How Mobility Increases Mobile Cloud Computing Processing Capacity

Anh-Dung Nguyen, Patrick Sénac, Victor Ramiro

How Mobility Increases Mobile Cloud Computing Processing Capacity

Anh-Dung Nguyen*†, Patrick Sénac*† and Victor Ramiro‡
anh-dung.nguyen@isae.fr, patrick.senac@isae.fr, vramiro@niclabs.cl
*ISAE/University of Toulouse, Toulouse, France
†LAAS/CNRS, Toulouse, France
‡NIC Chile Research Labs, Santiago, Chile

***Abstract*—In this paper, we address a important and still unanswered question in mobile cloud computing "how mobility impacts the distributed processing power of network and computing clouds formed from mobile ad-hoc networks ?". Indeed, mobile ad-hoc networks potentially offer an aggregate cloud of resources delivering collectively processing, storage and networking resources. We demonstrate that the mobility can increase significantly the performances of distributed computation in such networks. In particular, we show that this improvement can be achieved more efficiently with mobility patterns that entail a dynamic small-world network structure on the mobile cloud. Moreover, we show that the small-world structure can improve significantly the resilience of mobile cloud computing services.**

***Keywords*—Mobile cloud computing, mobility, quality of service**

I. Introduction

"Cloud computing" has recently appeared as a buzz word in many medias in which the term refers both to the technology advancement and also to the business model behind. The idea is not new but roots from already developed technologies such as distributed computing, autonomic computing, hardware virtualization and web services. It's the maturation and convergence of all these technologies that makes cloud computing viable today. By virtualizing the aggregated computing resources in order to offer to users the on-demand utility (e.g. computing, storage, software as service) in a pay-as-you-go fashion, much like the power distribution grid system, cloud computing appears as a main actor of information industry today. This can be seen through the explosion of cloud computing services deployed the Internet in recent years.

Besides, with the advances of electronic technologies, mobile wireless devices have gradually become more and more powerful in terms of processing, storage and communication capacity. This potentially leads to the emergence of mobile ad-hoc networks that deliver, without any infrastructure, computing resource complementary to the existing infrastructured networks. These "mobile clouds", which leverage on opportunistic contacts between users, can potentially deliver free communication, storage and processing services shared between users according to peer to peer resource sharing policies.

Although the application perspective sounds interesting, the underlying technology challenges are not negligible due to the difficulties raised by dynamic networks. The first obstacle comes from the mobile nature of such network and raises the question of "how the mobility impacts the distributed processing performances of the mobile clouds?". Indeed, the unstable network topology makes that continuous end-to-end communication unguaranteed and hence the service delivery may be disrupted. Indeed, in the context of spontaneous and infrastructureless networks, a kind of delay tolerant network, nodes must rely on intermittent contacts leading to use the store-carry-and-forward communication paradigm for inter-node communication. Therefore, if the role of mobility on communication performances such as end to end delay and bandwidth has been already studied, the impact mobility schemes on the global processing power delivered by a mobile network cloud has not been studied yet.

In this paper, we address this issue and show that the mobility can enhance significantly the computing capacity of network clouds composed of mobile nodes. Considering a dynamic network as an aggregate distributed computing resource, we use Particle Swarm Optimization (PSO) - an optimization method based on distributed autonomous agents-coupled with a generic mobility model to assess the impact of node mobility on distributed processing. The questions on service resilience against network churn are also discussed.

The rest of the paper is structured as follows. First, Section II discusses the state of the art of mobile cloud computing. In Section III, we study the impact of mobility on the quality of mobile cloud computing services. Section III studies of the impact of dynamic network structures on mobile cloud computing. In Section V, the question of service resilience is discussed. Finally, Section VI concludes the paper.

II. State of The Art

Mobile cloud computing is still a young field and there is still discussion on its definition. In its infancy, mobile cloud computing has been considered as a derived branch of cloud computing with two schools of thought (see [7] for a survey). The first refers to performing computing activities (data storage and processing) in infrastructured cloud and let mobile devices be simple terminals to access to service. This centralized approach has the advantage that mobile devices don't need to have a powerful computing capacity but the drawback is that users depend strongly on the infrastructure network and on its performances.

The second school of thought defines mobile cloud computing as performing computing activities on mobile platform.

Therefore a mobile cloud network is an infrastructureless extension of the traditional infrastructure based cloud networks. Mobile devices are clients of service but are also part of the cloud, providing hardware and software resources. The benefit of this distributed approach is the omnipresence and the speed of service accessibility, the support of mobility and locality, the freedom of deployment and use of new services as well as the reduced hardware maintenance costs. Although the approach is promising, its main challenge resides in the dynamic of network which poses difficulties in communication and hence service access. In this paper, we focus on this definition of mobile cloud computing.

To the best of our knowledge, very few contributions have been proposed for mobile cloud computing. Hyrax [2] is a mobile-cloud infrastructure that enables smart-phone applications that are distributed both in terms of data and computation. Hyrax allows applications to conveniently use data and execute computing jobs on smart-phone networks and heterogeneous networks of phones and servers. Its implementation is based on Hadoop and tested on Android platform. But since Android doesn't support ad-hoc network yet, the phones have to communicate through a WIFI central router.

Satyanarayanan et al. [4] present the cloudlet concept. In this approach a mobile client is seen as a thin client with respect to a service which is customized over a virtual machine in the wireless LAN. Hence the cloudlet is a proxy representation of a real service enhanced for the mobile device. The main motivation is how bandwidth limits and latency over wireless networks impacts over users services.

III. Impact of Mobility on Mobile Cloud Computing

In this section, we evaluate the impact of mobility on mobile cloud computing. Let us consider that a mobile cloud network created by several human portable devices offers a distributed processing service such as optimizing a function via a Particle Swarm Optimization (PSO) algorithm. According to PSO, each node in the network has a local solution to the optimization problem. Through intermittent contacts, mobile nodes learn others' solution to improve their local optimum and hence accelerate the convergence towards the global optimum. For the sake of simplicity, we make the following assumptions:

1) Each node in the network knows in advance the goal function and its solution.
2) An external system (e.g. WIFI hot-spots) is responsible for results retrieval from mobile nodes.
3) The service is considered delivered when the global optimum reaches a goodness predefined by user.

In practice, the goal function as well as its solution is usually unknown in advance and therefore we have to rely on a diffusion technique to disseminate the information of the goal function into the network. The obtained global optimum and stopping condition in that case will depend on the current solutions found by nodes (e.g. a node stops the computation when its local solution no longer changes for a while). In this theoretical work, we focus only on the impact of mobility on computing delay and therefore the previous assumptions seem reasonable.

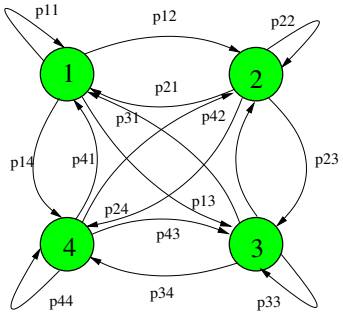


Fig. 1. STEPS markovian mobility model

A. Mobility model

In order to reproduce at the simulation level realistic human mobility patterns, we use the STEPS mobility model [3]. As we have shown in a previous work, this flexible parametric model can express a large spectrum of mobility patterns: from highly nomadic ones to localized ones. Therefore, STPES makes it possible to evaluate the impact of different mobility contexts on mobile cloud computing. In STEPS, the network area is modeled as a torus divided in several zones. The model implements the notion of preferential attachment usually observed in human mobility in which each node is attached to one or several preferential zones. Inside zones, mobile nodes move according to the Random Waypoint model. The movement of nodes between zones follows a Markov chain of which the transition probability is given by a power law distribution. This distribution is driven by a parameter of the STEPS model which allows the nodes nomadism to be enforced or reduced (i.e. the probability that a mode moving outside his preferential zone has to return to that zone). Figure 1 illustrates the underlying Markov chain of STEPS with 4 zones.

B. Particle Swarm Optimization

The Particle Swarm Optimization (PSO) algorithm [1] is an optimization method based on swarm intelligence - a sub-field of artificial intelligence which studies the collective intelligent behavior emerging from the interactions between individuals of a swarm of autonomous agents. Swarm intelligence considers intelligence as the combination of the knowledge acquired by individuals through experiences in the past and the knowledge acquired from the others through social interactions. In PSO algorithm, a set of candidate solutions called particles move around in the search space according to a simple mathematical formula over the particle's position and velocity. Each particle's movement is influenced by its local best known position and also by the global best known position found by other particles. The swarm is expected to move collectively towards the optimal solution. Besides, this method is able to solve a multimodal optimization problem.

In its simplest form, let $\vec{x}_i$ be the multidimensional vector of the particle i position, the position of the particle is updated according to the formula

$$\vec{x}_i(t+1) = \vec{x}_i(t) + \vec{v}_i(t) \quad (1)$$

where $\vec{x}_i(t)$ is the position of particle i at time t.

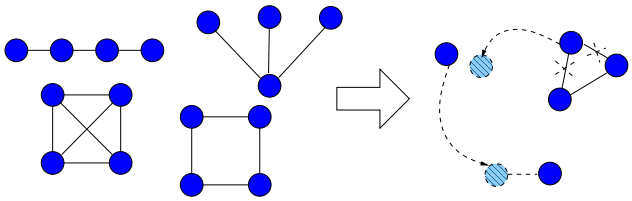


Fig. 2. Typical static neighborhood topologies vs dynamic neighborhood topology generated by STEPS

The velocity of the particle is updated according to the formula

$$\vec{v_i}(t) = \vec{v_i}(t-1) + \phi_1 \left[\vec{p_i} - \vec{x_i}(t-1)\right] + \phi_2 \left[\vec{p_g} - \vec{x_i}(t-1)\right] \quad (2)$$

where

- ϕ_1, ϕ_2 are uniform random variables taking values in $[0, 1]$. These variable represent the relativity between the effect of individual experience and of social influence.
- $\vec{p_l}$ denotes the best known position of particle i ("l" for local).
- $\vec{p_g}$ denotes the best known position of i's neighbors ("g" for global)

This formula entails wider and wider oscillations of particles in the search space. One solution to this issue is based on velocity damping, that is, if $v_{id} > V_{max}$ then $v_{id} = V_{max}$ else if $v_{id} < -V_{max}$ then $v_{id} = -V_{max}$ where v_{id} is the dimension d of $\vec{v_i}$. In consequence, the particles move only in a restricted search space.

In PSO, individual can be connected to one another according to a great number of neighborhood topologies (Figure 2 illustrates the most used schemes). Each neighborhood topology, traditionally considered as static conversely to our analysis, results in different behaviors and performances for the PSO algorithm.

In this work, since node moves, the neighborhood topology is no longer static but dynamic. Indeed, when the mobility degree is low, links between nodes is stable and the network is nearly static. On the contrary, when the mobility is high, links change rapidly over time and so does the neighborhood topology. Therefore the goal of this experiment is to evaluate the effect of mobility on the convergence delay of the algorithm.

C. Simulation Results

We implemented the mobility model and the PSO algorithm on MATLAB. At the beginning of each simulation, 100 nodes are uniformly distributed over the network area which is divided in 10×10 zones representing preferential attachment according to STEPS model. The movement of node between zones is driven by the locality degree parameter (α) of STEPS model. We vary α between 0 and 8 to obtain a large spectrum of mobility patterns. When $\alpha = 0$ nodes are highly nomadic, moving from a zone to one another in a random manner that makes the network highly dynamic. On the contrary, when $\alpha = 8$, nodes are highly localized (i.e.sedentary) and therefore there are less information exchange between distant zones.

Number of nodes/particles	100
Number of zones	10×10
Network size	100×100 m^2
Radio range	10 m
Node speed	$3 - 5$ km/h
De Jong function	Sphere
Number of dimensions	2
Stopping condition	error $<= 10^{-6}$

TABLE I. SIMULATION SETTINGS

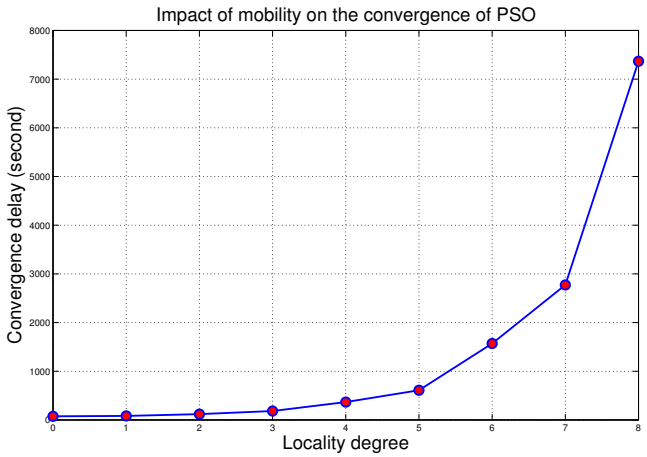


Fig. 3. Impact of mobility on the convergence delay of PSO algorithm

The PSO algorithm is implemented in every mobile nodes so that each node contains 1 particle. The position of particle is randomly initialized, taking values in range $[x_{max}, -x_{max}]$. Particle's position is updated at each contact with another node according to the formula 2.

As goal function, we used the Sphere function from the De-Jong test suite. This suite consists of goal functions with different difficulties to measure the performances of optimizers. The Sphere function is the first and easiest function of the suite. It is symmetric, unimodal and is often used to measure the general efficiency of optimizers. That is

$$f(\vec{x}) = \sum_{i=1}^{D} {x_i}^2$$

where D is the number of elements of $\vec{x}$. The Sphere function has a global optimum $f(\vec{x}) = 0$ at $\vec{x} = (0, 0, 0, \ldots, 0)$.

We used root-mean-square error to measure the goodness of the solution. The algorithm stops when the error is smaller than a predefined threshold. The simulation settings are summarized in Table I

Figure 3 shows the optimization convergence delay according to nods' locality degree. These results is are averaged over 10 simulation runs. On the figure, we can see that the more mobile nodes are, the smaller convergence delay is. This result shows that nodes mobility can increase dramatically the processing capacity of mobile cloud networks.

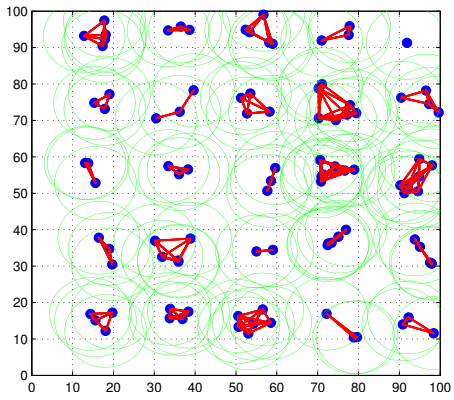


Fig. 4. Node spatial distribution

IV. Impact of Network Structure on Mobile Cloud Computing

In this section, we evaluate the processing capacity of mobile cloud computing under various dynamic network structures. With the same approach as introduced in Section III, we measure the convergence delay of a PSO algorithm implemented on a mobile cloud network and show that this delay can be significantly minimized if the network has a dynamic small-world structure.

The small-world phenomenon introduced by Watts and Strogatz [6] refers to static graphs with high clustering coefficient and low shortest path length. Through a process which consist in rewiring randomly edges of a graph, by varying the rewiring ratio, the authors showed that for an interval a rewiring ration the resulting static graph, exhibit a small world structure which cumulate short path observed in random graphs with high clustering coefficient intrinsic to regular lattices. In a previous work [3], we have shown that this small world behavior can be observed in dynamic graphs too. We have shown that in a dynamic networks, the analog of the rewiring process in static graph is done from varying the ratio and intensity of nomadic nodes. Moreover, we have shown that the STEPS model is capable of exhibiting this small-world phenomenon in dynamic networks.

Indeed, starting from the same network configuration as in Section III, we divide mobile nodes in 2 categories. The first consists in highly localized nodes which stay almost all their time in their preferential zones. The second category consists in highly nomadic modes which move constantly from zone to zone. At the beginning of simulation, the nodes are distributed over the network area so that nodes in different preferential zones cannot communicate each other (Figure 4 shows the node spatial distribution). We vary the fraction of mobile node p_m from 0 to 1. When p_m equals to 0, the network consists in disconnected islands with only intra-zone communications that entails a regular structure similar to the one in static graphs. On the contrary, when p_m equals to 1, the inter-zones movement of highly mobile nodes makes that the network topology changes constantly which entails a random network structure. Figure 5 shows the evolution of the clustering coefficient and shortest path length according to the fraction of highly mobile nodes.

We processed the PSO algorithm over all these network structures and then measured the resulting convergence delay.

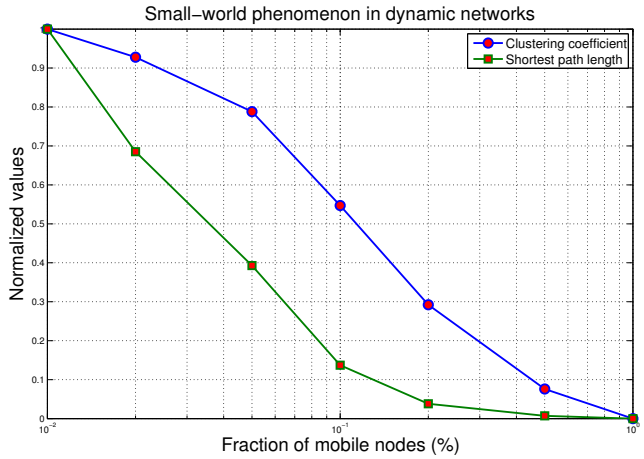


Fig. 5. Small-world phenomenon in dynamic network

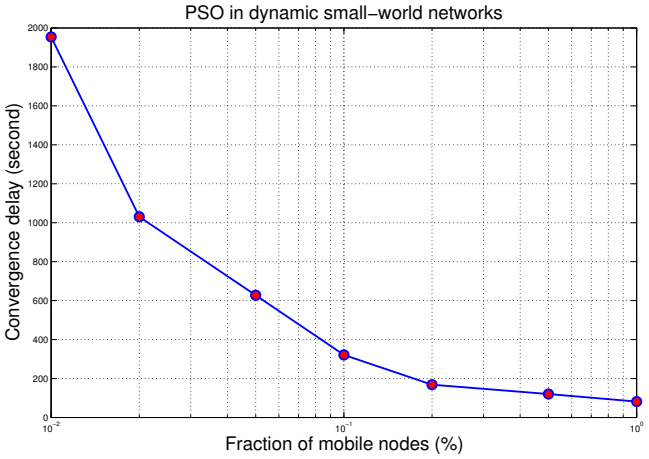


Fig. 6. Mobile cloud computing in small-world networks

Figure 6 shows the results averaged over 10 simulations. These simulation results show that the convergence delay of PSO decreases rapidly down to an asymptotic part started when the network exhibits a small world structure. This original result is significant because the small-world structure, which as shown in this paper improves distributed processing, was shown to emerge naturally in the great majority real dynamic networks [5].

V. Resilience of Mobile Cloud Computing Service

Nowadays, mobile devices still have limited energy capacity and communication as well processing are two important sources of energy waste. Therefore nodes' churn is intrinsic to dynamic netwoks clouds. Nodes running out of battery cannot contribute to distributed processing anymore and in consequence, mobile cloud networks may suffer impredictible nodes failures. Besides, mobile cloud networks may be the target of attacks, for instance DDOS, which can potentialy make unavailable parts of the network. In this section, we evaluate, under various mobility contexts, the resilience of distributed services deployed on such networks.

First, we assume that the evolution of number of inactive nodes (i.e. attacked or out of battery) follows a Poisson

process. Therefore, the number of inactive nodes during a time interval τ is distributed according to a Poisson distribution

$$P[(N(t+\tau) - N(t)) = k] = \frac{\exp -\lambda\tau(\lambda\tau)^k}{k!}$$

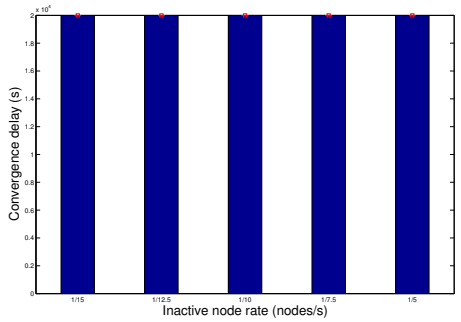
where $k = 0, 1, 2, \ldots$ and λ is the arrival rate of inactive nodes.

With the same simulation settings as introduced in Section IV, we perform simulations with various values of λ $(1/15, 1/12.5, 1/10, 1/5)$ and under various mobility contexts (i.e. by varying the fraction of mobile nodes). In these simulations, we stop the PSO algorithm when 95 % nodes reach the optimum. If this threshold is not reached before all the nodes become inactive, as there is no recovery possible in this case, the service will never be delivered and hence we assign the simulation duration time to the convergence delay.

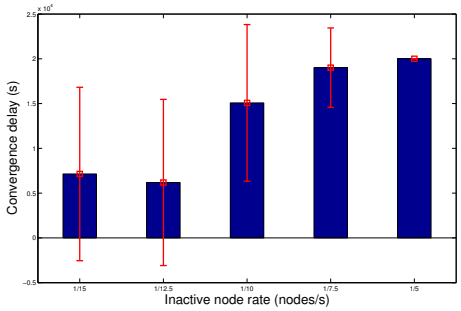
Figure 7 shows the results averaged over 20 simulations. These simulations show that with dynamic small-world networks (Figure 7(b) and 7(c)), the distributed service resists much better to departed nodes compared highly localized network (Figure 7(a)) and offers approximately the same resilience level than random networks (Figure 7(d)). These results suggest that a small-world structure not only contribute to enhancing distributed performances but also offers good resilience properties.

VI. Conclusion

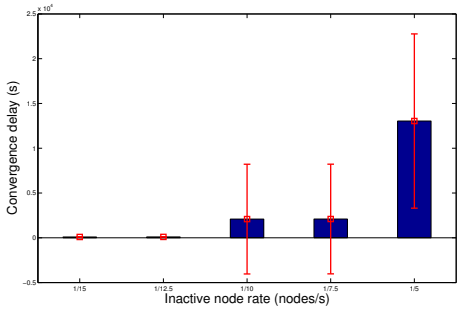
In this paper we not only show that nodes' mobility enhance the processing capacity of dynamic network cloud but we also showed how mobility impacts the performance and the resilience of these mobile clouds. In particular, we have shown that significant performance improvement can be obtained when dynamic networks exhibit a small-world structure and moreover, this particular structure can improve the resilience of the network against inactive nodes. This means that by introducing even a small percentage of highly mobile nodes in a high localized network, we can improve significantly the processing capacity and resilience of mobile cloud computing. These results open the way to adaptive strategies that would aim to adapt dynamic network topology and behavior according to their processing load and constraints. Moreover these strategies have to consider also storage and energy consumption which are critical in the context of handheld systems. Our current work investigates these research directions.



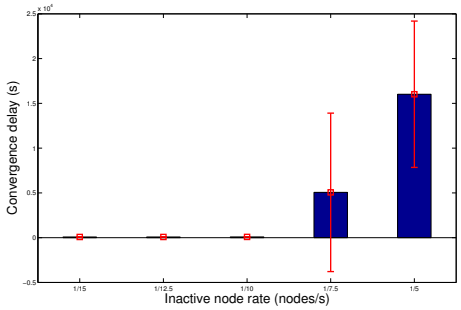
(a) 0 % of mobile nodes



(b) 10 % mobile nodes



(c) 50 % mobile nodes



(d) 100 % mobile nodes

Fig. 7. Resilience of mobile cloud network distributed services under various mobility contexts

References

[1] J. Kennedy and Russell C Eberhart. *Swarm intelligence*, volume 78. Springer, January 2006.

[2] E.E. Marinelli. Hyrax: cloud computing on mobile devices using MapReduce. Technical Report September, DTIC Document, 2009.

[3] A.D. Nguyen, Patrick Sénac, Victor Ramiro, and Michel Diaz. STEPS-an approach for human mobility modeling. In *NETWORKING 2011*, pages 254–265. Springer, 2011.

[4] Mahadev Satyanarayanan, Victor Bahl, Ramon Caceres, and Nigel Davies. The Case for VM-based Cloudlets in Mobile Computing. *IEEE Pervasive Computing*, 1, 2011.

[5] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora. Small-world behavior in time-varying graphs. *Physical Review E*, 81(5):5, May 2010.

[6] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442, June 1998.

[7] Xiaopeng Fan, Jiannong Cao and Haixia Mao. A Survey of Mobile Cloud Computing. *ZTE Corporation*, 2011.