Future Generation Computer Systems 64 (2016) 1–14

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Multi-device task offloading with time-constraints for energy efficiency in mobile cloud computing

Kaiyang Liu, Jun Peng *, Heng Li, Xiaoyong Zhang, Weirong Liu

School of Information Science and Engineering, Central South University, Changsha 410083, China

HIGHLIGHTS

- A task offloading decision method is proposed among multi-devices for energy saving.
- The problem is formalized as a 0–1 nonlinear integer programming problem.
- An iterative decoupling algorithm that combines with decision-variable relaxation and convex optimization is proposed for near-optimal decisions.

ARTICLE INFO

Article history:
Received 30 September 2015
Received in revised form
23 March 2016
Accepted 16 April 2016
Available online 20 May 2016

Keywords:
Mobile cloud computing
Energy efficiency
Time constraints
Convex optimization
Task offloading

ABSTRACT

Nowadays, in order to deal with the increasingly complex applications on mobile devices, mobile cloud offloading techniques have been studied extensively to meet the ever-increasing energy requirements. In this study, an offloading decision method is investigated to minimize the energy consumption of mobile device with an acceptable time delay and communication quality. In general, mobile devices can execute a sequence of tasks in parallel. In the proposed offloading decision method, only parts of the tasks are offloaded for task characteristics to save the energy of multi-devices. The issue of the offloading decision is formulated as an NP-hard 0–1 nonlinear integer programming problem with time deadline and transmission error rate constraints. Through decision-variable relaxation from the integer to the real domain, this problem can be transformed as a continuous convex optimization. Based on Lagrange duality and the Karush–Kuhn–Tucker condition, a solution with coupled terms is derived to determine the priority of tasks for offloading. Then, an iterative decoupling algorithm with high efficiency is proposed to obtain near-optimal offloading decisions for energy saving. Simulation results demonstrate that considerable energy can be saved via the proposed method in various mobile cloud scenarios.

1. Introduction

Nowadays, we have witnessed an explosive growth of mobile devices. According to the Cisco Visual Networking Index report, the number of mobile devices will exceed the number of people on earth by the end of 2015 [1]. Furthermore, numerous mobile applications have been emerging, changing the lifestyles of people and bringing us considerable convenience. Nevertheless, mobile devices have their inherent problems, such as finite computing power, low connectivity, and particularly limited battery life [2]. In fact, many applications operated on mobile devices are resource-intensive, such as navigation, high-definition image processing, face recognition, online mobile games, and sensor data processing [3–5]. In recent years, mobile cloud computing technologies have been envisioned as promising and challenging technologies to overcome these limitations [6].

As a newly emerging computing paradigm, mobile cloud computing brings a new idea to augment the capabilities of mobile devices by offloading computing tasks to resource-rich (e.g., CPU, storage, and network bandwidth) cloud servers [7]. Nowadays, various mobile cloud platforms have been deployed for energy-hungry applications. Reducing the energy consumption of mobile devices can lead to various benefits such as long battery life, no overheating, and increased system reliability. For instance, a mobile cloud service can improve the energy efficiency of wireless body sensors in e-Health systems [8].

In all recent mobile cloud schemes, the cloud offloading mechanism is treated as one of the most powerful and indispensable techniques that can potentially save energy for mobile users [9]. However, the current research efforts for the offloading mechanism are still limited and have defects that cannot be ignored [10]. The foremost challenge is that cloud offloading with multiple

* Corresponding author.
E-mail address: peng.pengj@gmail.com (J. Peng).

http://dx.doi.org/10.1016/j.future.2016.04.013
0167-739X/

mobile devices may cause considerable mutual communication interference [11], which will cause a reduction of the wireless communication quality. It is worth noting that an inferior wireless communication quality may potentially increase the running latency of the cloud offloading process [12]. Previous researches have shown that the existing cloud offloading systems, such as Gaikai, can cause communication time delays of up to 400 ms [13]. Moreover, a relatively long running latency and inferior wireless communication quality may result in increased energy consumption by mobile devices. In conclusion, if a considerably large number of mobile devices are offloading tasks to the cloud, with a relatively long time delay and inferior communication quality, cloud offloading will lose the advantage of energy conservation for mobile devices.

Hence, several important issues arise: Can all the above mentioned limitations be overcome to optimize the energy efficiency of multiple mobile devices? Can mobile devices obtain a cloud offloading service with a relatively low time latency and acceptable wireless communication quality? In general, the task offloading mechanism is leveraged to ease these issues [14]. In fact, most of the available mobile devices support multitasking. Thus, a mobile device may process parallel tasks simultaneously [15]. Recent advances in cloud offloading show that not all mobile applications are suitable to be offloaded for their different task characteristics, i.e., computation-intensive or communication-intensive [16]. The computational workload densities of computation-intensive tasks are relatively higher comparing with communication-intensive tasks. If more computation-intensive applications, e.g., mobile augmented reality [17], automatic target recognition [18] and 3D video games [19], are offloaded, more energy can be saved with cloud offloading service. Thus, to optimize the energy efficiency of mobile devices, a novel task offloading decision method that determines which tasks can be chosen for offloading is investigated in this study. The major contributions of this paper can be summarized as follows:

- To optimize the energy efficiency of mobile devices, we go deep into the task level to investigate the issue of energy conservation. The issue of the task offloading decision on a mobile device is formulated as a constrained 0–1 nonlinear integer programming problem, under the time delay and transmission error rate constraints.
- Because of its NP-hard complexity, we simplify this issue by relaxing assignment variables 0–1 from the integer to the real domain, transforming it as a standard convex optimization problem. Then, an equation with coupled terms is derived to determine the priorities of tasks for offloading by using the Lagrange duality method and the Karush–Kuhn–Tucker (KKT) techniques.
- By taking the worst case as the initial state, we have proposed an iterative decoupling algorithm with high efficiency in order to approximate the optimal offloading decisions for each mobile device. The proposed iterative algorithm combining linear relaxation and the convex optimization method can solve the NP-hard problem at a considerably lower cost, revealing near-optimal offloading decisions to minimize the energy consumption of mobile devices. The computation process is completed by the powerful cloud side, providing task offloading guidance for energy saving.

The rest of this paper is organized as follows: Section 2 briefly discusses the related work. Section 3 presents the preliminaries, defines the mobile cloud computing model, and introduces the task offloading flow path and the wireless fading channel. Then, Section 4 provides the problem formulation for the task offloading decision. Section 5 discusses efficient algorithms for solving the constrained optimization problem. Section 6 presents numerical simulations for various mobile cloud scenarios. Finally, Section 7 presents the conclusion and the future research direction.

2. Related work

2.1. Mobile cloud offloading

Recently, cloud offloading techniques have attracted considerable attention as promising and powerful techniques to augment the capabilities of mobile devices and improve user experience [20]. These works mainly focus on three aspects: (i) minimizing the energy consumption of mobile devices [9,21], (ii) minimizing the communication cost to the cloud [22,23], and (iii) minimizing the total application execution time [11,24,25].

As a pioneer research work, Kumar et al. [9] provided a tradeoff analysis between the energy consumption of a mobile device and the energy consumed by sending the input data to the cloud. However, the assumption in it may not meet the reality that simply lets a single value reflect the computing capacity for all applications given that each application has its own computational requirements. Zhang et al. [21] investigated a threshold offloading method to decide whether an entire application should be offloaded or locally executed on a mobile device, aiming at the reduction of the mobile device's energy consumption. However, this work only considers single-device offloading, ignoring the mutual interference caused by multiple mobile devices.

Furthermore, Barbera et al. [22] considered the bandwidth and the energy cost of mobile computation offloading, investigating the feasibility of both mobile cloud offloading and mobile data backup systems in a real setting. In vehicular cyber-physical systems, Wang et al. [23] modeled mobile data traffic offloading as a multi-objective optimization problem for the simultaneous minimization of mobile data traffic and QoS-aware service provision.

More recently, considering the overhead of cloud offloading in terms of both processing time and energy, Chen et al. [11] proposed a game-theoretic approach to solve the decentralized offloading decision making problem among multiple mobile devices. To save the energy of mobile devices and meet the requirement of application execution time, Huang et al. [24] proposed an effective dynamic offloading algorithm based on Lyapunov optimization, which provides a near-optimal solution with low complexity. Zhang et al. [25] investigated the collaborative task execution problem in the linear topology of a task model and formulated the offloading decision as a time deadline-constrained shortest path problem. An effective algorithm based on Lagrangian relaxation was proposed to minimize the energy consumption approximately.

Inspired by previous work and considering the mutual interference caused by multiple devices, in this study, we investigate the task offloading decision problem to minimize a mobile device's energy consumption with a time delay constraint and acceptable communication quality. The offloading priorities of all tasks are derived so that more computation-intensive but not communication-intensive tasks are offloaded to the cloud side for the minimization of the energy consumption of a mobile device.

2.2. Mathematical programming method

In the field of mobile cloud computing, mathematical programming methods have been adopted widely for the optimization of the offloading performance [26]. For instance, MAUI [27] allows code offloading to a resource-rich computing infrastructure for energy efficiency. The offloading decision issue is formulated as a 0–1 integer linear programming problem. However, the complexity of the solving process makes it unsuitable for practical application. Al-Kanj et al. [28] formulated the optimal cellular offloading problem as a mixed integer linear programming problem in device-to-device communication networks. To adapt the mobility of mobile devices, a dynamic programming approach is adopted for solving