A Design of Energy-efficient Resource Sharing Overlay Network in Mobile Cloud Computing

Wei Liu
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: liu@cube.kuee.kyoto-u.ac.jp

Ryoichi Shinkuma
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: shinkuma@i.kyoto-u.ac.jp

Tatsuro Takahashi
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: ttakahashi@i.kyoto-u.ac.jp

***Abstract*—Mobile Cloud Computing (MCC) is aimed at integrating mobile devices with cloud computing. Recently, along with the increase of resources in mobile devices, researchers have begun to take into consideration resource sharing among mobile devices themselves. How to establish an architecture of resource sharing is one of the most important issues need to be solved to achieve this goal. This paper presents an energy efficient method that establishes resource sharing overlay networks in wireless mobile networks. Theoretical models and a heuristic algorithm are presented. Simulation results demonstrated the effectiveness of the proposed models and heuristic algorithm.**

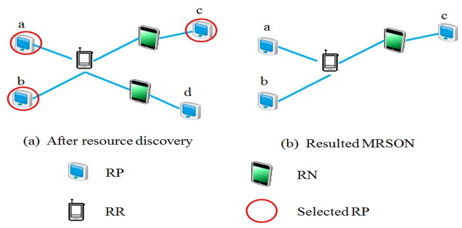


Fig. 1: System architecture

I. INTRODUCTION

Mobile cloud computing (MCC) is aimed at integrating mobile devices with cloud computing. It can be roughly divided into two different architectures: client-server based and cooperation based. In the traditional client-server based architecture, the cloud (data center) provides overall resource management for mobile devices. Along with the development of hardware and software technologies, modern mobile devices have many more resources than previously. Consequently, cooperation based architecture of MCC takes mobile devices as part of the cloud. Available resources in mobile devices are shared among themselves through wireless communications. How to select resource providers (RPs) in dynamic mobile environments is still an open issue need to be solved to achieve cooperation based MCC.

There is much research work on selecting RPs in the wired networks mainly in the field of web service [1]–[4]. [5] is the only work which considered about selecting a RP in mobile networks. However, it focused on a single RP while neglecting the possibility of selecting multiple RPs. We argue that three properties that were neglected by conventional work should be considered due to the unique characteristics of mobile wireless networks: (1) Resources from multiple devices need to be gatherd to facilitate task processing because resources in mobile devices are still limited compared with traditional devices. (2) Physical network conditions should be integrated, e.g., stable routes are prefered. (3) Energy efficiency is important since battery capacity can not cope with modern mobile applications. To satisfy previous requirements, this paper presents a design of energy-efficient mobile resource sharing overlay networks (MRSON, defined in Section.II) that select appropriate RPs to implement resource sharing in MCC.

In the rest of this paper, Section 2 defines the concept of the MRSON. Theoretical models and formal problem definitions are presented in Section 3. The proposed heuristic algorithm is introduced in Section 4. Section 5 presents simulation results. The paper concludes in Section 6.

II. MOBILE RESOUCE SHARING OVERLAY NETWORK

To utilize additional resources, a resource requester (RR) have to discover resources in other nodes at first. After that, the RR is aware of all the mobile nodes that have idle resources that are required (candidates of RPs) and routes to access these nodes. The multi-hop routes are also composed of mobile nodes in the mobile networks without any infrastructure. These nodes are called relay nodes (RN) in this paper. Then, the RR select a group of RPs among all candidates to utilize their resources. As a result, the RR, selected RPs and RNs along the routes from RR to RPs constitute a resource sharing overlay network. We call it mobile resource sharing overlay network (MRSON) because all nodes in it are mobile. Two characteristics of the MRSON need to be noticed: (1) An MRSON is created on-demand. (2) After the generated task finishes, the RR releases the MRSON to make resources in RPs available for other nodes. Fig. 1 outlines this scenario.

In this paper, we mainly focus on the sharing of computational and communication resources among mobile devices. The RR has to keep connected with RPs in the MRSON to utilize their resources (e.g., sends/receives data packages from RPs or coodinates and synchronizes computation process among different RPs). However, due to mobility of nodes, routes composed of RNs are dynamic. When a route becomes unavailable, the RR has to discover a new route to access the lost RP. Energy is consumed to discover a new route. It should be noted that this kind of energy consumed by maintaining the structure of an MRSON does not contribute anything to task processing. As a result, it is important to minimize this kind of extra energy consumption while preserving benefits from resource sharing when constructing an MRSON.

III. Modelling of the energy-efficient MRSON

In this section, theoretical models of an energy-efficient MRSON are analyzed. First, the concept of availability that represents the dynamics of an MRSON is defined. Then, task latency that measures the time consumed to process a task is explained. Finally, the problem of constructing an energy-efficient MRSON is formalized.

A. Availability of MRSON

The availability of a link is highly related to the mobility of its end nodes in mobile networks. If one node i moves out of the transmission range of its neighboring node j, the link between nodes i and j becomes unavailable. Availability of links expresses availability of routes. Consequently, the availability of an MRSON depends on the availabilities of links and routes in it.

1) Availability of links: The definition of availability of a link in this paper is the same as it was in [5]. We introduce it briefly here for completeness and convinience. N_i and N_j are two end nodes of a link. The transmission range of a node is R. Every node moves randomly and its moving range is a circle with a radius r. d is the distance between the two nodes. We assume that transmission range R is known and every node knows its location coordinates (e.g., through GPS). As a result, the distance d can be calculated by the Euclidean distance formula $d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, where (x_i, y_i) and (x_j, y_j) are the coordinates of node N_i and N_j. Finally, the radius r of moving range for a node is its moving speed v multiplied by moving period t, $r = v \times t$. The speed v of a mobile node can be calculated based on its moving distance during a period from t_1 to t_2, $v = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}/(t_2 - t_1)$. Based on previous definitions, the availability of the link composed of two nodes can be calculated by:

$$A_l = \frac{\beta R^2 + \alpha r^2 - (R^2 \sin\beta\cos\beta) + r^2 \sin\alpha\cos\alpha)}{\pi \times v^2 \times t^2} \quad (1)$$

where $\alpha = \arccos(\frac{r^2+d^2-R^2}{2\times r\times d})$ and $\beta = \arccos(\frac{R^2+d^2-r^2}{2\times R\times d})$. Proofs of previous results are available in [5].

2) Availability of routes: A route is available if and only if all the links in it are available. Therefore, the availability of a route A_r is the multiplications of availabilities of its links, $A_r = \prod_{i=1}^{k} A_{l-k}$.

3) Availability of MRSON: The MRSON generated by an RR is composed of routes to every selected RPs. As a result, the availability of an MRSON A_o is the multiplications of availabilities of its routes, $A_o = \prod_{i=1}^{q} A_{r-q}$.

B. Task latency

As presented in the last subsection, the period t in maintaining an MRSON is a key variable in its availability. We define the duration from when an MRSON for a task is constructed to when it is released as task latency since it illustrates the effectiveness of task processing from a user's perspective. Task latency includes three parts:

1) Time for task processing : Intuitively, the time of processing a task decreases if the utilized resources increases. We simplified the relationship into the following equation:

$$T_{proc} = \frac{S}{R} \quad (2)$$

for computational and communication resources, where S is the size of the task and R is the amount of resources.

2) Transmission delay of MRSON: To utilize idle resources in RPs, an RR distributes subtasks to different RPs at the beginning. After subtasks finish, RPs return results to the RR. Consequently, a period of round trip time (RTT) for each route is consumed. The transmission delay of an MRSON is equal to the maximum RTT value among its routes. We use D_{MRSON} to represent it.

3) Delay of maintaining routes : Task latency is also influenced by the time spent on re-discover a route from the RR to the lost RP when it became unaccessible. The calculation of this variable uses history-based method, namely, using average value of the past statistics. We use D_{route} to represent it.

C. Problem definition of energy-efficient MRSON

We assume K potential candidates of RPs are discovered after resource discovery phase and the i-th candidate have r_i units of resources that are required. We define x_i as an indicator (0 or 1) to show whether the i-th candidate is selected by the RR to be part of the MRSON. As a result, every possible choice of an MRSON is represented by a K dimentional vector X constituted by x_i. Then, the task latency for each choice of an MRSON is:

$$T_{latency} = \frac{S}{\sum_{i=1}^{K} x_i \times r_i} + D_{MRSON} + D_{route} \quad (3)$$

Substituting Eq. (3) into Eq. (1), the availability of every possible choice of an MRSON can be calculated. Therefore, the problem of consitituting an energy-efficient MRSON is defined as: constitue an MRSON with maximum availability from all possible choices with a constraint that selected MRSON have enough resources to finish the generated task within a time threshold T_{thresh}.

Objective: Maximize availability of the constituted MRSON.
Constraints: $T_{latency} \leq T_{thresh}$, $x_i = 0$ or 1 .

This is a non-linear 0-1 programming problem which is NP-hard. A greedy heuristic algorithm is introduced to solve it in the next section.

IV. Proposed heuristic algorithm

As presented in the previous section, if K potential candidates were discovered in the resource discovery phase, 2^K different combinations of RPs exists. It is not feasible to compare all choices when K is large. A greedy heuristic algorithm is proposed in this section to solve this problem.

In the algorithm, set C contains all discovered candidates of RP. The RR maintains a set S that includes selected RPs in the MRSON. It is initialized by an empty set. The availability of current MRSON S is presented by A_{o-S}. First, to satisfy

TABLE I: Parameters for simulations

Simulation area	1000 × 1000 m
Size of task	600
Number of SNs	10 nodes, 60 units of resources
Number of CNs	40 nodes, 10 units of resources
Number of RNs	50 nodes, no resources
Task latency	12 s
Energy of transmission	1 unit per package
Energy of reception	0.5 unit per package
Transmission range	250 m
Interval of task generation	5 s
Mobility model	Random way point
Simulation period	1000 s

the constraint of task latency, the RR continues to select the i-th candidate that makes A_{o-S} be largest among possible candidates in C. If the constraint of task latency can not be satisfied even if C became empty, the task fails due to a lack of resources. After the task latency constraint is satisfied, the RR continues to select candidates from C if and only if A_{o-S} increases. At last, the MRSON denoted by S is used to process the generated task.

V. Simulations

Simulation results are presented in this section. Nodes were distributed in a rectangular area. They were divided into three categories according to the amounts of their resources: Super nodes (SN), Common nodes (CN) and Relay nodes (RN). Key parameters used in the simulation are summarized in Table I.

The performce of the proposed heuristic algorithm was compared with other two methods:

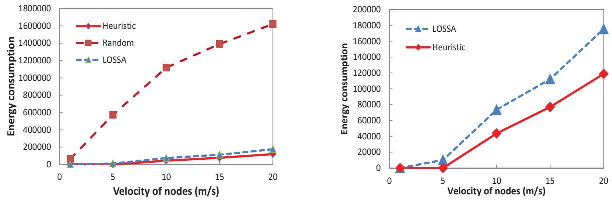
(1) Random: the RR selected RPs randomly until the contraint of task latency was satisfied.

(2) LOSSA: the RR selected a single RP with best availability among all candidates that had enough resources to satisfy the constraint of task latency. It was first proposed in [5] [1].

The energy consumed to maintain an MRSON under different velocities is plotted in Fig. 2. As Fig. 2(a) indicates, the random method consumed much more energy than the other two methods. This is because the selection of RPs in random method did not take into consideration the availablity of the resulting MRSON. As Fig. 2(b) indicates, the proposed heuristic method consumed less energy than the LOSSA method. This is because the proposed method utilized available resources in multiple RPs to reduce the lifetime of the resulting MRSON.

This is also proved by Fig. 3 that plots the task latency of different methods. The task latency of the proposed heuristic method was much less than the other two methods especially when the velocities of nodes was low. When the velocities of nodes was low, the availablity of an MRSON remained high even if more RPs were selected. Therefore, utilizing resources in these RPs accelerated the processing of task while

[1]It was a weighted average value in [5]. We eliminated ‘price of service’ and ’reliability of service’ by assign zero to their weights because these two properties are beyond the scope of this paper.



(a) (b)

Fig. 2: Energy consumed by diffent methods

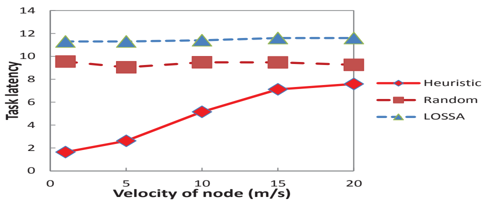


Fig. 3: Task latency of different methods

preventing from consuming extra energy.

VI. Conclusions

A design of energy-efficient mobile resource sharing overlay networks (MRSON) in MCC was proposed in this paper. The simulation results proved that the proposed model and method were efficient on both energy consumption and time. We plan to measure the gap between the proposed heuristic algorithm and the optimal bound and consider other factors that influence the choice of RPs (e.g., price of resources) in future.

Acknowledgment

This work was supported in part by the National Institute of Information and Communications Technology (NICT), Japan, under the new generation network R&D program for innovative network virtualization platform and its application.

References

[1] Liangzhao Zeng, Boualem Benatallah, Anne H. H. Ngu, Marlon Dumas, Jayant Kalagnanam, and Henry Chang, “Qos-aware middleware for web services composition,” *Software Engineering, IEEE Transactions on*, vol. 30, no. 5, pp. 311–327, 2004.

[2] E. Michael Maximilien and Munindar P Singh, “A framework and ontology for dynamic web services selection,” *Internet Computing, IEEE*, vol. 8, no. 5, pp. 84–93, 2004.

[3] Danilo Ardagna and Barbara Pernici, “Adaptive service composition in flexible processes,” *Software Engineering, IEEE Transactions on*, vol. 33, no. 6, pp. 369–384, 2007.

[4] Xiaohui Gu, Klara Nahrstedt, and Bin Yu, “Spidernet: An integrated peer-to-peer service composition framework,” in *High performance Distributed Computing, 2004. Proceedings. 13th IEEE International Symposium on*. IEEE, 2004, pp. 110–119.

[5] Kun Yang, Alex Galis, and Hsiao-Hwa Chen, “Qos-aware service selection algorithms for pervasive service composition in mobile wireless environments,” *Mobile Networks and Applications*, vol. 15, no. 4, pp. 488–501, 2010.