Offloading Mobile Computations To Cloud To Save Mobile Energy

First A. Author, Second B. Author, Jr., and Third C. Author,
Member, IEEE

Bhavatharini M G
M.Tech Student
Software Engineering
Channabasaveshwara Institute of Technology
Gubbi, Tumkur Dist, Karnataka
mgbhavatharini@gmail.com

Veena A
Asst. Professor
Computer Science Dept
Channabasaveshwara Institute of Technology
Gubbi, Tumkur Dist, Karnataka
veena.a@cittumkur.org

Abstract— **Mobile systems have limited resources, such as battery life, network bandwidth, storage capacity, and processor performance. These restrictions overcome by computation offloading. Energy efficiency is a fundamental consideration for mobile devices. Mobile cloud computing(MCC), with its promise to meet the urgent need for richer applications and services of resource-constrained mobile devices, is emerging as a new computing paradigm and has recently attracted significant attention. Mobile Cloud Computing (MCC) is a model for flexible growth of mobile device capabilities via universal wireless access to cloud storage and computing resources. Rather than conducting all computational and data operations locally, MCC takes advantage of the abundant resources in clouds to store and process data for mobile devices.**

This paper provides an overview of mobile cloud computing, computation offloading, offloading decision: performance, energy saving analysis, challenges of MCC, advantages, and applications.

Keywords-mobile cloud computing, computation offloading, energy analysis , performance.

I. INTRODUCTION

Together with an explosive growth of the mobile applications and emerging of cloud computing concept, mobile cloud computing (MCC) has been introduced to be a potential technology for mobile services. The tension between resource-hungry applications and resource-poor mobile devices is considered as one of the driving forces for the evolution of mobile platforms. Due to the limited physical size, mobile devices are inherently resource-constrained [1], equipped with a limited supply of resources in computation, bandwidth, energy and storage. In particular, the energy supply from the limited battery capacity [2] has been one of the most challenging design issues for mobile devices. The limited battery life has been found by market research as the biggest complaint for smart phones [3]. Therefore, resource limitations in the mobile devices should be considered for the design of the mobile applications.

MCC is a model for elastic augmentation of mobile device capabilities via ubiquitous wireless access to cloud storage and computing resources, with context-aware dynamic adaption to changes in the operating environment. Rather than conducting all computational and data operations locally, MCC takes advantage of the existing resources in cloud platforms to gather, store, and process data for mobile devices. Many popular mobile applications have actually employed cloud computing to provide enhanced services. Cloud computing and mobile applications have been the leading technology trends in recent years.

Energy efficiency has always been critical for mobile devices and the importance seems to be increasing. Use cases are developing towards always on-line connectivity, high speed wireless communication, high definition multimedia, and rich user interaction. Development of battery technology has not been able to match the power requirements of the increasing resource demand. The amount of energy that can be stored in a battery is limited and is growing only 5% annually [4]. Bigger batteries resulting into larger devices are not an attractive option. Also thermal considerations limit the power budged of the small devices without active cooling to about three watts [5]. Energy efficiency improvements can also always be traded for other benefits like device size, cost and R&D efficiency. Indeed, large part of the hardware technology benefits have been traded for programmability in mobile phone designs [6].

II. RELATED WORK

To extend the battery life of mobile devices computation offloading is necessary show by experiments that significant power can be saved through remote processing for several realistic tasks (up to 50% of battery life) decision making algorithm that learns and adapts its decision based on previous CPU time measurements [7]. Moreover, some literatures have studied the energy issues of cloud computing. [3] Presents an energy model to analyze whether to offload applications to the cloud, mainly considering computation energy in the mobile device the communication energy for offloading. Demonstrates that workload, data communication patterns and technologies used (i.e., WLAN and 3G) are the main factors that highly affect the energy consumption of mobile applications in cloud computing. This analysis is roughly based on statistical measurements and investigations.

Computation offloading has been the topic of a number of studies. However, only a subset of those studies focus on the effect offloading has on the energy consumption of the mobile device. In most cases the focus is on response time and other resource consumption. Large

part of the research uses modeling and simulation, like [8], which is an early investigation of offloading work from mobile to a fixed host concluding that under certain conditions 20% energy savings would be possible.

Compiler technology has been studied in, e.g., [10], where a program is partitioned to client and server parts. The client parts are run on a mobile device and the server part is offloaded. The main metrics evaluated are execution speed and energy consumption. Even though the measurements show that significant energy savings are possible, the outcome is shown to be sensitive to program inputs.

Middleware based approach has been studied in, e.g., [6]. The described framework performs resource accounting and uses execution time, energy usage and application fidelity as criteria for deciding between local, remote and hybrid execution.

Compared to these previous efforts, this paper provides the framework of mobile execution and cloud execution in order to conserve energy consumption.

III. OVERVIEW OF MOBILE CLOUD COMPUTING

In this section, we present computation energy model for the mobile execution and a transmission energy model for the cloud execution. First, we overview the mobile cloud computing. Following that, we describe the architecture of mobile cloud computing and service oriented cloud computing architecture

A. Mobile cloud computing (MCC)

The term 'mobile cloud computing' was introduced after the concept of 'cloud computing'. The attentions of entrepreneurs as a profitable business option that reduces the development and running cost of mobile applications, of mobile users as a new technology to achieve rich experience of a variety of mobile services at low cost, and of researchers as a promising solution for green IT [3].

Mobile cloud computing is an infrastructure where both the data storage and data processing happen outside of the mobile device. Mobile cloud applications move the computing power and data storage away from mobile phones and into the cloud, bringing applications and MC to not just smart phone users but a much broader range of mobile subscribers.

MCC as a new paradigm for mobile applications whereby the data processing and storage are moved from the mobile device to powerful and centralized computing platforms located in clouds. These centralized applications are then accessed over the wireless connection based on a thin native client or web browser on the mobile devices [5]. MCC combines both cloud and mobile applications this is most popular for mobile users to access applications and services on the Internet.

Briefly, MCC provides mobile users with the data processing and storage services in clouds. The mobile devices do not need a powerful configuration (e.g., CPU speed and memory capacity) because all the complicated computing modules can be processed in the clouds.

Virtualization technology is studied in [11]. Making it possible to run multiple operating systems and multiple applications on the same computer (or a set of computers) simultaneously, increasing Utilization and flexibility [3]. Since different types of virtual machines can be created, users can scale the number of virtual machines based on demand. Due to virtualization, these machines have separation and protection. Cloud computing uses virtualization to offer computing as a service based on their requirements.

B. Architectures of mobile cloud computing

The general architecture of MCC can be shown in Figure 1. In Figure 1, mobile devices are connected to the mobile networks via base stations (e.g., base transceiver station, access point, or satellite) that establish and control the connections (air links) and functional interfaces between the networks and mobile devices. Mobile users' requests and information (e.g., ID and location) are transmitted to the central processors that are connected to servers providing mobile network services.

Here, mobile network operators can provide services to mobile users as authentication, authorization, and accounting based on the home agent and subscribers' data stored in databases. After that, the subscribers' requests are delivered to a cloud through the Internet. In the cloud, cloud controllers process the requests to provide mobile users with the corresponding cloud services. These services are developed with the concepts of utility computing, virtualization, and service-oriented architecture (e.g., web, application, and database servers).

C. Service-Oriented Cloud Computing Architecture

The details of cloud architecture could be different in different contexts. In this paper, we focus on a layered architecture of CC (Figure 2). This architecture is commonly used to demonstrate the effectiveness of the CC model in terms of meeting the user's requirements [9].

• Data centers layer provides the hardware facility and infrastructure for clouds. In data center layer, a number of servers are linked with high-speed networks to provide services for customers. Typically, data centers are built in less populated places, with high power supply stability and a low risk of disaster.

• Infrastructure as a Service (IaaS) is built on top of the data center layer. IaaS enables the provision of storage, hardware, servers, and networking components. The client typically pays on a per-use basis.The examples of IaaS are Amazon Elastic Cloud Computing and Simple Storage Service (S3).
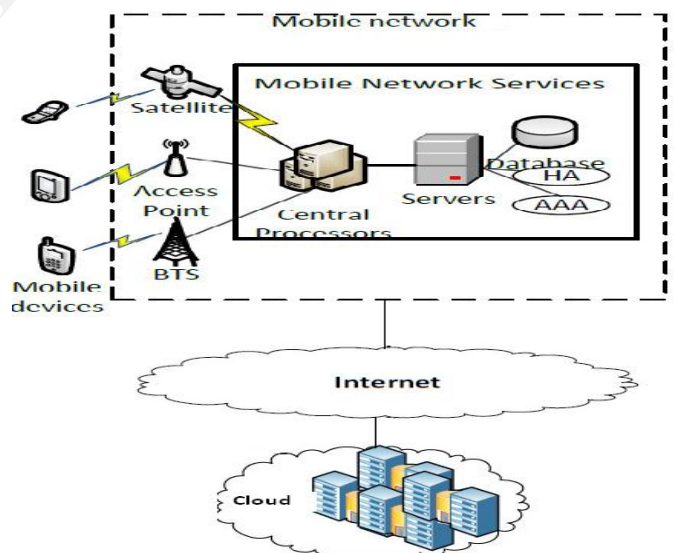

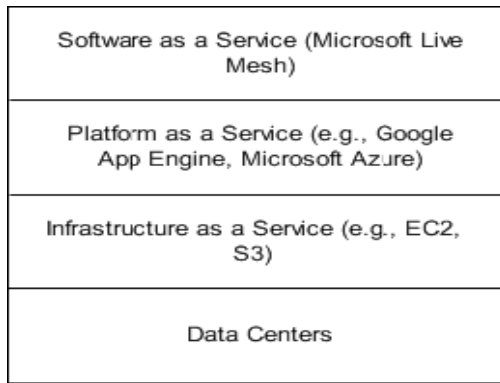Figure 1. Mobile cloud computing architecture

Figure 2. Service-Oriented Cloud Computing Architecture

- Platform as a Service (PaaS) offers an advanced integrated environment for building, testing, and deploying custom applications. The examples of PaaS are Google App Engine, Microsoft Azure, and Amazon Map Reduce/Simple Storage Service.

- Software as a Service (SaaS) supports a software distribution with specific requirements. The users can access an application and information remotely via the Internet and pay only for that they use. Salesforce is one of the pioneers in providing this service model. Microsoft's Live Mesh also allows sharing files and folders across multiple devices simultaneously.

Although the CC architecture can be divided into four layers as shown in Figure 2, it does not mean that the top layer must be built on the layer directly below it. For example, the SaaS application can be deployed directly on IaaS, instead of PaaS. Also, some services can be considered as a part of more than one layer. For example, data storage service can be viewed as either in IaaS or PaaS. Given this architectural model, the users can use the services flexibly and efficiently.

IV. SYSTEM MODELING AND PROBLEM FORMULATION

In this section, we describe computation offloading, offloading decision in that performances and energy saving analysis from mobile cloud computing.

A. Computation offloading

To overcome resource constraints on mobile devices, a general idea is to offload parts of resource-intensive tasks to the cloud. Since execution in the cloud is considerably faster than that on mobile devices, it is worth shipping code and data to the cloud and back to prolong the battery life and speed up the application.
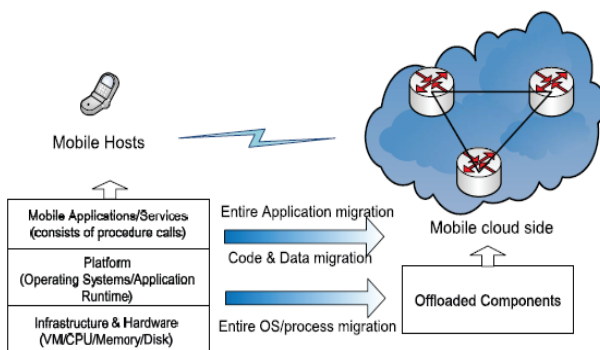

Figure 3. Computation offloading schemes in mobile wireless environment

Figure 3 shows general computation offloading schemes in mobile wireless environments. From the application layer perspective, the entire application or part of the code/data can be offloaded to the mobile cloud side to achieve better energy utility or better performance. The context and runtime of the application process is still maintained by operating system in mobile host. In the meantime, the whole execution platform (application runtime or infrastructure) can also be migrated to the cloud side, which indicates that the burden on programmers can be reduced.

B. Offloading decision

Offloading migrates computation to a more resourceful Computer (i.e., servers). It involves making a decision regarding whether and what computation to migrate. The body of research exists on offloading decisions for (1) improving performance and (2) energy saving analysis

1) Improving performance

The condition for offloading to improve performance can be shown below. Without loss of generality, we can divide a program into two parts: one part that must run on the mobile system and the other part may be offloaded. The first part may include user interface and the code that handles (such as the mobile system's camera). Let s_m be the speed of the mobile system. Suppose w is the amount of computation for the second part. The time to execute the second part on the mobile system is

$$w/ s_m \qquad (1)$$

If the second part is offloaded to a server, sending the input data d_i takes d_i /B seconds at bandwidth B. Here we ignore the initial setup time for the network. The program itself may also need to be sent to the server. We assume the size of the program is negligible, or the server may download the program from another site through a high-speed network [12]. Offloading can improve performance when execution, including computation and communication, can be performed faster at the server. Let s_s be the speed of the server. The time to offload and execute the second part is

$$d_i /B + w/ s_s \qquad (2)$$

Offloading improves performance when Eq. 1 > Eq. 2

$$w/ s_m > d_i /B + w/ s_s \Rightarrow w (1/ s_m +1/ s_s) > d_i /B \qquad (3)$$

This holds for
- large w: the program requires heavy computation.
- large s_s the server is fast.
- small a d_i: small amount of data is exchanged.
- large B: the bandwidth is high.

This inequality shows limited effects of the server's speed. If $w/ s_m < d_i /B$ even if the server is infinitely fast (i.e., $s_s \rightarrow \infty$), offloading cannot improve performance. Only tasks that require heavy computation (large w) with light data exchange (small di) should be considered. This requires analyzing programs to identify such tasks. Moreover, if we define $w(1/ s_m -1/ s_s)- d_i /B$ as the performance gain of offloading, the server's speed has diminishing return doubling s_s will not double the gain.

2) Energy saving Analysis

Energy is a primary constraint for mobile systems. A survey of 7,000 users across 15 countries showed that "75% of respondents said better battery life is the main feature they want" [17, 18]. Smartphones

are no longer used only for voice communication; instead, they are used for acquiring and watching videos, gaming, web surfing, and many other purposes. As a result, these systems will likely consume more power and shorten the battery life. Even though battery technology has been steadily improving, it has not been able to keep up with the rapid growth of power consumption of these mobile systems. Offloading may extend battery life by migrating the energy-intensive parts of the computation to servers [3].

The following analysis explains the conditions when offloading saves energy. Suppose p_m is the power on the mobile system and w is the amount of computation. Let s_m be the speed of the mobile system.

$$p_m \times w/s_m \qquad (4)$$

Let p_c be the power required to send data from the mobile system over the network. After sending the data, the system needs to poll the network interface while waiting for the result of the offloaded computation. During this time, the power consumption is p_i. Let s_s be the speed of the server. If task is offloaded to a cloud, sending the input data d_i takes d_i/B seconds at bandwidth B.

$$(p_c \times d_i/B) + (p_i \times w/s_s) \qquad (5)$$

Offloading saves energy when Eq. 4 > Eq. 5.

$$(p_m \times w/s_m) > [(p_c \times d_i/B) + (p_i \times w/s_s)]$$

To make offloading save energy, heavy computation (large w) and light communication (small d_i) should be considered. In equation, we assume that data must be transmitted from the mobile system to the server. Fortunately, this may not be true in many cases. For example, the data (such as photographs and videos) may also reside in servers with high-speed networks (such as Facebook.com and YouTube.com). Instead of transmitting the data from the mobile system to the server, the mobile system needs to provide links to the server and the server may download the data directly from the hosting sites. In this case, the bandwidth B can be substantially higher, allowing offloading to improve performance and save energy [13, 14].

V. CHALLENGES AND ADVANTAGES OF MCC

In this section, we describe challenges, advantages of mobile cloud computing and applications of mobile cloud computing.

A. Challenges of MCC

While cloud computing has tremendous potential to save energy, designers must consider several issues including privacy and security, reliability, handling real-time data and authentication.

(1) Privacy and security. In cloud computing, Web applications and data replace traditional stand-alone programs, which are no longer stored in users' computers. Shifting all data and computing resources to the cloud can have implications for privacy and security. Because the data is stored and managed in the cloud, security and privacy settings depend on the IT management the cloud provides.

(2) Communication. Another potential concern with mobile cloud computing is communication between cloud and mobile device [15] A mobile user performing computation in the cloud depends on the wireless network and cloud service. Dependence on the wireless network implies that cloud computing may not even be possible, let alone energy efficient, when connectivity is limited. This is typical in regions like national parks; a user may thus not be able to organize, retrieve, or identify any images captured in the park. Mobile cloud computing is also difficult in locations such as the basement of a building, interior of a tunnel, or subway.

(3) Real-time data. Some applications—including chess, searching newly captured images for content-based image retrieval, mobile surveillance, and context-aware navigation has real-time data. If the value of data to transfer to cloud is large, offloading may not save energy. In such cases, performing the computation on the mobile system may be more energy efficient. A possible solution is partitioning computation between the mobile system and the cloud to reduce energy consumption. Such a solution may include partially processing the real-time data on the mobile system. If the processed data are smaller in size, sending the processed data to the server reduces the wireless transmission energy.

(4) Authentication. In this method using CC to secure the data access suitable for mobile environments. The system requires input constraints that make it difficult for mobile users to use complex passwords. As a result, this often leads to the use of simple and short passwords or personal identification numbers (PINs).

B. Advantages of MCC

Cloud computing is known to be a promising solution for mobile devices because of many reasons (e.g., mobility, communication, and portability [16]. In the following, we describe how the cloud can be used to overcome obstacles in MC, thereby pointing out advantages of MCC.

(1) Extending battery lifetime. Battery is one of the main concerns for mobile devices. Several solutions have been proposed to enhance the CPU performance [19, 20] and to manage the disk and screen in an intelligent manner [21, 22] to reduce power consumption. However, these solutions require changes in the structure of mobile devices, or they require a new hardware that results in an increase of cost and may not be feasible for all mobile devices. Computation offloading technique is proposed with the objective to migrate the large computations and complex processing from resource-limited devices (i.e., mobile devices) to resourceful machines (i.e., servers in clouds). This avoids taking a long application execution time on mobile devices which results in large amount of power consumption.

(2) Improving data reliability. Storing data or running applications on clouds is an effective way to improve the reliability because the data and application are stored and backed up on a number of computers. This reduces the chance of data and application lost on the mobile devices. In addition, MCC can be designed as a comprehensive data security model for both service providers and users

(3) Improving data storage capacity and processing power. Storage capacity is also a constraint for mobile devices.MCC is developed to enable mobile users to store/access the large data on the cloud through wireless networks.

In addition, MCC also inherits some advantages of clouds for mobile services as follows:

- Dynamic provisioning. Dynamic on-demand provisioning of resources on a fine-grained, self-service basis is a flexible way for service providers and mobile users to run their applications without advanced reservation of resources.

- Scalability. The deployment of mobile applications can be performed and scaled to meet the unpredictable user demands due to flexible resource provisioning. Service providers can easily add and expand an application and service without or with little constraint on the resource usage.

- Multitenancy. Service providers (e.g., network operator and data center owner) can share the resources and costs to support a variety of applications and large number of users.

- Ease of integration. Multiple services from different service providers can be integrated easily through the cloud and Internet to meet the user demand.

C. *Applications of MCC*

Applications used for evaluation include multimedia [23, 24] text editors [27, 28] vision and recognition [25, 26] and gaming [29, 30] Text editors transfer relatively small amounts of data—i.e., text, and perform computations like spell check. Multimedia, vision, and recognition applications transfer large amounts of data in the form of images and videos. If all the multimedia is on the mobile system, the offloading decision depends primarily on bandwidth if most of the data is already on a server (such as YouTube), then offloading can be more beneficial. Gaming applications like chess are interesting candidates for offloading, because the amount of computation for the program to win the game depends on the skill level of the user thus computation *w* depends on how the user plays the game. The applications run on a wide range of clients with different computational capabilities, ranging from PDAs, laptops, and robots.

VI. CONCLUSION AND FUTURE ENHANCEMENT

In this paper we proposed a theoretical framework to Improving performance and save mobile energy by offloading computations to cloud. Our analysis suggests that cloud computing can potentially save energy for mobile users. However, not all applications are energy efficient when migrated to the cloud. Mobile cloud computing services would be significantly different from cloud services for desktops because they must offer energy savings. The services should consider the energy overhead for privacy, security, reliability, and data communication before offloading.

This paper only focuses on the theoretical parts of application offloading. For future work, we will evaluate the performance by running real applications with intensive computation and show the applicability of our theoretical framework. In addition, we will consider the multi-task offloading in a more fine-grained choice for mobile application execution.

REFERENCES

[1] M. Satyanarayanan, "Fundamental challenges in mobile computing," in *Proc. 1996 ACM Symp. Principles of Distributed Computing*, pp. 1–7.

[2] A. P. Miettinen and J. K. Nurminen, "Energy efficiency of mobile clients in cloud computing," in *Proc. 2010 USENIX Conference on Hot Topics in Cloud Computing*.

[3] K. Kumar and Y. H. Lu, "Cloud computing for mobile users: can offloading computation save energy?" *IEEE Computer*, vol. 43, no. 4, pp. 51–56, 2010.

[4] Robinson, S. Cellphone Energy Gap: Desperately Seeking Solutions. Tech. rep., 2009. Strategy Analytics.

[5] Neuvo, Y. Cellular phones as embedded systems. Digest of Technical Papers, IEEE Solid-State Circuits Conference (ISSCC) (2004), 32–37.

[6] Flinn, J., Narayanan, D., AND Satyanarayanan, M. Self-tuned remote execution for pervasive computing. In HOTOS '01: Proceedings of the Eighth Workshop on Hot Topics in Operating Systems (Washington, DC, USA, 2001), IEEE Computer Society, p. 61.

[7] Othman, M., AND Hailes, S. Power conservation strategy for mobile computers using load sharing. SIGMOBILE Mob. Comput. Commun. Rev. 2, 1 (1998), 44–51

[8] Kumar K, Lu YH (2010) Cloud computing for mobile users:can offloading computation save energy? IEEE Comput 43(4):51–56

[9]Tsai W, Sun X, Balasooriya J. Service-oriented cloud computing architecture, In Proceedings of the 7th International Conference on Information Technology: New Generations (ITNG), 2010; 684–689.

[10] Wang, C., AND Li, Z. A computation offloading scheme on handheld devices. J. Parallel Distrib. Comput. 64, 6 (2004), 740–746.

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

[12] Cuervo E, Balasubramanian A, Cho D, Wolman A, Saroiu S, Chandra R, Bahl P (2010) MAUI: making smartphones last longer with code offload. In: International conference on mobile systems, applications, and services, pp 49–62

[13] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," *IEEE Trans. Wireless Commun.*, vol. 8,no. 4, pp. 1866–1875, 2009.

[14] M. Zafer and E. Modiano, "Minimum energy transmission over a wireless fading channel with packet deadlines," in *Proc. 2007 IEEE Conference on Decision and Control*, pp. 1148–1155

[15] M. Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," tech. report UCB/EECS-2009-28, EECS Dept., Univ. of California, Berkeley, 2009.

[16] Forman GH, Zahorjan J. The Challenges of mobile computing. IEEE Computer Society Magazine 1994; 27(4): 38–47.

[17] Battery Life Concerns Mobile Users (2005). http://www cnn.com. Accessed 23 Sept 2005

[18] Mobile Phone Users Demand Decent Batteries (2005). http:// www.eetimes.com. Accessed 22 Sept 2005

[19] Kakerow R. Low power design methodologies for mobile communication, In Proceedings of IEEE International Conference on Computer Design: VLSI in Computers and Processors, 2003;

[20] Paulson LD. Low-power chips for high-powered handhelds. IEEE Computer Society Magazine 2003; 36(1):

[21] Davis JW. Power benchmark strategy for systems employing power management, In Proceedings of the IEEE Internationa Symposium on Electronics and the Environment, 2002; 117.

[22] Mayo RN, Ranganathan P. Energy consumption in mobil devices: why future systems need requirements aware energy scale-down, In Proceedings of the Workshop on Power-Aware Computing Systems, 2003.

[23] Chen G, Kang B-T, Kandemir M, Vijaykrishnan N, Irwin MJ, Chandramouli R (2004) Studying energy trade offsin offloading computation/compilation in java-enabled mobile devices. IEEE Trans Parallel Distrib Syst 15(9):795–809

[24] Hong YJ, Kumar K, Lu YH (2009) Energy efficient contentbased image retrieval for mobile systems. In: International symposium on circuits and systems, pp 1673–1676

[25] Balan RK (2006) Simplifying cyber foraging. PhD thesis, School of Computer Science, Carnegie Mellon University

[26] Kremer U, Hicks J, Rehg J (2003) A compilation framework for power and energy management on mobile computers. In: Proceedings of the 14th international conference on Language and compilers for parallel computing. Cumberland Falls, KY, USA, pp 115–131

[27] Gu X, Nahrstedt K, Messer A, Greenberg I, Milojicic D (2003) Adaptive offloading inference for delivering applications in pervasive computing environments. In: IEEE international conference on pervasive computing and communications, pp 107–114

[28] Huerta-Canepa G, Lee D (2008) An adaptable application offloading scheme based on application behavior. In: International conference on advanced information networking and applications - workshops, pp 387–392

[29] Chu H, Song H, Wong C, Kurakake S, Katagiri M (2004) Roam, a seamless application framework. J Syst Softw 69(3):209–226

[30] Perkins CE (1996) Handling multimedia data for mobile computers. In: Computer software and applications conference, pp 147–14