

## 基于部分可测马尔科夫决策过程业务感知的微基站休眠时长确定策略

陈前斌 何小强\* 吴攀 唐伦

(重庆邮电大学移动通信技术重点实验室 重庆 400065)

**摘 要:** 针对密集组网场景中业务不确定性引起的基站休眠周期难以确定的问题, 该文提出一种基于部分可测马尔可夫决策过程(Partially Observed Markov Decision Process, POMDP)业务感知的微基站休眠时长确定策略。该策略将周期分为长周期和短周期, 每个周期由轻度和深度两个阶段构成。通过 POMDP 感知到达基站的业务状态, 动态调整周期时长, 进而选取适合当前周期的时长。仿真结果表明, 该策略可以根据业务感知提前确定微基站关断时长, 与基于业务门限值的基站关断机制相比节能效果更好。

**关键词:** 密集组网; 关断机制; 部分可测马尔可夫过程; 业务感知; 长/短休眠周期; 动态调整

**中图分类号:** TN929.5

**文献标识码:** A

**文章编号:** 1009-5896(2018)01-0130-07

**DOI:** 10.11999/JEIT170274

## Micro Base Station Sleeping Cycle Determination Strategy Based on Partially Observed Markov Decision Process Traffic Aware

CHEN Qianbin HE Xiaoqiang WU Pan TANG Lun

(Key Laboratory of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** In order to solve the problem that the sleeping cycles are difficult to be determined due to the traffic uncertainty in dense network scenarios, this paper proposes a Micro base station sleeping cycle determination strategy which based on the Partially Observed Markov Decision Process (POMDP) traffic-aware. In this strategy, the sleeping cycle is divided into long cycle and short cycle, and each cycle consists of deep and light stage. Based on the POMDP traffic-aware, it can dynamic adjusting the cycle and determine the proper length of cycle. Both the analytical and simulation results show that compare with sleeping strategy based on the traffic threshold, the base station sleeping strategy based on traffic awareness can effectively reduce the energy consumption of the micro base stations in the dense network by adjusting the sleeping time of the micro base stations in real time.

**Key words:** Dense network; Sleeping strategy; Partially Observed Markov Decision Process (POMDP); Traffic-aware; Long/short sleeping cycle; Dynamic adjusting

### 1 引言

随着 4G 时代的到来, 更快、更可靠的无线网络已经融入了我们的生活。随着智能终端的普及, 用户对数据流量的需求更是成指数倍增长, 这也意味着需要部署更多的微基站来满足对日益增长的数据流量的需求, 直接增加了密集组网的能耗<sup>[1]</sup>。在密集网络能耗中, 基站的能耗占据了总能耗的 60%~80%<sup>[2]</sup>, 而在基站本身的能耗中有 70%是来自功率放大器和空调系统。为了降低基站的能耗, 引入了微

基站关断机制<sup>[3,4]</sup>, 由于无线通信业务在时间和空间上的不均匀性, 使得微基站的关断机制可行<sup>[5-8]</sup>。微基站在业务很少或者没有业务到达的时候, 能效会明显降低, 此时可以允许微基站进入休眠模式, 当到达的业务较多时开启微基站, 以便能够及时处理业务保证用户的 QoS。

目前业界的学者发表了一些关于基站关断相关的学术文章, 文献[9]提出了基于集中式和非集中式算法来决定基站关断的算法, 能够根据网络低负载的时候动态关断一些低负载基站, 但是作者并没有实现对业务的预测。文献[10]从用户接入和基站关断相结合来降低能效, 并且提出了简单的贪婪开启和贪婪关闭算法来实现对基站开关操作的最优化处理。文献[11]提出了一种基于门限值的关断策略: 基站当没有用户的时候处于关断状态, 当有  $N$  个用户

收稿日期: 2017-03-31; 改回日期: 2017-08-09; 网络出版: 2017-09-14

\*通信作者: 何小强 724331631@qq.com

基金项目: 国家高科技研究发展计划(2014AA01A701), 国家自然科学基金(61571073)

Foundation Items: The National High Technology Research and Development Program of China (2014AA01A701), The National Natural Science Foundation of China (61571073)

到达或者在一段时间后再开启基站。作者在此基础上权衡了功耗和时延对能效的影响。文献[3]根据已知的业务模型提出了基站关断策略,并且以网络覆盖率作为约束条件结合基站关断对网络能效进行了优化。然而上述文献都没有实现对业务的预测。文献[12,13]为了降低智能终端能耗的关断算法,并用部分可测马尔科夫决策过程(Partially Observed Markov Decision Process, POMDP)来预测当前阶段的业务状态实现对算法的优化。然而,目前还没有相关文献在对网络业务进行预测的前提下提出周期性开启和关闭基站。为此,本文提出了一种基于业务感知的动态调整微基站深度休眠周期时长的关断机制。该机制有如下两个特点:(1)将休眠状态下的基站分为长周期休眠和短周期休眠,并将每个休眠周期细分为深度休眠和轻度休眠;(2)本文通过POMDP对业务状态建模,能够实现对业务的感知,从而可以动态调整基站深度休眠时长。

## 2 基于业务感知的动态调整微基站深度休眠周期时长的策略

### 2.1 长短休眠周期的基站关断机制

现有的基站关断策略中大多采取的都是定周期性开启和关闭基站,这种定周期开启和关闭基站策略因为不能够感知请求业务的数量,在到达业务数量较多的情况下会带来较高的时延。针对现有基站关断策略中关断时长选取不灵活的问题,本文提出了一种长短休眠周期的基站关断机制,原理图如图1所示。将休眠周期分为短周期和长周期两类,每个休眠周期都由深度休眠周期和轻度休眠周期组成。处于激活状态下的基站在连续一段时间 $\tau$ 内没有业务到达,将进入休眠状态。基站只有当连续几

个周期业务到达量较多时才会处于激活状态,否则一直处于休眠状态。处于激活状态和轻度休眠状态下的基站都可以处理业务请求,只是前者功耗和服务速率要大于后者。在休眠状态的每个休眠周期中,基站首先处于短周期休眠,若连续 $N_1$ 个短休眠周期内都没有业务到达或者到达的业务很少,基站将进入长周期休眠,在长周期休眠内如果到达的业务较多则下一个休眠周期将转变到短周期休眠。当轻度休眠周期内不能处理完深度休眠周期内缓存的业务,则可以适当延长轻度休眠时长以处理完缓存的业务。如果连续 $N_2$ 个短周期休眠内到达的业务数量都很多,则开启基站。

当基站处于深度休眠时,基站关闭绝大部分功能,只保留接收功能模块用于保存到达基站的服务请求。当基站处于轻度休眠时,基站以较低的服务速率处理业务,其固定功耗和发射功率都低于激活状态下基站的固定功耗和发射功率。基站的功耗主要来自于功率放大器、天线接口单元、无线射频单元、基带处理单元、DC-DC单元、AC-DC单元和自身散热单元<sup>[14]</sup>。当基站处于轻度休眠时,由于基站仍然要处理业务,但业务量相对较少,因此关闭自身散热单元和功率放大器模块;当基站处于深度休眠时,基站将关闭功率放大器、基带处理单元和自身散热单元。

### 2.2 业务感知

由上述长短休眠周期基站关断机制可知,基站休眠时长的选取是由到达基站的业务数量所决定。因此,要选取合适的基站休眠时长就必须预知业务数量,从而可以提前确定基站的状态以及基站休眠周期的时长。然而实际中每一个周期内业务的状态都很难直接得到,但可以统计当前周期内到达的业

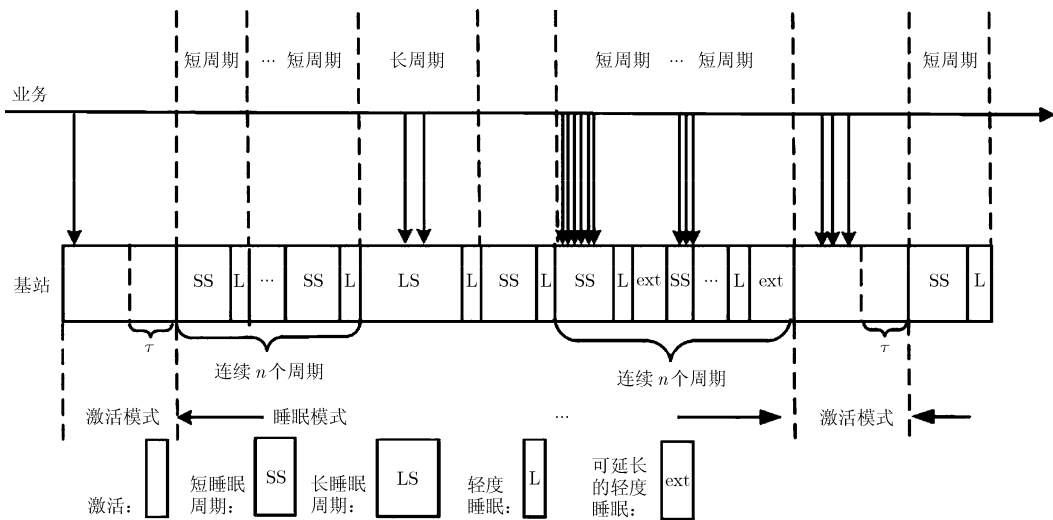


图1 一种长短休眠周期的基站关断机制

务数量, 这为实现业务预测提供了必要条件。基于此问题, 本文采用 POMDP 对业务状态进行建模, 通过观测当前周期缓存的业务数量来估计下一周期的业务到达率, 从而预估下个周期可能到达的业务数量, 实现对业务的预测。

系统建模如下: 由于网络的负载变化具有随机性和波动性, 业务流利用离散时间马尔科夫调制泊松过程 (discrete-time Markov-Modulated Poisson Process, dMMPP) 进行建模<sup>[15]</sup>, dMMPP 过程是一种比传统泊松模型适用面更广的用于描述因特网流量到达规律的随机过程, 它可被看成若干到达率不同的泊松流的加权和, 也就是说如果该 dMMPP 的调制马尔科夫链共有  $M$  个状态, 那么可以看成是  $M$  个不同到达率  $\lambda_m, (m \in M)$  的泊松流的加权和。

本文提出的一种基于业务感知的动态调整微基站休眠周期时长的策略使用 POMDP<sup>[16]</sup>进行建模, 能够准确地估计业务状态从而在每个决策时刻判断出下一时刻的系统状态。POMDP 模型一般由 6 个多元组组成  $\{S, A, P, Z, O, R\}$ , 其中  $S$  表示系统状态集合,  $A$  表示行动集合,  $P$  表示系统状态转移概率空间,  $Z$  表示观测值,  $O$  表示观测值概率空间,  $R$  表示回报函数。

系统状态: 系统在每个周期末端都会有一个决策时刻, 系统会在每个决策时刻执行相应的行动, 然后转移到下一个状态。用  $D = \{d_1, d_2, \dots, d_m, d_{m+1}, \dots, d_M\}$  表示系统的决策时刻集合。因此上面提出的 POMDP 模型中, 状态  $S$  就由 dMMPP 中不同到达率  $\lambda_m, (m \in M)$  表征, 定义  $d_t$  时刻系统状态集合为  $S = \{s_1, s_2, \dots, s_m, s_{m+1}, \dots, s_M\}, (s_m)_{d_t} \in S, \forall m \in M$  代表在第  $d_t$  个时刻基站所处的状态, 其到达率为  $(\lambda_m)_{d_t}$ 。

系统行动: 本文中系统行动表示基站选取的深度休眠时长, 定义系统行动集合为  $A = \{a_1, a_2, \dots, a_m, a_{m+1}, \dots, a_M\}, (a_m)_{d_t} \in A, \forall m \in M$  代表在选择深度休眠时长, 用  $(T_{ds})_{d_t}$  表示。

观测状态: 由于系统实际状态是隐藏不可直接获取的, 因此只能通过观察得到。然而系统在上一个周期内业务的到达数量是可以得到的, 因此将业务的数量作为观测状态, 表示为  $Z = \{z_{d_1}, z_{d_2}, \dots, z_{d_i}, z_{d_{i+1}}, \dots, z_{d_M}\}$ , 其中  $z_{d_i}$  表示第  $i-1$  到第  $i$  个决策时刻间业务到达的数量。

观测值函数: 用  $O$  表示观测值概率空间,  $O = \{o(z_{d_1}), o(z_{d_2}), \dots, o(z_{d_i}), o(z_{d_{i+1}}), \dots, o(z_{d_M})\}$ , 其中

$o(z_{d_i})$  表示在时刻  $d_i$  业务到达数量为  $z_{d_i}$  的概率。由于到达的业务服从泊松分布, 观测值函数可以根据泊松分布的概率密度函数由式(1)计算得出:

$$\begin{aligned} & o(z_{d_t}, (a_m)_{d_{t-1}}, (s_m)_{d_t}) \\ &= \Pr(z_{d_t} | (a_m)_{d_{t-1}}, (s_m)_{d_t}) \\ &= \frac{((\lambda_m)_{d_t} (T_{ds})_{d_{t-1}})^{z_{d_t}}}{z_{d_t}!} \exp^{-(\lambda_m)_{d_t} (T_{ds})_{d_{t-1}}} \quad (1) \end{aligned}$$

式(1)表示的含义是在  $d_{t-1}$  时刻执行动作系统状态转移到  $d_t$  时刻状态  $(s_m)_{d_t}$  得到观测值  $z_{d_t}$  的概率。

信念状态: 在 POMDP 环境中, 决策者只能通过观察来预测当前状态, 但是这些观察并不足以确定当前状态, 还需要借助历史信息。为此引入信念状态, 它可以代替历史序列来计算其长远回报, 它是系统状态空间上的概率分布, 是决策者对环境的一个信念反映, 是一个代表系统状态上概率分布的向量。信念状态是对系统可能处于某个状态概率的直接反映。在这里定义  $d_t$  时刻信念状态为  $B_{d_t} = \{b(s_1)_{d_t}, b(s_2)_{d_t}, \dots, b(s_m)_{d_t}, b(s_{m+1})_{d_t}, \dots, b(s_M)_{d_t}\}$ , 其中  $b(s_m)_{d_t} \in B_{d_t}$  表示在决策时刻  $d_t$  业务状态为  $s_m$  的概率, 并且  $\sum_{\forall s_m \in S} b(s_m)_{d_t} = 1$ 。每当系统转移到新的状态后, 都需要更新系统的信念状态。系统信念状态的更新取决于上一个决策时刻采取的行动, 以及执行该行动后得到的观测值。根据贝叶斯法则, 可以得出信念状态的更新式子如式(2):

$$\begin{aligned} & b(s_m)_{d_t} \\ &= \Pr(s_m | B_{d_{t-1}}, a_{d_{t-1}}, z_{d_t}) \\ &= \frac{\Pr(s_m, B_{d_{t-1}}, a_{d_{t-1}}, z_{d_t})}{\Pr(B_{d_{t-1}}, a_{d_{t-1}}, z_{d_t})} \\ &= \frac{o(z_{d_t}, (a_m)_{d_{t-1}}, (s_m)_{d_t}) \cdot \sum_{\forall s_n \in S} q_{n,m} b(s_n)_{d_{t-1}}}{\sum_{\forall (s_m)_{d_t} \in S} o(z_{d_t}, (a_m)_{d_{t-1}}, (s_m)_{d_t}) \sum_{\forall (s_n)_{d_t} \in S} q_{n,m} b(s_n)_{d_{t-1}}} \quad (2) \end{aligned}$$

其中,  $q_{nm}$  表示系统从第  $n$  个状态转移到第  $m$  个状态的状态转移概率。

回报函数: 为了更加直观地看到基站关断所带来的节能效果, 将基站在  $d_t$  时刻下系统的平均能耗定义为系统的回报函数  $r((s_m)_{d_t}, (a_m)_{d_t})$ , 其物理含义是在决策时刻  $d_t$  状态为  $s_m$  时采取行动  $(a_m)_{d_t}$  的平均能耗, 很显然平均能耗越小意味着系统的性能越好。

## 2.3 评估指标

**2.3.1 平均功耗** 为了评估基站在处于休眠模式下的功率消耗，定义基站在  $d_t$  时刻周期内的平均功耗为  $\bar{E}((s_m)_{d_t}, (a_m)_{d_t})$ ，此处的平均功耗由 3 部分组成：深度休眠周期的能耗、轻度休眠周期的能耗和深度休眠与轻度休眠之间转换的能耗。下面给出平均功耗的表达式：

$$\bar{E}((s_m)_{d_t}, (a_m)_{d_t}) = E_{to}/T_{ex} \quad (3)$$

$$E_{to} = 2E_{ch} + U(E_{ds}T_{ds} + E_{ls}T_{ls}) + (1-U)E_{ac}T_{ac} \quad (4)$$

$$T_{ex} = T_{ds} + T_{ls} \quad (5)$$

其中，当处于休眠状态时， $U = 1$ ，否则  $U = 0$ 。 $E_{ch}$  表示深度休眠状态和轻度休眠状态之间切换消耗的能量， $E_{ds}$  表示处于深度休眠期间基站的功耗， $E_{ls}$  表示处于轻度休眠期间基站的功耗， $E_{ac}$  表示处于激活模式下基站的功耗。 $T_{ds}$  和  $T_{ls}$  分别表示深度休眠时长和轻度休眠时长。

**2.3.2 平均时延** 系统的平均时延是由等待服务时长和服务时长两部分组成，在基于长短周期的休眠时长关断机制下，由于在这种机制中基站有激活状态和休眠状态，所以分别计算出两种模式下时延。

(1) 激活周期中的时延：由于在激活状态下，到达的业务根据先到先服务的规则，依次被基站服务，所以时延就是队列里的等待时长与服务时长之和。根据  $M/G/1$  排队系统中相关定义，利用 P-K 公式可以得到在激活周期中的平均时延：

$$E[D_{ac}] = E[T_{M/G/1}] = \frac{\rho_a}{\lambda} + \frac{\rho_a^2 + \lambda^2 \sigma^2}{2\lambda[1 - \rho_a]} \quad (6)$$

其中， $\rho_a = \lambda/\mu_{ac}$  表示系统处于激活状态下的业务强度， $\sigma^2$  是服务时间的方差。

(2) 休眠周期中的时延：由于在休眠周期中分为深度休眠和轻度休眠两种小模式，而且业务是在轻度休眠周期中才被开始服务，因此业务在休眠周期中的平均时延除了在队列中的等待时长和服务时长之外，还要加上在周期中平均假期时长  $E[T_{SV}]$ ，即

$$E[D_{sl}] = E[T_{M/G/1}] + E[T_{SV}] \quad (7)$$

根据有休假的  $M/G/1$  排队系统定义，可以得到  $E[T_{SV}]$  表达式为

$$E[T_{SV}] = T_{ds}/2 \quad (8)$$

代入式(7)可得

$$\begin{aligned} E[D_{sl}] &= E[T_{M/G/1}] + E[T_{SV}] \\ &= \frac{\rho_s}{\lambda} + \frac{\rho_s^2 + \lambda^2 \sigma^2}{2\lambda(1 - \rho_s)} + \frac{T_{ds}}{2} \end{aligned} \quad (9)$$

其中  $\rho_s = \lambda/\mu_{sl}$  表示系统处于休眠状态下的业务强度。

## 2.4 POMDP 深度休眠时长选取策略

基于前面分析的评估指标，本文设计了 POMDP 深度休眠时长策略，其原理框图如图 2 所示，图中  $s_{d_t}$  表示在  $d_t$  时刻系统状态， $z_{d_t}$  表示在  $d_t$  时刻系统观测状态， $B_{d_t}$  表示  $d_t$  时刻系统的信念状态， $a_{d_t}$  表示  $d_t$  时刻该系统采取的行动。在  $d_t$  时刻，系统统计得到观测状态  $z_{d_t}$ ，同时更新系统信念状态  $B_{d_t}$ ，然后从行动集合里面选取相应的行动  $a_{d_t}$ ，系统就由状态  $s_{d_t}$  转移到下一个状态  $s_{d_{t+1}}$ 。基于 POMDP 深度休眠时长调整策略的优点在于通过对业务数量的预知，能够提前选取合适的深度休眠时长，当预知业务数量少时可以提前选择长周期休眠，在不影响用户 QoS 的前提下最小化能耗；当预知业务数量较多时，可以提前选择短周期休眠，以便能够及时处理数据，保证用户的 QoS。

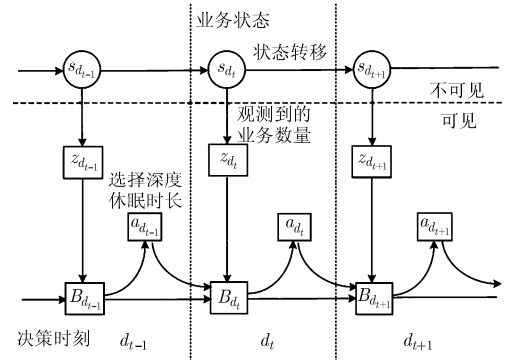


图 2 基于 POMDP 深度休眠时长选取策略原理图

该策略以平均时延作为  $\delta_{th}$  约束条件，用回报函数直接体现出基于 POMDP 深度休眠时长策略的效果。回报函数的赋值过程如表 1 所示，如果平均时延满足规定的时延约束，那么系统就会获得与之相应的回报函数，否则回报函数则会被赋予事先规定好的最大值。

通过回报函数赋值算法计算出回报函数后，再结合信念状态，可以求解出最优的深度休眠时长，即在满足用户 QoS(时延的约束)下，最小化回报函数即最小化平均功耗。利用 POMDP 建模可以通过计算出信念状态来预估业务状态，再结合回报函数，可以采用近似求解来求解出最优策略<sup>[17]</sup>。近似求解的思路是，只在部分信念空间上更新值函数。这是因为对于绝大多数的 POMDP 问题，能到达的信念状态点也只是信念空间里的小部分。并且最优值函

表 1 回报函数赋值

|                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------|
| 输入: $d_t$ 时刻的系统状态 $(s_m)_{d_t}$ 、行动 $(a_m)_{d_t}$ 、平均时延门限 $\delta_{th}$                                                       |
| 输出: $d_t$ 时刻系统的回报函数值 $r((s_m)_{d_t}, (a_m)_{d_t})$                                                                            |
| (1) 对于 $d_t$ 时刻每一个系统状态 $(s_m)_{d_t}$                                                                                          |
| (2) 对于 $d_t$ 时刻每一个行动 $(a_m)_{d_t}$                                                                                            |
| (3) 计算当前休眠周期平均时延 $E[D_{sl}]$                                                                                                  |
| $E[D_{sl}] \leftarrow \frac{\rho_s}{\lambda} + \frac{\rho_s^2 + \lambda^2 \sigma^2}{2\lambda(1 - \rho_s)} + \frac{T_{ds}}{2}$ |
| (4) 计算平均功耗 $\bar{E}((s_m)_{d_t}, (a_m)_{d_t})$                                                                                |
| $\bar{E}((s_m)_{d_t}, (a_m)_{d_t}) = \frac{2E_{ch} + U(E_{ds}T_{ds} + E_{ls}T_{ls}) + (1 - U)E_{ac}T_{ac}}{T_{ds} + T_{ls}}$  |
| (5) if $(E[D_{sleep}] < \delta_{th})$ $r((s_m)_{d_t}, (a_m)_{d_t}) \leftarrow \bar{E}((s_m)_{d_t}, (a_m)_{d_t})$              |
| else $r((s_m)_{d_t}, (a_m)_{d_t}) \leftarrow (\bar{E}((s_m)_{d_t}, (a_m)_{d_t}))_{\max}$                                      |
| endif                                                                                                                         |

数是分段线性凸函数, 当一个点的值被更新以后, 该信念点邻近的信念状态的值也会同时得到改进。这里引用一种确定性信念选择算法, 该算法是在基于点的值迭代基础上改进而来的。基于确定性信念选择算法主要分为两大步: 第 1 步是在给定的信念状态集合上更新值函数; 第 2 步是扩充信念状态集合。这两个步骤交替执行, 直到值函数收敛或者满足终止条件为止。

(1)更新值函数: 采用式(10)更新值函数:

$$\begin{aligned}
 V_n(b(s_n)) &= \max_{a_n \in A} \left[ r(b(s_n), a_n) \right. \\
 &\quad \left. + \gamma \sum_{o \in O} P(o|b(s_n), a_n) V_{n+1}(b'(s_n)) \right] \\
 &= \max_{a_n \in A} \left[ \sum_{s \in S} b(s_n) r(s_n, a_n) \right. \\
 &\quad \left. + \gamma \sum_{o \in O} P(o|b(s_n), a_n) V_{n+1}(b'(s_n)) \right] \quad (10)
 \end{aligned}$$

其中,  $\gamma$  表示折扣因子,  $P(o|b(s_n), a_n)$  表达式如式(11)

$$\begin{aligned}
 P(o|b(s_n), a_n) &= \sum_{s'_n \in S} P(o, s'_n | a_n, b(s_n)) \\
 &= \sum_{s'_n \in S} P(s'_n | a_n, b(s_n)) P(o | s'_n, a_n, b(s_n)) \\
 &= \sum_{s'_n \in S} \sum_{s_n \in S} P(s'_n | a_n, s_n) b(s_n) P(o | s'_n, a_n) \\
 &= \sum_{s'_n \in S} O(s'_n, a_n, o) b(s_n) \sum_{s_n \in S} P(s_n, a_n, s'_n) b(s_n) \quad (11)
 \end{aligned}$$

此时, 最优的基站深度休眠时长可由式(12)得到:

$$\begin{aligned}
 a_n^*(b(s_n)) &= \arg \max_{a_n \in A} \left[ \sum_{s \in S} b(s_n) r(s_n, a_n) \right. \\
 &\quad \left. + \gamma \sum_{o \in O} P(o|b(s_n), a_n) V_{n+1}(b'(s_n)) \right] \quad (12)
 \end{aligned}$$

(2)扩充信念状态集合: 初始时, 信念状态集合只包括一个初始的信念状态  $b_0$ , 假设当前的信念状态集合为  $B_t$ , 而且已经在这些信念点上更新了值函数, 那么算法通过选择代理可以到达其他信念状态来扩充当前的信念集合  $B_t$ , 扩充时的基本原则是使新的信念状态集合在整个信念空间中的分布密度尽量均匀。为使信念集合在信念空间中的分布尽可能均匀, 算法使用 1-范数距离来衡量信念状态之间的距离。两个信念状态  $b_1$  和  $b_2$  的 1-范数距离由式(13)表示:

$$D(b_1, b_2) = \|b_1 - b_2\| = \sum_{s \in S} |b_1(s) - b_2(s)| \quad (13)$$

确定性信念选择算法扩充信念点集合的具体过程如下:

(a)把  $B_t$  中的所有信念状态赋给新的信念状态集合  $B_{t+1}$ ;

(b)对  $B_t$  中每个信念点  $b$ , 根据每一个行动和采取行动后获得的观察计算出下一时刻能到达的信念状态并计算该信念状态的确定性, 并存储到一个集合中;

(c)从可能到达的信念状态集合中, 选择到  $B_{t+1}$  中已经包含的信念点 1-范数距离最远且不在集合  $B_{t+1}$  中的一个信念状态, 并将这个信念状态添加到状态集合  $B_{t+1}$  中。

基于业务感知的动态调整基站休眠周期时长的总体算法如表 2 所示。

### 3 结果评估

为了验证算法的性能, 本文对基于业务感知的动态调整微基站休眠周期时长的策略进行了仿真验证。仿真参数设置如表 3 所示。

Eunsung 等人<sup>[18]</sup>给出了流量在一天内 24 h 的变化情况, 图 3 选取的是夜间流量相对较少的时间段中连续 400 个周期内业务的实际到达量和预测到达量对比图。从图中可以看出, 使用 dMMPP 对业务流进行建模, 并用 POMDP 对业务流量进行预测后, 能够较准确地与实际业务相吻合, 图 4 是对本文提出的基于业务感知的动态调整微基站深度休眠周期时长概率的验证情况, 从图中可以看出在这 11 个周期中实际业务到达量与预测业务到达量变化趋势以及数值基本是一致的, 两者之间的差异很小。

图 5 表示的是基站在这 11 个周期内基于业务门

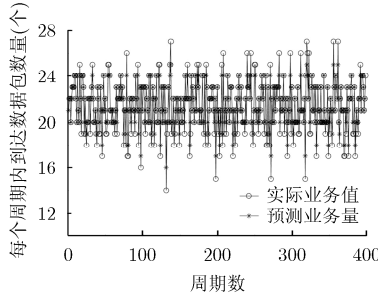


图 3 连续 400 个周期实际业务到达量和预测业务到达量对比图

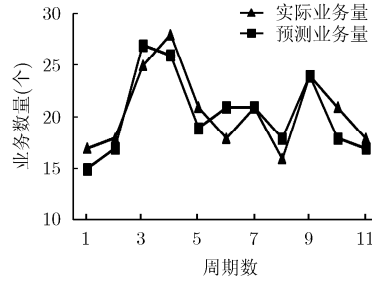


图 4 连续 11 个周期实际业务到达量和预测业务到达量对比图

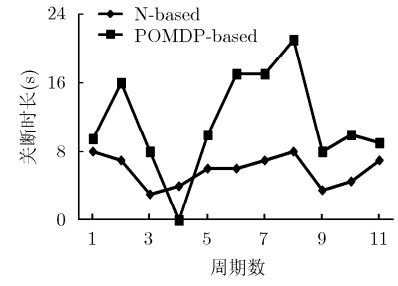


图 5 基站关断时长对比图

表 2 基于业务感知的动态调整基站深度休眠时长策略

输入： $d_t$  时刻的系统状态  $(s_m)_{d_t}$ 、 $d_{t-1}$  时刻的行动  $(a_m)_{d_t}$ 、平均时延门限  $\delta_{th}$

输出： $d_t$  时刻系统的最优基站深度休眠时长

(1) 根据统计到的观测值，计算出观测值函数  $o(z_{d_t}, (a_m)_{d_t}, s_m(d_t))$

$$o(z_{d_t}, (a_m)_{d_t}, s_m(d_t)) \leftarrow \frac{\left( (\lambda_m)_{d_t} (T_{ds})_{a_{d_{t-1}}} \right)^{z_{d_t}}}{z_{d_t}!} \exp^{-(\lambda_m)_{d_t} (T_{ds})_{a_{d_{t-1}}}}$$

(2) 更新信念状态  $b(s_m)_{d_t}$

$$b(s_m)_{d_t} \leftarrow \frac{o(z_{d_t}, (a_m)_{d_{t-1}}, (s_m)_{d_t}) \cdot \sum_{s_n \in S} q_{n,m} b(s_n)_{d_{t-1}}}{\sum_{\forall (s_m)_{d_t} \in S} o(z_{d_t}, (a_m)_{d_{t-1}}, (s_m)_{d_t}) \sum_{\forall (s_n)_{d_t} \in S} q_{n,m} b(s_n)_{d_{t-1}}}$$

(3) 计算当前休眠周期平均时延  $E[D_{sl}]$  和平均功耗

$$\bar{E}((s_m)_{d_t}, (a_m)_{d_t})$$

(4) if  $(E[D_{sl}] < \delta_{th})$

$$r((s_m)_{d_t}, (a_m)_{d_t}) \leftarrow \bar{E}((s_m)_{d_t}, (a_m)_{d_t})$$

else

$$r((s_m)_{d_t}, (a_m)_{d_t}) \leftarrow \left( \bar{E}((s_m)_{d_t}, (a_m)_{d_t}) \right)_{\max}$$

endif

(5) 更新值函数  $V_n(b(s_n))$

$$V_n(b(s_n)) \leftarrow \max_{a_n \in A} \left[ \sum_{s \in S} b(s_n) r(s_n, a_n) + \gamma \sum_{o \in O} P(o | b(s_n), a_n) V_{n+1}(b'(s_n)) \right]$$

(6) 扩充信念状态集合

(7) 获取最优基站深度休眠时长

$$a_n^*(b(s_n)) = \arg \max_{a_n \in A} \left[ \sum_{s \in S} b(s_n) r(s_n, a_n) + \gamma \sum_{o \in O} P(o | b(s_n), a_n) V_{n+1}(b'(s_n)) \right]$$

表 3 仿真参数

|                            |                     |
|----------------------------|---------------------|
| 深度休眠时长 $T_{ds}$ (s)        | 8/16                |
| 轻度休眠时长 $T_{ls}$ (s)        | 6                   |
| 切换消耗的能量 $E_{ch}$ (J)       | 10                  |
| 深度休眠期间基站的能耗 $E_{DS}$ (W)   | 30                  |
| 轻度休眠期间基站的能耗 $E_{LS}$ (W)   | 50                  |
| 激活状态下基站的能耗 $E_{ac}$ (W)    | 100                 |
| 折扣因子 $\gamma$              | 0.5                 |
| 休眠状态下服务速率 $\mu_{sl}$ (包/s) | 5                   |
| 激活状态下服务速率 $\mu_{ac}$ (包/s) | 10                  |
| 业务到达率 $\lambda$ (包/s)      | 0.5, 1, 1.5, 2, 2.5 |
| $N_1$                      | 2                   |
| $N_2$                      | 1                   |

提出的策略关断时长为零。关断时长越长意味着节约的能量越多，两种策略下的能耗如图 6 所示。

图 6 表示的是 11 个周期内不同基站休眠策略的能耗对比，图中以基站始终处于开启状态(BS-on)的能耗为基准。从图中可见本文提出的基于 POMDP-based 方案在基站处于休眠状态下的能耗要低于基于业务门限值  $N(N=8)$  基站休眠策略的能耗。这也体现出了基于业务预测的基站休眠策略的优势，能够根据预测的业务量选择最为合适的基站休眠时长。在基站处于激活状态下，基于业务门限值的基站休眠策略能耗要低于本文提出的方案，这是因为前者会有一段休眠时间直到等到业务超过规定的业务门限值才开启基站。总的来说，当基站

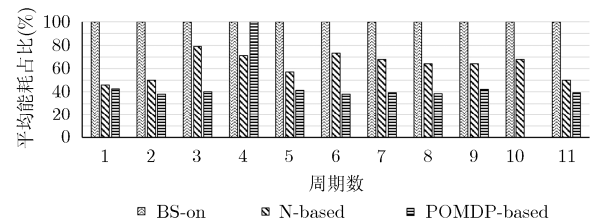


图 6 不同基站休眠策略间基站能耗对比

限值 N-based ( $N=8$ )<sup>[11]</sup>和本文提出的基于业务预测 POMDP-based 的基站休眠策略关断时长对比图。从图中可以看出，N-based 基站休眠时长在每个周期的关断时长都要低于当到达基站业务量较少时，第 4 个周期由于基站处于激活状态，因此本文

处于休眠状态时本文提出的基于业务感知的基站休眠策略节能效果更好。

#### 4 结束语

本文提出了一种基于 POMDP 业务感知的微基站休眠周期时长确定策略, 将休眠周期分成深度休眠和轻度休眠两类, 基站首先进入深度休眠, 当深度休眠持续时间结束时则转换到轻度休眠。基站在处于深度休眠时到达的业务请求暂缓到请求队列, 直到转换到轻度休眠时才处理队列中等待服务的业务请求。同时利用 POMDP 对业务状态进行建模, 实现对业务较准确的预测。仿真结果表明, 该策略可以根据业务感知提前确定微基站关断时长, 与基于业务门限值的基站关断机制相比节能效果更好。

#### 参考文献

- [1] ISMAIL Muhammad, ZHUANG Weihua, SERPEDIN Erchin, *et al.* A survey on green mobile networking: From the perspectives of network operators and mobile users[J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(3): 1535–1556. doi: 10.1109/COMST.2014.2367592.
- [2] WU Jingjin, ZHANG Yujing, ZUKERMAN Moshe, *et al.* Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey[J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(2): 803–826. doi: 10.1109/COMST.2015.2403395.
- [3] COMBES R, ELAYOUBI S E, ALI A, *et al.* Optimal online control for sleep mode in green base stations[J]. *Computer Networks*, 2015, 78: 140–151. doi: 10.1016/COMNET.2014.10.031.
- [4] SAKER L and ELAYOUBI S E. Sleep mode implementation issues in green base stations[C]. 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. Istanbul, Turkey, 2010: 1683–1688. doi: 10.1109/PIMRC.2010.5671926.
- [5] ALSHARIF M H, NORDIN R, and ISMAIL M. Cooperation management among base stations based on cells switch-off for a green LTE cellular network[J]. *Wireless Personal Communications*, 2015, 81(1): 303–318. doi: 10.1007/S11277-014-2129-9.
- [6] HU Jinming, HENG Wei, ZHANG Guodong, *et al.* Base station sleeping mechanism based on traffic prediction in heterogeneous networks[C]. Telecommunication Networks and Applications Conference, Sydney, NSW, Australia, 2015: 83–87.
- [7] JIN Zhe, PAN Zhiwen, LIU Nan, *et al.* Dynamic pico switch on/off algorithm for energy saving in heterogeneous networks [C]. 2015 IEEE 81st Vehicular Technology Conference, Glasgow, UK, 2015: 1–5.
- [8] EBRAHIM Aysha and ALSUSA Emad. Interference minimization through sleep mode based resource allocation for future femtocell networks[C]. 2015 IEEE International Conference on Communications (ICC). London, UK, 2015: 1679–1684.
- [9] ZHOU S, GONG J, YANG Z, *et al.* Green mobile access network with dynamic base station energy saving[OL]. <http://network.ee.tsinghua.edu.cn/papers/BSpowerSaving090830.pdf>. 2009.
- [10] SON Kyuho, KIM Hongseok, YI Yung, *et al.* Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks[J]. *IEEE Journal on Selected Areas in Communications*, 2011, 29(8): 1525–1536. doi: 10.1109/JSAC.2011.110903.
- [11] WU Jian, ZHOU Sheng, and NIU Zhisheng. Traffic-aware base station sleeping control and power matching for energy-delay tradeoffs in green cellular networks[J]. *IEEE Transactions on Wireless Communications*, 2013, 12(8): 4196–4209. doi: 10.1109/TWC.2013.071613.122092.
- [12] YU Yuping and FENG Kaiteng. Traffic-based DRX cycles adjustment scheme for 3GPP LTE systems[C]. Vehicular Technology Conference, Yokohama, Japan, 2012: 1–5.
- [13] HSU Chunghsien, FENG Kaiteng, and CHANG Chungju. Statistical control approach for sleep-mode operations in IEEE 802.16 m systems[J]. *IEEE Transactions on Vehicular Technology*, 2010, 59(9): 4453–4466. doi: 10.1109/TVT.2010.2070086.
- [14] GUNTHER A, OLIVER B, VITO G, *et al.* D2. 3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown[OL]. [https://bscw.ict-earth.eu/pub/bscw.cgi/d71252/EARTH\\_WP2\\_D2.3\\_v2.pdf](https://bscw.ict-earth.eu/pub/bscw.cgi/d71252/EARTH_WP2_D2.3_v2.pdf). 2010.
- [15] WANG Ke, LI Xi, and JI Hong. Traffic-based queue-aware scheduling for 3GPP LTE system[J]. *The Journal of China Universities of Posts and Telecommunications*, 2014, 21(2): 63–68. doi: 10.1016/S1005-8885(14)60287-9.
- [16] CASSANDRA A R and KAEHLING L P. Learning policies for partially observable environments: Scaling up[C]. Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, California, 1995. 362.
- [17] 冯奇, 周雪忠, 黄厚宽, 等. POMDP 基于点的值迭代算法中一种信念选择方法[J]. 北京交通大学学报, 2009, 33(5): 77–80. FENG Qi, ZHOU Xuezhong, HUANG Houkuan, *et al.* POMDP iterative algorithm based on the value of the point of a belief selection method[J]. *Journal of Beijing Jiaotong University*, 2009, 33(5): 77–80.
- [18] EUNSUNG Oh, KYUHO Son, and BHASKAR Krishnamachari. Dynamic base station switching-on/off strategies for green cellular networks[J]. *IEEE Transactions on Wireless Communications*, 2013, 12(5): 2126–2136. doi: 10.1109/TWC.2013.032013.120494.

陈前斌: 男, 1967年生, 教授, 博士生导师, 研究方向为个人通信、多媒体信息处理与传输、下一代移动通信网络、异构蜂窝网络等。

何小强: 男, 1992年生, 硕士生, 研究方向为绿色通信、基于能效的网络资源管理。

吴攀: 女, 1992年生, 硕士生, 研究方向为宽带无线接入。

唐伦: 男, 1973年生, 教授, 博士, 研究方向为新一代无线通信网络、异构蜂窝网、软件定义无线网络等。