A Non-convex ADMM Approach for Data Offloading in Mobile Edge Computing

Tri Nguyen Dang, Choong Seon Hong

Department of Computer Science and Engineering, Kyung Hee University, 446-701,Korea

Email: {trind, cshong}@khu.ac.kr

Abstract

The objective of this research article is to address the problem of data offloading in Mobile Edge Computing(MEC). Currently, the Internet of Things (IoT) allowed various type of devices connected and transmitted data through the core network. The more data traffic will be increasing the chance congestion occurs, reduce quality of services, and increasing delay for the end user. In this paper, we first proposed data offloading model, and formulated as an optimization problem. We then apply non-convex Alternating Direction Method of Multipliers (ADMM) mechanism to solving the problem. The numerical results have shown that the effectiveness of our scheme in terms of convergence speed and achieving the optimal solution.

1. INTRODUCTION

The rapid development of IoT devices with fully connected to the internet has brought huge pressure on the core network [1]. Lots of data transmitted through the core will be made the network congestion and reduced quality of services [2]. A number of research has focus more in this challenging [3], [4], [5]. Each of those study has some disadvantage and advantage, respectively. One of the best solution is bring server near the end user that can improved performance of the core network, significantly. This solution named as Mobile Edge Computing that deployed a strong enough server at BS side. The communication between end users and cloud server will be temporary replaced by the end user and MEC servers. With the collaborating from MEC, Cloud server may reduce a lot of data traffic in the core, also improve its performance, and service qualities. The MEC is a promising technic on the future internet and the Internet of Things. Because, the heaviest data and consumed more bandwidth in the core is media files. Instead of transfer that kind of data and storage in the Cloud, it can be storage temporary at the MEC with some period and upload to the Cloud at free time.

2. NETWORK MODEL

In this section, we consider a system model of data offloading in MEC. In the Figure 1, there are many users requested to using the storage services from MEC servers.

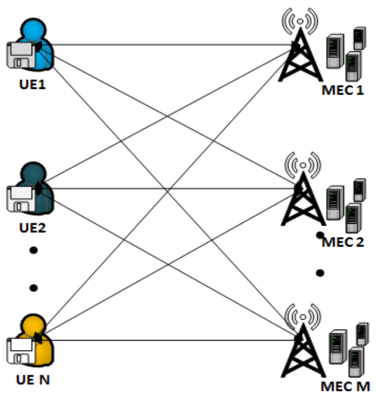

Figure 1: System model

In this model we assuming that each of the end user has different amount of data upload to cloud server. The network condition is depending on the base station that user associated. Denoted the set of N users as $\mathcal{N} = \{1, 2 \ldots, N\}$. For each user $i \in N$ has an amount of data to be offload is Π_i. For example, user UE1 has $\Pi_i = 3(GB)$ data needed to storage in the Cloud server. That amount can be requested to temporary storage at multiple MECs by fractioning of the original UE1's demand. Let $M=\{m_1,\cdots,m_M\}$ denoted the set of M MECs server.

3. PROBLEM FORMULATION

For each amount of data offloaded to MEC will cost some amount of sensitive define by function $f_i(x_i)$ where $f(\bullet)$ can be non-decreasing function, linear function depend on x_i. Whenever one user association with one BS, it will get an amount of bandwidth allocation for its channel to transfer the data. In this model, we also consider the transmission cost for each of the amount transfer from the user to the BS. The cost function determines as h(x) with following $h_j(x_{ij}) = x_{ij} / a_{ij}$ when a_{ij} is positive otherwise infinity. The bandwidth allocation for the system defined as the matrix:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{pmatrix},$$

a_{ij} is amount of bandwidth allocated for user i of BS j. Consequently, we formulated the optimization problem as follows:

$$\underset{x}{\text{minimize}}: \sum_{i=1}^{N}\sum_{j=1}^{M}\left(f_i(x_{ij})+h_i(x_{ij})\right)$$

$$\text{subject to}: \sum_{i=1}^{N} x_{ij} \le \Gamma_j, \forall j \in M \qquad (1)$$

$$\sum_{j=1}^{M} x_{i,j} = \Pi_i, \forall i \in N \qquad (2)$$

$$x_{ij} \ge 0, \forall i \in N, \forall j \in M \qquad (3)$$

The constraints number (1) and (2) are represented quality of service and user requirement that total amount of data upload to MEC cannot exceed its capacity, the demand of user should be served, respectively. The last constraint (3) is guarantee the amount of data offload nonnegative. The optimization problem is non-convex non-smooth problem cannot be solving in polynomial time. We proposed a non-convex ADMM to solve this challenging. Algorithm 1 states the proposal solution in detail.

4. EVALUATION

In this section, we conducted simulation results from our proposed Algorithm 1. In this algorithm, we first randomly generating matrix A. We then used technic transformation the original problem to an equivalent problem in order to solving in polynomial time and distributed way. The algorithm will be stop if the time is greater than predefined time or the objective value has converged.

In order to solving the optimization problem, we introduce a new variable y that associated with x via constraint x = y, then the original problem become following model:

$$\underset{x,y}{\text{minimize}}: \sum_{i=1}^{N}\sum_{j=1}^{M}\left(f_i(x_{ij})+h_i(y_{ij})\right)$$

subject to:

$$x_{ij} = y_{ij}, \forall i \in N, \forall j \in M \qquad (6)$$

$$x_{ij} \in X, \forall i \in N, \forall j \in M \qquad (7)$$

Then we redefined the feasible set of this problem as $X=\{x \mid (1),(2),(3)\}$, in term of the ADMM framework we defined the updating for each variable following:

$$x^{k+1} = \arg\min_{x\in\mathbb{R}^N}\{f(x)+\lambda^T(x-y^k)+\frac{\rho}{2}\|y^k-x\|_2^2\} \quad (8)$$

$$y^{k+1} = \arg\min_{x\in\mathbb{R}^N}\{h(y)+\lambda^T(x^{k+1}-y)+\frac{\rho}{2}\|x^{k+1}-y\|_2^2\} \quad (9)$$

$$\lambda^{k+1} = \lambda^k + \rho(x^{k+1}-y^{k+1}) \qquad (10)$$

Algorithm 1: Minimizing cost of data offloading and transmission(MONT)

1. Initial information for any N, M, and A
2. Randomly choosing x0, y0, and lambda0
3. For each time slot $t \in \{1, 2, \ldots, T\}$
   - **a.** Calculating (8), (9), (10)
4. Return results.

In simulation, we analyze the algorithm by Julia language, convex.jl package, and had shown the numerical results. We have used three different algorithms for performance comparison: Augmented Lagrangian method, centralized algorithm, greedy algorithm and our proposal.

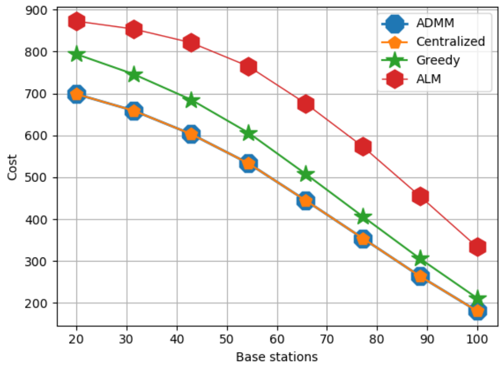

Figure 2: Comparison with Centralized, greedy, and ALM method.

In the Figure 2, the simulation result shown that the proposed algorithm has achieved optimal solution and outperform the other method except centralized. But the performance is better than the centralized algorithm is computational time and resources.

In the Figure 3, we have shown the time get to optimal of our algorithm under twenty iterations. Furthermore, our proposals have shown that the objective value is minimal and better than all others which was our main goal.

5. CONCLUSION

Data offloading can help the core network reduce the network congestion, also improving the service qualities of cloud server. This study shows us that by applying a non-convex ADMM to determine whether the data will be offloading, or decision which server to upload not only depending on network condition also the offloading cost, can significantly enhance the performance. Simulations have shown that proposed scheme significantly enhances the performance and achieves very fast convergence. Finally, a detailed study into the actual operating environment is required, as a future course, to validate our proposal and its performance

6. ACKNOWLEDGEMENT

This work was supported by Institute for Information &communications Technology Promotion(IITP) grant funded by the Korea government(MSIP)(B0190-16-2017, Resilient/Fault-Tolerant Autonomic Networking Based on Physicality, Relationship and Service Semantic of IoT Devices) *Dr. CSHong is the corresponding author

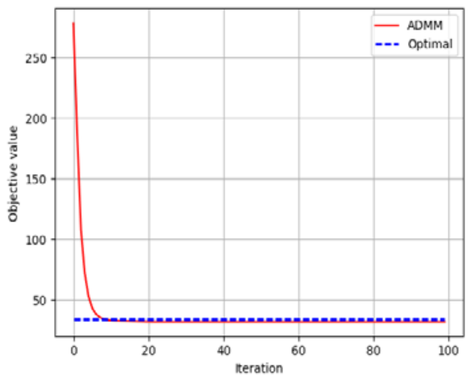

Figure 3: Convergences of ADMM

REFERENCES

[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi, "Internet of Things for Smart Cities," in *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32, Feb. 2014.

[2] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," in *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397-413, Aug 1993..

[3] C. Joo, S. Bahk and S. S. Lumetta, "A hybrid active queue management for stability and fast adaptation," in *Journal of Communications and Networks*, vol. 8, no. 1, pp. 93-105, March 2006.

[4] Byeong Kil Lee, L. K. John and E. John, "Architectural support for accelerating congestion control applications in network processors," *2005 IEEE International Conference on Application-Specific Systems, Architecture Processors (ASAP'05)*, 2005, pp. 169-175..

[5] H. Wang, W. Wei, Y. Li, C. Liao, Y. Qiao and Z. Tian, "Two-Degree-of-Freedom Congestion Control Strategy against Time Delay and Disturbance," *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, Miami, FL, 2010, pp. 1-5.