

移动边缘计算的移动性管理研究

Mobility Management of Mobile Edge Computing

王秋宁/WANG Qiuning
谢人超/XIE Renchao
黄韬/HUANG Tao

(北京邮电大学, 北京 100876)
(Beijing University of Posts and
Telecommunications, Beijing 100876, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 01-0037-005

摘要: 提出了移动边缘计算(MEC)中基于马尔可夫决策过程(MDP)的虚拟机迁移策略。该策略可以在保证用户体验的前提下,能使得系统成本达到最优。每次用户进入新的 MEC 区域后,都通过该策略决定是否进行迁移。通过与常规迁移策略对比,仿真结果表明本策略总能达到最优情况,有效地降低了系统成本。

关键词: MEC; MDP; 移动性管理; 成本最优

Abstract: In this paper, a virtual machine migration strategy based on Markov decision process (MDP) in mobile edge computing (MEC) is proposed, which can minimize the system cost while ensuring the user experience. After entering a new MEC area, the users can be determined whether or not to migrate by this policy. Compared with the conventional migration strategy, the simulation results show that this strategy can always achieve the optimal situation and effectively reduce the system cost.

Keywords: MEC; MDP; mobility management; cost optimization

据 Cisco 最新发布的预测报告:全球移动数据流量在 2016—2021 年间将增长 7 倍,移动数据流量将在 2016—2021 年保持 47% 的年增长率,到 2021 年达到每月 49.0 EB^[1]。除了爆炸性的流量增长,虚拟现实所需的接近实时响应时间和物联网所需的海量数目连接等要求也都超出了现有网络的承受能力。为了应对上述挑战,业界在 5G 移动网络中采用了移动边缘计算(MEC)技术。MEC 的核心思想是将计算能力从移动网络的数据中心转移到无线接入网边缘,向移动边缘应用程序提供 IT 服务环境和计算功能,从而可以将业务本地化,在接入网处理用户请求。这减少了用户等待时间,确保了高效的网络运行和服务交付,也缓解了网络流量的回传需求,降低了网络运营成本。

尽管 MEC 给 5G 带来巨大的潜在收益,但将 MEC 应用于实践仍面临许多挑战,移动性管理正是关键瓶颈之一。移动性管理的功能是跟踪移动用户设备(UE)并将其与适当的基站(BS)相关联,使得移动系统能够

交付数据和服务。这一技术已经广泛应用于传统的异构蜂窝网络^[2],能够实现动态移动性管理并保证高数据速率和低误码率。但是现有的移动性管理技术并不能直接应用到 5G 网络中,因为它忽略了 MEC 服务器上的计算资源对切换策略的影响。当 UE 在 MEC 处进行卸载计算时,确保服务的连续性是非常重要的,比如车联网中,车辆需要实时地上传位置信息,并且 MEC 需对周围的车辆信息进行汇总计算并给出指导或警告。当一个移动用户从一个区域移动到另一个区域时,既可以继续在前一个区域的 MEC 上运行服务,并通过回程网络将数据传输给用户,也可以将承载应用程序的虚拟机或数据迁移到新区域中的 MEC。这两种情况都会有成本:第 1 种情况是数据传输成本,第 2 种情况是迁移成本,并

且两种情况中用户获得服务的延迟也不相同。因此,虚拟机迁移决策既要考虑到系统的成本,也要保证服务质量。我们分析蜂窝网络中用户移动规律并归纳了马尔可夫链,在此基础上加入成本回报函数,将成本优化问题等效为马尔可夫决策过程,并通过仿真证明本策略能达到节约成本、能量高效的目的。

1 移动边缘计算

根据欧洲电信标准化协会(ETSI)的定义,MEC 通过在无线接入网部署通用服务器,为无线接入网提供 IT 和云计算能力^[3]。MEC 系统允许移动设备将计算任务卸载到网络边缘节点,如基站和无线接入点等,既缓解了云服务器远离用户带来的高延迟问题,又增强了移动设备处理数据的能力。因而 MEC 迅速成为

收稿日期: 2017-12-13
网络出版日期: 2018-01-11

5G 的一项关键技术,使得接入网具有了高带宽和低延迟地处理信息,感知网络上下文信息和向第三方边缘应用开放等能力,有助于 5G 网络达到低时延、高能效、高容量和高可靠等技术指标。

MEC 的系统架构如图 1 所示,关键部件包括移动设备、演进型节点(eNB)、汇聚节点和 MEC 服务器。MEC 服务器是小型数据中心,由运营部署在靠近最终用户的位置,并且可以和基站共同放置,但通常都是通过汇聚节点控制几台基站。通过网关,服务器通过 Internet 连接到数据中心。移动设备和基站之间通过空中接口通信,使用先进的无线通信技术建立无线链路。

MEC 平台可分为 3 层:最底层为基础设施层,通过网络功能虚拟化的方式为上层提供计算、存储和转发等物理资源;第 2 层为应用平台层,它是移动边缘应用程序所需的基本功能集合,能提供移动边缘服务的环境,它具有服务注册、流量管理和数据分析等功能;最顶层为移动边缘应用程序,它在 MEC 平台上的虚拟机(VM)运行,并且可以通过应用程序编程接口(API)与移动边缘平台进行交互以调用 MEC 的功能。

在 MEC 中,需要考虑的关键问题

之一是如何保证用户在移动过程中获得服务的连续性。当 UE 正在使用 MEC 虚拟机的边缘移动应用程序功能时,如果 UE 的地理位置从原来的 MEC 区域移动到新的 MEC 区域,为了保证 MEC 服务的可持续性,有 3 种选择:第 1 种选择是增加基站的传输功率从而保证了用户的服务质量,这种方法适用于用户移动速度慢且移动距离较近的情况;第 2 种选择是原 MEC 通过回程链路与 UE 进行通信,这适用于用户和原 MEC 的距离不太远的情况,以免用户获得服务的延迟过高;第 3 种选择是进行虚拟机迁移。原 MEC 将用户正在使用的虚拟机数据发送到新的 MEC,并关闭虚拟机,新的 MEC 开启虚拟机接收数据,继续向用户提供计算服务。相比于第 2 种选择,传输虚拟机数据的成本较高,但却降低了用户获得服务的延迟。当然对于 UE 来说,不是所有的计算任务都需要卸载到 MEC,并且也不是所有的移动边缘应用都会受到 UE 移动性的影响,比如对 UE 上传到云端数据进行简单处理并上传的应用程序,或者能够在 UE 切换 MEC 之后快速重建相关状态的应用程序。考虑到文中的研究情景,接下来提到的移动边缘应用程序都是指 UE 需要全部卸载到 MEC 进行计算的,并且

会受到 UE 移动性影响的应用程序。

2 移动性管理的相关工作

文献[4]将蜂窝网络中的用户移动描述为马尔可夫链模型,分析了各状态之间的转化关系,并在此基础上计算了虚拟机迁移到最佳 MEC 之后用户和最佳 MEC 的平均距离、用户获得服务的平均时延、进行虚拟机迁移的平均成本和虚拟机迁移的平均时延等数据。仿真结果表明:与虚拟机 50% 和 10% 的部分迁移相比,虚拟机全体迁移需要成本最高,但用户服务延迟最低。

文献[5]将蜂窝网络简化为一维移动模型,进一步将 VM 迁移策略制定为连续时间马尔可夫决策过程(CTMDP),并尝试在启动 VM 迁移时找到最佳阈值策略。该策略允许在用户体验质量和服务迁移产生的成本之间取得良好平衡。仿真结果表明:与其他两个基本策略相比,所提出的服务迁移决策机制总是达到最大的期望收益。

文献[6]通过研究设备移动性的模式来解决辅助移动性的机会计算卸载问题。首先利用定义为接触时间和接触率的移动性的统计特性制定最优的机会卸载模型,然后利用凸优化的方法确定要卸载到其他设备的计算量。人机交互下的仿真证明了此方案的效率在各种设置下比基准情况下能获得更高的计算成功率。

文献[7]通过预测用户的移动进一步优化了虚拟机迁移策略,提出了基于移动性的服务迁移预测方案(MSMP),在成本和服务质量之间采取了折中。该方案有 3 部分:(1)提前估计用户在整个网络漫游时可以从各个 MEC 服务器接收到的吞吐量;(2)估计用户执行切换时的时间窗;(3)执行 VM 的迁移,根据吞吐量选择最优的 MEC 服务器。仿真结果表明:该方案相对于文献[5]提出的方案能够将延迟降低 35%;但是迁移成本更高并且会需要采集信息,并预估

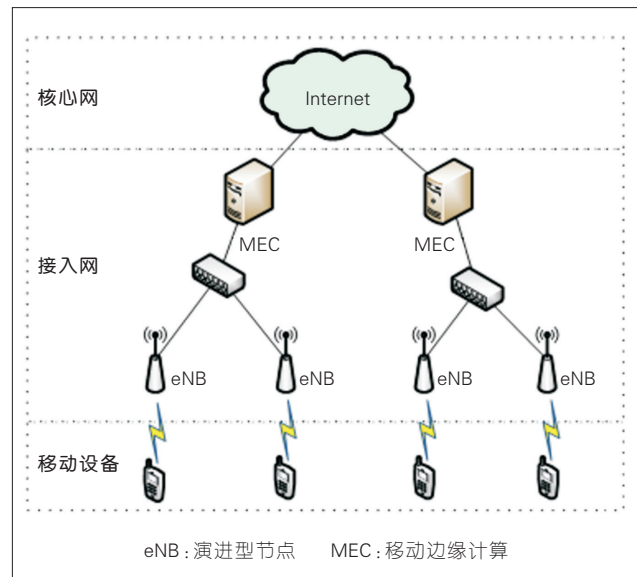


图 1
MEC 系统架构

吞吐量。

现有的方案多是建立在用户移动的一维模型,或者是通过距离因素简化后的模型,并未考虑用户在蜂窝网络中移动的实际情况。借助于文献[4]提出的蜂窝网络模型,文章中我们将研究用户的平面移动时的成本最优策略,同时保证用户体验。相比于上述文献所述方案,文中我们的研究贡献主要有以下几点:

(1)对蜂窝网络中的用户移动规律进行研究,发现按照用户和MEC的距离对蜂窝小区进行分类不够准确,我们对每层蜂窝网络进行了更细化的分类并给出了状态转移概率。

(2)在此二维模型基础上,从成本考虑将虚拟机迁移建模成了马尔可夫决策过程,同时为了保证用户服务的质量,加入了MEC与UE依靠回程链路的最远传输距离这一新的约束条件。

(3)给出了策略迭代算法以求得最优解,并对文中的马尔可夫决策过程进行仿真,结果显示相比于常规策略,本策略能够取得最优成本。

3 移动边缘计算的模型构建

本节主要在蜂窝网络的用户移动情况基础上,确定马尔可夫决策过程的决策时刻、行动集、系统状态集、状态转移概率和回报函数,进而确定优化目标函数和解决方法。

第3代合作伙伴(3GPP)网络采用正六边形小区形式覆盖所有区域,每个基站的覆盖范围被近似视为一个正六边形,如图2所示。实际中MEC通常会控制几个基站的区域,我们为了突出研究模型,简化为每个MEC控制一个基站,也即每个正六边形区域各自属于一个MEC节点。为了简化叙述,称用户进行边缘应用程序计算所在的MEC为通信MEC。我们研究的前提约束条件是保证对用户的服务质量,也即将用户与通信MEC的延迟控制在一定范围内。这可以通过限定用户和通信MEC的最

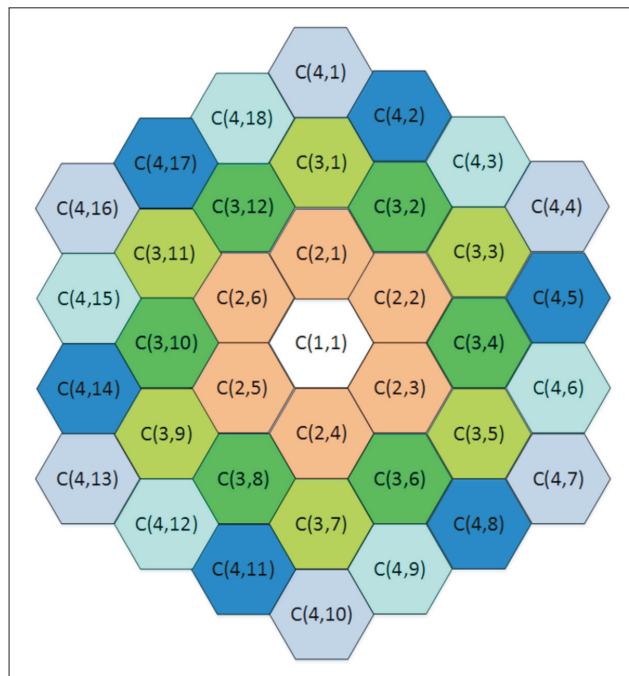


图2
3GPP蜂窝网络

大距离来保证,我们限定用户和通信MEC的最大距离为 k 个小区。比如 $k=3$ 时,在图2中,如果此时通信MEC在小区C(1,1),那么用户与其通信的最远位置为小区C(4,1-18);如果用户进入更远的小区,就会跳过策略判断,直接将虚拟机迁移至用户所在小区的MEC,以保证服务质量。

3.1 决策时刻集和行动集

决策时刻集 $T=\{0,1,2,\dots\}$,每当用户进入到新的小区后,开始进行判决策略决定是否需要进行虚拟机迁移。用户在小区的停留时间服从参数为 λ 的指数分布,也即决策时刻的差服从指数分布。

系统的行动集 $A=\{a_1, a_2\}$, a_1 表示进行虚拟机迁移, a_2 表示UE仍和原MEC进行通信,采用的行动 $a \in A$ 。行动集 A 取决于系统状态,但不受决策时刻的影响,也不受历史状态影响。

3.2 系统状态和转移概率

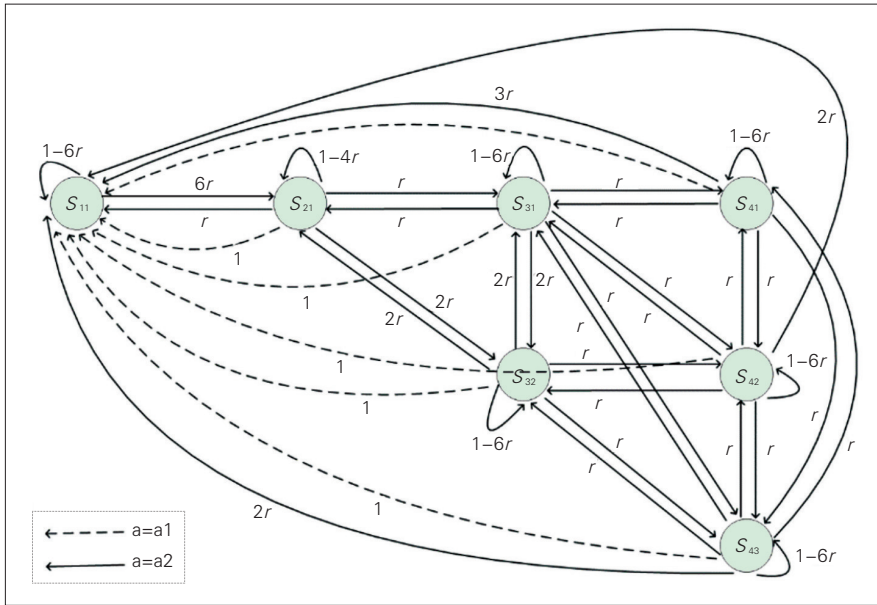
在图2中,固定通信MEC的位置为C(1,1),如果执行策略后将虚拟机迁移到新的小区,就把新的MEC编

号为C(1,1),其余小区编号做出相应改变。同时默认用户的初始位置为C(1,1),即恰好进行完一次虚拟机迁移。处于蜂窝网络中的UE,周围有6个小区。假设UE移动到任意一个邻居小区的概率均是 r ,则UE在小区不移动的概率是 $(1-6r)$ 。

按照小区由内而外的顺序分为 $(k+1)$ 层,第 n 层($n>1$)共有 $6(n-1)$ 个小区。将第 n 层的小区分为 $(n-1)$ 类: $S_{nm}=\{C(n, m+i(n-1)) | i=0,1,2,3,4, 1 \leq m \leq (n-1), m \in N\}$,设 $S_{11}=\{C(1,1)\}$ 。所有类的集合即为系统状态 S : $S=S_{11} \cup \{S_{nm} | 2 \leq n \leq (k+1), 1 \leq m \leq (n-1), n \in N, m \in N\}$ 。

之所以要把第 n 层划分为 $(n-1)$ 类是因为同层不同的类节点之间的状态转移概率是不同。 $k=3$ 时的状态转移如图3所示,易知这是一个马尔可夫链。

结合图3和蜂窝网络结构,可知系统转移概率是一个和状态 S 及行动集 A 有关的函数。设当前状态为 $s=s_{nm}$,下一状态为 s' 。这里 $4 \leq n \leq (k+1)$, n 更小的情况可以在图3中直接读出。我们可以推出转移概



▲ 图3 4层蜂窝网络的状态转移示意图

率为：

$$P(s'|s_{nm}, a) = \begin{cases} 1 & s' = s_{11}, a = a_1 \\ rs' = s_{n(m \pm 1)}, s' = s_{(n \pm 1)m}, s' = s_{(n+1)(m+1)}, a = a_2 \\ rs' = s_{(n+1)m}, m = 1, a = a_2 \\ rs' = s_{(n-1)(m-1)}, m \neq 1, a = a_2 \\ 3rs' = s_{11}, m = 1, a = a_2 \\ 2rs' = s_{11}, m \neq 1, a = a_2 \\ 1-6rs' = s_{11}, n = k+1, a = a_2 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

3.3 回报函数

本策略的主要目标是能量高效，因此回报函数设定为传输耗能。传输耗能有两类：第1类为虚拟机迁移耗能 C_m ，主要是将虚拟机数据传输到新的 MEC 区域的耗能，关闭和开启虚拟机服务的耗能，前者与数据量大小和传输距离有关，后者是一个常量，综合后将 C_m 当作常量；第2类为 UE 和通信 MEC 的数据传输耗能 C_t ，是一个和传输距离有关的变量，因此表示为 $C_t = c(n-1)$ ， n 表示用户所在小区的层数， c 表示单位距离的传输耗能。当虚拟机迁移完成后，由于用户和通信 MEC 的距离很近，所以这时的传输耗能近似为 0。由于每次决定虚拟机迁移都是在用户刚进入新小区时，在虚拟机迁移过程中用户和通信 MEC 的传输耗能可以忽略不计，所以迁移耗能和传输耗能不会同

时存在。当虚拟机迁移完成后，由于用户和通信 MEC 的距离很近，所以这时的传输耗能近似为 0。综合两类情况可知回报函数是一个和系统状态及行动集有关的函数，当系统状态 $s = s_{nm}$ ，推得回报函数为：

$$R(s_{nm}, a) = \begin{cases} C_m a = a_1 \\ c(n-1) a = a_2 \end{cases} \quad (2)$$

3.4 目标函数

结合前 3 节的推导，基于行动集、状态集、转移概率和汇报函数，定义目标函数 $V(s, a)$ 为系统在状态 s 和策略 a 时的期望回报：

$$V(s, a) = E \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a) \right] \quad (3)$$

目标函数表示的是系统的总成本，其中 γ 是折扣因子，表示对未来自报的重视程度； s_t 表示 t 时刻的系统状态，以递归形式表示为：

$$V(s, a) = R(s, a) + \sum_{s' \in S} \gamma s' |s, a| V(s') \quad (4)$$

目标函数由当前回报和一定比重的未来回报组合而成。假设策略集合为 π ，那么由上述分析可知 $A = \pi(S)$ 。此模型下，行动集和系统状态都是有限的，而时间长度是无限

的，所以采用马尔可夫决策的无限阶段折扣模型。最优的目标函数为：

$$V^*(s) = \min_{\pi} V(s, \pi(s)) \quad (5)$$

最优策略为：

$$\pi^*(s) = \arg \min (V(s)) \quad (6)$$

3.5 策略迭代法

本模型中，虽然时间集 T 是无限的，但是系统的状态集 S 是有限的。可以使用策略迭代法，在有限的时间内求得最优策略。具体方法如下：

步骤 1：确定一个决策策略 f ， $f \in \pi$ ；

步骤 2：对于所有的 $s \in S$ ，求解下面 $|S|$ 个方程

$$V(s) = R(s, f(s)) + \sum_{s' \in S} \gamma s' |s, f(s)| V(s') \quad (7)$$

步骤 3：将上一步求得的 $V(s)$ 带入式 (8)，求得对于所有 $s \in S$ ，满足式子条件的新决策策略 f' （若有多个满足条件的则任取一组）：

$$\begin{aligned} \min_{a \in A} \left\{ R(s, a) + \sum_{s' \in S} \gamma s' |s, a| V(s') \right\} = \\ R(s, f'(s)) + \sum_{s' \in S} \gamma s' |s, f'(s)| V(s') \\ \geq R(s, f(s)) + \sum_{s' \in S} \gamma s' |s, f(s)| V(s') \end{aligned} \quad (8)$$

步骤 4：如果对于所有的 $s \in S$ ，式 (8) 的等号都成立，则返回最优策略 f' ；否则令 $f = f'$ 并返回步骤 2。

4 仿真结果

与文中我们提出的马尔可夫决策策略进行对比的是频繁迁移策略和不迁移策略。频繁迁移策略是当用户进入新的小区后就进行虚拟机迁移；不迁移策略是指除非用户距通信 MEC 为 k ，否则不进行虚拟机迁移。假设仿真中用到的数值如下：用户和通信 MEC 的最大距离 $k=9$ ，蜂窝网络层数 n 的最大值为 10，折扣因子 $\gamma=0.5$ ，虚拟机迁移的耗能 $C_m=100$ ，用户和通信 MEC 的单位距离耗能 $c=30$ ，用户向其他小区移动概率 $r=0.1$ 。初始状态用户位于通

信 MEC 所在小区。

如图 4 所示,横坐标是用户和通信 MEC 的距离,即不同小区之间的距离;纵坐标是系统耗能。图中共有 3 条曲线,分别代表总是进行虚拟机迁移策略、不进行虚拟机迁移策略和文中所提策略。对比之后可以看出:总是迁移策略的每次耗能是固定的,即为虚拟机迁移的耗能;不迁移策略中用户需要和原 MEC 进行通信,耗能随距离增加;文中所提策略可在两种策略中取得最优情况,即最小耗能。

图 5 中,横坐标用户移动次数,用户每移动一次都会进行一次小区选择。纵坐标是累计的系统总耗能,分别展示了 3 种策略的耗能情况。对比这 3 种情况,可以看出虽然前期不迁移策略耗能和本文策略相同,但随着用户的移动距离逐渐变远,文中策略的节能效果就体现了出来。

须考虑的问题,我们从能量高效的角度给出了一种虚拟机迁移策略。

为了使模型更加贴近实际,我们首先对蜂窝网络中的用户移动规律进行研究,发现仅按照用户和 MEC 的距离对蜂窝小区进行分类是可以改善的,并对每层蜂窝网络进行了更细化的分类并给出了状态转移概率;之后分析以能量高效为前提的马尔可夫决策模型,主要是通过对 MEC 移动管理中的耗能进行分析量化,反映到回报函数中,进而能通过解决马尔可夫决策模型达到能量高效的目的。通过仿真实验证明:本策略相比于常规策略能够有效地节约能量。接下来我们的研究方向可以考虑结合用户移动性调整模型,或是设计 MEC 移动性管理系统模块。

参考文献

- [1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper[R/OL].(2017–03–28)[2017–12–13].<https://www.cisco.com/c/en/us/solutions/>

collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

- [2] DAMNANOVIC A. A Survey on 3GPP Heterogeneous Networks[J].IEEE Wireless Communications,2011,18(3):10–21. DOI: 10.1109/MWC.2011.5876496
- [3] Mobile Edge Computing—A key technology towards 5G[EB/OL].(2017–09–03)[2017–12–13].http://10.3.200.202/cache/9/03/www.etsi.org/96365783d6e7ab98a99488bf5507dfbe/etsi_wp11_mec_a_key_technology_towards_5g.pdf
- [4] TALEB T, KSENTINI A. An Analytical Model for Follow Me Cloud[C]//Proceeding of IEEE Global Communications Conference (GLOBECOM). USA: IEEE, 2013:1291–1296
- [5] KSENTINI A, TALEB T, CHEN M. A Markov Decision Process–Based Service Migration Procedure for Follow Me Cloud[C]//2014 IEEE International Conference on Communications (ICC). USA:IEEE, 2014: 1350–1354. DOI: 10.1109/ICC.2014.6883509
- [6] WANG C, LI Y, JIN D. Mobility-Assisted Opportunistic Computation Offloading [J]. IEEE Communications Letter, 2014, 18(10): 1779–1782. DOI: 10.1109/LCOMM.2014.2347272
- [7] NADEMBEGA A, HAFID A S, BRISEBOIS R. Mobility Prediction Model–Based Service Migration Procedure for Follow Me Cloud to Support QoS and QoE [C]// 2016 IEEE International Conference on Communications (ICC). USA:IEEE, 2016:1–6. DOI: 10.1109/ICC.2016.7511148

5 结束语

MEC 中的用户移动性管理是必

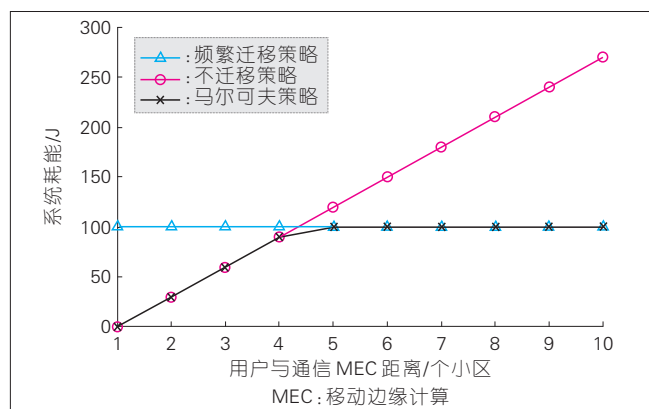


图 4
用户与通信 MEC 不同
距离的系统耗能

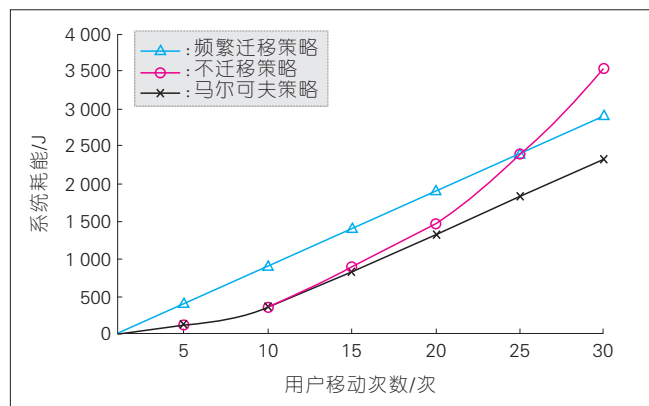
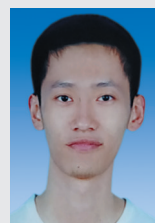


图 5
系统总耗能

作者简介



王秋宁, 北京邮电大学在读硕士研究生; 主要研究方向为移动边缘计算和内容分发。



谢人超, 北京邮电大学副教授、硕士生导师; 主要研究方向为未来网络体系架构、信息中心网络与边缘计算等; 主持国家与省部级项目 4 项; 已发表论文 40 余篇, 申请授权专利 10 余项。



黄韬, 北京邮电大学教授、博士生导师; 主要研究方向为未来网络体系架构、软件定义网络、信息中心网络等; 主持国家与省部级项目 10 余项; 已发表论文 100 余篇, 申请授权专利 40 余项。