# Can Mobile Cloudlets Support Mobile Applications?

Yujin Li and Wenye Wang
Email: yli27@ncsu.edu, wwang@ncsu.edu
Dept. of Electrical and Computer Engineering, NC State University, USA

***Abstract*—A *mobile cloudlet* is a set of resource-rich mobile devices—referred to as *cloudlet nodes*—that an initiator mobile device can connect to for services. In this paper, we examine the fundamental mobile cloudlet properties that unfold *whether and when a mobile cloudlet can provide mobile application service*. Specifically, we investigate the *cloudlet size*, *cloudlet node's lifetime and reachable time*. Traces and mathematical analysis demonstrate that 1) the more frequently mobile devices meet, the larger the pool of computing resources an initiator can access; 2) intermittent connection between devices has little adverse effect on the optimal computing performance of mobile cloudlet in the long run; 3) the ratio $E(T_C)/[E(T_I)+E(T_C)]$ indicates the connection likelihood of an initiator and a cloudlet node (i.e., reachability of the cloudlet node), where $T_C$ and $T_I$ are their contact and inter-contact time. We further derive upper and lower bounds on computing capacity and computing speed of a mobile cloudlet. An initiator can use both bounds to decide whether to offload its task to remote clouds or local mobile cloudlets for better mobile application services.**

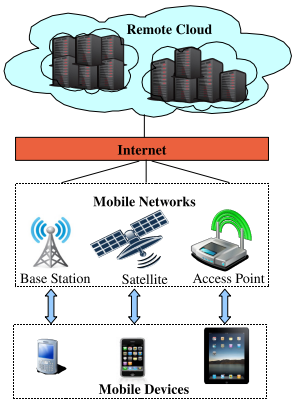



Fig. 1. MCC uses remote cloud.

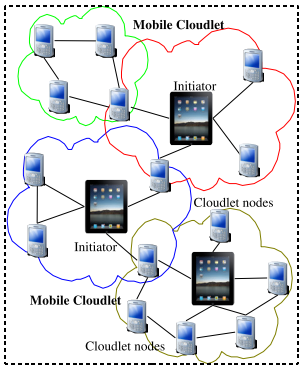



Fig. 2. MCC uses mobile cloudlet.

## I. Introduction

With the emerging cloud computing [1] and the explosive growth of mobile applications, mobile cloud computing (MCC) has become a promising technology for mobile services. In MCC, mobile devices, such as smart phones and tablets, can offload data storage and computational task onto the cloud through wireless communications, thereby overcoming their limited capabilities regarding processing power, storage capacity, and battery lifetime [2].

Traditionally, MCC uses the *client-server communication model* [2, 3] shown in Fig. 1. The *remote cloud* provides data storage and computing service while mobile devices are clients to access the service through wireless networks, mainly cellular and WLAN (Wireless Local Area Network). When accessing a remote cloud is costly due to long WAN (Wide Area Network) latencies, a mobile user can exploit nearby computers [4, 5] that are well-connected to the Internet and uses cloud service over a high-bandwidth WLAN. The vast computation resources on remote cloud servers can enable computation intensive applications, such as image processing for video games, optical character recognition (OCR), and augmented reality, run on mobile devices.

On the other hand, a *peer-to-peer communication model* for MCC [6–9] is proposed in light of the increasing memory and computational power of mobile devices [10]. As shown in Fig. 2, *a group of nearby mobile devices can connect by WiFi or Bluetooth to form a mobile cloudlet*, in which mobile devices (referred to as cloudlet nodes) can be computing service providers as well as clients of the service. By dividing the task among cloudlet nodes, the initiator mobile device could speedup computing and conserve energy. Users can get direct cloud computing access instantly through interactions among one another, eliminating the communication latency and data roaming charges introduced by the cellular networks. Mobile cloudlet is appealing to users pursuing a common goal in group activities, such as multimedia sharing for audience at an event and language translation for a group of tourists in a foreign country. The major concern of using mobile cloudlet resides in the limited computing power of mobile devices and the unstable connections between cloudlet nodes due to node mobility. Hence, a fundamental question for mobile cloudlet development is *whether and under what conditions mobile cloudlet is feasible for providing mobile application services*.

To answer this question, we study the computing performance of a mobile cloudlet through investigating the properties of a mobile cloudlet with respect to *cloudlet size*, cloudlet node's *lifetime* and *reachable time*. Cloudlet size is defined as the number of mobile devices that a mobile initiator encounters within time $\tau$. Cloudlet size decides the potential computing resources in a mobile cloudlet. A cloudlet node's lifetime is from its first contact to its last contact with an initiator before the task expires. A cloudlet node's lifetime is its maximum time to perform computation for an initiator, thus relates to the optimal computing performance of the mobile cloudlet. Reachable time is the total contact duration between a cloudlet node and the initiator within time $\tau$, reflecting the cloudlet node's connection likelihood for task dissemination and retrieval. These properties not only determine how much

This work is supported by the NSF Career Award CNS-0546289 and NSF Award CNS-1018447.

computing resource a mobile cloudlet can provide but also implicate how reliable a mobile cloudlet is.

The definitions of cloudlet properties imply that they are determined by the contact and inter-contact between mobile devices. Hence, we examine the mobile cloudlet properties using contact-based mobility traces and mathematical analysis of the alternating contact and inter-contact process between mobile users. Traces and analysis results together prove that the mobile cloudlet size follows negative exponential growth with parameter $1/E(T_I)$, and the expected lifetime and reachable time grow linearly with $\tau$ with the increase rates 1 and $E(T_C)/[E(T_C)+E(T_I)]$, where $E(T_C)$ and $E(T_I)$ are the expectations of contact and inter-contact time. Our results suggest that 1) the more frequently mobile devices meet each other, the larger the pool of computing resources the initiator can access; 2) a cloudlet node's lifetime is close to $\tau$ when $\tau$ is large, indicating little adverse effect of intermittent connectivity on the optimal computing performance of mobile cloudlet; 3) the ratio $E(T_C)/[E(T_C)+E(T_I)]$ indicates the probability that a cloudlet node can be reached by the initiator (i.e., the connection likelihood) for task dissemination and retrieval, thus can be used for choosing reliable cloudlet nodes.

Based on the above properties of mobile cloudlet, we study the computing capacity and speed of a mobile cloudlet. The upper and lower bounds of mobile cloudlet computing capacity are derived based on the lifetime and reachable time of cloudlet nodes, respectively. As time $\tau$ goes to infinity, the long-term computing speed (the computing capacity over time $\tau$) is bounded by two constants $(n-1)V$ and $\frac{(n-1)VE(T_C)e^{-\rho/E(T_C)}}{E(T_C)+E(T_I)}$, where $n$ is the number of nodes in the network, $V$ is the computing speed of each device, $\rho$ is the maximum transmission time for task dissemination and retrieval. The bounds on computing capacity and speed of mobile cloudlet can both be used for an initiator to decide whether to execute a task in mobile cloudlet or remote cloud. When the task's computational demand C (or $C/\tau$) is larger than computing capacity (or speed) upper bound, the task needs to be offloaded to remote clouds for computing; when C (or $C/\tau$) is less than computing capacity (or speed) lower bound, the task can be executed by mobile cloudlet during contacts of the initiator and cloudlet nodes; otherwise, the task can be offloaded to either remote clouds or mobile cloudlet.

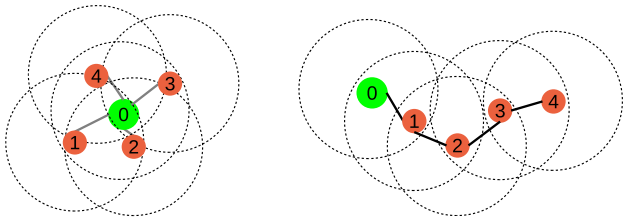
The rest of this paper is organized as follows: we present models and definitions in Section II; we study mobile cloudlet properties using two mobility traces in Section III; in Section IV, we mathematically analyze mobile cloudlet properties to validate and augment our trace observations; we derive bounds on mobile cloudlet computing capacity and speed in Section V; we conclude this paper in Section VI.

## II. Models and Definitions

### A. Network Model

We consider a mobile cloud computing network of $n$ mobile nodes on a torus surface $\Omega_n = [0, \sqrt{\frac{n}{\lambda}}]$, where $\lambda$ is the spatial density of mobile users. Suppose each mobile device has a transmission radius $r$. Denote by $\mathcal{X}_t = \{X_1(t), \ldots, X_n(t)\}$ the positions of users at time $t$, two nodes are in contact if $||X_i(t) - X_j(t)|| \leq r$ and out of contact otherwise. We assume that the mobility process of a node is stationary and ergodic that a node's location $X_i(\cdot)$ has uniform stationary distribution in the network area [11]. Mobility processes of nodes are independent and identically distributed (i.i.d.).

Without loss of generality, we assume that a mobile user needs to offload a task to nearby mobile devices at time 0. As shown in Figs. 3(a) and 3(b), initiator 0 can connect to cloudlet nodes $1, 2, 3, 4$ by direct communication links or multi-hop communication paths. In the *one-hop mobile cloudlet* (Fig. 3(a)), direct connections between the initiator and the cloudlet nodes ensure short delay in task transfer and easy management of task distribution and retrieval. In the *multi-hop mobile cloudlet* (Fig. 3(b)), employing mobile devices in *multi-hop* range provides the potential to utilize more devices in a large area. However, multi-hop communications incur longer delay and unreliable task dissemination and retrieval due to node mobility. Fesehaye et al. [5] show that when the maximum number of wireless hops in a cloudlet is larger than two, accessing cloudlet nodes incurs longer data transfer delay than directly accessing remote cloud through 3G/4G network. Hence, we only consider the more practical one-hop mobile cloudlet in this paper.




(a) One-hop mobile cloudlet (b) Multi-hop mobile cloudlet

Fig. 3. The initiator device 0 can distribute tasks to cloudlet nodes $1, 2, 3, 4$ through one-hop or multi-hop communications.

Apparently, node mobility affects structure of mobile cloudlet [12]. Specially, how frequent nodes meet and how long they stay in contact affect the size and stability of a mobile cloudlet. In turn, contact and inter-contact time between nodes influence the computing capacity and performance of a mobile cloudlet as tasks can only be distributed and retrieved when there are communication links between an initiator and cloudlet nodes.

Contact time $T_C$ is also called link lifetime or link duration. Zhao et al. [13] find that the PDF of link lifetime can be approximated by exponential distribution with parameter characterized by the ratio of average node speed to effective transmission range. Hence, in this paper, we assume that $T_C$ follows *exponential* distribution with parameter $\lambda_C$. Inter-contact time $T_I$ has been shown to exhibit exponential tail decay under many mobility models (such as random waypoint and Brownian motion) [14]. Analysis of a diverse set of mobility traces [15] also reveals that $T_I$ follows a power law decay

up to a characteristic time, beyond which $T_I$'s distribution decays exponentially. For the simplicity of analysis, we assume that $T_I$ has an *exponential* distribution with parameter $\lambda_I$. Note that our analysis and results can be easily extended to the case when $T_C$ and $T_I$ follow other distributions.

### B. Mobile Cloudlet Properties

Suppose the delay requirement of an initiator's task is $\tau$, mobile devices that meet the initiator before the task expires have the potential to provide computing services, thus can form a mobile cloudlet for the task computation. We assume that all nodes are willing to support cloudlet computing. Hence, *a mobile cloudlet is dynamically formed by the nodes that the initiator encounters over a period of time* $\tau$. Formally, we define a mobile cloudlet as follows.

**Definition 1.** (*Mobile Cloudlet*) For $\tau \in \mathbb{R}_+$, let $\mathcal{C}_\tau$ be the mobile cloudlet for an initiator $v_i$ with a task to compute within delay $\tau$. $\mathcal{C}_\tau$ is the set of nodes $v_i$ encountered within time $\tau$, where cloudlet node $v_j \in \mathcal{C}_\tau$ if and only if $v_i \neq v_j$ and there exists a link between $v_i$ and $v_j$ at a time $0 \leq t \leq \tau$.

The task dispatching, computing, and retrieving can only be performed after the first contact between an initiator and a cloudlet node and before their last contact within time $\tau$. We find the following definition useful.

**Definition 2.** (*Lifetime*) For any cloudlet node $v_j \in \mathcal{C}_\tau$ for an initiator $v_i$, $v_j$'s *lifetime* $LT(\tau) = exit - entr$, where its *entrance time* to $\mathcal{C}_\tau$

$$entr \triangleq \inf_{0 \leq t \leq \tau} \{t : ||X_i(t) - X_j(t)|| \leq r\},$$

and its *exit time* from $\mathcal{C}_\tau$

$$exit \triangleq \inf_{0 \leq t \leq \tau} \{t : \forall t' >= t \text{ and } t' \leq \tau, ||X_i(t') - X_j(t')|| > r\}.$$

In an optimal situation, an initiator utilizes a cloudlet node's whole lifetime for computing. A cloudlet node receives tasks at its entrance time; then it can compute the tasks during its lifetime even when it is not in contact with the initiator; it sends back the tasks right before its exit time. Hence, the lifetime of a cloudlet node can be used to provide an *upper bound* on the computing capacity of mobile cloudlet.

Nevertheless, the task dissemination and retrieval can only be performed during the contact period of an initiator and a cloudlet node. The percentage of time that a cloudlet node is in contact with the initiator shows how likely the initiator can reach it . In order to study the reachability of cloudlet nodes and reliability of a mobile cloudlet, we define the *reachable time* $RT(\tau)$ *as the total contact duration between a cloudlet node and an initiator within time* $\tau$.

Based on these mobile cloudlet properties, we can study mobile cloudlet computing performance and find out whether and when a mobile cloudlet can serve mobile applications.

## III. Mobile Cloudlet in Traces

The cloudlet size and cloudlet node's lifetime and reachable time are determined by contacts and inter-contacts between the initiator and cloudlet nodes, which have been studied using mobility traces in mobile wireless networks. Hence, we start examining the mobile cloudlet properties using mobility traces.

### A. Mobility Traces

Mobility traces record mobile users' access to base stations or access points (i.e., infrastructure based traces), or GPS locations (i.e., GPS based traces), or contact and inter-contact time (i.e., direct contact based traces). Because mobile cloudlet exploits contacts among nodes for computing, we choose the direct contact based traces. Moreover, since mobile cloudlet is promising for a social group sharing common tasks, mobility traces of users in social groups are preferred. Therefore, we select the *Cambridge/haggle2009* dataset [16] that includes several traces of Bluetooth sightings by groups of users carrying small devices (iMotes) for several days in office and conference.

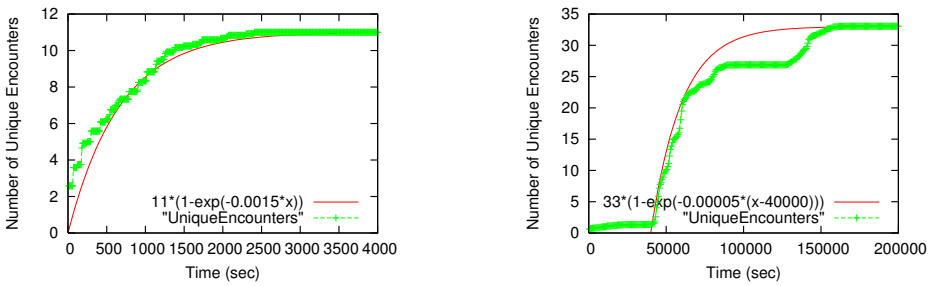
In Cambridge/haggle2009 data collection, experiment 2 distributed iMotes to 19 graduate students from the System Research Group at University of Cambridge for around 5 days in 2005. Number of contacts, contact and inter-contact time among nodes were collected. Only 12 iMotes were used to produce trace file *Exp2*, while others were discarded because of hardware resets. Similarly, experiment 3 distributed iMotes to 50 students attending the student workshop at the IEEE Infocom Conference in Grand Hyatt Miami from March 7th to March 10th, 2005. Only 41 iMotes delivered useful contact information for trace file *Exp3*.

Exp2 and Exp3 represent node contact on campus and in conference environments, respectively. In both scenarios, mobile users are likely to work on a common task due to their common social activities (i.e., working in the same lab and attending the same conference). Nearby mobile users can create computing communities in which mobile devices can collaboratively execute shared tasks. Thus, properties of mobile cloudlet extracted from these two trace files can characterize the real mobile cloudlet system.

### B. Cloudlet Size

According to the definition of mobile cloudlet, we analyze the size of $\mathcal{C}_\tau$ by calculating the average number of encountered nodes over time $\tau$ in traces Exp2 and Exp3. As shown in Fig. 4, size of $\mathcal{C}_\tau$ increases as $\tau$ increases. Using curve fitting, we can see that negative exponential distributions approximately fit the data in Fig. 4. In other words, the size of $\mathcal{C}_\tau$ is a negative exponential function of $\tau$ and the number of nodes $n$ in the network (e.g., the maximum cloudlet size is 11 in trace Exp2).

Besides the network size $n$ and time $\tau$, the exponential growth rate is vital for cloudlet size. A large growth rate means that an initiator frequently meets resource-rich devices, thus likely acquires a large pool of potential computing resources. On the contrary, a small growth rate means that an initiator seldom encounters new nodes and can only acquires computing resources from a small portion of nodes in the network, which may lead to poor computing performance. We will




(a) Exp2 (b) Exp3

Fig. 4. Sizes of mobile cloudlet $\mathcal{C}_\tau$ follow negative exponential growth with $\tau$.

further study the increase rate and how it affects the computing performance of mobile cloudlet through mathematical analysis in Sections IV and V.

### C. Lifetime

Based on Definition 2, node $v_j$'s lifetime in a mobile cloudlet for initiator $v_i$ is from $v_i$ and $v_j$'s first contact to their last contact within time $\tau$. Fig. 5 shows the average $LT(\tau)$ in traces Exp2 and Exp3. $LT(\tau)$ increases as $\tau$ increases. Lifetime increases slowly when $80000 < \tau < 132000$ (about 14-hour period) in trace Exp2 and when $90000 < \tau < 129000$ (about 11-hour period) in trace Exp3. This is probability because users have little contact during nights, which is also observed in cloudlet node's reachable time in Fig. 6.

From Fig. 5, it is difficult to determine the lifetime when $\tau$ is small because of the randomness of inter-contact time $T_I$. But, it agrees with our intuition that a node's lifetime increases linearly with rate 1 when $\tau$ is large as shown in Fig. 5. This implies that for delay tolerant application (i.e., large $\tau$), the optimal computing performance of mobile cloudlet— achieved by exploiting the cloudlet nodes' whole lifetime for computing— is hardly influenced by intermittent connections between an initiator and cloudlet nodes.

### D. Reachable Time

The reachable time $RT(\tau)$ of a cloudlet node is its total contact duration with the initiator within time $\tau$. Fig. 6 shows the average reachable time in traces Exp2 and Exp3, which are piecewise linear functions. The piecewise linearity is due to different mobility patterns of users at different times (daytime and night time). For instance, in trace Exp2, students are working in the lab during $50000 < \tau < 80000$ (about 8-hour period), producing long contact time and short inter-contact time. This leads to high growth rate of reachable time, as shown in Fig. 6(a). On the other hand, they have short contacts and long inter-contacts during off time, which lead to small growth rate of reachable time. Examining Fig. 6 more closely, we discover that the increase rate depends on the average contact time and inter-contact time between two nodes over the corresponding period of time. For instance, the slope of segment $\tau \in [44400, 85200]$ in Fig. 6(b) is 0.01882, which is very close to average $T_C/$(average $T_C$ + average $T_I$)=0.01877 during $44400 < \tau < 85200$.

*Remark* 1. In traces Exp2 and Exp3, cloudlet nodes' reachable times increase linearly with $\tau$ and the increase rates are approximately average $T_C/$(average $T_C$ + average $T_I$), which are varying according to users' mobility patterns. The increase rate indicates the connection likelihood between an initiator and a cloudlet node. If an initiator can connect to devices with high likelihood, it could receive omnipresent and reliable mobile cloudlet computing service.

## IV. Mobile Cloudlet Properties Analysis

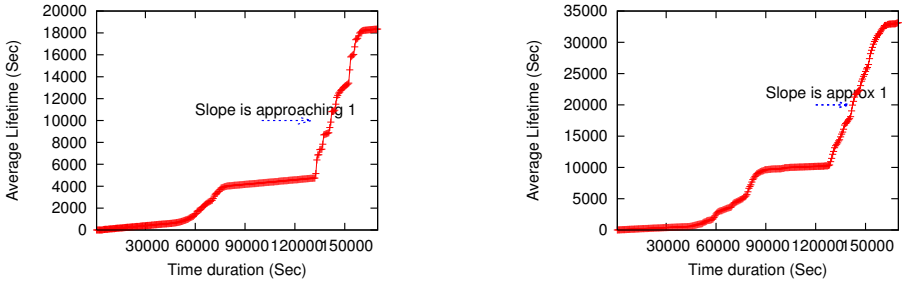
In this section, we mathematically analyze the properties of mobile cloudlet. Our analysis not only confirms our previous observations but also enables us to investigate the computing performance of a mobile cloudlet, such as computing capacity and speed, which determine when a mobile cloudlet is competent for executing mobile applications.

### A. Cloudlet Size Analysis

Denote by $\mathcal{N}_{\mathcal{C}_\tau}(t)$ $(0 \le t \le \tau)$ the number of devices that an initiator encounters within in time $t$. Then, $\mathcal{N}_{\mathcal{C}_\tau}(0) = N^*$, where $N^*$ is the number of the initiator's neighbors at $t = 0$, and cloudlet size $N_{\mathcal{C}_\tau} = \mathcal{N}_{\mathcal{C}_\tau}(\tau)$. We can study $\mathcal{N}_{\mathcal{C}_\tau}(t)$ by superposing multiple 0-1 processes that are sequences of on (contact) and off (inter-contact) times between two nodes.

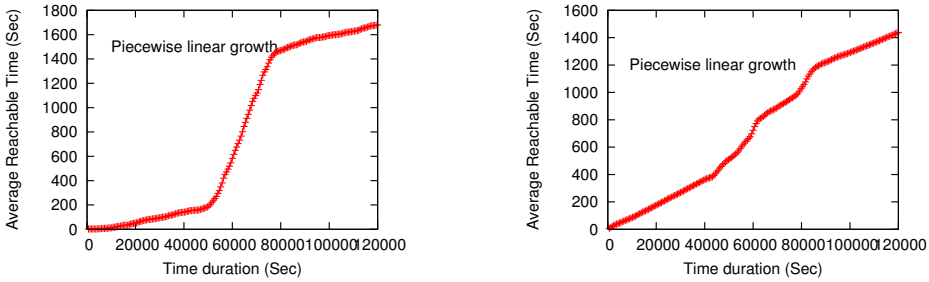
Over a period of time $t$, a cloudlet node and the initiator will be in contact (on) and inter-contact (off) states alternately. If this process begins at the origin of a contact or an inter-contact, it can be modeled as the conventional alternating renewal process. However, in practice, the choice of time origin does not always coincide with the beginning of a contact or an inter-contact. Thus, we adopt the following modified alternating renewal process.

**Definition 3.** Define a stochastic process $\{\chi(t), 0 \le t < \infty\}$ with values in an abstract space $X = A + B$. The process $\{\chi(t)\}$ assumes the states $A$ and $B$ alternately. Denote by $\xi^1, \eta^1, \xi^2, \eta^2, \cdots$ the successive sojourn times spent in states $A$ and $B$, respectively, where $\xi^2, \xi^3, \cdots$ are i.i.d., $\eta^1, \eta^2, \cdots$ are i.i.d., while $\xi^1$ has a different distribution. Define $S_0 = 0$, $S_1 = \xi^1$, and $S_n = \xi^1 + \eta^1 + \xi^2 + \eta^2 + \cdots + \eta^{n-1} + \xi^n$ for $n \ge 2$, the process $N(t) = \max_{n \ge 0}\{n | S_n \le t\}$ is called *modified alternating renewal process*. This process is also called *equilibrium alternating renewal process*, if $\xi^1$ has




(a) Exp2 (b) Exp3

Fig. 5. Average lifetimes of cloudlet nodes increase approximately linearly with $\tau$ when time $\tau$ is large.




(a) Exp2 (b) Exp3

Fig. 6. Average reachable times of cloudlet nodes are piecewise linear functions of time $\tau$ with slope depending on contact and inter-contact time.

the PDF $[1 - F_\xi(x)]/E(\xi)$, where $F_\xi(x)$ and $E(\xi)$ are the cumulative distribution function (CDF) and expectation of $\xi^i$ for all $i > 1$.

Let $A$ and $B$ represent contact and inter-contact states. When two node are in contact at $t = 0$, $\xi^1 = \widetilde{T_C^1}$ (the residual time of $T_C^1$), $\xi^i = T_C^i$ $(\forall i > 1)$, and $\eta^i = T_I^i$ $(\forall i \geq 1)$; when two nodes are *not* in contact at $t = 0$, $\xi^1 = \widetilde{T_I^1}$ (the residual time of $T_I^1$), $\xi^i = T_I^i$ $(\forall i > 1)$, and $\eta^i = T_C^i$ $(\forall i \geq 1)$. It is reasonable to assume that the process of contact and inter-contact between nodes has been running for a long time before it is first observed. Then $\xi^1$ will have the PDF of $[1 - F_\xi(x)]/E(\xi)$ [17]. Therefore, this process can be seen as an equilibrium alternating renewal process.

**Theorem 1.** *The expectation of cloudlet size is*

$$E(N_{\mathcal{C}_\tau}) = (n-1)\left[1 - \left(1 - \frac{\pi r^2}{n/\lambda}\right) e^{-\lambda_I \tau}\right]. \quad (1)$$

*Proof:* Assume at $t = 0$, there are $N^*$ nodes in the initiator's transmission range, $\mathcal{N}_{\mathcal{C}_\tau}(t) - \mathcal{N}_{\mathcal{C}_\tau}(0)$ $(t > 0)$ can be seen as superposition of $n - N^* - 1$ number of 0-1 processes $1_{\{N_i(t)>0\}}$, where $N_i(t)$ has the same distribution as $N(t)$.

$$\mathcal{N}_{\mathcal{C}_\tau}(t) = N^* + \sum_{i=1}^{n-N^*-1} 1_{\{N_i(t)>0\}} = n - 1 - \sum_{i=1}^{n-N^*-1} 1_{\{\xi_i^1 > t\}}.$$

It is worth noting that $N^*$ is a random variable depending on initial node distribution in the network. Rigorously, $P(\mathcal{N}_{\mathcal{C}_\tau}(t) = k) = E(P(\mathcal{N}_{\mathcal{C}_\tau}(t) = k|N^*))$ and $E(\mathcal{N}_{\mathcal{C}_\tau}(t)) = E(E(\mathcal{N}_{\mathcal{C}_\tau}(t)|N^*))$. Therefore, $\mathcal{N}_{\mathcal{C}_\tau}(t)$ is determined by the initial node distribution and the residual inter-contact time between two nodes. In homogeneous network, $N^*$ satisfies $P(N^* = m) = \binom{n-1}{m}\left(\frac{\pi r^2}{n/\lambda}\right)^m \left(1 - \frac{\pi r^2}{n/\lambda}\right)^{n-1-m}$. Then, $\mathcal{N}_{\mathcal{C}_\tau}(t)$ has a binomial distribution with parameters $n-1$ and $\frac{\pi r^2}{n/\lambda} + F_{\widetilde{T_I}}(t)\left(1 - \frac{\pi r^2}{n/\lambda}\right)$, where $F_{\widetilde{T_I}}(t) = P(\widetilde{T_I} \leq t)$. Thus,

$$E(\mathcal{N}_{\mathcal{C}_\tau}(t)) = (n-1)\left(\frac{\pi r^2}{n/\lambda} + \left(1 - \frac{\pi r^2}{n/\lambda}\right) F_{\widetilde{T_I}}(t)\right). \quad (2)$$

In equilibrium alternating renewal process, the density function of $\widetilde{T_I}$ is $\lambda_I[1 - F_{T_I}(x)]$, where $F_{T_I}(x)$ is the CDF of $T_I$ and $\lambda_I^{-1} = \int_0^\infty x * F_{T_I}(dx) = \int_0^\infty (1 - F_{T_I}(x))dx$. When the inter-contact time $T_I$ follows exponential distribution with parameter $\lambda_I$, $\widetilde{T_I}$ is identically distributed with $T_I$. By $E(N_{\mathcal{C}_\tau}) = E(\mathcal{N}_{\mathcal{C}_\tau}(\tau))$, we prove Eq. (1). ■

*Remark* 2. Theorem 1 shows that the expected cloudlet size follows negative exponential growth with $\tau$, which is consistent with our observation in Fig. 4. Moreover, Theorem 1 gives the negative exponential growth rate as $\lambda_I = 1/E(T_I)$. The smaller $E(T_I)$ is, the more nodes an initiator encounters within time $\tau$, and vice versa. This implies that the more frequently nodes meet one another and the larger the network size $n$ is, the more devices are in the mobile cloudlet to provide computing resources and the better computing performance is likely to be achieved, and vice versa.

### B. Lifetime Analysis

We study a cloudlet node's lifetime based on the modified alternating renewal process in Definition 3, and deduce the following theorem.

**Theorem 2.** *The expected lifetime of a cloudlet node is approximately* $\tau - \frac{1}{\lambda_I}(1 - \frac{\pi r^2}{n/\lambda} + \frac{\lambda_C}{\lambda_I+\lambda_C})$ *when* $\tau$ *is large.*

*Proof:* In equilibrium alternating renewal process in Definition 3, $\chi(t) = 1$ when two nodes are in contact at time $t$; $\chi(t) = 0$, otherwise.

(i) When $\chi(0) = \chi(\tau) = 1$, the node's entrance time is 0 and its exit time is $\tau$. Clearly, its lifetime $LT(\tau) = \tau$.

(ii) When $\chi(0) = 0, \chi(\tau) = 1$, the node's entrance time is $\xi^1$ and its exit time is $\tau$. Then, $LT(\tau) = \tau - \xi^1 \cdot 1_{\{\xi^1 < \tau\}}$, where $\xi^1 = \widetilde{T_I}$ is the forward recurrence time of $T_I$.

(iii) When $\chi(0) = 1, \chi(\tau) = 0$, the node's entrance time is 0 and its exit time is $S_{N(\tau)}$. Thus, $LT(\tau) = S_{N(\tau)} = \tau - \xi^{N(\tau)} \cdot 1_{\{\xi^{N(\tau)} < \tau\}}$, where $\xi^{N(\tau)} = \widehat{T_I}$ is the backward recurrence time of $T_I$.

(iv) When $\chi(0) = 0, \chi(\tau) = 0$, if $N(\tau) = 0$, the node's lifetime is 0; if $N(\tau) > 0$, the node's entrance time is $\xi^1$ and its exit time is $S_{N(\tau)} + \eta^{N(\tau)}$. Hence, lifetime equals $[\tau - (\xi^1 + \xi^{N(\tau)+1}) \cdot 1_{\{\xi^1 + \xi^{N(\tau)+1} < t\}})] \cdot 1_{\{\xi^1 < t\}}$, where $\xi^1 = \widetilde{T_I}$ and $\xi^{N(\tau)+1} = \widehat{T_I}$.

Denote $\pi_{ij}(t)$ as the equilibrium probability, given that $\chi(0) = i$ and $\chi(t) = j$ $(i, j = 0, 1)$. Let $p_0$ and $p_1$ denote $P(\chi(0) = 0)$ and $P(\chi(0) = 1)$, respectively. Because $T_I$ and $T_C$ are exponential random variables with parameters $\lambda_I$ and $\lambda_C$, respectively, $\widetilde{T_I}$ and $\widehat{T_I}$ have the same distribution as $T_I$ and $\widetilde{T_I} + \widehat{T_I}$ follows Erlang-2 distribution Erlang(2,$\lambda_I$). Thus,

$$\begin{aligned} E(LT(\tau)) &= \tau^2 \lambda_I e^{-\lambda_I \tau} \pi_{00}(\tau) p_0 \\ &+ \tau[1 + (\pi_{01}(\tau)p_0 + \pi_{10}(\tau)p_1 + \pi_{00}(\tau)p_0)e^{-\lambda_I \tau}] \\ &- \frac{1}{\lambda_I}(1 - e^{-\lambda_I \tau})(\pi_{01}(\tau)p_0 + \pi_{10}(\tau)p_1 + 2\pi_{00}(\tau)p_0), \end{aligned} \tag{3}$$

where $p_1 = \frac{\pi r^2}{n/\lambda}$ and $p_0 = 1 - p_1$. The equilibrium probability $\pi_{ij}(\tau)$ can be derived based on Cox's Renewal Theory (Chapter 7.4) [17]: $\pi_{00}(\tau) = \beta + \gamma e^{-\beta\tau/\lambda_C}$, $\pi_{01}(\tau) = \gamma - \gamma e^{-\beta\tau/\lambda_C}$, $\pi_{10}(\tau) = \beta - \beta e^{-\beta\tau/\lambda_C}$, and $\pi_{11}(\tau) = \gamma + \beta e^{-\beta\tau/\lambda_C}$, where $\beta = \frac{\lambda_C}{\lambda_I+\lambda_C}$ and $\gamma = \frac{\lambda_I}{\lambda_I+\lambda_C}$. When $\tau$ is large, $e^{-\lambda_I \tau}$ and $e^{-(\lambda_I+\lambda_C)\tau}$ approach 0,

$$E(LT(\tau)) \approx \tau - \frac{1}{\lambda_I}(1 - \frac{\pi r^2}{n/\lambda} + \frac{\lambda_C}{\lambda_I + \lambda_C}), \tag{4}$$

i.e., the expected lifetime grows linearly with time $\tau$. ■

To better understand $E(LT(\tau))$, we have numerical analysis of $E(LT(\tau))$ to show how $E(LT(\tau))$ changes with $\tau$ in Fig. 7. We set $p_0 = 0.9$, $p_1 = 0.1$, $\lambda_I = 0.0001$, and $\lambda_C = 0.01$. Parameters $\lambda_I$ and $\lambda_C$ are set to be approximately equal to $1/E(T_I)$ and $1/E(T_C)$ in trace Exp3. Fig. 7 shows that when $\tau > 4 \times 10^4$, $E(LT(\tau))$ grows linearly with slope 1, which is consistent with Eq. (4) and Fig. 5 (b). When $\tau$ is small $(0 < \tau < 1000)$, the close-up figure shows that $E(LT(\tau))$ is mainly influenced by $\tau^2$.

*Remark* 3. When $\tau$ is small, $E(LT(\tau))$ exhibits *quadratic growth*. As $\tau$ increases, $E(LT(\tau))$ tends to grow *linearly* with slope 1 and the gap between $\tau$ and $E(LT(\tau))$ is a constant $\frac{1}{\lambda_I}(1 - \frac{\pi r^2}{n/\lambda} + \frac{\lambda_C}{\lambda_I+\lambda_C})$. This indicates that for application with long delay tolerance $\tau$, intermittent connectivity between an initiator and cloudlet nodes has a small constant negative effect on the optimal computing performance achieved when cloudlet nodes compute task for the initiator throughout their lifetimes.

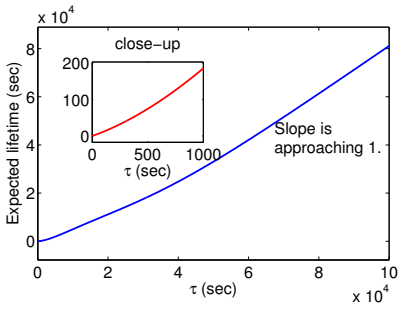



Fig. 7. Expected lifetime of a cloudlet node grows linearly with slope 1 when $\tau$ is large.

### C. Reachable Time Analysis

In order to study the reachable time $RT(\tau)$, i.e., total contact duration between two nodes within time $\tau$, we rewrite Definition 3 as follows.

**Definition 4.** Let $\{(I_n, Z_n\}_{n=0}^{\infty}$ be a bivariate stochastic process on a probability space $(\Omega, \chi, P)$ such that $Z_0 = 0$ and $I_n = 1 - I_{n-1}$ for all $n \geq 1$, where $p_0 = P(I_0 = 0)$ and $p_1 = P(I_0 = 1)$ satisfying $p_0 + p_1 = 1$. Assume that, conditional on $I_{n-1}$, the $Z_n - Z_{n-1}$ for all $n \geq 1$ are mutually independent. Let the conditional distribution of $Z_1$ be $F^1_{\widetilde{T_C}}$ $(F^1_{\widetilde{T_I}})$ if $I_0 = 1$ $(0)$, and the distributions of $Z_n - Z_{n-1}$ conditioned on $I_{n-1}$ be $F_{T_C}$ $(F_{T_I})$ if $I_{n-1} = 1$ $(0)$, for all $n \geq 2$. Distributions of $F^1_{\widetilde{T_C}}$ and $F^1_{\widetilde{T_I}}$ have density $\lambda_C[1 - F_{T_C}]$ and $\lambda_I[1 - F_{T_I}]$, respectively, where $\lambda_C^{-1} = \int_0^{\infty} F_{T_C}(dx)$ and $\lambda_I^{-1} = \int_0^{\infty} F_{T_I}(dx)$. Then, the point process characterized by $\{(I_n, Z_n)\}_{n=0}^{\infty}$, in which $I_n$ is the point type and $Z_n$ is the waiting time until the $n$th event, is an *equilibrium alternating renewal process*.

The total contact duration between two nodes within time $\tau$ depends on not only distribution of contact time but also the number of contacts, which can be represented as follows.

**Definition 5.** Let $\{(I_n, Z_n\}_{n=0}^{\infty}$ be an equilibrium renewal alternating process. For all $\tau > 0$, let $K(\tau) = \sum_{n=1}^{\infty} 1_{Z_{2n-1} \leq \tau}$, and $L(\tau) = \sum_{n=1}^{\infty} 1_{Z_{2n} \leq \tau}$. Then, $K(\tau)$ $(L(\tau))$ is the odd (even) counting random variable of the process at time $\tau$.

When $I_0 = 1$, the number of contacts is $K(\tau)$ (or $K(\tau)+1$) if $K(\tau) = L(\tau)$ (if $K(\tau) \neq L(\tau)$). When $I_0 = 0$, the number of contacts is $L(\tau)$ (or $L(\tau) + 1$) if $K(\tau) = L(\tau)$ (if $K(\tau) \neq L(\tau)$). To find out $RT(\tau)$, we define the sojourn time of the equilibrium renewal alternating process as follows.

**Definition 6.** Let $\{(I_n, Z_n\}_{n=0}^{\infty}$ be an equilibrium renewal alternating process with odd (even) counting random variable $K(\tau)$ $(L(\tau))$. Let $S_0 = T_0 = 0$, $S_n = \sum_{i=1}^{n}(Z_{2i-1} - Z_{2i-2})$, and $T_n = \sum_{i=1}^{n}(Z_{2i} - Z_{2i-1})$, for all $n \geq 1$. Then $S_n$ $(T_n)$ is called the $n$th sum of the odd (even) states of the process,

$n \geq 0$. For all $\tau > 0$, let

$$\alpha_0(\tau) = T_{L(\tau)} \cdot 1_{\{K(\tau)=L(\tau)\}} + [\tau - S_{K(\tau)}] \cdot 1_{\{K(\tau)\neq L(\tau)\}},$$
$$\alpha_1(\tau) = S_{K(\tau)} \cdot 1_{\{K(\tau)\neq L(\tau)\}} + [\tau - T_{L(\tau)}] \cdot 1_{\{K(\tau)=L(\tau)\}}.$$

Then $\alpha_0(\tau)$ ($\alpha_1(\tau)$ ) is called the sojourn time of the even (odd) states of the process. Clearly, $\alpha_0(\tau) + \alpha_1(\tau) = \tau$. The sojourn time in the on state of the process (i.e., total contact duration) during $(0, \tau]$ is given by

$$RT(\tau) = \alpha_0(\tau) \cdot 1_{\{I_0=0\}} + \alpha_1(\tau) \cdot 1_{\{I_0=1\}}. \quad (5)$$

**Lemma 1.** *For an equilibrium alternating renewal process, the total contact duration (i.e., $RT(\tau)$) satisfies*

$$E(RT(\tau)) = E(\alpha_0(\tau))p_0 + E(\alpha_1(\tau))p_1, \quad (6)$$

*where $E(\alpha_0(\tau))$ and $E(\alpha_1(\tau))$ are inverse Laplace transform of Eqs. (7) and (8), respectively, $p_1$ and $p_0$ represent the probabilities that two nodes are originally in contact and out of contact, respectively.*

*Proof:* In [18], M.H. Rossiter derived the sojourn time distribution in on state for a two-state system by applying Laplace transform and double Laplace transform. For an equilibrium renewal alternating process, the Laplace transform of the expected sojourn time in on state conditioning on $I_0 = 0$

$$L(E[\alpha_0(\tau)]; s) = \frac{[1 - F_{T_C}(s)]F_{\widetilde{T_I}}(s)}{s^2[1 - F_{T_C}(s)F_{T_I}(s)]}, \quad (7)$$

while conditioning on $I_0 = 1$

$$L(E[\alpha_1(\tau)]; s) = \frac{1 - F_{\widetilde{T_C}}(s) - (F_{T_C} - F_{\widetilde{T_C}})(s)F_{T_I}(s)}{s^2[1 - F_{T_C}(s)F_{T_I}(s)]}, \quad (8)$$

where $L\{\ \cdot\ ; s\}$ represents the Laplace transform, and $F_X(s)$ is the Laplace transform of $X$, i.e., $F_X(s) = \int_0^\infty e^{-sx}F(dx)$.

We have $E(\alpha_0(\tau))$ and $E(\alpha_1(\tau))$ by taking inverse Laplace transform of Eqs. (7) and (8). Substituting $E(\alpha_0(\tau))$ and $E(\alpha_1(\tau))$ in the expectation of Eq. (5) completes our proof. ■

**Theorem 3.** *In homogeneous network with uniform node distribution, if $T_C$ and $T_I$ follow exponential distributions with parameters $\lambda_C$ and $\lambda_I$, respectively, the expected reachable time of a cloudlet node*

$$E(RT(\tau)) = \frac{\lambda_I \tau}{\lambda_I + \lambda_C} + \frac{\lambda_C p_1 - \lambda_I p_0}{(\lambda_I + \lambda_C)^2}(1 - e^{-(\lambda_I+\lambda_C)\tau}), \quad (9)$$

*where $p_1 = \frac{\pi r^2}{n/\lambda}$ and $p_0 = 1 - p_1$.*

*Proof:* For exponential random variable $T_C$ ($T_I$), its forward recurrence time $F_{\widetilde{T_C}}$ ($F_{\widetilde{T_I}}$) also has exponential distribution with parameter $\lambda_C$ ($\lambda_I$) because of the memoryless property of exponential random variable. Then, $F_{T_C}(s) = F_{\widetilde{T_C}}(s) = \frac{\lambda_C}{s+\lambda_C}$ and $F_{T_I}(s) = F_{\widetilde{T_I}}(s) = \frac{\lambda_I}{s+\lambda_I}$.

Based on results in Eqs. (7) and (8) from [18],

$$L(E[\alpha_1(\tau)]; s) = \frac{1 - \frac{\lambda_C}{s+\lambda_C}}{s^2[1 - \frac{\lambda_C}{s+\lambda_C}\frac{\lambda_I}{s+\lambda_I}]},$$

and $L(E[\alpha_0(\tau)]; s) = L(E[\alpha_1(\tau)]; s)\frac{\lambda_I}{s+\lambda_I}$. Performing inverse Laplace transform, we have

$$E[\alpha_0(\tau)] = \frac{\lambda_I \tau}{\lambda_I + \lambda_C} + \frac{\lambda_I}{(\lambda_I + \lambda_C)^2}(e^{-(\lambda_I+\lambda_C)\tau} - 1), \quad (10)$$

$$E[\alpha_1(\tau)] = \frac{\lambda_I \tau}{\lambda_I + \lambda_C} + \frac{\lambda_C}{(\lambda_I + \lambda_C)^2}(1 - e^{-(\lambda_I+\lambda_C)\tau}). \quad (11)$$

In our homogeneous network model, $p_0 = \frac{\pi r^2}{n/\lambda}$ and $p_1 = 1 - p_0$. Substituting them into Eq. (6) completes our proof. ■

Numerical results of Theorem 3 are shown in Fig. 8. The parameter settings are the same as those in Fig. 7. Theorem 3 and Fig. 8 show that when $T_C$ and $T_I$ are exponentially distributed, the reachable time of a cloudlet node grows linearly with slope $\frac{\lambda_I}{\lambda_I+\lambda_C}$, i.e., $\frac{E(T_C)}{E(T_I)+E(T_C)}$ as $\lambda_I = \frac{1}{E(T_I)}$ and $\lambda_C = \frac{1}{E(T_C)}$. The linear growth in Fig. 8 is different from the piecewise linear growth in Fig. 6 because depending on people's schedules, $T_I$ and $T_C$ follow different distributions during different time in traces Exp2 and Exp3.

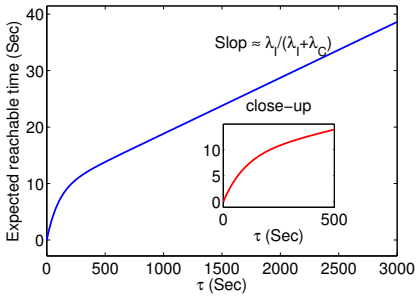



Fig. 8. Expected reachable time of a cloudlet node grows linearly with slope approximately $\frac{\lambda_I}{\lambda_I+\lambda_C}$ when $\tau$ is large.

*Remark* 4. The increase rate of $E(RT(\tau))$ shows that the mean reachable time within time $\tau$ is mainly determined by the ratio $\frac{E(T_C)}{E(T_I)+E(T_C)}$ when $\tau$ is large. Nodes that meet an initiator frequently and have long contact time have high connection likelihood and can provide reliable computing services while still support mobility of the initiator. The initiator can estimate the ratio $\frac{E(T_C)}{E(T_I)+E(T_C)}$ based on its contact histories and use it as an indicator for whether an encountered node is suitable for providing mobile application services.

## V. Computing Performance of Mobile Cloudlet

The amount of computation that the cloudlet nodes can provide for the initiator not only depends on the computing capabilities of cloudlet nodes and how the task is partitioned for parallel processing but also depends on the cloudlet node's lifetime and reachable time. Evaluating the computing capability of mobile processor and designing application partition schemes [19] are beyond the scope of this paper. In this paper, we simply assume that computing speed of each device is a constant $V$, and the initiator can partition the task into $M$ subtasks that can be computed on cloudlet nodes in parallel. The data sizes of each subtask before and after computing are $D_d^i$ and $D_r^i$ ($1 \leq i \leq M$), respectively.

Assume a total of $B$ Hz spectrum is shared by all nodes and each node has a fixed transmission power $P$. The noise $N$—including ambient and interference noise—is constant everywhere in the network. We characterize the wireless link using a pass loss model with attenuation exponent $\alpha \geq 2$. The capacity of a wireless link is $B\log_2(1+\frac{P}{N}d^{-\alpha})$, where $d$ is the Euclidean distance between the sender and the receiver. Assume that advanced error control coding is used such that the available link bandwidth is equal to its capacity. For a task contains $D_d^i$ $(1 \leq i \leq M)$ bits, the transmission time of dispatching a task is

$$0 < t_d = \frac{D_d^i}{B\log_2(1+\frac{P}{N}d^{-\alpha})} \leq \frac{D_d}{B\log_2(1+\frac{P}{N}r^{-\alpha})}, \quad (12)$$

where $D_d \triangleq \sum_{1\leq i\leq M} D_d^i$. The transmission time of retrieving a task is

$$0 < t_r = \frac{D_r^i}{B\log_2(1+\frac{P}{N}d^{-\alpha})} \leq \frac{D_r}{B\log_2(1+\frac{P}{N}r^{-\alpha})}, \quad (13)$$

where $D_r \triangleq \sum_{1\leq i\leq M} D_r^i$. Then the total transmission time is

$$0 < t_d + t_r \leq \frac{D_d + D_r}{B\log_2(1+\frac{P}{N}r^{-\alpha})} \triangleq \rho. \quad (14)$$

If tasks are computed on cloudlet nodes during their whole lifetime, the optimal mobile cloudlet capacity is achieved. If an initiator only employs cloudlet nodes when they are in contact, i.e., the computing times equal to reachable times minus transmission times, we have a lower bound on the computing capacity of a mobile cloudlet. Based on this methodology and results in Section IV, we have the following theorem on the computing capacity of a mobile cloudlet.

**Theorem 4.** *A mobile cloudlet $\mathcal{C}_\tau$'s expected computing capacity is upper bounded by $C^u_{\mathcal{C}_\tau}$ and lower bounded by $C^l_{\mathcal{C}_\tau}$ in Eqs. (15) and (18), respectively.*

*Proof:* During a cloudlet node's lifetime, the maximum time a cloudlet node can use for computing is its lifetime minus the transmission time of the task. Hence, the computing capacity of a mobile cloudlet $\mathcal{C}_\tau$ satisfies

$$C_{\mathcal{C}_\tau} \leq \sum_{i=1}^{N_{\mathcal{C}_\tau}} (LT^i(\tau) - t_r - t_d)V.$$

$$E(C_{\mathcal{C}_\tau}) < C^u_{\mathcal{C}_\tau} \triangleq E(N_{\mathcal{C}_\tau})E(LT(\tau))V, \quad (15)$$

where $E(N_{\mathcal{C}_\tau})$ and $E(LT(\tau))$ can be found in Eqs. (1) and (3), respectively.

To employ cloudlet nodes in contact, a cloudlet node's total computing time over time $\tau$ is

$$CT(\tau) = RT(\tau) - (p_0 + p_1\pi_{10}) \sum_{i=1}^{N(\tau)} \overline{CT^i} - p_1\pi_{11}P(T_C < \tau) \sum_{i=1}^{N(\tau)+1} \overline{CT^i}, \quad (16)$$

where $\overline{CT^i} \triangleq T_C^i \cdot 1_{T_C^i<\rho} + \rho \cdot 1_{T_C^i\geq\rho}$, $\pi_{i,j}$ $(i,j=0,1)$ are the equilibrium probabilities and $N(\tau)$ is number of renewals within time $\tau$ in the equilibrium alternating renewal process.

Denote $N^0(\tau) = N(\tau)|\{I_0 = 0\}$ and $N^1(\tau) = N(\tau)|\{I_0 = 1\}$. According to the renewal equation for modified renewal process,

$$E(N^0(\tau)) = F_{\widetilde{T_I}}(\tau) + \int_0^\tau E(N^0(\tau))dF_{T_I+T_C}(s),$$

$$E(N^1(\tau)) = F_{\widetilde{T_C}}(\tau) + \int_0^\tau E(N^1(\tau))dF_{T_I+T_C}(s).$$

Taking the Laplace transform on both sides,

$$L_{E(N^0(\tau))}(s) = L_{\widetilde{T_I}}(s)/s + L_{N^0(\tau)}(s)L_{T_I+T_C}(s),$$

$$L_{E(N^1(\tau))}(s) = L_{\widetilde{T_C}}(s)/s + L_{N^1(\tau)}(s)L_{T_I+T_C}(s),$$

where $L_{\widetilde{T_I}}(s) = \lambda_I/(s+\lambda_I)$ and $L_{\widetilde{T_C}}(s) = \lambda_C/(s+\lambda_C)$. $T_I + T_C$ has density function $\frac{\lambda_I\lambda_C}{\lambda_C-\lambda_I}(e^{-\lambda_I t} - e^{-\lambda_C t})$, which gives the Laplace transform $L_{T_I+T_C}(s) = \lambda_I\lambda_C/(s+\lambda_I)(s+\lambda_C)$. Thus, $L_{E(N^0(\tau))}(s) = \lambda_I(s+\lambda_C)/[s^2(s+\lambda_I+\lambda_C)]$ and $L_{E(N^1(\tau))}(s) = \lambda_C(s+\lambda_I)/[s^2(s+\lambda_I+\lambda_C)]$. Taking the inverse Laplace transform, we then have

$$E(N^0(\tau)) = \gamma\lambda_C\tau + \gamma^2(1 - e^{-(\lambda_I+\lambda_C)\tau}),$$

$$E(N^1(\tau)) = \beta\lambda_I\tau + \beta^2(1 - e^{-(\lambda_I+\lambda_C)\tau}),$$

where $\beta = \frac{\lambda_C}{\lambda_I+\lambda_C}$ and $\gamma = \frac{\lambda_I}{\lambda_I+\lambda_C}$. Subsequently,

$$E(N(\tau)) = E(N(\tau)|I_0=0)p_0 + E(N(\tau)|I_0=1)p_1, \quad (17)$$

$$= \frac{\lambda_I\lambda_C\tau}{\lambda_I+\lambda_C} + \frac{p_0\lambda_I^2 + p_1\lambda_C^2}{(\lambda_I+\lambda_C)^2}(1 - e^{-(\lambda_I+\lambda_C)\tau}),$$

where $p_1 = \frac{\pi r^2}{n/\lambda}$ and $p_0 = 1 - p_1$.

The computing capacity of a mobile cloudlet $\mathcal{C}_\tau$ satisfies

$$C_{\mathcal{C}_\tau} \geq \sum_{j=1}^{N_{\mathcal{C}_\tau}} CT^j(\tau)V.$$

Therefore, $E(C_{\mathcal{C}_\tau})$ is lower bounded by

$$C^l_{\mathcal{C}_\tau} \triangleq E(N_{\mathcal{C}_\tau})V\left\{E(RT(\tau)) - \left[(1 - p_1\pi_{11}e^{-\lambda_C\tau})E(N(\tau)) + p_1\pi_{11}(1 - e^{-\lambda_C\tau})\right]E(\overline{CT})\right\}, \quad (18)$$

where $E(\overline{CT}) = E(T_C \cdot 1_{T_C<\rho}) + \rho P(T_C \geq \rho) = \frac{1-e^{-\lambda_C\rho}}{\lambda_C}$. ■

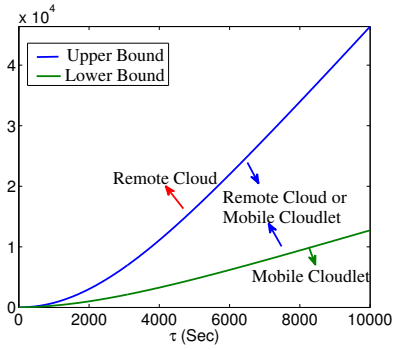



Fig. 9. Bounds on computing capacity of mobile cloudlet where $\lambda_I = 0.0002$, $\lambda_C = 0.001$, $p_0 = 0.9$, $p_1 = 0.1$, $n = 10$, $V = 1$, $\rho = 0.1$.

Fig. 9 shows the numerical results of this theorem by setting $\rho = 0.1$ second (typical packet transmission time in mobile wireless networks), $n = 10$, $E(T_I) \approx 83$ minutes, and $E(T_C) \approx 17$ minutes. This parameter setting reflects

a scenario with 10 students in a team or 10 colleagues at a conference working on the same project. They meet for about 17 minutes between classes or conference sessions to compute a common task by sharing computing resources on their mobile devices. Two curves in Fig. 9 divide MCC into three categories: MCC relying on i) remote cloud, ii) remote cloud or mobile cloudlet, iii) mobile cloudlet. When this group of users need to execute a task with computational demand $C$ and delay requirement $\tau$, if $C \geq C^u_{\mathcal{C}_\tau}$, they should offload the task to remote cloud; if $C^l_{\mathcal{C}_\tau} < C < C^u_{\mathcal{C}_\tau}$, they can use remote cloud or mobile cloudlet; if $C \leq C^l_{\mathcal{C}_\tau}$, they can simply share the task computation during their contacts with each other for mobile cloudlet computing.

The average computing speed of a mobile cloudlet is $E(C_{C_\tau})/\tau$. When $\tau \to \infty$, we have the long-term computing speed $CS = \lim_{\tau\to\infty} \frac{E(C_{C_\tau})}{\tau}$.

**Theorem 5.** *The long-term computing speed of a mobile cloudlet is upper bounded by* $CS^u = (n-1)V$ *and lower bounded by* $CS^l = \frac{(n-1)V\lambda_I e^{-\lambda_C \rho}}{\lambda_C+\lambda_I}$.

*Proof:* Based on Theorem 4, we have the upper bound,

$$CS \leq \lim_{\tau\to\infty} \frac{C^u_{\mathcal{C}_\tau}}{\tau} = (n-1)V \triangleq CS^u. \tag{19}$$

Similarly, the lower bound of $CS$ is $\lim_{\tau\to\infty} \frac{C^l_{\mathcal{C}_\tau}}{\tau}$. As in modified renewal process,

$$\lim_{\tau\to\infty} \frac{E(N(\tau))}{\tau} = \frac{1}{E(T_C+T_I)} = \frac{\lambda_C\lambda_I}{\lambda_C+\lambda_I}.$$

Accordingly, we have

$$CS \geq \lim_{\tau\to\infty} \frac{C^l_{\mathcal{C}_\tau}}{\tau} = \frac{\lambda_I e^{-\lambda_C\rho}}{\lambda_C+\lambda_I}(n-1)V \triangleq CS^l. \tag{20}$$

■

*Remark* 5. The bounds on long-term computing speed can also be used by an initiator mobile device to decide where to offload its task for computing service. Suppose the initiator has a task with computational demand $C$ and delay requirement $\tau$, if $C/\tau \geq (n-1)V$, the initiator needs to offload its task to a remote cloud; if $C/\tau \leq \frac{\lambda_I e^{-\lambda_C\rho}}{\lambda_C+\lambda_I}(n-1)V$, the initiator can distribute its task to nearby devices in a mobile cloudlet; otherwise, the initiator can choose either remote cloud or mobile cloudlet based on other constraints, such as battery life and quality of wireless communication.

## VI. Conclusions

In this paper, we have explored the domain of mobile cloudlet in MCC through studying cloudlet properties and computing performance. The negative exponential growth of cloudlet size shows that the number of resource-rich devices an initiator can connect to for computing service is determined by the number of nodes in the network and how frequently they meet. When the task is delay-tolerant, the intermittent connection has little negative effect on optimal performance of a mobile cloudlet. Furthermore, $E(T_C)/(E(T_C)+E(T_I))$ implies the connection likelihood of a cloudlet node to an initiator, thus can be used by the initiator to choose reliable cloudlet node. Based on cloudlet properties, we have also derived upper and lower bounds on the computing capacity and long-term computing speed of a mobile cloudlet. An initiator can use these bounds to decide whether to upload a task to remote clouds or utilize nearby mobile cloudlet. In our future work, we would like to design and implement mobile applications on a mobile cloudlet system to investigate the feasibility and performance of mobile cloudlet computing.

## References

[1] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *Journal of Internet Services and Applications*, vol. 1, pp. 7–18, May 2010.

[2] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wireless Communications and Mobile Computing*, 2011.

[3] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84 – 106, 2013.

[4] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, no. 4, pp. 14 –23, Oct.-Dec. 2009.

[5] D. Fesehaye, Y. Gao, K. Nahrstedt, and G. Wang, "Impact of cloudlets on interactive mobile cloud applications," in *The Proc. of IEEE EDOC*, 2012.

[6] E. E. Marinelli, "Hyrax: Cloud computing on mobile devices using mapreduce," Master's thesis, Carnegie Mellon University, 2009.

[7] G. Huerta-Canepa and D. Lee, "A virtual cloud computing provider for mobile devices," in *ACM Workshop on Mobile Cloud Computing & Services (MCS'10)*, 2010.

[8] N. Fernando, S. Loke, and W. Rahayu, "Dynamic mobile cloud computing: Ad hoc and opportunistic job sharing," in *Proc. of IEEE UCC*, 2011.

[9] C. Shi, V. Lakafosis, M. H. Ammar, and E. W. Zegura, "Serendipity: enabling remote computing among intermittently connected mobile devices," in *Proc. of ACM MobiHoc*, 2012.

[10] E. Koukoumidis, D. Lymberopoulos, K. Strauss, J. Liu, and D. Burger, "Pocket cloudlets," in *The Sixteenth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, Mar. 2011.

[11] L. Sun and W. Wang, "On latency distribution and scaling: from finite to large cognitive radio networks under general mobility," in *Proc. of IEEE INFOCOM*, 2012.

[12] ——, "On the connectivity of large multi-channel cognitive radio networks," in *Proc. of IEEE ICC*, 2012.

[13] M. Zhao, Y. Li, and W. Wang, "Modeling and analytical study of link properties in multihop wireless networks," *IEEE Trans. on Communications*, vol. 60, no. 2, pp. 445–455, 2012.

[14] H. Cai and D. Y. Eun, "Toward stochastic anatomy of inter-meeting time distribution under general mobility models," in *Proc. of the ACM MobiHoc*, 2008.

[15] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic, "Power law and exponential decay of inter contact times between mobile devices," in *Proc. of the ACM MobiCom*, Sept. 2007.

[16] A. Chaintreau, P. Hui, J. Scott, and et al., "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. on Mobile Computing*, vol. 6, no. 6, pp. 606–620, Jun. 2007.

[17] D. Cox, *Renewal Theory*. Methuen & Co, 1962.

[18] M. Rossiter, "The sojourn time distribution of an alternating renewal process," *Australian Journal of Statistics*, vol. 31, no. 1, pp. 143–152, 1989.

[19] E. Cuervo, A. Balasubramanian, D.-k. Cho, and et al., "Maui: making smartphones last longer with code offload," in *Proc. of MobiSys*, 2010.