_j^i 用户,计算出当前 AP 的 APAV 值与 APEI 值:若满足

$$Current_APEI > \delta$$

则以当前 AP 作为该 MU_j^i 的 Best_AP,为 MU_j^i 的通信链接建立提供服务;否则,建立决策准则,实现终端服务感知的动态负载均衡机制,为 MU_j^i 选择 Best_AP。

准则建立过程中,终端根据各网络提供的信息,建立切换选择的决策规则,分为两个阶段:

1)计算候选网络的 APEI。若非上述情况, MU_j^i 用户获取终端接口管理器中的所有候选 AP,采用上一节介绍的模糊逻辑推理算法,计算各个备选 AP 的 APAV 向量与用户期望指数 APEI 值。再根据各个 AP 的 APEI 值的大小, MU_j^i 用户将大于给定 APEI 阈值 δ 的若干候选 AP 进行降

序排列,建立一个基于 APEI 值的 AP 降序列表 $List[apei]$ 。

2) 建立决策准则。根据 APEI 值得降序列表,从第一个具有最大 APEI 值的 AP 开始遍历整个列表:若用户预选 AP_i 作为自己的最佳接入 AP,则应使得 AP_i 与列表中除 AP_i 外的所有 AP 中具有最小业务阻塞的 $AP_j (i \neq j)$ 之间的业务阻塞概率差的绝对值取得最小值,以确保由 AP_i 为该用户提供业务链接服务后不会造成网络自身的资源饱和或拥塞加剧,从而实现局部业务负载的动态均衡。因此,异构无线网络中,为实现终端服务感知的动态负载均衡机制,可建立如下的切换规则:

$$\begin{aligned} & \text{MIN} |P_{block,i}(t) - P_{block,min}(t)| \\ \text{s. t.} & \begin{cases} APEI_i > APEI_{i+1} > \delta, i=1, 2, \dots, l_{List[apei]} \\ P_{block,min}(t) = \underset{j=1, j \neq i}{\text{MIN}} P_{block,j}(t) \end{cases} \end{aligned}$$

且 AP_i 的业务阻塞概率基于 $M|M|n_i(n_i, N_i)$ 排队服务模型计算^[8],如式(11)所示:

$$P_{block,i} = \frac{[\rho_i(t)]^{n_i} / [(n_i)!]}{\sum_{r=0}^{n_i} [\rho_i(t)]^r / (r!)} \quad (11)$$

式中, n_i 表示 AP_i 内的截止队列长度。

决策执行过程中,终端自主选择符合自身需要的接入 AP。对于满足式(11)的 AP_i ,则 MUU 用户使用该候选 AP 作为其 $Best_AP$,为 MUU 的业务链接预留通信资源,完成资源预留条件下的网络动态负载均衡。整个负载均衡过程采用 NAH(Network Assisted Handover)的切换模式。

终端服务感知的动态负载均衡机制以用户为核心,在网络协助下实现终端对无线接入点的自适应选择,完成网络资源的提前部署。在选 $Best_AP$ 过程中,充分考虑了终端用户的接收信号强度的环境信息、终端电池剩余能量、服务花费及服务质量。用户进入网络单元覆盖区域后,根据模糊算法推断每个 AP 的 APAV 值和 APEI 值,在保证局部负载均衡时,筛选出高 APEI 值的 AP 作为终端用户的 $Best_AP$ 。该机制采用类似贪心算法的思想,不仅缓解了网络中局部热点 AP 之间的负载过重带来的网络性能下降,在一定程度上降低了热点地区的业务切换率,而且以用户的接入期望值作为约束条件,提高了终端用户的网络服务体验。后文将通过仿真验证该机制的有效性。

4 性能评价

本节首先介绍实验评估模型与参数设置,然后对实验结果进行分析。

4.1 性能评估模型与实验设计

为了对本文提出的服务感知动态负载均衡机制 DLB-SA 的性能进行全面的分析和评价,采用高斯-马尔科夫移动模型^[12] (Gauss-Markov Model)对重叠覆盖区域内的移动用户进行模拟,刻画终端用户位置变化。另外,实验场景不妨假设如下:结合现实状况,考虑由 1 个 3G 网络、1 个 WiMAX 网络与 3 个 WLAN 网络在松耦合模式下形成交织重叠覆盖,并规定 3G 网络在区域内实现全境覆盖,WiMAX 在区域内大面积覆盖,WLAN 网络随机小范围覆盖。在某个高峰期,大量用户进入该无线异构网络覆盖区域,并随意活动。用户在重叠区域内,基于本文提出的自主决策机制,选取最合适的接入网

络。表 4 给出了仿真参数及其取值。

表 4 实验参数设置

	3G	WiMAX	WLAN		
	gAP	wAP	AP1	AP2	AP3
RSSI	-20~-125dBm	-30~-110dBm	-45~-95dBm		
E_s	25				
CS	0.9	0.5	0.2		
dw	0~2.5Mbps	4Mbps	0~11Mbps		
bl	0~200ms	0~200ms	0~50ms		
λ_0	0.5				
$\theta_i(t)$	1.2	1.6	1.5		
n_i	32	20	10		
δ	0.4				
F_th	0.35				

这里分别考虑 3G/WLAN 网络在不同的重叠覆盖半径和移动用户数下,使用自主决策机制前后的用户满意度及网络资源变化。

4.2 结果分析

通过仿真实验的对比分析,验证使用服务感知动态负载均衡机制 SA-DLB 前后,重叠覆盖区域内的用户满意度变化及各个无线接入点的网络拥塞变化趋势。

图 6 给出了运行终端自主决策前后,终端用户期望指数统计分布的仿真结果。图(a)反映了服务感知前后用户 $apei$ 的概率密度变化。从图中可以看出,按照 Non-APEI、APEI 的顺序,其边缘函数的“主跳”从左向右移动;终端用户在选择切换网络时,如果采用 Non-APEI 算法,则用户所选 AP 的 $apei$ 值的峰值区间分布在较宽泛的 0.2~0.9 之间,且相对平稳, $apei$ 值在 0-1 区间内的分布趋于均衡。采用 APEI 算法,基于终端用户的期望指数 $apei$ 进行切换网络的自适应选择,建立服务感知的动态负载均衡机制,用户对其所选择 AP 的 $apei$ 最高峰值区间集中分布在 0.5~0.85 之间,并且在该区间内其主跳斜率更陡峭,以保证更高的用户满意度。图(b)中的 CDF 图反映了用户期望指数的概率分布变化,采用 Non-APEI、APEI 算法的 CDF 曲线分别在 0.2、0.5 左右开始逐渐上升,说明采用 APEI 算法后终端用户获得高 $apei$ 值的概率更大,相较于网络切换时未考虑用户满意度 APEI 算法的情形,大量用户可以获得具有较大 $apei$ 值的最佳接入 AP。如当 $apei=0.7$ 时,采用 SA-DLB 机制时,所占的用户数约 50%,与未考虑 APEI 算法的 DLB 机制 30% 的用户数相比,服务用户数大约提高了 20%。同时,在相对平稳的期望指数值区间内,与未考虑 APEI 算法的 DLB 机制相比,采用 SA-DLB 机制可以服务大量终端用户,为用户提供了高网络服务体验。

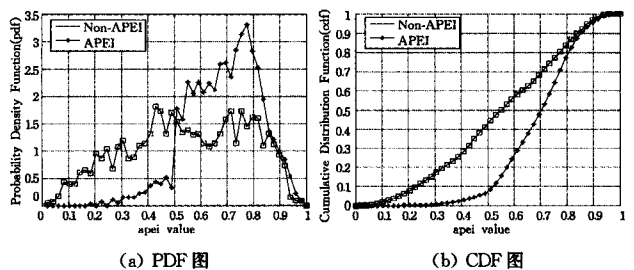


图 6 自主决策前后用户期望指数统计分布图

图 7 中的(a)、(b)分别刻画了 3G 与 WiMAX 网络资源饱和和发生状况。从图中可以看出,使用 DLB 机制实现松耦合网络动态负载均衡以选择合适的接入 AP,与未使用任何机制相比,在一定程度上降低了网络拥塞,均衡了网络的潜在用户

负载;将运行终端自主决策的动态负载均衡机制后的效果与DLB机制的效果相比较可以发现,网络拥塞概率得到了更大程度的降低,网络工作效率得到了提高。使用自主决策态负载均衡机制,用户根据自己的个性化需求,在实现局部网络负载均衡条件下,选择具有最佳满意度值的网络作为最佳接入AP,在获得高用户满意度的同时,延长了网络服务周期,提高了网络的无线资源利用率。

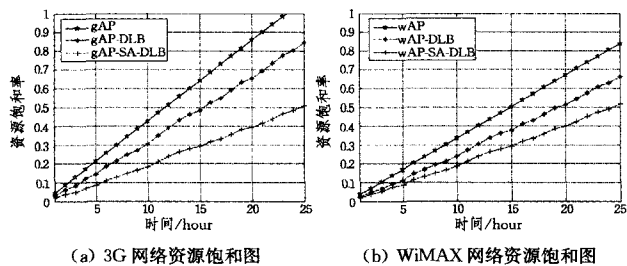


图7 3G、WiMAX资源饱和图

图8给出了WLAN的各个AP的网络资源拥塞状况的仿真结果。从图中可以发现,相比未使用任何机制的情形,只使用DLB机制实现基于资源预留的异构网络切换明显降低了WLAN局部热点网络工作负荷;引入SA-DLB机制,在实现负载均衡时,通过选择具有较大用户接入指数值的网络作为用户的最佳接入AP,提高了用户的网络服务体验;并且,对位于WLAN局部热点中的某个用户,由于AP的APEI较小,用户转而选择重叠覆盖区域内具有更高APEI值的网络,为其预留资源预留,从而明显降低了各无线接入点AP的工作负荷。结合图7的资源饱和图可知,使用终端自主决策的DLB机制,在选取高用户满意度值的接入网络的同时有效降低了网络的工作负荷,实现了局部的网络负载均衡,提高了网络资源利用率,改善了网络服务性能。

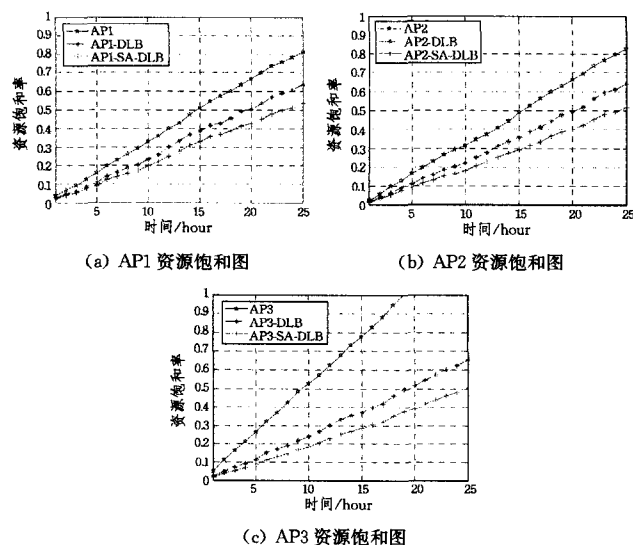


图8 WLAN各个AP资源饱和图

通过上面的仿真实验结果可知,终端在建立基于资源预留的DLB机制时,结合终端对各个网络的评价指标,即各个网络的APEI值,可以有选择性地获取具有较高较平的用户满意的AP,达到用户负载的动态均衡。较未考虑用户满意度的DLB机制,终端SA-DLB机制进一步缓解了各个AP的资源拥塞,推迟了资源饱和点出现的时间,同时保障了用户对所选的接入网络的满意度。

结束语 移动互联网中,负载均衡是改善网络性能的

一种有效方式。然而,现有的负载均衡机制大多从网络自身利益的角度出发,没有考虑终端用户偏好。针对这一问题,本文提出了一种终端服务感知态负载均衡机制,即基于资源预留与网络负载预先均衡思想,采用模糊算法对终端用户的网络接口中的候选网络进行预先评估并推算出其APEI值,根据各AP的资源拥塞状况和其APEI值,由终端选择最佳切换网络实现负载均衡机制,完成基于终端满意度的网络通信资源预先规划,达到网络中用户负载的动态均衡。重叠覆盖区域内的移动用户,根据各网络的业务阻塞概率与终端对各网络的期望值,在实现局部负载均衡时,尽可能选择具有一定期望值的网络作为其业务链接预留通信资源,从而较好地保障用户的网络服务体验,增强网络服务能力。仿真结果表明,该机制不仅保证了用户的接入网络满意度,而且减少了融合网络中的各个无线接入点的网络工作负荷,有效降低了网络接入点的资源拥塞概率,延长了网络服务生命周期。

参考文献

- [1] Yan X, Sekercioglu Y A, et al. A survey of vertical handover decision algorithms in forth generation heterogeneous wireless networks[J]. *Computer Networks*, 2010, 54(11): 1848-1863
- [2] Zukri M, Jouaba B, Zeghlache D. A review on mobile management and vertical handover solutions over heterogeneous wireless networks[J]. *Computer Communications*, 2012, 35(17): 2055-2068
- [3] Nkansah-Gyekye Y, Abginya J I. Vertical handoff decision algorithm for UMTS-WLAN[C]// (AusWireless 2007) The 2nd International Conference on Proc. Wireless Broadband and Ultra Wideband Communications, 2007. 2007: 37-42
- [4] Singhrova A, Prakash D N. A Review of Vertical Handoff Decision Algorithm in Heterogenous Networks[C]// Proc. the 4th ACM International Conference on Mobile Technology, Applications, and Systems. New York, ACM Digital Library, 2007: 68-71
- [5] Shahgholi B, Ghahafarokhi, Movahhedinia N. A context-aware handover decision based on user perceived quality of service trigger[J]. *Wireless Communication and Mobile Computing*, 2011, 11(6): 723-741
- [6] He Qing. A fuzzy logic based vertical handoff decision algorithm between WWAN and WLAN[C]// 2010 2nd International Conference on Proc. International Conference on Networking and Digital Society (ICNDS). Wenzhou, 2010: 561-564
- [7] 马飞, 杨飞霞, 徐光宪. 网络与终端协同选择及切换机制的研究[J]. *计算机科学*, 2012, 39(1): 77-81
- [8] 刘琪, 袁坚, 山秀明, 等. 3G/WLAN网络中基于终端移动与业务认知的动态负载均衡机制[J]. *计算机学报*, 2010, 33(9): 1569-1579
- [9] Ciubotaru B, Manteau G M. Sasha-A quality-oriented handover algorithm for multimedia content delivery to mobile users[J]. *IEEE Transactions on Broadcasting*, 2009, 55(2): 437-450
- [10] Liao Hong-wei, Tie Ling, Du Zhao. A Vertical Handover Decision Algorithm Based on Fuzzy Control Theory[C]// Proc. The First International Multi-Symposiums on Computer and Computational Science. Hanzhou, 2006: 309-313
- [11] Shen Wei, Zeng Qing-an. Cost-Function-Based Network Selection Strategy in integrated Wireless and Mobile Networks[J]. *IEEE Transactions on vehicular thechnology*, 2008, 57(6): 3778-3788
- [12] 童超, 牛建伟, 龙翔, 等. 移动模型研究综述[J]. *计算机科学*, 2009, 36(10): 5-11