# The Unheralded Power of Cloudlet Computing in the Vicinity of Mobile Devices

Yujin Li and Wenye Wang
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC, USA
Email: {yli27, wwang}@ncsu.edu

***Abstract*—With the popularity of smartphones and explosion of mobile applications, mobile devices are becoming the prevalent computing platform for convenient communication and rich entertainment. Because mobile devices still have limited processor power, computing-intensive applications need to be offloaded to either *remote clouds* or *nearby cloudlets* for processing. But, remote cloud computing is hindered by the long latency and expensive roaming charges of cellular radio access. Therefore, *cloudlet computing* becomes appealing to provide instant and low-cost service through resource-rich devices (e.g., desktops) in the vicinity of mobile devices. It is evident that cloudlet computing is challenged by the intermittent connection between cloudlets and mobile devices due to user mobility. The question is *how to evaluate the impact of user mobility on cloudlet computing performance*. In this paper, we examine the cloudlet access probability, task success rate, and task execution speed to measure the impact of mobility. We discover that the cloudlet access probability is $\mu_{T_C}/(\mu_{T_I}+\mu_{T_C})$ determined by mean connection time $\mu_{T_C}$ and mean inter-connection time $\mu_{T_I}$ between the mobile device and the cloudlet. Furthermore, we find that the task success rate and execution speed depend on not only task computation demand and cloudlet computing speed but also cloudlet access probability. Our findings reveal that the ratio $\mu_{T_C}/(\mu_{T_I}+\mu_{T_C})$ quantifies the impact of node mobility on both cloudlet access probability and cloudlet computing performance.**

## I. Introduction

Mobile devices (such as smartphones and tablets) are becoming an inseparable part of our lives for convenient communication and entertainment. The number of smartphones in use worldwide reached 1.038 billion units during the third quarter of 2012, and smartphone users are expected to be over 2 billion by 2015 [1]. With the popularity of mobile devices, there is also an explosion of mobile applications, such as terrestrial navigation, email and web browsing, and mobile games, indicating that mobile devices are quickly becoming the dominant computing platform.

Although mobile devices are rapidly gaining more computing power and memory resources (up to 64 GB flash memory on smartphones currently) [2], they are still limited by their available processor power, memory size and battery life [3]. There is still a large gap in computing speed between mobile devices and desktops. Thus, computing-intensive mobile applications, such as video decoding, speech recognition, and augmented reality, need to be offloaded to the cloud for processing, namely *mobile cloud computing*.

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

In mobile cloud computing, the cloud is formed by data centers and servers, which are distributed in different geographic locations. Mobile devices can offload their computing-intensive applications to the cloud in order to save energy and improve application performance. The mobile devices access the cloud service through *wireless networks* (mainly cellular and WLAN networks). But, remote cloud access through wireless communication could be unavailable or costly [2, 4].

Specifically, cellular connection is known to suffer from very *long latency*, which may make mobile application offloading expensive. For instance, a search query on a high-end smartphone with a 3G connection can take 3 to 10 seconds depending on location, device and operator used. When 3G radio is not connected or only EDGE connectivity is available, this delay can be doubled or even tripled. Although the throughput of radio links on mobile devices and channel bandwidth will continue to increase over time, the long latency of remote cloud access through cellular radio is unlikely to improve dramatically, because the radio link needs 1.5 to 2 seconds to wake up from standby mode. Moreover, using 3G/4G connection can quickly *drain battery*, which is the bottleneck of mobile device development. In addition, network provider's (e.g., AT&T and Verizon) service is very *expensive*. As an example, AT&T charges a mobile user $20 for 300MB data plan (or $30 for 3GB data plan) per month and $20 per 300MB (or $10 per 1GB) overage fee [5].

In order to overcome the limitations of accessing remote clouds through cellular radio, paper [4] proposes to use *cloudlet*, which offloads mobile applications to resource-rich devices in the vicinity of mobile devices. Mobile applications can be seamlessly off-loaded from mobile devices onto nearby cloudlets using virtual machine technology [6]. Fig. 1 illustrates an example of using cloudlets. Suppose mobile user Bob needs to do mobile commerce (e.g., mobile transactions and payments, mobile ticketing) using his smartphone. In order to avoid data overage charge and preserve battery on his phone, Bob offloads the task to a nearby cloudlet that includes resource-rich devices, such as desktops, laptops, even tablets and high-end smartphones. As Bob moves around, he exploits different cloudlets during different periods of time.

The benefits of utilizing cloudlet are the speed of service accessibility, the support of mobility, the enhanced application performance, the elongated battery life, and the reduced roaming data cost [7]. First, as computations and information reside

978-1-4799-1353-4/13/$31.00 ©2013 IEEE

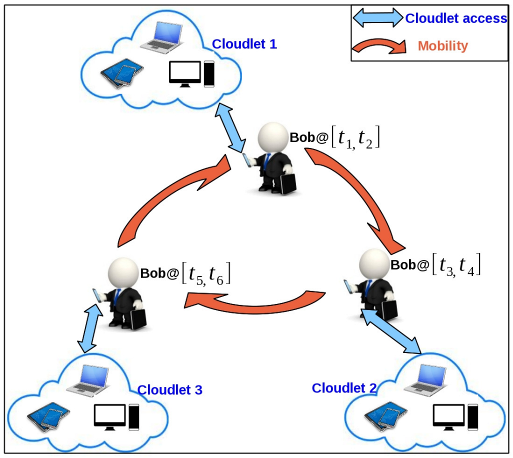


Fig. 1. Mobile cloud computing through cloudlets in the vicinity of a mobile device: Bob uses cloudlet 1 during $[t_1, t_2]$, cloudlet 2 during $[t_3, t_4]$, and cloudlet 3 during $[t_5, t_6]$ to execute mobile applications on his phone.

on nearby devices, users can get direct access instantly through interactions with cloudlets, eliminating the communication latency introduced by the cellular network. Second, cloudlet still preserves the conventional offloading benefit as that in cloud computing, which allows applications that can not be processed on a single mobile device to be executed in the cloudlet. Application performance can be improved if the execution sequence of an application can be reordered to increase the level of parallel processing, thus the mobile user's experience can be significantly enhanced. Third, compared with a complete local execution, offloading the task (e.g. construct augmented reality tourist guide) to a group of nearby devices saves mobile device's energy. Finally, offloading to nearby devices saves money, because expensive data charging in roaming situation is avoided.

The main challenge of utilizing cloudlet resides in the intermittent connectivity between mobile devices and cloudlets due to the mobility of mobile devices and even devices in cloudlets. Mobile devices are held by human that move to different locations for different social activities. Sometimes, a mobile user may move to an area that has no cloudlet in the transmission range of his/her mobile device, while at other times, a mobile user may be at a place with cloudlets in vicinity and can connect to the cloudlets for mobile application computing. For example, in Fig. 1, Bob can't connect to any cloudlet when he is moving from the location of Cloudlet 1 to the location of Cloudlet 2. Only when Bob's smartphone is within the transmission range of Cloudlet 1 during $[t_1, t_2]$ or that of Cloudlet 2 during $[t_3, t_4]$, Bob can access the computing resources on devices in Cloudlets 1 or 2, respectively. Such intermittent connection to cloudlet greatly affects the performance of cloudlet computing for mobile applications [8].

Therefore, a fundamental question for cloudlet computing is: *what is the impact of node mobility on cloudlet computing?* In this paper, we investigate the basic scenario where a mobile device needs to offload computational tasks to cloudlets. Because the mobile device has its intrinsic mobility pattern, the connection to cloudlets is intermittent and limited in duration. We examine *the cloudlet access probability* that a mobile device can connect to at least one cloudlet, *the probability of a successful task execution*, *the number of executed tasks* and *the average task execution speed* over time $t$. Studying them will not only reveal the power of cloudlet computing but also identify how mobility affects cloudlet computing's performance.

As the intermittent cloudlet connectivity is due to node mobility, we model the connection and inter-connection between a mobile device and a cloudlet as an *alternating renewal process*. Based on renewal theory, we find that a cloudlet's access probability equals to $\mu_{T_C}/(\mu_{T_C} + \mu_{T_I})$, where $\mu_{T_C}$ is the expectation of connection time $T_C$ and $\mu_{T_I}$ is the expectation of inter-connection time $T_I$ between the mobile device and the cloudlet. Furthermore, we show that the probability of successfully executing a task by a cloudlet is determined by cloudlet access probability, probability distribution of connection time, task computation demand, cloudlet computing speed, and channel bandwidth. Finally, we prove that the number of executed tasks over time $t$ and the average task execution speed are both determined by cloudlet access probability and task completion time (i.e., the sum of task transmission and computing time). In summary, node mobility affects not only cloudlet access probability but also cloudlet computing performance and its impact can be represented by $\mu_{T_C}/(\mu_{T_C} + \mu_{T_I})$.

The remainder of this paper is organized as follows: we define the network and cloudlet connection models in Section II; using renewal theory, we derive cloudlet access probability and computing performance in Section III; we conclude this paper with insights from our analysis as well as future work in Section IV.

## II. Models and Definitions

In reality, mobile users visit many community locations to perform social activities, such as working in office, shopping in mall, and staying at home. At these community sites, likely there are resource-rich devices, such as desktops and laptops, that can provide cloudlet computing for a mobile device. Therefore, we focus on a heterogeneous network environment that is composed of a set of cloudlets with different computing capabilities at different locations.

**Network Model:** Assume that a mobile device is moving in a network $\Omega_m$ with $m$ cloudlets. The locations of cloudlets can be community locations extracted from real map or points generated according to a random process. The network is partitioned into a Voronoi diagram with $m$ Voronoi cells, and there is one cloudlet in each region. Fig., 2 shows 19 cloudlets in the network and the mobile device is connected to cloudlet $C_7$ for mobile application computing.

Suppose the computational task on the mobile device requires $C$ instructions. Let $S_i$ be the computing speed, in instructions per time slot (e.g., second), of the cloudlet $C_i$, $\forall i = 1, 2, \ldots, m$. This task thus takes $C/S_i$ time slots to

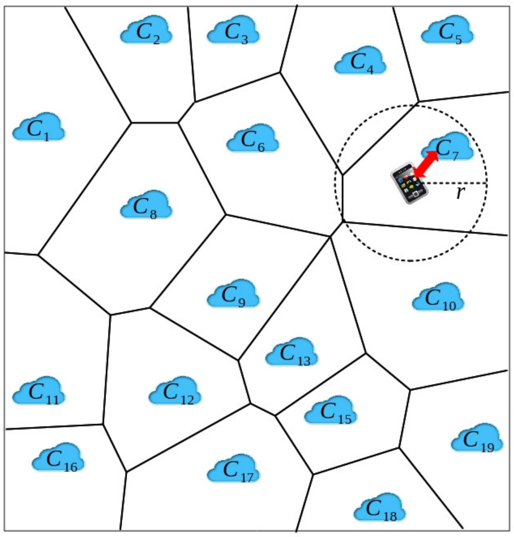


Fig. 2. Cloudlet network model

compute on cloudlet $C_i$, $\forall i = 1, 2, \ldots, m$. Denote $B$ as the network bandwidth. If the mobile device needs to send $D_{out}$ bytes of task data to the cloudlet while the cloudlet needs to send back $D_{in}$ bytes of executed task data to the mobile device, it takes $D_{out}/B$ and $D_{in}/B$ time slots to transmit and receive data, respectively. Define $\delta_i = C/S_i + (D_{in} + D_{out})/B$ as the *task completion time*, which is the sum of task transmission time and task computing time at cloudlet $C_i$.

**Cloudlet Connection Model:** When cloudlet $C_i$ ($1 \leq i \leq m$) is within the mobile device's transmission range, the mobile device can access the computing resources in $C_i$; otherwise, the mobile device is disconnected from $C_i$. In Fig. 2, connection to $C_7$ is available to the mobile device. Formally, suppose mobile device is moving in the network according to a mobility process $\mathcal{M}$. Denote by $X(t)$ and $X_{C_i}(t)$ the locations of the mobile device and cloudlet $C_i$, respectively. Let the transmission range of the mobile device be $r$. Connection to $C_i$ is available at time $t$ if and only if $||X(t) - X_{C_i}(t)|| \leq r$, where $||\cdot||$ is the Euclidean norm in 2-Dimension. Further, the *connection and inter-connection time* between a mobile device and a cloudlet are defined as follows.

**Definition 1.** The connection time $T_C$ of the mobile device and cloudlet $C_i$ ($\forall i = 1, 2, \ldots, m$) is defined as

$$T_C^i \triangleq \inf_{t>0}\{t : ||X(t) - X_i(t)|| > r\}, \quad (1)$$

given that $||X(0) - X_i(0)|| > r$ and $||X(0^+) - X_i(0^+)|| \leq r$. The inter-connection time (i.e., time between two consecutive connections) of the mobile device and cloudlet $C_i$ is defined as

$$T_I^i \triangleq \inf_{t>0}\{t : ||X(t) - X_i(t)|| \leq r\}, \quad (2)$$

given that $||X(0) - X_i(0)|| \leq r$ and $||X(0^+) - X_i(0^+)|| > r$.

Let $F_{T_C}^i$ and $F_{T_I}^i$ denote the distribution functions of the connection time $T_C^i$ and inter-connection time $T_I^i$, respectively, and suppose that they have finite expectations $\mu_{T_C}^i$ and $\mu_{T_I}^i$ and their density functions $f_{T_C}^i$ and $f_{T_I}^i$ exist and are continuous on $(0, \infty)$, respectively. In reality, distributions of $T_C^i$ and $T_I^i$ can be estimated based on movement history of mobile users as human tend to repeat their everyday schedules [9].

In this paper, we study the performance of using cloudlet for mobile applications. Due to node mobility, cloudlet connection is intermittent, which poses challenges for utilizing cloudlet computing. In order to identify power and node mobility of cloudlet computing, we examine the probability that a mobile device can connect to at least one cloudlet, which is called *cloudlet access probability*, the *success rate of task execution*, the total number of tasks executed by cloudlets and *average task execution speed* over time $t$.

## III. Cloudlet Computing Performance

Because of node mobility, the connection between a mobile device and a cloudlet can be intermittent. In order to study cloudlet computing performance, we start with modeling the connection and inter-connection process between a mobile device and a cloudlet.

**Definition 2.** Let $\{\eta(t), 0 \leq t < \infty\}$ be a stochastic process with state space $\{0, 1\}$. If a mobile device can connect to a cloudlet at time $t$, $\eta(t) = 1$; otherwise, $\eta(t) = 0$. Denote by $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots$ the lengths of successive intervals spent in states 0 and 1, respectively, in time $(0, \infty)$, where $\alpha_1, \alpha_2, \ldots$ are i.i.d. and $\beta_1, \beta_2, \ldots$ are i.i.d.. The process $\{\eta(t)\}$ assumes the states 0 and 1 alternately, as shown in Fig. 3. The process $\{\eta(t)\}$ is called *alternating renewal process*.

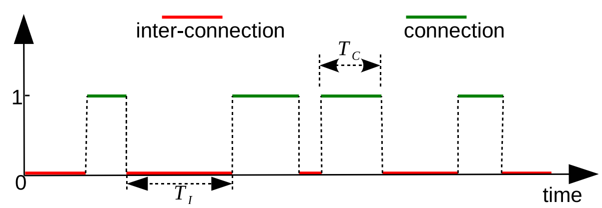


Fig. 3. The connection and inter-connection process of a mobile device and a cloudlet is an alternating renewal process.

With only a slight loss of generality, we assume that the time origin of the process $\{\eta(t)\}$ is an arbitrary connection or inter-connection. When $\eta(0) = 0$, the mobile device is initially disconnected from the cloudlet; when $\eta(0) = 1$, the mobile device is initially connected with the cloudlet. If $\eta(0) = 0$, $\alpha_i = T_I^i$ and $\beta_i = T_C^i$, i.e., $i$th inter-connection and connection time, respectively; if $\eta(0) = 1$, $\alpha_i = T_C^i$ and $\beta_i = T_I^i$, i.e., $i$th connection and inter-connection time, respectively. The former case is shown in Fig. 3. Based on the alternating renewal process of the connection between a mobile device and a cloudlet, we study the cloudlet access probability, task success rate and execution speed.

### A. Cloudlet access probability

A mobile device's connection to cloudlets is intermittent due to node mobility. If there is no cloudlet in the vicinity of mobile device, cloudlet computing is unavailable. Hence, it is important to find out the *cloudlet access probability*, which is defined as the probability that a mobile device can connect to at least one cloudlet in the network.

Note that if the mobile device can connect to multiple resource-rich devices at the same time, these devices can be seen as belonging to one cloudlet. In other words, a mobile device can connect to at most one cloudlet at any time $t$, and a mobile device's connections with different cloudlets are exclusive. This assumption is reasonable because different cloudlets probably locate at different community sites. Based on this assumption and the connection and inter-connection process of a mobile device and a cloudlet, we have the following theorem for the cloudlet access probability.

**Theorem 1.** *The limiting cloudlet access probability is*

$$CA = \sum_{i=1}^{m} \frac{\mu_{T_C}^i}{\mu_{T_C}^i + \mu_{T_I}^i}. \tag{3}$$

*where $\mu_{T_C}^i$ and $\mu_{T_I}^i$ are expectations of connection time $T_C^i$ and inter-connection $T_I^i$ between the mobile device and cloudlet $C_i$ ($i = 1, 2, \ldots, m$), respectively.*

*Proof:* The movements of mobile device result in alternating connection and inter-connection with a cloudlet $C_i, 1 \leq \forall i \leq m$, which is modeled in Definition 2. The probability that the connection between the mobile device and cloudlet $C_i$ is available at time $t$, conditional on the initial state, is given by Cox in Renewal Theory (1962, p.83) [10]. When the mobile device is initially connected to cloudlet $C_i$,

$$CA_1^i(t) = 1 - F_{T_C}^i(t) + \int_0^t h_1^i(u)[1 - F_{T_C}^i(t-u)]du,$$

where $h_1(u)$ is the inverse Laplace transform of

$$H_1^i(s) = \frac{f_{T_C}^i(s) f_{T_I}^i(s)}{s(1 - f_{T_C}^i(s) f_{T_I}^i(s))};$$

when the mobile device is disconnected from cloudlet $C_i$ at $t = 0$,

$$CA_0^i(t) = F_{T_I}^i(t) + \int_0^t h_1^i(u)[1 - F_{T_I}^i(t-u)]du.$$

It is reasonable to assume that the process of connection and inter-connection between nodes has been running for a long time before it is first observed. The limiting connection probability of cloudlet $C_i$ is

$$CA^i = \lim_{t\to\infty} CA_1^i(t) = \lim_{t\to\infty} CA_0^i(t) = \frac{\mu_{T_C}^i}{\mu_{T_C}^i + \mu_{T_I}^i}. \tag{4}$$

As the mobile device's connections to different cloudlets are exclusive, the cloudlet access probability is $CA = \sum_{i=1}^m CA^i$. Thus, we finish our proof. ∎

*Remark* 1. Connection probability $CA^i$ of a cloudlet $C_i$ is determined by the average connection and inter-connection time, i.e., mobility pattern of a mobile user. The more frequent visit and the longer sojourn time at the location of a cloudlet, the more likely a mobile user can connect to this cloudlet. The mobile device's isolation probability is $1 - CA$, which is determined by the percentage of time that the mobile user is at locations without any cloudlet (i.e., user mobility pattern).

### B. Task Success Rate

Cloudlets available at a point in time can not guarantee a successful task execution. In order to successfully compute a task, a cloudlet has to maintain a connection with the mobile device during the task transmissions and computation. In other words, cloudlet $C_i$ can successfully execute a task for a mobile device if a connection is available between them for at least $\delta_i$ period of time, where $\delta_i = C/S_i + (D_{in} + D_{out})/B$. We derive the task success rate by applying results on interval availability of an alternating renewal process [11].

**Theorem 2.** *The task success rate is*

$$SR = \sum_{i=1}^{m} CA^i \left(1 - \int_0^{\delta_i} [1 - F_{T_C}^i(x)]/\mu_{T_C}^i dx\right), \tag{5}$$

*where cloudlet $C_i$'s access probability $CA^i = \frac{\mu_{T_C}^i}{\mu_{T_C}^i + \mu_{T_I}^i}$.*

*Proof:* Define $SR^i(t, \delta_i)$ as the probability that a connection between the mobile device and cloudlet $C_i$ is available at least $\delta_i$ period of time starting at $t$. Based on the interval availability of an alternating renewal process derived by Barlow and Hunter (1961) [11], we have that conditioning on initial state $\eta(0) = 1$,

$$SR_1^i(t, \delta_i) = 1 - F_{T_C}^i(t+\delta_i) + \int_0^t h_1^i(u)[1 - F_{T_C}^i(t+\delta_i-u)]du;$$

while conditioning on initial state $\eta(0) = 0$,

$$SR_0^i(t, \delta_i) = \int_0^t h_2^i(u)[1 - F_{T_C}^i(t+\delta_i-u)]du,$$

where $h_2^i(u)$ is the inverse Laplace transform of

$$H_2^i(s) = \frac{f_{T_I}^i(s) f_{T_I}^i(s)}{s(1 - f_{T_C}^i(s) f_{T_I}^i(s))}.$$

It is readily seen

$$SR^i = \lim_{t\to\infty} SR_1^i(t, \delta_i) = \lim_{t\to\infty} SR_0^i(t, \delta_i) \tag{6}$$
$$= \frac{\int_{\delta_i}^{\infty}(1 - F_{T_C}^i(u))du}{\mu_{T_C}^i + \mu_{T_I}^i} = \frac{\mu_{T_C}^i(1 - \int_0^{\delta_i}[1 - F_{T_C}^i(u)]/\mu_{T_C}^i du)}{\mu_{T_C}^i + \mu_{T_I}^i}.$$

Note that $SR^i$ is the product of the limiting cloudlet access probability $CA^i$ and the limiting probability that it survives an interval of duration $\delta_i$. As the limiting success rate is $SR = \sum_{i=1}^m SR^i$, we complete our proof. ∎

The connection time between a mobile device and a cloudlet is also called contact time or link lifetime, which has been formally described to be *exponential* random variable under various mobility models [12, 13]. When $T_C^i$ ($i = 1, 2, \ldots, m$) follows exponential distribution with parameter $1/\mu_{T_C}^i$, we have the following corollary.

**Corollary 1.** *When $\{T_C^i, i = 1, 2, \dots, m\}$, are exponential random variables with rates $\{1/\mu_{T_C}^i, i = 1, 2, \dots, m\}$, the limiting task success rate is*

$$SR = \sum_{i=1}^{m} CA^i e^{-\frac{\delta_i}{\mu_{T_C}^i}}. \tag{7}$$

*Remark* 2. The probability that a task can be executed successfully by cloudlet $C_i$ not only depends on the *cloudlet access probability* $CA^i$ *and the* probability distribution of connection time $T_C^i$ *but also depends on the* task completion time $\delta_i$, *which is determined by computation demand* $C$, *sizes of task data* $D_{in}$ *and* $D_{out}$, *cloudlet computing speed* $S_i$, *and channel bandwidth* $B$.

### C. Task Execution Speed

In mobile cloud computing, mobile applications, such as mobile learning, health monitoring, and map navigation, require recurrent services. For example, in mobile gaming, mobile users offload game engine (e.g., graphic rendering) to the servers in the cloudlet and users need to access the cloudlet repeatedly as the game refreshes during the game playing time. In general, a mobile device has a large amount of tasks to compute, and each task is sent to a cloudlet for computing after the previous task is finished. It is important to find out how many tasks can be executed successfully over time $t$ and what is the average task execution speed.

Because of node mobility, the connection between a mobile device and a cloudlet is unstable. In order to maintain high cloudlet computing reliability, it is reasonable to assume that for recurrent task computing, a mobile device only utilizes a cloudlet when they are connected. Accordingly, the average number of executed tasks over a fixed time $t$ depends on the total connection time between a mobile device and its encountered cloudlets as well as task completion time. We derive the following theorem using renewal theory.

**Theorem 3.** *The average number of executed tasks over time* $t$*, denoted as* $N(t)$*, satisfies*

$$E(N(t)) = \sum_{i=1}^{m} \left\lfloor \frac{E(N_C^i(t))\mu_{T_C}^i}{\delta_i} \right\rfloor, \tag{8}$$

*where* $\lfloor\cdot\rfloor$ *is the floor function, and* $E(N_C^i(t))$ *is the number of connections between a mobile user and cloudlet* $C_i$ *within time t. Formally,* $E(N_C^i(t))$ *is the inverse Laplace transform of* $F_{T_C+T_I}^i(s)/[s(1-F_{T_C+T_I}^i(s))]$ *and* $F_{T_C+T_I}^i(s)$ *is the Laplace transform of random variable* $T_I^i + T_C^i$*.*

*Proof:* In the connection and inter-connection process of a mobile device and cloudlet $C_i$, define $S_0^i = 0$ and $S_n^i = \alpha^{i1} + \beta^{i1} + \alpha^{i2} + \beta^{i2} + \cdots + \alpha^{in} + \beta^{in}$ for $n \geq 1$. The process $N_C^i(t) = \max_{n\geq 0}\{n|S_n^i \leq t\}$ is the number of renewals over time $t$. The total connection time between a mobile device and cloudlet $C_i$ over time $t$ is approximately

$$CT_i(t) \approx \sum_{k=1}^{N_C^i(t)} T_C^{ik}, \tag{9}$$

where $T_C^{ik}$ is the $k$th connection time between a mobile device and cloudlet $C_i$. As $\{T_C^{ik}, k = 1, 2, \dots\}$ are i.i.d. and have the same distribution as $T_C^i$,

$$E(CT_i(t)) = E(N_C^i(t))E(T_C^i) = E(N_C^i(t))\mu_{T_C}^i, \tag{10}$$

where $E(N_C^i(t))$, by renewal theory, is the inverse Laplace transform of $F_{T_C+T_I}^i(s)/[s(1 - F_{T_C+T_I}^i(s))]$ and $F_{T_C+T_I}^i(s)$ is the Laplace transform of random variable $T_I^i + T_C^i$. The number of tasks executed by cloudlet $C_i$ is $\lfloor E(CT_i(t))/\delta_i \rfloor$. Accordingly, the total number of executed tasks over time $t$ is sum of $\lfloor E(CT_i(t))/\delta_i \rfloor$ over all cloudlets $C_i$, $i = 1, 2, \dots, m$. ■

Similar to connection time $T_C$, inter-connection time $T_I$, also called inter-contact time, has been shown to exhibit exponential tail decay under many mobility models [14]. Under the special case when $T_C^i$ and $T_I^i$ $(i = 1, 2, \dots, m)$ are exponential random variables, we can derive the closed form for the average number of renewals $E(N_C^i(t))$ over time $t$, thus $E(N(t))$ in the following corollary.

**Corollary 2.** *If* $T_C^i$ *and* $T_I^i$ $(1 \leq \forall i \leq m)$ *are exponential random variables with rates* $1/\mu_{T_C}^i$, $1/\mu_{T_I}^i$*, respectively,*

$$E(N(t)) = \sum_{i=1}^{m} \left\lfloor \frac{CA^i t + CA^i(1 - CA^i)\mu_{T_C}^i \left(1 - e^{-\frac{t}{CA^i \mu_{T_I}^i}}\right)}{\delta_i} \right\rfloor, \tag{11}$$

*where* $CA^i = \mu_{T_C}^i/(\mu_{T_C}^i + \mu_{T_I}^i)$.

*Proof:* When $T_C^i$ and $T_I^i$ are exponentially distributed with rates $1/\mu_{T_C}^i$ and $1/\mu_{T_I}^i$, respectively, $T_I^i + T_C^i$ has density function $\frac{1}{\mu_{T_I}^i - \mu_{T_C}^i}(e^{-t/\mu_{T_I}^i} - e^{-t/\mu_{T_C}^i})$, which gives the Laplace transform $L_{T_I+T_C}(s) = \frac{1}{\mu_{T_C}^i \mu_{T_I}^i (s+1/\mu_{T_C}^i)(s+1/\mu_{T_I}^i)}$. Then,

$$\mathcal{L}(E(N_c^i(t)), s) = \frac{1}{\mu_{T_C}^i \mu_{T_I}^i s^2 (s + \frac{1}{\mu_{T_C}^i} + \frac{1}{\mu_{T_I}^i})}.$$

Performing inverse Laplace transform, we have

$$E(N_c^i(t)) = \frac{t}{\mu_{T_C}^i + \mu_{T_I}^i} + CA^i(1 - CA^i)(1 - e^{-(\frac{1}{\mu_{T_C}^i} + \frac{1}{\mu_{T_I}^i})t}).$$

Substituting this equation into Eq. (8), we finish the proof. ■

To understand how the number of tasks executed by cloudlets increases over time $t$, we give some numerical results in Fig. 4. We set the scenario that a mobile user mainly stays at work place and home. Let there be two cloudlets in the network (i.e., $m = 2$). Cloudlet $C_1$ locates at the mobile user's home, and Cloudlet $C_2$ locates at the mobile user's office. The mean connection and inter-connection time between the mobile user with these two cloudlets are $\mu_{T_C}^1 = 12$ hours and $\mu_{T_I}^1 = 12$ hours, $\mu_{T_C}^2 = 8$ hours and $\mu_{T_I}^2 = 16$ hours, respectively. Let the task completion time $\delta_1 = \delta_2 = 1$ minute. Fig. 4 shows that number of tasks executed by cloudlet $C_1$ (or $C_2$) increases linearly with rate approximately equal to cloudlet $C_1$'s (or $C_2$'s) access probability $CA^1 = \frac{1}{2}$ (or $CA^2 = \frac{1}{3}$). The total number of executed tasks by cloudlets

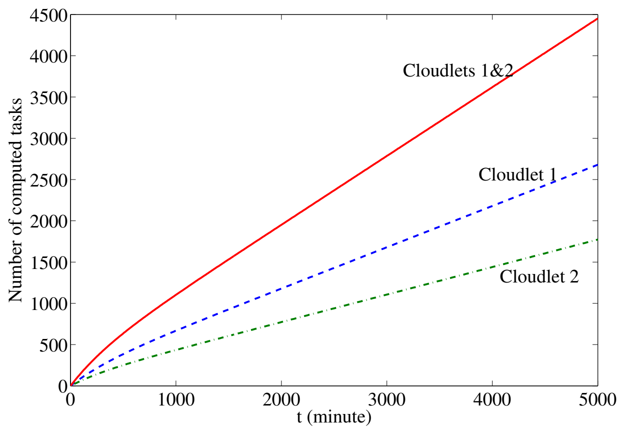


Fig. 4. Number of tasks computed by cloudlet 1, cloudlet 2, and both cloudlets over time $t$.

in the network increases linearly with rate approximately equal to cloudlet access probability $CA = CA^1 + CA^2 = \frac{5}{6}$.

*Remark* 3. Average number of executed tasks over time $t$ grows linearly with rate approximately equal $\sum_{i=1}^{m} CA^i/\delta^i$. In other words, number of executed tasks is mainly determined by cloudlet access probability and task completion time.

**Theorem 4.** *The limiting average speed of task execution satisfies*

$$CS = \lim_{t\to\infty} \frac{E(N(t))}{t} = \sum_{i=1}^{m} \frac{CA^i}{\delta_i}. \tag{12}$$

*Proof:* The average speed of task execution is $CS(t) = E(N(t))/t$. Based on Theorem 3, we have limiting task execution speed when $t \to \infty$,

$$CS = \lim_{t\to\infty} \frac{E(N(t))}{t} = \sum_{i=1}^{m} \lim_{t\to\infty} \frac{E(N_c^i(t))\mu_{T_C}^i}{t}.$$

According to the elementary renewal theorem,

$$\lim_{t\to\infty} \frac{E(N_c^i(t))}{t} = \frac{1}{(\mu_{T_C}^i + \mu_{T_I}^i)}.$$

Hence, we complete our proof. ■

*Remark* 4. The higher the cloudlet access probability $CA^i$ $(1 \leq \forall i \leq m)$ and the shorter the task completion time $\delta_i$ $(1 \leq \forall i \leq m)$ are, the faster the task execution speed $CS$ is. Findings in this paper reveal that mobility pattern of a mobile user determines its connection and inter-connection time to cloudlets, which in turn affect not only the cloudlet access probability, but also success rate and speed of task execution.

## IV. Conclusion

In this paper, we study the cloudlet computing for mobile applications, in which mobile users offload tasks to nearby resource-rich devices for instant service access and saving on roaming charges. As a cloudlet locates at a community site that a mobile user visits, its access probability for this mobile user is $\mu_{T_C}/(\mu_{T_I} + \mu_{T_C})$ determined by the mobile user's mean connection time and inter-connection time with the cloudlet. Moreover, cloudlet access probability affects the cloudlet computing performance, such as probability of successful task execution, average number of tasks executed over time $t$, and the limiting task execution speed. In summary, a mobile user's mobility pattern has significant impact on its cloudlet access probability and cloudlet computing performance, which can be measured by $\mu_{T_C}/(\mu_{T_I} + \mu_{T_C})$.

In the future work, we will integrate real mobility traces and mobile applications to evaluate the cloudlet computing performance, such as task completion time or execution speed, and energy consumption on local mobile devices. By comparing computing performance using cloudlets with using remote clouds, we could glean insights on the advantages and disadvantages of cloudlet and remote cloud, respectively, thus provide guidelines on whether or when cloudlets or remote clouds is suitable for mobile application computing.

## References

[1] S. Bicheno, "Global smartphone installed base forecast by operating system for 88 countries: 2007 to 2017," tech. rep., Strategy Analytics, Oct. 2012.

[2] 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.

[3] 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.

[4] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, pp. 14 –23, Oct.-Dec. 2009.

[5] A. Inc., "At&t individual & mobile share data plans." http://www.att.com/shop/wireless/data-plans.html, 2013.

[6] B.-G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "Clonecloud: elastic execution between mobile device and cloud," in *Proceedings of the sixth conference on Computer systems*, EuroSys '11, pp. 301–314, 2011.

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

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

[9] C. Song, N. B. Z. Qu, and A.-L. Barabási, "Limits of predictability in human mobility," *Science magazine*, 2010.

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

[11] R. E. Barlow and L. C. Hunter, "Reliability of analysis of a one-unit system," *Operations Research*, vol. 9, 1961.

[12] I. Gruber and H. Li, "Link expiration times in mobile ad hoc networks," in *Proceedings of 27th Annual IEEE Conference on Local Computer Networks (LCN)*, pp. 743 – 750, Nov. 2002.

[13] W. Wang and M. Zhao, "Joint effects of radio channels and node mobility on link dynamics in wireless networks," in *Proceedings of The 27th IEEE Conference on Computer Communications (INFOCOM)*, pp. 933 –941, Apr. 2008.

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