2011 10th IEEE/ACIS International Conference on Computer and Information Science

# A Survey of Research on Mobile Cloud Computing

Le Guan
Beijing University of Posts and Telecommunications
Beijing, China
optimism1226@gmail.com

Xu Ke
Beijing University of Posts and Telecommunications
Beijing, China
permit@263.net

Meina Song
Beijing University of Posts and Telecommunications
Beijing, China
mnsong@bupt.edu.cn

Junde Song
Beijing University of Posts and Telecommunications
Beijing, China
jdsong@bupt.edu.cn

*Abstract*—**The rapid development of mobile computing and cloud computing trigger novel computing paradigm-----Mobile Cloud Computing. This paper review current research effort towards Mobile Computing. First, we present several challenges for the design of Mobile Cloud Computing service. Second, a concept model has been proposed to analyze related research work. Third, we survey recent Mobile Cloud Computing architecture, application partition & offloading , and context-aware service.**

***Keywords-mobile cloud computing, application partition, offloading, context-aware***

## I. INTRODUCTION

Nowadays, the market of mobile phone has expanded rapidly. By the end of 2009, less than 20 years later, the number of mobile cellular subscriptions worldwide reached approximately 4.6 billion, 370 times the 1990 number [1]. The widely use of mobile phone lead to the prosperity of mobile service. Dream of “Information at your fingertips anywhere, anytime” has become true. However, mobile devices still lack in resources compared to a conventional information processing device such as PCs and laptops. Also, the limitation of battery restricts working time. How to augment capability of mobile phone has become the important technical issue for mobile computing.

The paradigm of cloud computing brings opportunities for this demand. Cloud computing provide new supplement, consumption, and delivery model for IT service. Cloud-based services are on-demand, scalable, device-independent and reliable. Thus, there comes Mobile Cloud Computing, which aims at using cloud computing techniques for storage and processing of data on mobile devices, thereby reducing their limitations. According to ABI Research, by 2015, more than 240million business customers will be leveraging cloud computing services through mobile devices, driving revenues of $5.2billion[2].

To deliver cloud service in mobile environment, we might face several problems. Device may hand off among different wireless transmission district, and transport channels are not so reliable to guarantee cloud service delivery. Furthermore, mobile devices can’t handle complicated applications due to their innate characters. Also, it is impossible that mobile device always online, that is, we should consider the offline solution for mobile device. What’s more, the absence of standards, security and privacy, elastic mobile applications requirement may obstruct the development of Mobile Cloud Computing as well.

Researchers provide various solutions for Mobile Cloud Computing service. Some proposal application partition & offload schemes to leverage the working load of Cloud and Client, which may reduce processing burden on the mobile client. Several researchers focus on the feature of “Mobile”, to provide context-aware service for users, which may triggers new applications for mobile environment. Contexts include geo-location and social activities.

This paper introduces the basic model of Mobile Cloud Computing, and surveys state-of-art of systems. First, we describe technical challenges of Mobile Cloud Computing. Then, after introducing concept model and basic architecture, we survey key technologies, e.g. partition & offloading and context-based service. At last, we conclude the recent research activities of Mobile Cloud Computing.

The rest of paper is organized as follows: Section Ⅱ shows challenge of Mobile Cloud Computing service. Section Ⅲ presents basic model and architecture of Mobile Cloud Computing systems. Section Ⅳ talks about partition & offloading schemes and Section Ⅴ describes context-aware service. Section Ⅵ conclude the whole paper.

## II. CHALLENGES OF MOBILE CLOUD COMPUTING

Mobile Cloud Computing services are implemented in mobile wireless environment, incorporating several challenges such as the dependency on continuous network connections. Also Mobile Cloud Computing concepts rely on an always-on connectivity and will need to provide a scalable and high quality mobile access.

### *A. Network latency and limited bandwidth in the mobile network*

First, Mobile Cloud Computing may face the challenge from the transmission channel due to the intrinsic nature and constraints of wireless networks and devices. This is especially true when it comes to rich-internet and immersive mobile applications, e.g. online gaming and augmented reality that require high-processing capacity and minimum network latency. These will most probably continue to be processed

978-0-7695-4401-4/11 $26.00 © 2011 IEEE
DOI 10.1109/ICIS.2011.67

locally on powerful smart phones and mobile tablets. Mobile broadband networks generally require longer execution times for a given application to run in the cloud and network latency issues may deem certain applications and services unfit for the mobile cloud.

### B. *Various access scheme in mobile envinroment*

Mobile Cloud Computing would be deployed in a heterogeneous access scenario with different radio access technologies such as GPRS, 3G, WLAN, WiMax. Mobile Cloud Computing requires wireless connectivity with the following features:

- Mobile Cloud Computing requires an "always-on" connectivity for a low data rate cloud control signaling channel.
- Mobile Cloud Computing requires an "on-demand" available wireless connectivity with a scalable link bandwidth.
- Mobile Cloud Computing requires a network selection and use that takes energy-efficiency and costs into account.

The most critical challenge of Mobile Cloud Computing is probably to guarantee a wireless connectivity that meets the requirements of Mobile Cloud Computing with respect to scalability, availability, energy- and cost-efficiency [3].

### C. *Elastic application models*

Cloud Computing services are scalable, via dynamic provisioning of resources on a fine-grained, self-service basis near real-time, without users' consideration for peal loads. This requirement is particularly important towards mobile cloud computing scenario. Mobile applications can be launched on the device or cloud, and can be migrated between them according to dynamic changes of the computing context or user preferences. Also, limited resource of mobile device will restrict application processing. Thus, elastic application model should be proposed to solve fundamental processing problem

### D. *Security and Privacy*

Cloud computing users prove their identities with digital credentials, typically passwords and digital certificates. If an attacker could fake or steal these credentials, the cloud computing system will suffer from spoofing attacks. In mobile cloud computing, the problem is even severe because mobile devices often lack of computing power to execute sophisticated security algorithms. Moreover, it is difficult to enforce a standardized credential protection mechanism due to the variety of mobile devices [4].

## III. CONCEPT MODEL AND ARCHITECTURE

.In this section, we present concept model to analyze mobile cloud computing technology, and then provide several architecture model to organize Mobile Cloud Computing systems.

### A. *Concept model*

As well known, cloud computing service can be divided into three types according to delivery manner: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS). However, Mobile Cloud Computing would not separate into these types. Mobile Cloud Computing focuses on the connection between client and cloud, which may differ from common features of cloud computing.

In architectural considerations of creating next generation mobile applications, Jason H Christensen [5] proposal three component archetype: the combination of smart mobile device, REST based cloud computing, and context enablement. This three component model matches with transmission model of Mobile Cloud ---"Client-Connection-Cloud".

We can reconstruct concept model on vertical view, as shown in Figure.1. The left and right entities are respectively client and cloud. Between client side and cloud side there is "Transmission Channel" component. Upon this entity are "Resource Scheduling" and "Context Management" components, both of which occupy client and cloud sides. The prerequisite of this model is that: a) The client is context-aware; b) Cloud side should deliver elastic, on-demand service for client. Next, we explain three middle part of the model with down-top approach.

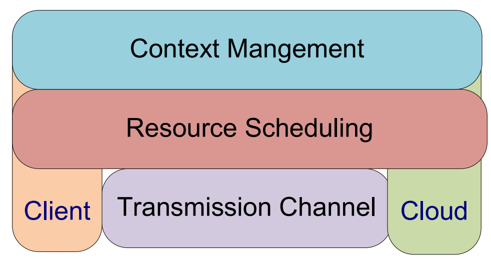


Figure 1. Concept model of Mobile Cloud Computing

#### 1) *Transmission Channel*

Transmission channel refers to various wireless transport protocols. The wireless connection between client and cloud is double-edged sword for mobile cloud computing applications: For one thing, the weak transmission channels degrade performance of stable cloud service; for another, dynamic characters of connection produce various contexts, which trigger prosperous mobile applications.

#### 2) *Resource Scheduling*

Resource scheduling component address the schedule o resource, such as computing resource and storage resource. In this level, virtual machines will be introduced to handle of resource dispatch. Nevertheless, we can view this problem in another view. Resource may be stable but applications may transmit to other places. In mobile cloud computing scenario we often consider to decompose complex application and handle application with parallel methods. Usually, *application partition and offloading* may contribute to usage of mobile device. Partition and offloading approach will be studied in Section Ⅳ.

*3) Context Management*

Context Enabled features of mobile device allow us to ascertain additional information from the computing device itself without the need for explicit user input. Context Management module can track context parameters and adapt to modification of context conditions. This capability has enabled a number of new application spaces such as Location Based Services (LBS), spatial augmented reality (SAR), and explicit spatial contexts using Bluetooth or WiFi.

Typically, context can be classified into two types:

*a) Spatial contexts*

Spatial contexts are contexts that are based on position, proximity. They allow context-aware applications to provide input for Location Based Service. For example, my iPhone can get location information and provide to Foursquare software, then I can play online games.

*b) Social contexts*

Social contexts are contexts that have been explored in social network analysis threads. In the context of mobile computing these contexts are particular ones that out of inherent characters of mobile computing but encourage user to group interaction.

Context management technology will be surveyed in Section Ⅴ.

*B. Architecture*

The architecture of Mobile Cloud Computing refers to the organization of Mobile Cloud Computing systems. Generally, most researchers want to enhance capability of mobile devices with cloud technology. Also, some researchers explore the use of cloud computing to execute mobile applications in behalf of the device. Thus, architecture scheme contains two types: agent-client scheme and collaborated scheme.

*1) Agent-client scheme*

In this scheme, cloud side provides overall resource management for mobile devices, to help to overcome limitations of mobile devices in particular of the processing power and data storage. As is shown in Figure 2, cloud side generate agent for each device. Mobile device communicate with its agent to contact with other entities outside this domain.

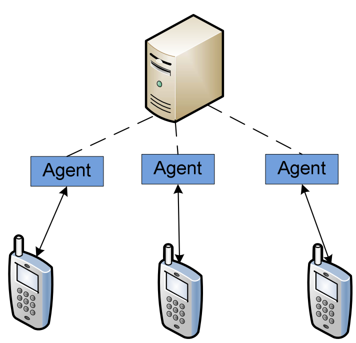


Figure 2. Agent-client architecture for Mobile Cloud Computing

Mahadev Satyanarayanan [6] provided cloudlet-based, resource-rich, mobile computing. In this architecture, a mobile user exploits virtual machine technology to customize service software on a nearby cloudlet and then uses that service over a wireless LAN; the mobile device typically functions as a thin client with respect to the service. A cloudlet is decentralized and widely-dispersed Internet infrastructure whose compute cycles and storage resources can be leveraged by nearby mobile computers. The natural implementation is to extend Wi-Fi access points to include substantial processing, memory and persistent storage for use by associated mobile devices.

Xinwen Zhang [7] built elastic applications which augment resource-constrained platforms. An elastic application can consist of one or more weblets, which function independently, but communicate with each other. When the application is launched, an elasticity manager running on the device monitors the resource requirements of the weblets of the application, and makes decisions where they should be launched. Computation intensive weblets usually strain the processors of mobile devices, therefore they can be launched on one or more platforms in the cloud; while user interface components (UI) or those needing extensive access to local data may be launched on the device.

*2) Collaborated scheme*

Collaborated schemes regard device as a part of cloud. This approach utilize remain resource of mobile device. The function of cloud server may be the controller and scheduler for collaboration among devices.

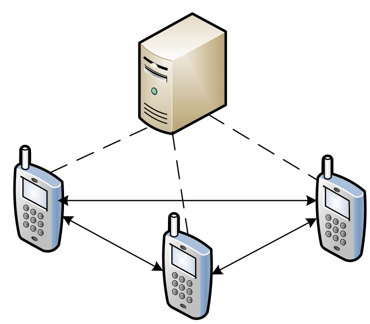

Figure 3. Collaborated architecture for Mobile Cloud Coputing

Hyrax [8] is a platform derived from Hadoop that supports cloud computing on Android smart phones. Hyrax allows client applications to conveniently utilize data and execute computing jobs on networks of smart phones and heterogeneous networks of phones and servers. By scaling with the number of devices and tolerating node departure, Hyrax allows applications to use distributed resources abstractly, oblivious to the physical nature of the cloud. In Hyrax, several traditional machines play the role of NameNode and JobTracker.

Black and Edgar [9] demonstrated the feasibility and value of enabling mobile devices within a grid computing framework by implementing the BOINC client on an Apple iPhone. Work units are downloaded from a BOINC server and executed on the iPhone via a virtual machine emulating an x86 processor, and results are uploaded to the server. The world of mobile devices brings renewed challenges to the problem of grid client design in the areas of network bandwidth, processor capability, storage, and energy consumption.

## IV. Application Partition and Offloading

Application partition and offloading technology play an important role for the implementation of elastic applications. Application partition decompose complex workload to atomic ones, thus can be processed concurrently. Offloading application can free burden of mobile devices and save their energy consumption.

### *A. Partition*

To achieve seamless and transparent migration and offloading, each application should be partitioned into components. Application partition should consider resource consumption and data dependency.

Ioana Giurgiu [10] el. proposed two-step approach to optimally partition an application between a mobile phone and a server. First, they abstract an application's behavior as a data flow graph of several inter-connected software modules. Given this graph, in the second step, a partitioning algorithm finds the optimal cut that maximizes (or minimizes) a given objective function. They propose two types of partitioning algorithms: ALL and K-step. In the first case, the best partitioning is computed offline by considering different types of mobile phones and network conditions. In the second case, the partitioning is computed on-the-fly, when a phone connects to the server and specifies its resources and requirements. ALL fits the first scenario, while K-step the second one.

Xun Luo [11] presented "Cloud-Mobile Convergence for Virtual Reality (CMCVR)" concept. In CMCVR, to take advantage of the better load balancing inherent in by-region application partitioning, the author proposed Hybrid Application Partitioning Strategy, also in two steps: The first stage breaks down the workload with the by-scale strategy; large workloads at high scale levels are further partitioned in the second stage which uses the by-region strategy. The partitioning process is completed when an optimized overall system performance is achieved.

Byung-Gon and Petros [12] introduced the notion of dynamic partitioning of applications between weak devices and clouds and argue that it is the key to addressing heterogeneity problems. The author found that partitioning applications statically does not provide optimal user experience as more and more applications are used in diverse environments and inputs. So it is decision to demand particular purpose. The decision may be impacted not only by the application itself, but also by the expected workload and the execution conditions, such as network connectivity and CPU speeds of both weak and cloud devices. After formalizing the dynamic partitioning problem, and sketch how to construct a system that supports dynamic partitioning.

### *B. Offloading*

Offloading task from client to cloud can reduce energy consumption of mobile device [13]. The problem is the choice of offloading percent and methods. Should we put all applications to cloud?

Byung-Gon and Petros [14] first introduced offloading execution from the smart phone to a computational infrastructure hosting a cloud of smart phone clones. The idea is simple: clone the entire set of data and applications from the smart-phone onto the cloud and selectively execute some operations on the clones, reintegrating the results back into the smart-phone. There are five types of augmentation, each of which uses special method to offloading. One can have multiple clones for the same smart-phone, clones pretending to be more powerful smart-phones, etc.

Eduardo Cuervo [15] el. presented MAUI, a system that enables fine-grained energy-aware offload of mobile code to the infrastructure. It maximizes the potential for energy savings through fine-grained code offload while minimizing the changes required to applications. First, MAUI uses code portability to create two versions of a smart phone application, one of which runs locally on the smart phone and the other runs remotely in the infrastructure. Second, MAUI uses programming reflection combined with type safety to automatically identify the remote methods and extract only the program state needed by those methods. Third, MAUI profiles each method of an application and uses serialization to determine its network shipping costs

## V. Context-aware Service

It is context that distinguishes mobile cloud computing from common concepts. Context leads to advent of various mobile applications. As mentioned in Section III, contexts can be classified into two types, spatial contexts and social contexts. First we introduce common context management methods, and then discuss

### *A. Common context management*

Andreas Klein [16] el. presented a framework for the use of context information for the Heterogeneous Access Management (HAM) provided by the Mobile Cloud as a service for the mobile terminals. A formal method assessing link quality based on available context information has been developed for triggering handover mechanisms. The proposed Context Management Architecture (CMA) is responsible for acquiring, processing, managing, and delivering context information. Context Quality Enabler (CQE) controls the provision of context information according to the requirements of the Mobile Cloud Controller. Finally, based on the outlined HAM concept, the author presented a context-aware radio network simulator (CORAS) that is able to model context availability, accuracy, and delay, thus enabling an evaluation of the impact of different levels of context relevance, confidence, and quality on simulation results.

Hyun Jung La [17] presented a framework for enabling context-aware mobile services. The framework enables tasks of capturing context, determining what context-specific adaptation is needed, tailoring candidate services for the context, and running the adapted service. The net result of context-aware services is for consumers to receive better services which fit to the current context of the consumers.

Aaron Beach [18] presented a vision of mobile-cloud computing in which context-aware services are organized and integrated by a Context-Aware Intention Compiler (CAIC). Run-time creation of these programs allows contextual information from a mobile phone and the environment to be integrated in real-time. Furthermore, the mobile device can look up context-aware services using a Contextual Lookup Service, which maps context and intention to the appropriate Context-Aware Intention Compiler. Use of the CAwbWeb framework allows mobile-cloud challenges to be divided into four major concerns: specifying intention, describing context, identifying appropriate actions, and efficient actuation of those actions.

### *B. Spatial contexts service*

Patrick Stuedi [19] el. presented WhereStore, a location-based data store for Smartphones interacting with the cloud. The key property of WhereStore is that it uses the phone's location history to determine what data to replicate locally. The main goal of caching cloud data on the phone is to decrease the overall data access latency and also reduce the probability of data becoming unavailable in periods of no connectivity. Furthermore, WhereStore is a shared resource for different applications and exchanges data with the cloud in batches, thus potentially reducing the overall energy consumption on the phone.

Pelin Angin [20] el. proposed a mobile-cloud collaborative approach for context-aware navigation by exploiting the computational power of resources as well as location-specific resources available on the Internet. The author proposes an extensible system architecture that minimizes reliance on infrastructure, thus allowing for wide usability.

### *C. Social contexts service*

Dejan Kovachev [21] el. proposed Mobile Community Cloud Platform (MCCP) as a cloud computing system that can leverage the full potential of mobile community growth. Also, the author analyses the requirements of mobile communities, proposes a cloud computing model for mobile communities, and discusses the technical settings of this cloud infrastructure.

Lan Zhang [22] el. designed and constructed a multi-hop networking system named MoNet based on WiFi, and a privacy-aware geo-social networking service. Also the author designs a distributed content sharing protocol which can significantly shorten the relay path, reduce conflicts and improve data persistence and availability. A role strategy is designed to encourage users to collaborate in the network. Furthermore, a key management and an authorization mechanism are developed to prevent some attacks and protect privacy.

Eric Jung [23] el. proposed to exploit the potential of smart phones in proximity cooperatively, using their resources to reduce the demand on the cellular infrastructure. The author introduces RACE (Resource Aware Collaborative Execution), a Markov Decision Process (MDP) optimization framework that takes user profiles and user preferences to determine the degree of collaboration. Then RACE can enable the use of other mobile devices in the proximity as mobile data relays.

## VI. CONCLUSION

This paper surveys recent research activities on Mobile Cloud Computing. Mobile Cloud Computing aims to utilize cloud computing techniques for storage and processing of data on mobile devices, thereby reducing their limitations. Several problems would challenge the development, including intrinsic nature of mobile device and wireless connection. Then we proposal concept model for Mobile Cloud Computing systems and analyze typical architecture. After that, we discuss the detail of technology, application partition & offloading and context-aware services.

## ACKNOWLEDGMENT

This work is supported by the National Key project of Scientific and Technical Supporting Programs of China (Grant Nos.2008BAH24B04, 2008BAH21B03); the National Natural Science Foundation of China (Grant No.61072060); the Program of the Co-Construction with Beijing Municipal Commission of Education of China.

## REFERENCES

[1] Mobile Phone’s wikipedia. http://en.wikipedia.org/wiki/Mobile_phone

[2] Mobile Cloud Applications. ABI Research Report. http://www.abiresearch.com/research/1003385-Mobile+Cloud+Computing

[3] Chetan S., Gautam Kumar, K. Dinesh, Mathew K., and Abhimanyu M.A. Cloud Computing for Mobile World. 2010; Available from: http://chetan.ueuo.com/projects/CCMW.pdf.

[4] Sheng Xiao and Weibo Gong. Mobility Can Help: Protect User Identity with Dynamic Credential. in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. 2010.

[5] Jason H. Christensen, Using RESTful web-services and cloud computing to create next generation mobile applications, in *Proceeding of the 24th ACM SIGPLAN conference companion on Object oriented programming systems languages and applications*. 2009, ACM: Orlando, Florida, USA. p. 627-634.

[6] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, The Case for VM-Based Cloudlets in Mobile Computing. Pervasive Computing, IEEE, 2009. 8(4):14-23.

[7] Xinwen Zhang, Joshua Schiffman, Simon Gibbs, Anugeetha Kunjithapatham, and Sangoh Jeong, Securing elastic applications on mobile devices for cloud computing, in *Proceedings of the 2009 ACM workshop on Cloud computing security*. 2009, ACM: Chicago, Illinois, USA. p. 127-134

[8] EE Marinelli, Hyrax: Cloud Computing on Mobile Devices using MapReduce. 2009, Carnegie Mellon University.

[9] M. Black and W. Edgar. Exploring mobile devices as Grid resources: Using an x86 virtual machine to run BOINC on an iPhone. in *Grid Computing, 2009 10th IEEE/ACM International Conference on*. 2009.

[10] Ioana Giurgiu, Oriana Riva, Dejan Juric, Ivan Krivulev, and Gustavo Alonso, Calling the cloud: enabling mobile phones as interfaces to cloud applications, in *Proceedings of the 10th ACM/IFIP/USENIX International Conference on Middleware*. 2009, Springer-Verlag New *York, Inc.: Urbanna, Illinois. p. 1-20.*

[11] Luo Xun. From Augmented Reality to Augmented Computing: A Look at Cloud-Mobile Convergence. in *Ubiquitous Virtual Reality, 2009. ISUVR '09. International Symposium on*. 2009

[12] Byung-Gon Chun and Petros Maniatis, Dynamically partitioning applications between weak devices and clouds, in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. 2010, ACM: San Francisco, California. p. 1-5.

[13] K. Kumar and Lu Yung-Hsiang, Cloud Computing for Mobile Users: Can Offloading Computation Save Energy? Computer, 2010. 43(4):51-56

[14] Byung-Gon Chun and Petros Maniatis, Augmented smartphone applications through clone cloud execution, in *Proceedings of the 12th conference on Hot topics in operating systems (HotOS)*. 2009

[15] Eduardo Cuervo, Aruna Balasubramanian, Dae-ki Cho, Alec Wolman, Stefan Saroiu, Ranveer Chandra, and Paramvir Bahl, MAUI: making smartphones last longer with code offload, in *Proceedings of the 8th international conference on Mobile systems, applications, and services*. 2010, ACM: San Francisco, California, USA. p. 49-62.

[16] Andreas Klein, Christian Mannweiler, Joerg Schneider, and Hans D. Schotten. Access Schemes for Mobile Cloud Computing. in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. 2010.

[17] La Hyun Jung and Kim Soo Dong. A Conceptual Framework for Provisioning Context-aware Mobile Cloud Services. in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*. 2010.

[18] Aaron Beach, Mike Gartrell, Richard Han, and Shivakant Mishra. CAwbWeb: Towards a Standardized Programming Framework to Enable a Context-Aware Web. Technical Report CU-CS-1063-10, March 2010.

[19] Patrick Stuedi, Iqbal Mohomed, and Doug Terry, WhereStore: location-based data storage for mobile devices interacting with the cloud, in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. 2010

[20] Pelin Angin, Bharat Bhargava, and Sumi Helal. A Mobile-Cloud Collaborative Traffic Lights Detector for Blind Navigation. in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. 2010.

[21] Dejan Kovachev, Dominik Renzel, Ralf Klamma, and Yiwei Cao. Mobile Community Cloud Computing: Emerges and Evolves. in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. 2010

[22] Lan Zhang, Xuan Ding, Zhiguo Wan, Ming Gu, and Xiang-Yang Li, WiFace: a secure geosocial networking system using WiFi-based multi-hop MANET, in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. 2010

[23] Eric Jung, Yichuan Wang, Iuri Prilepov, Frank Maker, Xin Liu, and Venkatesh Akella, User-profile-driven collaborative bandwidth sharing on mobile phones, in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. 2010