

基于马尔可夫决策模型的异构无线网络切换选择算法

梁潇 钱志鸿 田洪亮 王雪

Markov decision model based handoff selection algorithm for heterogeneous wireless networks

Liang Xiao Qian Zhi-Hong Tian Hong-Liang Wang Xue

引用信息 Citation: *Acta Physica Sinica*, 65, 236402 (2016) DOI: 10.7498/aps.65.236402

在线阅读 View online: <http://dx.doi.org/10.7498/aps.65.236402>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2016/V65/I23>

您可能感兴趣的其他文章

Articles you may be interested in

简谐振子免疫优化算法求解异构无线网络垂直切换判决问题

Simple harmonic oscillator immune optimization algorithm for solving vertical handoff decision problem in heterogeneous wireless network

物理学报.2012, 61(9): 096401 <http://dx.doi.org/10.7498/aps.61.096401>

混沌量子克隆算法求解认知无线网络频谱分配问题

Chaos quantum clonal algorithm for spectrum allocation of cognitive wireless network

物理学报.2011, 60(6): 068803 <http://dx.doi.org/10.7498/aps.60.068803>

简谐振子免疫优化算法求解异构无线网络垂直切换判决问题

Simple harmonic oscillator immune optimization algorithm for solving vertical handoff decision problem in heterogeneous wireless network

物理学报.2012, 61(9): 096401 <http://dx.doi.org/10.7498/aps.61.096401>

大型水轮发电机冷却方式综合评价方法的研究

Comprehensive evaluation method in the cooling mode of large-scale hydro-generators

物理学报.2015, 64(17): 178802 <http://dx.doi.org/10.7498/aps.64.178802>

基于马尔可夫决策模型的异构 无线网络切换选择算法*

梁潇¹⁾ 钱志鸿^{1)†} 田洪亮¹⁾²⁾ 王雪¹⁾

1)(吉林大学通信工程学院, 长春 130012)

2)(东北电力大学信息工程学院, 吉林 132012)

(2016年7月26日收到; 2016年9月7日收到修改稿)

针对异构环境下不同业务类型用户对于接入网络的不同服务质量 (quality of service, QoS) 需求, 该文提出了一种基于马尔可夫决策模型的切换选择算法. 建立基于软件定义网络 (software defined network, SDN) 的异构无线网络架构, 以实现异构网络的通透控制. 利用马尔可夫过程预测下一时刻的网络状态以得到采取动作后的一次回报, 依据网络的不同状态属性针对实时用户和非实时用户分别构建立即回报函数, 并采用层次分析法确定属性权重; 基于状态动作对构建期望回报函数, 采用逐次逼近的迭代方式得到使长期期望回报最大的切换策略. 仿真结果表明, 该方法针对不同业务类型用户均能选取最优切换策略, 同时降低阻塞率, 提高了用户的 QoS 和无线网络的资源利用率.

关键词: 异构无线网络, 切换选择, 马尔可夫过程, 层次分析法

PACS: 64.60.aq, 89.70.-a, 02.50.Ga, 87.55.kd

DOI: 10.7498/aps.65.236402

1 引言

未来无线通信网络正朝着泛在化、扁平化、全 IP 化的方向快速发展, 随着移动通信和信息技术等不同领域技术的相互结合, 信息将渗透到人们日常生活的方方面面. 下一代无线网络的发展趋势已不再是一种技术一统天下, 而是多种无线接入技术 (radio access technologies, RAT) 并存, 且能满足终端用户业务多样性需求的异构无线网络^[1,2]. 不同类型网络的融合为人们提供了多样化的接入方案, 使“在任何时间、任何地点与任何人进行某种类型的信息交换”逐渐成为可能. 然而, 由于异构无线网络覆盖情况的复杂性、接入技术的差异性、终端用户业务类型的多样性, 这必将给异构网络的有效融合与协同工作带来极大的挑战. 异构

环境下的垂直切换能有效实现网络间的互通, 是异构网络融合的重要基础, 也迅速成为了研究热点之一^[3]. 由于不同无线网络在覆盖范围、负载能力、业务支持以及传输速率等方面都存在较大差异, 如何随时随地地依据用户需求选择最合适的切换网络, 以最大化用户的服务质量 (quality of service, QoS) 和无线网络的资源利用率, 是一个亟待解决的关键问题^[4-6].

目前, 对于异构无线网络中的切换选择问题已有广泛研究, 文献^[7, 8]中提出了一种基于预测接收信号强度 (received signal strength, RSS) 的垂直切换算法, 当备选网络 RSS 值接近已设定阈值时进行预切换判决, 并设置自适应驻留时间以防止网络频繁切换. 文献^[9, 10]中将异构无线网络中的网络选择问题建模为一个博弈模型, 文献^[9]中博弈的参与者是在不同网络覆盖下的用户群, 并分别采用

* 国家自然科学基金 (批准号: 61371092)、国家自然科学基金青年科学基金 (批准号: 61401175)、吉林省重点科技攻关项目 (批准号: 20140204019GX) 和长春市重大科技攻关计划 (批准号: 2014026/14KG021) 资助的课题.

† 通信作者. E-mail: dr.qzh@163.com

种群进化和强化学习算法来达到博弈的动态均衡, 以避免因网络拥塞而导致选择性能的下降; 文献 [10] 中则认为博弈的参与者之间信息封闭, 并建立不完全信息的博弈模型, 使网络选择策略收敛于此博弈的纳什均衡. 文献 [11] 中以影响网络选择的多个属性为判决指标, 采用基于多属性决策 (multiple attribute decision making, MADM) 的方法进行垂直切换选择, 并利用数值模拟对七种不同 MADM 方法进行性能评估以确定不同环境下的最优评价准则. 文献 [12, 13] 中提出了一种基于模糊神经网络的垂直切换算法, 分别结合粒子群算法及能效效用函数为模糊神经网络的输入设定初值, 由训练好的神经网络自适应地作出决策以达到异构无线网络的负载均衡. 文献 [14] 中提出了一种基于简谐振子免疫优化算法的垂直切换判决方案, 能有效平衡网络负载、增加终端电池的生存时间. 上述文献分别从不同角度提高了异构无线网络的切换性能, 但都只以当前时刻的网络状态为判决指标, 而未从长远角度来选取最优切换网络以提高整个异构无线网络的资源利用率.

针对以上不足, 文献 [15] 中将异构网络中的垂直切换问题建模为一个马尔可夫决策过程, 以切换收益和开销来构建每个连接的有限阶段总期望回报, 并以数值迭代算法确定切换选则策略. 文献 [16, 17] 中考虑网络状态的实时变化, 以马尔可夫过程来预测下一时刻的网络状态, 并采用模糊逻辑的方法构建效用函数, 以最大化效用函数为网络切换准则. 但上述文献均未考虑终端用户的业务多样性需求, 且只从网络状态出发来进行切换选择决策. 因此, 本文提出了一种基于马尔可夫决策模型的异构无线网络切换选择算法, 在基于软件定义网络 (software defined network, SDN) 的异构无线网络架构下, 利用马尔可夫过程对下一时刻的网络状态进行预测. 考虑到不同业务类型用户对于切换选择具有不同的 QoS 需求, 分别针对不同用户构建立即回报函数并采用层次分析的方法确定各状态属性权重, 从而得到采取动作后的一次回报. 从长远角度来看, 以立即回报函数序列组成的目标函数来衡量将来长期回报, 从状态动作对出发逐次对期望回报进行迭代, 得到使长期期望回报最大的切换策略. 仿真结果表明此方法可以有效降低终端用户的阻塞率, 并针对不同类型用户均能选择最优切换策略, 有效提高异构无线网络的资源利用率.

2 系统模型

采用实际部署中典型的 LTE/WLAN 异构无线网络架构, 同时引入 SDN 这种数据控制分离、软件可编程的新型网络体系 [18,19], 通过控制-转发接口实现控制平面的集中式控制, 并由软件编写的方式灵活定义网络设备的转发功能, 以实现异构无线网络通信的通透性. 系统架构如图 1 所示, 由 SDN 控制器掌握异构无线网络的实时状态信息, 根据各网络性能特点并结合自身业务需求动态地调度网络资源, 实现网络的最佳切换选择. 其中, SDN Agent 软件模块, 遵守 SDN 协议并负责和 SDN 控制器进行信息交互, 通过基站与无线 AP 实时地将本地备选网络拓扑、状态信息和终端用户信息向 SDN 控制器上报, 由 SDN 控制器更新全局视图数据库, 对上报信息进行存储. 同时, 网络信息处理模块提取出相应 RAT 的网络参数和状态信息, 如: 网络带宽、时延、时延抖动、误码率和网络负载等, 并对其进行分析处理, 根据终端用户访问的业务类型和网络状态等条件进行切换选择并执行.

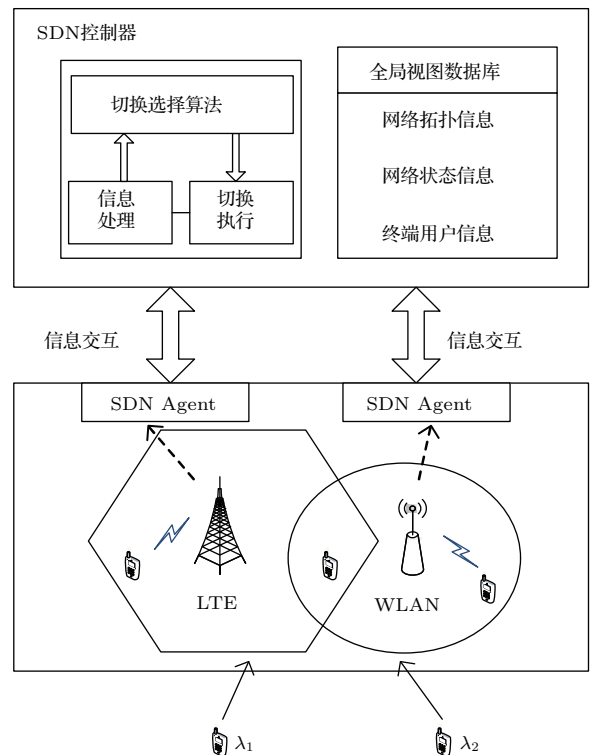


图 1 基于 SDN 的异构无线网络架构

Fig. 1. Heterogeneous wireless network architecture based on SDN.

移动终端用户具有多模接入能力, 可接入到存在不同 RAT 的异构无线网络中, 采用随机服务系

统理论来描述此排队现象 [20,21]. 假设实时业务用户到达率服从参数为 λ_1 的泊松分布, 即实时业务的第 k_1 个呼叫发起概率为: $P_1^{k_1} = (\lambda_1^{k_1} e^{-\lambda_1})/k_1!$; 非实时业务用户到达率服从参数为 λ_2 的泊松分布, 即非实时业务的第 k_2 个呼叫发起概率为: $P_2^{k_2} = (\lambda_2^{k_2} e^{-\lambda_2})/k_2!$, 实时业务用户与非实时业务用户接入异构无线网络的驻留时间分别服从参数为 μ_1 和 μ_2 的负指数分布. 若异构无线网络中存在 M 个备选网络, 因备选网络占用而导致不同业务类型用户被阻塞的状态概率分别为:

$$P_1 = \frac{\alpha_1^M / M!}{\sum_{k_1=0}^M \frac{\alpha_1^{k_1}}{k_1!}}, \quad P_2 = \frac{\alpha_2^M / M!}{\sum_{k_2=0}^M \frac{\alpha_2^{k_2}}{k_2!}},$$

其中 $\alpha_1 = \lambda_1 / \mu_1$, $\alpha_2 = \lambda_2 / \mu_2$.

3 基于马尔可夫决策模型的切换选择算法

当不同业务类型的用户分别到达异构无线网络区域时, SDN 控制器实时根据各网络的状态信息和用户的业务类型做出相应的切换选择, 包括是否阻塞该用户或切换至何种网络. 若以 τ 为时间间隔对网络状态信息进行采样, 下一时刻网络状态只与当前网络状态与行动有关, 而与历史状态无关, 则将异构无线网络中的切换选择问题建模为一个时间离散、状态连续的马尔可夫过程. 选取多个网络参数来描述采取动作后的立即回报函数, 根据各个属性参数在不同业务类型用户到来时对切换选择的影响程度, 确定属性权重, 从而得到衡量切换选择性能的一次回报. 为了更好地优化切换选择的 QoS, 考虑由每个采样时刻的立即回报函数构成回报函数序列, 最大化回报函数序列之和的数学期望, 以得到从长远角度来看的最优切换决策.

3.1 马尔可夫过程

马尔可夫过程是一个离散时间的随机过程 [22,23], S 为有限的网络状态空间; T 为所有决策时刻的点集; A 为有限的系统动作空间; $P(s, a, s') \in [0, 1]$ 为状态转移概率, 表示在状态 s 下选择动作 a 后使网络状态转移到状态 s' 的概率.

假设异构无线网络区域内并列 M 个网络, 考虑影响切换选择 QoS 的五个因素: 时延 D 、时延抖

动 P 、网络带宽 B 、误码率 E 和网络负载 L , 则网络状态空间 S 可以表示为

$$S = \{1, 2, \dots, M\} \times D^1 \times P^1 \times B^1 \times E^1 \times L^1 \times D^2 \times P^2 \times B^2 \times E^2 \times L^2 \times \dots \times D^M \times P^M \times B^M \times E^M \times L^M, \quad (1)$$

其中 $\{1, 2, \dots, M\}$ 表示 M 个可用网络的集合, \times 表示笛卡尔积, 网络状态空间由可用网络集合和网络状态属性作笛卡尔积构成. 为了减少状态空间的维数, 避免维数灾难的发生, 把每个状态属性在其取值范围内离散化, 即:

$$\begin{aligned} D^m &= \{d_1^m, d_2^m, \dots, d_{\max}^m\}, \\ P^m &= \{p_1^m, p_2^m, \dots, p_{\max}^m\}, \\ B^m &= \{b_1^m, b_2^m, \dots, b_{\max}^m\}, \\ E^m &= \{e_1^m, e_2^m, \dots, e_{\max}^m\}, \\ L^m &= \{l_1^m, l_2^m, \dots, l_{\max}^m\}. \end{aligned}$$

其中 $m = 1, 2, \dots, M$, 集合内元素分别表示网络 m 的时延、时延抖动、网络带宽、误码率和网络负载经离散化后的数值. 令决策时刻 t 为作出决策并选择动作的时间点, 则决策时刻的点集为 $T = \{0, 1, 2, \dots\}$. 在每一个决策时刻, SDN 控制器根据当前的网络状态和终端用户的业务类型进行切换选择判决, 决定用户是继续连接在当前网络还是切换到其他网络. 终端用户在决策时刻 t 进行备选网络的切换选择, 即采取的动作:

$$a_t = \{a_t^1, a_t^2\}, \quad (2)$$

其中 $a_t^{1(2)} \in A = \{0, 1, \dots, M\}$. 若 $a_t^{1(2)} \in A = \{1, \dots, M\}$, 表示在决策时刻 t 实时业务用户 (非实时业务用户) 切换到网络 m 中; 若 $a_t^{1(2)} = 0$, 表示在决策时刻 t 实时业务用户 (非实时业务用户) 被拒绝切换到备选网络中. 若在当前状态下按照一定的策略采取行动, 会获得转移到下一特定状态的状态转移概率. 由于当前状态 s 向下一时刻状态 s' 转移的概率只取决于当前状态 s 和采取的动作 a , 而与历史状态和动作无关, 具有马尔可夫性, 即

$$\begin{aligned} P(s = s_t, a = a_t, s' = s_{t+1}) \\ = p\{s_{t+1} | s_t a_t s_{t-1} a_{t-1} \dots s_0 a_0\} = p\{s_{t+1} | s_t a_t\}, \end{aligned}$$

若已知当前网络状态 s 并采取动作 a , 则转移到下一网络状态 s' 的概率为

$$P(s, a, s') = \prod_{m=1}^M p\{d^{m'}, p^{m'}, b^{m'}, e^{m'}, l^{m'} | d^m, p^m, b^m, e^m, l^m\},$$

$$\sum_{s' \in S} P(s, a, s') = 1, \quad (3)$$

其中 $m = 1, 2, \dots, M$ 且 $m' = 1, 2, \dots, M$, 分别表示当前时刻和下一时刻的接入网络. 为了简化问题, 假设异构无线网络区域内 M 个网络的不同状态属性间相互独立.

3.2 一次回报

任意一个决策时刻 t , 在网络状态 s 下采取动作 a 会获得一次回报 $r(s, a)$, 若异构无线网络区域存在 M 个备选网络, 终端用户采取动作 a 切换至备选网络 m 后所带来的一次回报越大, 则表示在决策时刻切换选择的 QoS 越高且性能越好, 定义 $r(s, a) : S \times A \rightarrow \mathfrak{R}$ 为终端用户在网络状态 s 下采取动作 a 后所获得的一次回报函数:

$$r(s, a) = \sum_{s' \in S} r(s, a, s') P(s, a, s'), \quad (4)$$

其中 $r(s, a, s')$ 表示终端用户在网络状态 s 下采取动作 a 转移到状态 s' 后获得的立即回报函数, $P(s, a, s')$ 表示状态转移概率. 当实值函数 $r(s, a)$ 为正时, 表示收入; 当其为负时, 表示费用.

考虑无线网络的五个状态属性, 采用加权求和的方式构建立即回报函数 $r(s, a, s')$, 即:

$$r(s, a, s') = \omega_D r_D(s, a, s') + \omega_P r_P(s, a, s') + \omega_B r_B(s, a, s') + \omega_E r_E(s, a, s') + \omega_L r_L(s, a, s'),$$

$$\sum_{x \in X} \omega_x = 1, \quad (5)$$

其中 $x \in X = \{D, P, B, E, L\}$ 表示无线网络的状态属性, $r_x(s, a, s')$ 表示状态属性 x 的立即回报, ω_x 表示回报函数所对应的权重因子. 权重因子反映了无线网络各状态属性在构建立即回报函数中的地位或作用, 直接影响综合决策的结果. 而在备选网络负载均衡的条件下, 不同业务类型的用户对切换选择的 QoS 具有不同的要求, 实时业务用户对于切换网络的时延和时延抖动较为敏感, 而非实时业务用户对切换网络的传输速率和可靠性更为关注. 因此, 不同业务类型的用户对于构建回报函数的权重分配存在差异, 本节采用层次分析法确定状

态属性权重 [24], 将网络选择问题按目标、评价准则和具体备选方案的顺序分解为不同的层次结构, 层次分析结构模型如图 2 所示.

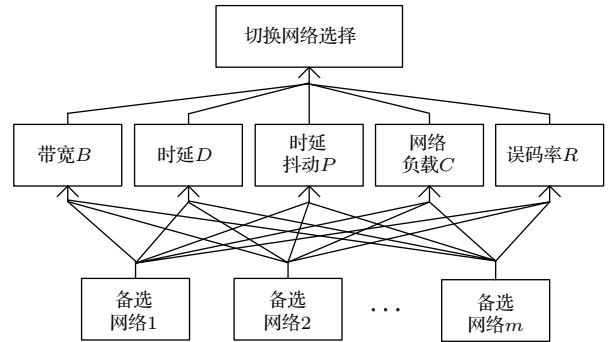


图2 层次分析结构模型

Fig. 2. Analytic hierarchy model.

依据图 2 的层次分析结构模型, 对不同场景下的实时业务用户和非实时业务用户分别构造两两比较的判断矩阵 $Z = (z_{ij})_{5 \times 5}$. 其中, z_{ij} 是用状态属性两两重要性程度之比的形式表示出两个方案的相应重要程度等级, 它表示任意两个状态属性相比较对网络切换的影响程度, $z_{ij} z_{ji} = 1$ 且 $z_{ii} = 1, i, j = 1, 2, \dots, 5$. 按照 1—9 标度法为 z_{ij} 取值, 由不同业务类型用户对切换选择 QoS 的不同需求, 可分别得到实时业务用户和非实时业务用户的判断矩阵, 如表 1 和表 2 所示.

表1 实时业务用户的判断矩阵

Table 1. Comparison matrix of real-time service users.

Z_1	B	D	P	C	R
B	1	1/6	1/5	1/3	3
D	6	1	2	5	7
P	5	1/2	1	4	6
C	3	1/5	1/4	1	4
R	1/3	1/7	1/6	1/4	1

表2 非实时业务用户的判断矩阵

Table 2. Comparison matrix of non real-time service users.

Z_2	B	D	P	C	R
B	1	5	7	3	2
D	1/5	1	2	1/4	1/5
P	1/7	1/2	1	1/5	1/6
C	1/3	4	5	1	1/3
R	1/2	5	6	3	1

依据建立的两两比较判断矩阵计算状态属性的权重:

$$\omega_i = \frac{\left(\prod_{j=1}^5 z_{ij}\right)^{1/n}}{\sum_{i=1}^5 \left(\prod_{j=1}^5 z_{ij}\right)^{1/n}}, \quad i, j = 1, 2, \dots, 5, \quad (6)$$

其中 $\omega_i (i = 1, 2, \dots, 5)$ 分别表示无线网络五个状态属性的权重因子. 由于判断矩阵的构造带有一定的主观性, 对判断矩阵 Z 进行一致性检验, 计算一致性比例 CR :

$$CR = \frac{\lambda_{\max} - 5}{4RI}, \quad (7)$$

其中 λ_{\max} 表示判断矩阵的最大特征值, RI 表示平均一致性指标, 可通过查表得到. 当 $CR < 0.10$ 时, 认为判断矩阵的一致性是可以接受的; 否则需重新调整 Z , 直到其具有满意的一致性.

3.3 最优切换决策

每一个决策时刻 t 都会产生一个立即回报函数, 随着时间的推移, 将会产生一个立即回报函数序列, 寻找最优切换决策即是最大化此序列之和的数学期望, 找到使将来长期回报最大的切换策略. 因此, 决策目标函数 J 表示为

$$J = E \left[\sum_{t=0}^{\infty} \gamma^t r_{t+1} \right], \quad (8)$$

其中 $r_t = r(s, a, s')$ 表示终端用户在决策时刻 t 得到的立即回报. $\gamma \in (0, 1]$ 表示折扣因子, 用于权衡得到的立即回报与将来长期回报二者间的重要性, 当 $\gamma = 1$ 时, 决策目标

$$J = \lim_{N \rightarrow \infty} \frac{1}{N} E \left[\sum_{t=0}^N r_{t+1} \right]$$

表示决策时刻 t 得到的平均回报. 若终端用户采取动作的一个确定性策略定义为 $\pi : S \rightarrow A$, 表示在网络状态 s 下采取动作 $\pi(s)$ 的概率为 1, 策略 π 不随时间的变化而变化. 则状态动作函数 $V^\pi(s, a)$ 表示从状态动作对 (s, a) 出发, 终端用户在网络状态 s 下依据策略 π 采取动作后所获得的长期期望回报, 它是对回报函数的一种预测, 表示从长远角度来看选择动作的优劣, 期望回报函数 $V^\pi(s, a)$ 表示为

$$V^\pi(s, a) = E^\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a \right], \quad (9)$$

其中 $E^\pi[\cdot]$ 表示在策略 π 和状态转移概率 P 分布上的数学期望. 对于任意网络状态 s 和采取动作 a , 在策略 π 下的期望回报函数 $V^\pi(s, a)$ 都可以表示为

$$\begin{aligned} & V^\pi(s, a) \\ &= E^\pi \{ r_{t+1} + \gamma V^\pi(s_{t+1}, \pi(s)) | s_t = s, a_t = a \} \\ &= \sum_{s' \in S} P(s, a, s') (R(s, a, s') + \gamma V^\pi(s', a')) \\ &= R(s, a) + \sum_{s' \in S} \gamma P(s, a, s') V^\pi(s', a'), \end{aligned} \quad (10)$$

其中 $s' = s_{t+1}$ 表示 $(t + 1)$ 时刻的网络状态, $a' = \pi(s)$ 表示在决策时刻 t 依据策略 π 采取的动作, $V^\pi(s')$ 表示在 $(t + 1)$ 时刻得到的状态函数, $R(s, a, s') = E[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots]$ 表示终端用户在网络状态 s 下采取动作 a 转移到状态 s' 后获得的期望回报, $R(s, a)$ 表示终端用户在网络状态 s 下采取动作 a 后所获得的一次期望回报.

在每个决策时刻 t 找到最优切换策略, 使终端用户在当前网络状态 s 和动作 a 下, 依据策略 π 进行切换选择后可以得到最大的期望回报, 即对于任意 $s \in S, a \in A$ 和策略 π 满足 $V^*(s, a) \geq V^\pi(s, a)$, 则最优期望回报可以表示为

$$\begin{aligned} & V^*(s, a) = \max_{\pi} V^\pi(s, a) \\ &= \max_{a' \in A} (R(s, a) + \sum_{s' \in S} \gamma P(s, a, s') V^*(s', a')). \end{aligned} \quad (11)$$

对于此优化问题的求解, 可以采用逐次逼近的迭代方法. 用迭代的方式逼近最优期望回报函数 $V^*(s, a)$, 从而得到最优切换策略 π^* , 对任意 $s \in S, a \in A$, 其迭代方程为

$$\begin{aligned} & V_{t+1}(s, a) \leftarrow \max_{a' \in A} (R(s, a) \\ & \quad + \sum_{s' \in S} \gamma P(s, a, s') V_t(s', a')). \end{aligned} \quad (12)$$

算法流程如下:

步骤1 初始化: 令 $\Delta = 0$, 对 $\forall s \in S, a \in A$, 给定初值 $V(s, a) = 0$;

步骤2 迭代算法: 执行 $v \leftarrow V(s, a)$,

$$V(s, a) \leftarrow \max_{a' \in A} (R(s, a) + \sum_{s' \in S} \gamma P(s, a, s') V(s', a')),$$

$$\Delta \leftarrow \max(\Delta, |V(s, a) - v|);$$

步骤3 若 $\Delta < \theta$, 进入步骤4, 否则重新回到步骤2;

步骤4 输出策略: $\pi = \arg \max_{a' \in A} (R(s, a) + \sum_{s' \in S} \gamma P(s, a, s') V(s', a'))$.

输出策略为设定 $\Delta < \theta$ 下的最优切换策略. 由于(12)式中迭代方程是一个压缩映射^[25], 保证了此迭代算法的收敛性, 其一步收敛速度为 γ . 每进行一次迭代, 算法的最大计算复杂度为 $O(|S|^2|A|)$.

4 仿真结果与分析

对本文所提出的基于马尔可夫决策模型的切换选择算法在 Matlab 平台中进行仿真验证, 采用的系统架构如图1所示. 在 LTE/WLAN 异构无线网络中, LTE 网络和 WLAN 网络覆盖半径分别为 1000 和 300 m. 在仿真环境中随机部署 100 个不同业务类型的终端用户节点, 且其移动速度为 3 m/s. 实时业务用户和非实时业务用户均按照泊松分布到达异构无线网络区域, 其到达率分别为 1—8 (呼叫/s) 和 2—18 (呼叫/s), 并且速率均在 50—100 kb/s, 终端用户进行业务连接所能忍受的最大时延为 400 ms, 若时延超过此阈值则视为阻塞. 假设相邻两个决策时刻的时间间隔 τ 为 10 s, 对无线网络的时延、时延抖动、网络带宽、误码率和网络负载五个状态属性进行采样分析. 由于不同业务类型用户对网络性能指标的关注不同, 在构建终端用户的立即回报函数时, 实时业务用户的权重因子为 [0.4577 0.3102 0.0693 0.0394 0.1235]; 非实时业务用户的权重因子为 [0.0655 0.0428 0.4172 0.3066 0.1680]. 本节针对不用业务类型用户分别得到期望回报和最优切换策略, 然后对终端用户的阻塞率进行仿真验证, 并与其他切换选择方法进行了对比分析.

图3和图4分别显示了在折扣因子的不同取值下实时业务用户和非实时业务用户得到的期望回报. 将本文提出的MDM算法与MADM-SAW算法^[11]进行对比, 随着折扣因子 γ 从0.88增加到0.98, 两种算法得到的期望回报都呈递增趋势. 当 $\gamma = 0.98$ 时, 实时业务用户基于MDM算法得到的期望回报要比MADM-SAW算法高63.98%; 非实时业务用户基于MDM算法得到的期望回报要比MADM-SAW算法高36.40%, 这是由于本文提出的MDM算法采用层次分析法针对不同业务类型用户分别确定网络状态属性的权重, 并取得了满意的一致性.

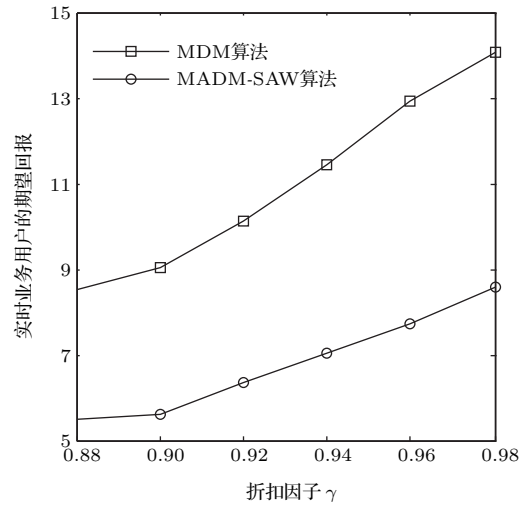


图3 实时业务用户的期望回报

Fig. 3. Expected reward of real-time service users.

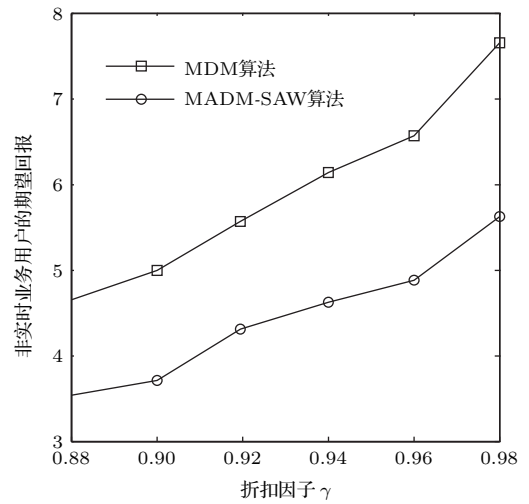


图4 非实时业务用户的期望回报

Fig. 4. Expected reward of non real-time service users.

图5和图6分别显示了实时业务用户和非实时业务用户的最优切换策略. 为方便构造不同业务类型用户的最优切换策略, 均设定折扣因子 $\gamma = 0.9$, LTE网络和WLAN网络的带宽资源是变化的, 同时固定其他网络状态属性. 对于实时业务用户, 若设定 $b^1 = 1$, 当 $b^2 < 2$ 时, 最优切换策略是选择动作 $a_t^1 = 1$, 即接入LTE网络; 对于非实时业务用户, 若设定 $b^1 = 5$, 当 $b^2 > 4$ 时, 最优切换策略是选择动作 $a_t^2 = 2$, 即接入WLAN网络. LTE网络时延和时延抖动较小, 更适合实时业务用户接入; WLAN网络带宽较大且误码率较小, 更适合非实时业务用户接入, 因此在本文提出的MDM算法下, 不同业务类型用户均能选取最优切换策略.

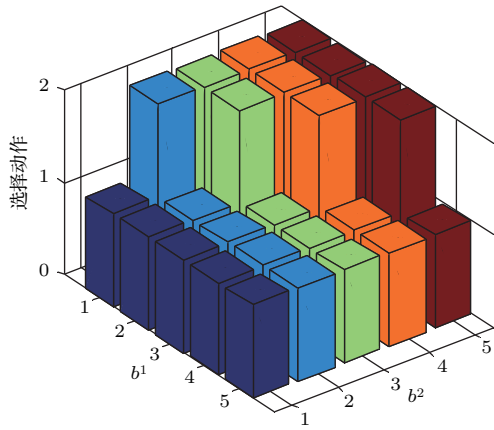


图5 (网刊彩色) 实时业务用户的最优切换策略
Fig. 5. (color online) Optimal handoff strategy for real-time service users.

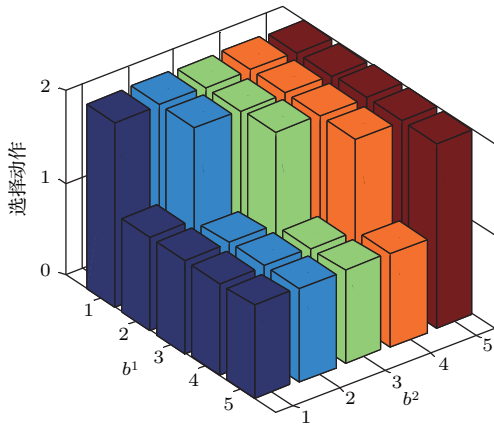


图6 (网刊彩色) 非实时业务用户的最优切换策略
Fig. 6. (color online) Optimal handoff strategy for non real-time service users.

图7显示了在到达率的不同取值下实时业务用户的阻塞率, 当到达率 λ_1 从1(呼叫/s)增加到8(呼叫/s)时, 3种算法的阻塞率都呈递增趋势; 图8显示了在到达率的不同取值下非实时业务用户的阻塞率, 当到达率 λ_2 从2(呼叫/s)增加到18(呼叫/s)时, 3种算法的阻塞率也都呈递增趋势, 表示单位时间内随着实时业务用户或非实时业务用户数量的增多, 因备选网络占用而导致用户被阻塞的概率也逐渐增大. 从图中可以看出, 本文提出的MDM算法对于不同业务类型用户阻塞率均小于MADM-SAW算法和不区分业务类型的切换选择方法. 这是由于本文从长远角度出发, 以立即回报函数序列组成的目标函数来衡量将来长期回报, 得到使长期期望回报最大的切换策略并依此进行切换选择, 较其他算法能有效降低阻塞率.

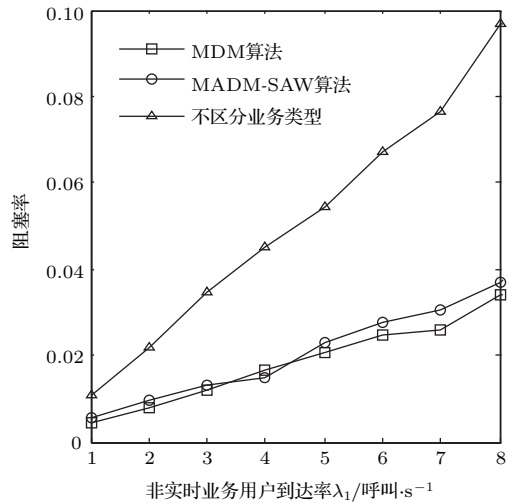


图7 实时业务用户的阻塞率

Fig. 7. Blocking rate of real-time service users.

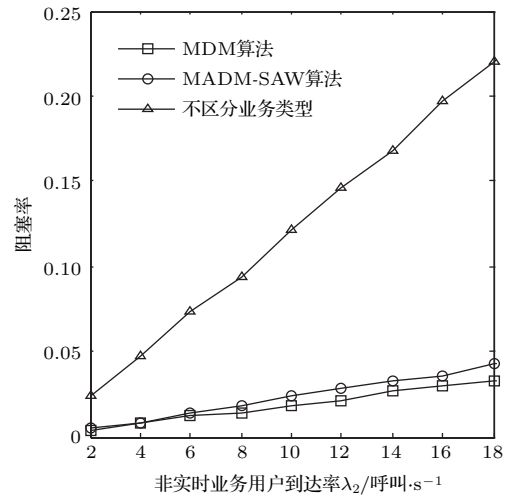


图8 非实时业务用户的阻塞率

Fig. 8. Blocking rate of non real-time service users.

5 结 论

本文在异构无线网络环境下提出了一种基于马尔可夫决策模型的切换选择算法. 建立基于SDN的异构无线网络架构, 通过对网络资源的通透控制实现异构网络的有效互通. 将异构网络中的切换选择问题建模为一个马尔可夫过程, 利用采取动作后的一次回报来衡量切换性能, 并采用层次分析法分别对不同业务类型用户构建立即回报函数, 以满足不同用户的QoS需求. 最后从网络的状态动作对出发, 通过逐次逼近的迭代方式最大化长期期望回报来选取满足条件的最优策略. 仿真结果显示此算法可以提高不同业务类型用户的QoS和无线网络的资源利用率, 同时降低阻塞率. 但算法只考

虑了实时用户与非实时用户的不同需求,未来将通过大数据分析所有用户的网络偏好参数,以针对用户制定个性化切换选择方案.

参考文献

- [1] Falowo O E, Chan H A 2012 *Eurasip J. Wirel. Comm.* **221**
- [2] Zhu K, Niyato D, Wang P 2010 *Proceedings of IEEE Wireless Communications and Networking Conference* Sydney, Australia, April 18–21, 2010 p1
- [3] Yan X, Şekercioğlu Y A, Narayanan S 2010 *Comput. Networking* **54** 1848
- [4] Liu J, Xiong Q Y, Shi X, Wang K, Shi W R 2015 *Chin. Phys. B* **24** 076401
- [5] Ahmed A, Boulahia L M, Gaiti D 2014 *IEEE Commun. Surv. Tutorials* **16** 776
- [6] Hasib A, Fapojuwo A 2008 *IEEE Trans. Veh. Technol.* **57** 2426
- [7] Kunarak S, Suleesathira R, Dutkiewicz E 2013 *Proceedings of IEEE International Conference of Region 10* Xi'an, China, October 22–25, 2013 p1
- [8] Salem M, Ismail M, Misran N 2011 *J. Appl. Sci.* **11** 336
- [9] Niyato D, Hossain E 2009 *IEEE Trans. Veh. Technol.* **58** 2008
- [10] Naghavi P, Rastegar S H, Shah-Mansouri V, Kebriaei H 2016 *IEEE Wirel. Commun. Lett.* **5** 52
- [11] Stevens-Navarro E, Martinez-Morales J D, Pineda-Rico U 2012 *J. Appl. Res. Technol.* **10** 534
- [12] Wang N, Shi W X, Fan S S, Liu S X 2011 *Proceedings of 2nd International Conference on Challenges in Environmental Science and Computer Engineering* Haikou, China, December 14–15, 2011 p55
- [13] Liu K M 2014 *J. Inf. Comput. Secur.* **11** 3373
- [14] Zhu S F, Liu F, Chai Z Y, Qi Y T, Wu J S 2012 *Acta Phys. Sin.* **61** 096401 (in Chinese) [朱思峰, 刘芳, 柴争义, 戚玉涛, 吴建设 2012 物理学报 **61** 096401]
- [15] Ning Z L, Song Q Y, Liu Y J, Wang F Z, Wu X Y 2014 *Comput. Electr. Eng.* **40** 456
- [16] Ma B, Deng H, Xie X Z, Liao X F 2015 *China Commun.* **12** 106
- [17] Ma B, Xie X Z, Liao X F 2015 *J. Electron. Inform. Technol.* **37** 874 (in Chinese) [马彬, 谢显中, 廖晓峰 2015 电子与信息学报 **37** 874]
- [18] Chen T, Matinmikko M, Chen X F, Zhou X, Ahokangas P 2015 *IEEE Commun. Mag.* **53** 126
- [19] Wang H C, Chen S Z, Xu H, Ai M, Shi Y 2015 *IEEE Network* **29** 16
- [20] Shen Y 2013 *Chin. Phys. B* **22** 058902
- [21] Yang X L, Tan X Z, Guan K 2015 *Acta Phys. Sin.* **64** 108403 (in Chinese) [杨小龙, 谭学治, 关凯 2015 物理学报 **64** 108403]
- [22] Tsai C, Yang F N 2013 *J. Hydraul. Eng.* **139** 1265
- [23] Fei R, Cui D W 2009 *Acta Phys. Sin.* **58** 5133 (in Chinese) [费蓉, 崔杜武 2009 物理学报 **58** 5133]
- [24] Alavipoor F S, Karimi S, Balist J, Khakian A H 2016 *Global. J. Environ. Sci. Manage.* **2** 197
- [25] Marco W, Martijn V O 2012 *Reinforcement Learning: State of the Art* (Berlin: Springer) pp223–229

Markov decision model based handoff selection algorithm for heterogeneous wireless networks*

Liang Xiao¹⁾ Qian Zhi-Hong^{1)†} Tian Hong-Liang¹⁾²⁾ Wang Xue¹⁾

1) (College of Communication Engineering, University of Jilin, Changchun 130012, China)

2) (College of Information Engineering, Northeast Dianli University, Jilin 132012, China)

(Received 26 July 2016; revised manuscript received 7 September 2016)

Abstract

Coexistence of multiple wireless access technologies will be an indicator of next-generation wireless network, and the integration of heterogeneous wireless networks will meet the needs of high-performance services for mobile users. According to unique quality of service (QoS) requirements of different service type users in heterogeneous environment, the Markov decision model based handoff selection algorithm is proposed in this paper. A heterogeneous wireless network architecture based on the software defined network (SDN) is established to realize the transparency control of heterogeneous networks. Network state information of heterogeneous wireless networks is mastered by SDN controller. It is responsible for scheduling network resources dynamically according to the performance characteristics of each network. If the network state information in equal interval is sampled, the next moment state of network is only related to the current network state and action, but it is not related to the historical state. The problem of handoff selection for heterogeneous wireless networks is modeled as a Markov process with discrete time and continuous state. To predict the next moment state of network by Markov process to obtain a reward, when the reward is positive, it represents the income; when it is negative, it represents the cost. An immediate reward function is constructed for real-time service and non real-time service users respectively according to their different state attributes of the network. Considering five state attributes of wireless network as follows: delay, delay jitter, bandwidth, error rate and network load, the immediate reward function is constructed with weighted summation. Due to the difference in attribute weight distribution among different service type users, the attribute weights are determined by the analytic hierarchy process. In the long term, the objective function which consists of immediate reward function sequence is used to measure future long-term rewards. Then expected reward function based on the state action pair is constructed to obtain the handoff strategy of the maximum expected return by the iterative method of successive approximation. The proposed Markov decision model based handoff selection algorithm is used in simulation of the Matlab platform. The simulation results show that the proposed method can select the optimal handoff strategy for different service type users and reduce the blocking rate, thereby improving the QoS of users and resource utilization of wireless networks.

Keywords: heterogeneous wireless networks, handoff selection, Markov process, analytic hierarchy process

PACS: 64.60.aq, 89.70.-a, 02.50.Ga, 87.55.kd

DOI: [10.7498/aps.65.236402](https://doi.org/10.7498/aps.65.236402)

* Project supported by the National Natural Science Foundation of China (Grant No. 61371092), the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 61401175), the Key Science and Technology Program of Jilin Province, China (Grant No. 20140204019GX), and the Key Science and Technology Program of Changchun City, China (Grant No. 2014026/14KG021).

† Corresponding author. E-mail: dr.qzh@163.com