Accepted Manuscript

DMPO: Dynamic mobility-aware partial offloading in mobile edge computing

Fangxiaoqi Yu, Haopeng Chen, Jinqing Xu

PII: S0167-739X(17)32981-3
DOI: https://doi.org/10.1016/j.future.2018.07.032
Reference: FUTURE 4350

To appear in: *Future Generation Computer Systems*

Received date : 28 December 2017
Revised date : 20 March 2018
Accepted date : 17 July 2018

Please cite this article as: F. Yu, H. Chen, J. Xu, DMPO: Dynamic mobility-aware partial offloading in mobile edge computing, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.07.032

DMPO: Dynamic Mobility-Aware Partial Offloading in Mobile Edge Computing

Fangxiaoqi Yu[a], Haopeng Chen[a], Jinqing Xu[a]

[a]*School of software, Shanghai Jiao Tong University, Shanghai, China*

Abstract

Mobile Edge Computing (MEC) brings computation closer to user equipment (UE) and reduces the latency as well as energy consumption for computation offloading. However, with the development of Ultra-Dense Network (UDN), traditional offloading algorithms and mobility management (MM) approaches are not sufficient anymore with a frequent handover between base stations. It becomes far from optimal to execute offloading algorithms only at the beginning of a task when UE moves frequently or in the environment of UDN. Also, simple dynamic adjustment for offloading proportion of data input during UE's movement cannot solve this problem very well, because this approach executes global decision with information from only one position. Exploiting short-term mobility prediction in MM, we propose a novel dynamic mobility-aware partial offloading (DMPO) algorithm to figure out the amount of data for offloading dynamically, together with the decision of communication path in MM, minimizing the energy consumption while satisfying the delay constraint. This proposed algorithm predicts the time to next handover as well as the moving of UE in this phase and assigns data size to each time slot in this phase to achieve our goal of minimizing energy consumption while satisfying the delay constraint. In the simulation, we obverse the average delay and energy consumption with different delay constraint, moving speed of UE and base station density. The results demonstrate that our proposed algorithm saves up to 70% of energy while having a better performance in satisfying the delay constraint compared to traditional approaches.

Keywords:
Mobile Edge Computing, Mobility Management, dynamic offloading, task size assignment

1. Introduction

With the popularity and development of smartphones, many kinds of mobile phone applications emerge. Although the ability of smartphone is increasing, it is still hard to meet the requirements of some kinds of applications which need huge processing in a limited time like object recognition, because user equipment (UE) has limited computation power and battery capacity. In order to deal with this problem, a strategy to offload the calculation to the cloud is raised, which is called Mobile Cloud Computing (MCC)[1].

However, offloading tasks to cloud experiences high latency because of the allocation of backhaul resources and cannot fit the requirements of real-time applications, such as augmented reality (AR), video games, etc. Some researchers propose to deploy computational resources at the edge of the network [2] and offload tasks closer to users such as regional server, macro base stations (eNB) or small cell base stations (SCeNB). Computational resources can be maintained either by mobile network operators accessing through a base station, which is called mobile edge computing (MEC) [2, 3], or by any casual user at a fog server accessing through a wireless access point (AP), which is called fog computing [2, 4]. The concept fog of things allows different devices easy to work together by improving the communication protocol [5].

With several edge servers to choose from, user equipment has to decide the place for offloading and the parts of the task to offload. This decision can be made to minimize energy consumption, minimize task latency, minimize energy consumption while satisfying execution delay constraint [6] or minimize task latency with the limitation of battery capacity [7]. In addition, because of the mobility of user equipment, mobility management (MM) is also a part of the MEC architecture. This part includes the decision of whether to change serving base station (handover) and whether to change

Email addresses: yufxq@sjtu.edu.cn (Fangxiaoqi Yu), chen-hp@sjtu.edu.cn (Haopeng Chen), jinqingxu@outlook.com (Jinqing Xu)