© 2017 IJSRSET | Volume 3 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Multi-Resource Allocation for Cloudlet-Based cloud Computing

S. K. Manigandan, S Arputhaprasanth

Department of MCA, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi, Chennai, Tamil Nadu, India

ABSTRACT

Mobile cloud computing utilizing cloudlet is a develop technology to improve the quality of mobile services. In this paper, to improve overcome the main bottlenecks of the computation capability of cloudlet and the wireless capacity between mobile devices and cloudlet, we consider the multi-wealth allocation problem for the cloudlet environment with resource-intensive and latency sensitive mobile applications. The proposed multi-resource allocation planning increase the quality of mobile cloud service, in terms of the system throughput (the number of admitted mobile applications) and the service dormancy. We formulate the resource allocation copy as a semi-Markov decision process under the average cost criterion, and solve the optimization problem using linear programming technology. From reproduction result, it is marked that the system adaptively adjusts the allocation policy about how much resource to allocate and whether to utilize the distant cloudlet. Mobile Cloud Computing is an evolving technology that integrates the notation of cloud computing into the mobile environment. Smart phones are boon in the world of technology but they have certain protocols (e.g. battery life, network bandwidth, storage, energy) when working complex applications which require large computations. Using Cloud Computing in mobile phones, these rules can be addressed. Certain frameworks have been proposed past the years that can address the issues in cloudlet.
Keywords: Cloud Computing, Mobile Cloud Computing, MCC, Wide Area Networks, Virtual Machines, SMD, Quality of Service, MATLAB, API

I. INTRODUCTION

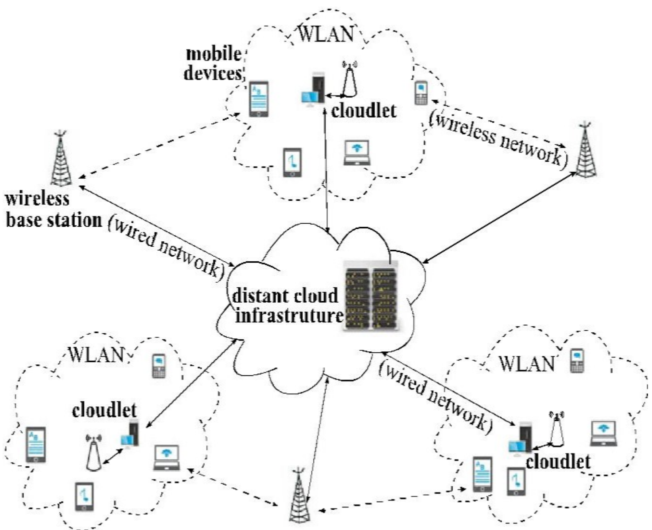
MOBILE Cloud Computing (MCC) is a gifted system that introduces powerful cloud computing into a mobile computing environment, where mobile devices connect to the Internet through wireless network and then communicate with the distant cloud.

Cloud computing (CC) provides a wide variety of computing resources from servers and storage to enterprise applications. Cloud computing is a i receive environment that is immediate, flexible, scalable, secure and available. The computing resources from cloud can be easily and quickly accessed and released after Use with very less management effort. Compared to mobile devices, the cloud server in MCC can provide huge storage, high computation power, as well as strong security. By send subcomponents of mobile application to the cloud slave for execution, the performance of mobile applications can be greatly improved and the energy consumption of mobile devices can be much reduced. Furthermore, MCC extends the array of mobile applications that are too resource-intensive to execute solely on mobile devices, such as virus scanning. However, for latency-sensitive mobile applications, such as augmented matter with real-time constraints, offloading to the remote cloud is insufficient, because of the high latency of Wide Area Networks (WAN). The requirement of real-time interactive can be face by one-hop high-bandwidth wireless access to the cloudlet. Based on the mobile user's request, one or more rule virtual machines (VMs) can be instantiate immediately on capabilities. Through providing a resource rich server/cluster with it hence arrange of the mobile users, the requirement of real-time interactive response can be face by one-hop high-bandwidth wireless access to the cloudlet. Based on the mobile user's request, one or more rule virtual machines (VMs) can be instantiated immediately on the cloudlet for remote execution of applications in a thin client fashion. A cloudlet is usually set up at the public place, like shopping center, theater, office building, and assembly room to set up convenient access for mobile devices. Compared to the conventional MCC, the efficiency of the cloudlet system is greatly improved by providing the capable computing resource closer through one-hop wireless network. Whereas in reality, the computing resource of the cloudlet server cannot be treated as ample as the remote cloud cluster, and the wireless bandwidth that attack mobile mechanism and the cloudlet server is limited and raced. There is a high chance that the cloudlet runs out of resources so that no new mobile request can be accept, when an excessive number of mobile buyer offloading their applications for execution at the cloudlet .specially, when the offloaded applications are mainly resource-intensive and latency-sensitive, such as collective high-definition video gaming, the resource of computing capability and wireless bandwidth will become weak rapidly and the quality of service (QoS) will be seriously degraded. Therefore, the coordinated allocation of computing resource and wireless bandwidth is a cutting issue, in order to improve the quality of experience from mobile users' point of view. In contrast to traditional network resource allocation problems, new challenges are come into the cloudlet-based MCC system. One is how to utilize the. The other is how to perform multi-resource allocation jointly considering both the wireless bandwidth and the figure resource. Computation offloading allows SMDs to become more capable. In contrast to traditional client-server architecture, where clients always empty the computation to the slave and are completely dependent on it, the computational offloading migrates programs to servers which are outside of the user's computing environment. The term "surrogate computing" or "virtual foraging" is also used for computation offloading. Before offloading the computation, various parameters such as battery life, current memory capacity of SMD, chain bandwidth and latency, computation execution time, In this paper, we propose a joint multi-resource allocation framework in the cloudlet system based on semi-Markov Decision Processing (SMDP). The objective of the framework is to obtain the optimal decision of computing and wireless bandwidth resource allocation among multiple mobile users in cloudlet-based mobile cloud computing environments, by maximizing the overall benefits of the whole system to consequently enhance the quality of service for mobile customer (i.e., low service veto probability and short service time). The work of this paper is outline as follows: We develop the system reward model for resource allocation with wireless bandwidth, and computing talent of both cloudlet and distant cloud. The accolade model considers the system benefits or impacts in accepting or rejecting the new request of using resource according to the current system traffic, the availability of the system resources, and the Quality of Service guarantee of mobile users. Based on the reward model, we propose a multi resource allocation planning, which can complete whether to get a new mobile service request for execution at the cloudlet or the distant cloud. Furthermore, the strategy can adaptively determine the optimal load of wireless bandwidth and computing resource to allocate to the accepted request, and thus attain the optimal system benefits. We formulate the multi-resource allocation problem as a semi-Markov decision processing, which is fix as a linear programming problem using the solver tool. Our approach has the divining ability of the future state. The predictive feature in our approach lies in the development probabilities from the modern state to all potential next states upon receiving a new request. In order to verify the efficiency of our proposed multi-resource allocation in the cloudlet based MCC system, we meet the simulations of two different eager polices and compare their blocking probabilities and average service time with our proposed algorithm's. We also examine the shock of various reward parameters in our model. Our extensive performance simulations show that the proposed resource allocation mechanism provides a minor request rejection relative compared to those of eager policies. At the same time, a short time latency of mobile service is also guaranteed. The proposed multi-resource allocation algorithm can be utilized in practice by being executed offline given the various traffic parameters, the amount of system resource, and the resource price in the reward model according to the importance of the related resources.

II. METHODS AND MATERIAL

1. Related Work

One of the main branches of research on the efficiency of mobile cloud computing focuses on application offloading. The mobile application is partitioned into multiple tasks/components, some of which are apt loading executions. The work in proposed mobile cloud computing model. What makes this model different is the scalability feature. It can be expanded to have the design number of cloudlets in the covered area. The real implementation results of this model showed a great efficiency and reasonable skill consumption. It is known that the mobile devices are power keen specially when using them to run excessive applications. Motivated by this fact that optimizing power is very important in mobile cloud system; the researchers in the offloaded computing tasks need to be chosen fully according to the characteristics of the mobile application, the computing ability of the cloud, and the network condition. In general, the complex tasks that require above computing capability are offloaded to the remote cloud for execution while the less complex ones running on mobile devices . We also proposed a dynamic programming based partitioning and offloading scheme that can obtain the optimal partitioning much more quickly than traditional methods used in. Another important factor that determines the efficiency of mobile cloud computing is resource management, which is recently investigated in from different aspects. In, several types of virtual machines are configured at the distant cloud to provide the service to different mobile users give to their requirements and the availabilities of computing wealth, presents how to manage the cloud resources across multiple cloud domains to support continuous cloud service. Then beginning with captures the dynamic advent and departures of resource requests for decision making of their resource allocation.



However, these works only address the computing resource allocation of the far cloud without the consideration of the cellular bandwidth resource, which is an indispensible factor in mobile computing environments. The allotment of computing and radio (bandwidth) resources are jointly considered in within the scheme of wireless base stations and cloud service providers, instead of the cloudlet. Liang et al. propose a cooperation scheme among different service providers and obtains the maximum number of applications that can run under the shared computing and base station bandwidth resources. However, the bandwidth resources investigated in are from the wireless base station, not the high-speed WLANs connecting the mobile mechanism and the cloudlet server. Cloudlet is proposed as a practical platform for accelerating mobile cloud computing. Preprocessing, caching and scheduling approaches are planned by for efficient usage of the potent and resourceful cloudlet. A bandwidth aware admission control policy is developed in for cloudlet; where a mobile service is always allowing a fixed amount of system resource without the flexibility of alter to request traffic and resource availability of the time-varying system. Furthermore, the ability of the distant cloud connected with the Internet is not taken into account in their studies. Xia et al. address the problem of online request admission in a cloudlet with an objective of maximizing the system throughput for a certain time period, without the knowledge of future request advent rate. In each time slot, requests are admitted/rejected one by one using a selfish strategy, according to the resources availability and admission cost which need to be updated after admitting a request. The computing capability of the far cloud is not considered in the models of and. What they studied are actually admission control problem, and they cannot adaptively adjust the allocated wireless radio band and computing resource of the cloudlet and the remote cloud for mobile service requests according to the availability of the current system resource. On the other hand, boost of networked sensors inspired the researchers to use them to collect different type of data in different useful aspects of life including health, military, crowd management, and in quick cities applications [14]. The big amount of collected data needs to be stored effectively, so it will be transferred to the cloud servers for storage and further processing.

2. System Description

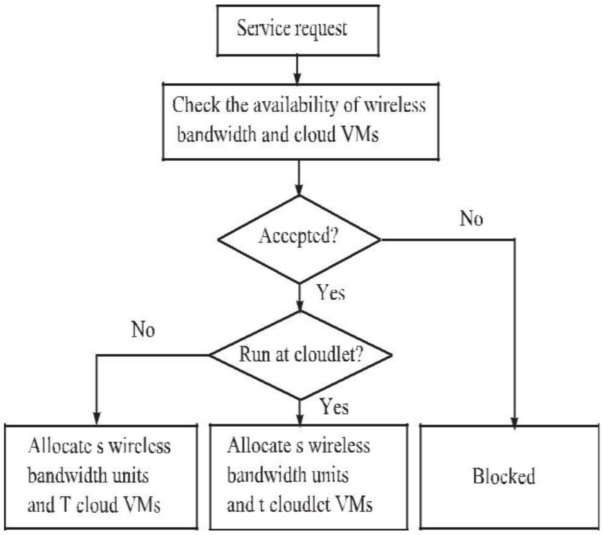
We consider the cloudlet-based MCC environments as 1.The cloudlet provides wireless LAN (WLAN) connections for mobile devices within its working range, and the far cloud infrastructure is connected with cloudlets through the high-speed connected network. The mobile device can run mobile applications locally, or offload some workload to the cloudlet or to the distant cloud for hot execution. In our model, compared to the conventional way of accessing to the distant cloud through wireless base site (e.g., 4G or LTE), the mobile devices connect to the far cloud through cloudlets as long as WLAN is possible between the mobile device and cloudlet, which enables much faster data transmission already being, or users are out of the cloudlet range.. For the task component of a mobile application that is determined to be offloaded, a light-weight virtual machine will and allow to it at either the cloudlet server or the remote cloud data center (distant cloud). In the mean while, the execution completion of the application can be advance up if more bandwidth or more VMs are allocated for data transmission or cloud computation, respectively. The computing resource of the cloudlet is adequate for running multiple mobile applications simultaneously, but not as sufficient as that of the remote cloud data center, such as Amazon, Google Compute Engine ,and Microsoft Azure , where the computing resource is always available as long as the mobile users purchase/subscribe the service. Thus, the number of possible VMs at the cloudlet server is usually much less than that at the distant cloud data center, and there will be cases that the tasks offload to the cloudlet need to be further offloaded to the distant cloud data center via wired network because no more possible VMs available at the cloudlet. The wireless bandwidth resource in the model refers to the WLAN connections between the mobile devices and the cloudlet, and one wireless bandwidth unit assign to the minimum bandwidth required to support mobile computing offloading, for example, 50, 100 Kbps, etc. Then, the total bandwidth available can be expressed as everyone multiple of the bandwidth unit. For the manageability of the model computation, we assume a single mobile service requires at least one basic crew of WLAN bandwidth or channel, and only the integral numbers of basic bandwidth units are allocated for wireless resource. Note that, a longtime running service or continuously alive service can be separated into multiple tasks/computing modules in our model. The tasks requiring low data transmission rate and high computing capability will be offloaded to the cloudlet, and each such task is considered as a mobile request. In this way, the power of mobile device can be saved and the application execution time shortened. For each task/computing module, the wireless bandwidth to transmit data and them computing resource to process data will not undergo obvious change over the relatively short period of task execution time. The minimum resource requirement of each task/ request is thus relatively fixed. In this model, the number of VMs that can be supported at the cloudlet is denoted as M, and the number of VMs at the distant cloud data center is assumed to be infinite compared with the small scale of the cloudlet. The number of wireless bandwidth units provided by the cloudlet is denoted as B. The minimum requirement for cloud computing is to be one cloud VM and one wireless bandwidth unit. The notations used in this paper are summarized.

3. Traffic Model

The arrival of mobile request for offloading application med to follow a Poisson process with mean rate. If a mobile request is accepted by the cloudlet, then the service departure from the cloudlet is assumed to follow exponential distribution with amount , 1 and the mean service time at the cloudlet for this mobile request is, where stands for the allocated number of wireless bandwidth units and the number of assigned VMs at the cloudlet server. Here, we use the maximal number of bandwidth crew the system provides to one mobile request, while T is the maximal number of VMs grant for one mobile request. On the other hand, if the system determines to offload the task to the distant cloud data inside, T VMs will be assigned to fully support the fast execution of the mobile application. Then, the service escape from the distant cloud is assumed to follow exponential distribution with rate, and the mean service time at the distant cloud is, where i is the number of allocated wireless bandwidth units. Note that, we assume the far cloud can allocate the maximal number (T) of VMs that the cloudlet can provide to one mobile service, since data inside at the distant cloud usually owns much more available server machines as long as the service is paid by mobile users.

4. Problem Statement

The decision creating procedure of multi-resource allocation is delineate in. once a brand new request arrives, the system determines whether or not to just accept it or not per the current request traffic and also the utilization of the wireless bandwidth and computing resources at the cloudlet. If the request is suitable, the system can assign this new service request to the cloudlet or the distant cloud with an explicit number of wireless resource and VMs of the cloudlet/ distant cloud. The target of our multi-resource allocation system for MCC is to create associate best call regarding whether to just accept the mobile service request, and wherever to run the mobile application with what proportion allotted wireless bandwidth and computing resources if the request is accepted so as to maximize the system advantages and to guarantee the QoS of mobile users.



5. State and Action

In the cloudlet-based MCC system, the total number of ongoing services covering wireless bandwidth crew and j cloudlet VMs is show as, while the total number of ongoing services occupying wireless bandwidth units and the demanded computing resource at the distant cloud is denoted. The sum of the system wireless bandwidth units and cloudlet VMs being used by all the ongoing services should be balance to or less than the total bandwidth (B) and cloudlet VMs (M) the system can provide, respectively. Stand for a escape event of a service occupying i bandwidth units and j cloudlet VMs, and a departure event of a service occupying i bandwidth crew and T distant cloud VMs, respectively. An event in state s is defined. The decision process is on a state space S, where each state s (s 2 S) describes the numbers of ongoing services occupying various resources and the current event in the system. Represents an action to accept the request by allocating wireless bandwidth unit sand j cloudlet VMs, and represents acceptance of the request by allocating i wireless bandwidth units and T distant cloud VMs. Note T is the maximum number of VMs allowed to allot to one mobile service request. The cumulative event rate is the sum of rates of all constituent processes from state s to others after selecting action a and the expected sojourn time is the average time duration from the current state s to others after selecting action a. For each possible combination of event and selected action can be calculated.

6. State Transition Probability

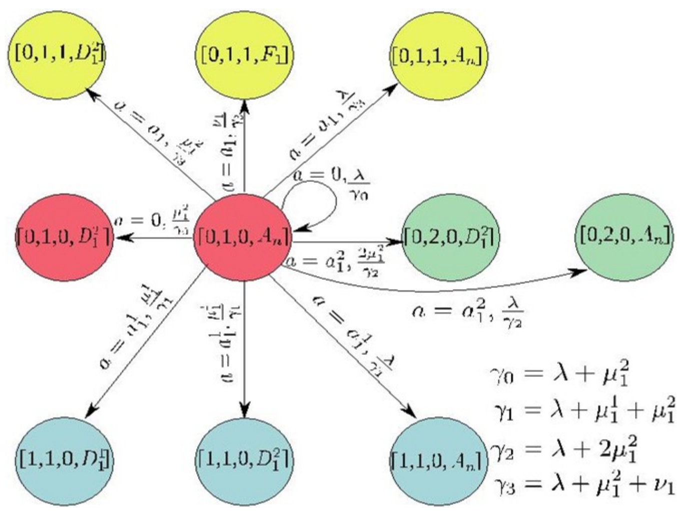
The state transition probability is defined as the probability that the system will be in state k at the next decision epoch, if action is chosen at the current state s. There are three cases to consider depending on the event type an (new request arrival), service departure from cloudlet), and (service departure from distant cloud) at the current state s, respectively. For the current state with event an, the candidate action can be to reject the request or to allocate a certain number of system resources. According to the selected action and the next state, the transition probability to the next state.

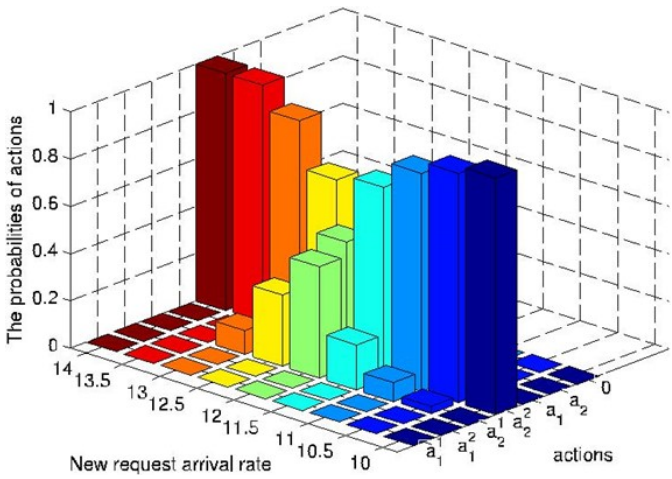
System Reward In order to find the optimal resource allocation policy that maximizes the MCC system benefits; we define a real valued function as the system reward for select a state. Following the definition of system reward in, can be calculated as the sum of the lump income of decision making and the continuous cost of resource usage in our model the lump reward portion and is the system cost per time unit for selecting action at states. In the definitions of and we consider the information including the mobile service offer traffic, the usage wireless and cloudlet computing resource, and the significance of accepting/rejecting one single request. The detailed definitions of are described next.

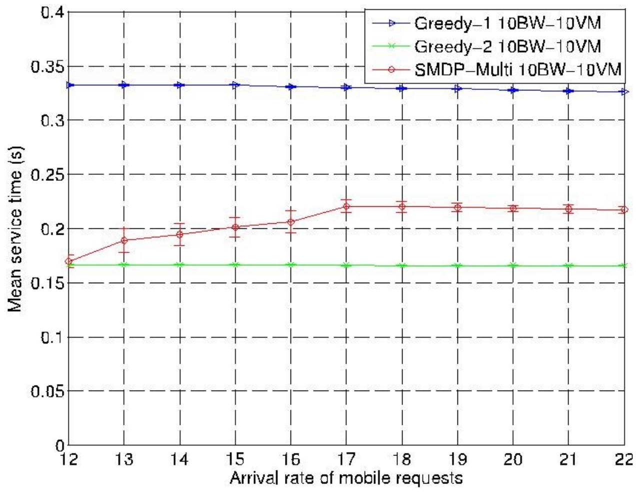
III. RESULTS AND DISCUSSION

1. Performance Evaluation

In this section, we evaluate the efficiency of the proposed SMDP-based multi-resource allocation strategy by investigating the reproduction results of the obtained optimal policy(i.e., set of action selecting decisions), the blocking probability, the mean service time, and the system reward. The performance is evaluated under various mobile service request traffics, wireless network conditions and system model parameters. The advantages of the proposed multi-resource allocation algorithm are clearly revealed by comparing with two different kinds of greedy admission control methods. The simulations are written using MATLAB, in which a free API of lp_solve_5.5.2.0 is embedded for solving the linear programming problem. The system parameter assumptions in our reproduction are: the maximum number of VMs that the cloudlet can provide to one mobile service request is two, and the number of wireless bandwidth units that the system can assign to one mobile service request is up to two. Hence, the state s in our model is in the form of the default values of other parameters used in the system model are listed in Table 2. Here, the system reward framework (Ea, Err, and Ct) values having direct impact on the system reward are selected with the purpose of mimicking the ratio values that could be used in the real system. And if the weights of the framework are considered differently by the system, the values should be set accordingly. For example, if the system considers the event of rejecting a mobile service request having much more native impact, the value of Err could be the upper bound of blocking probability of the system, if the system does not consider the blocking probability that seriously, the optimization constraint can be relaxed by setting a higher value of. The default request arrival rate is set to a relative high value of12.0. From our simulation we found, when is low (e.g., less than 6.0), the blocking probability approaches 0 (since the resource is too sufficient to refuse any service request). The heavy service request traffic helps to investigate and verify the efficiency of our approach under a simulation environment with insufficient resource. As the criteria to evaluate the QoS of the proposed resource allocation algorithm, the service request blocking probability (blocking) and the mean service time are calculated as below to show their simulation values: where all the statistical data is the average of ten times simulation results and each simulation lasts 10,000 seconds.







IV. CONCLUSION AND FUTURE WORK

The ability to produce cloud service is vital for contemporary mobile cloud computer system. During this paper, we tend to present a unique multi-resource allocation approach for cloudlet-based Mobile Cloud computer system. in step with the present traffic of mobile requests and also the handiness of the system resource, our formula permits the system to assign AN best quantity of allotted wireless information measure, cloudlet computing resource, and distant cloud computing resource for cloud computing of mobile applications. Because the service request from mobile user will increase, the approach ensures the prime quality of cloud service isn't affected abundant by with efficiency victimization each native cloudlet and distant cloud resource. Moreover, we tend to conjointly show the way to alter the designed reward model for service latency sensitive application to satisfy numerous system necessities. Regarding the service request from mobile users, we tend to victimization the Poisson method as several others adopted. we tend to will still study more resource allocation problem in MCC once real traffic knowledge is offered, as an example, a correct mathematical model or knowledge primarily based} like bar chart based model. Another future work is to analyze the admission management approach with the thought of user quality. During a physically larger venue with multi-cloudlet setup, mobile service transfer for inter domain are going to be a significant issue. The way to assign the resourof a cluster of cloudlets considering the football play caused by the user mobility is a stimulating topic.

V. REFERENCES

[1]. H. Rim, S. Kim, Y. Kim, and H. Han, "Transparent method off-loading for slim execution," in Proc. Int. Symp. Wireless Pervasive Compute. Jan. 2006, pp. 1–6.

[2]. Y. Liu and M. J. Lee, "Security-aware resource allocation for mobile cloud computing systems," in Proc. IEEE Int. Conf. Compute. Common. New. Aug. 2015, pp. 1–8.

[3]. Y. Liu and M. J. Lee, "An adaptive resource allocation algorithm for partitioned services in mobile cloud computing," in Proc. IEEE Sump. Service-Oriented Syst. Eng., Mar. 2015, pp. 209–215.

[4]. H. Wu, Q. Wang, and K. Welter, "Tradeoff between performance improvement and energy saving in mobile cloud offloading sys-teems," in Proc. IEEE Int. Conf. Common., Jun. 2013, pp. 728–732.

[5]. K. Yang, S. Our, and H. H. Chen, "On effective offloading services for Resource-constrained mobile devices running heavier mobile Internet applications," IEEE Common. Mag., vol. 46, no. 1, pp. 56– 63, Jan. 2008.

[6]. C. Xian, Y. Lu, and Z. Li, "Adaptive computation offloading for energy conservation on battery-powered systems," in Proc. Int. Conf. Parallel Diatribe. Syst., 2007, pp. 1–8.

[7]. T. Verboten, P. Simons, F. D. Truck, and B. Doted, "Cloudlets: Bringing the cloud to the mobile user," in Proc. 3rd ACM Workshop Mobile Cloud Compute. Services, 2012, pp. 29–36.