EFFICIENT RESOURCE ALLOCATION FOR ON DEMAND MOBILE EDGE CLOUD COMPUTING

[1]Zohra Jabeen, [2]Mrs.Y. Susheela,
[1]mtech, Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA, Ht no:20S41D5822, zohrajabeensoha@gmail.com
[2]Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA, chiduralasusheela@gmail.com

ABSTRACT: Mobile-edge cloud computing is a new paradigm to provide cloud computing capabilities at the edge of pervasive radio access networks in close proximity to mobile users. Aiming at provisioning flexible on-demand mobile-edge cloud service, in this paper we propose a comprehensive framework consisting of a resource-efficient computation offloading mechanism for users and a joint communication and computation (JCC) resource allocation mechanism for network operator. Specifically, we first study the resource-efficient computation offloading problem for a user, in order to reduce user's resource occupation by determining its optimal communication and computation resource profile with minimum resource occupation and meanwhile satisfying the QoS constraint. We then tackle the critical problem of user admission control for JCC resource allocation, in order to properly select the set of users for resource demand satisfaction. We show the admission control problem is NP-hard, and hence develop an efficient approximation solution of a low complexity by carefully designing the user ranking criteria and rigourously derive its performance guarantee. To prevent the manipulation that some users may untruthfully report their valuations in acquiring mobile-edge cloud service, we further resort to the powerful tool of critical value approach to design truthful pricing scheme for JCC resource allocation. Extensive performance evaluation demonstrates that the proposed schemes can achieve superior performance for on-demand mobile-edge cloud computing.

Keywords *– Computation offloading, edge computing, joint communication and computation optimization, on-demand resource allocation.*

1. INTRODUCTION

As smartphones are gaining enormous popularity, more and more new mobile applications such as face recognition, natural language processing, interactive gaming, and augmented reality are emerging and attract great attention [1]–[3]. This kind of mobile applications are typically resource-hungry, demanding intensive computation and real-time responsiveness. Due to the physical size constraint, however, mobile devices are in general resource-constrained, having limited computation resources. The tension between resource-hungry applications and resource-constrained mobile devices hence poses a significant challenge for the future mobile platform development. Mobile cloud computing is envisioned as a promising approach to address such a challenge. By offloading the computation via wireless access to the resource-rich cloud infrastructure, mobile cloud computing can augment the capabilities of mobile devices for resource-hungry applications. Currently, one common approach for mobile cloud computing is to offload the computation-intensive tasks to the remote public cloud infrastructure(e.g., Amazon EC2 and Windows Azure), in order to utilize the powerful computing and processing capabilities by the public clouds. As a matter of fact, the current public cloud architecture - built around static Internet-based installments of cloud resources not integrated with wireless networks - is already starting to show its limits in terms of computation-intensive mobile application support, since mobile users would experience long latency for data exchange with the public cloud through the wide area network (WAN), which risks to become the major impediment in satisfying the real-time interactive response requirement of mobile applications.

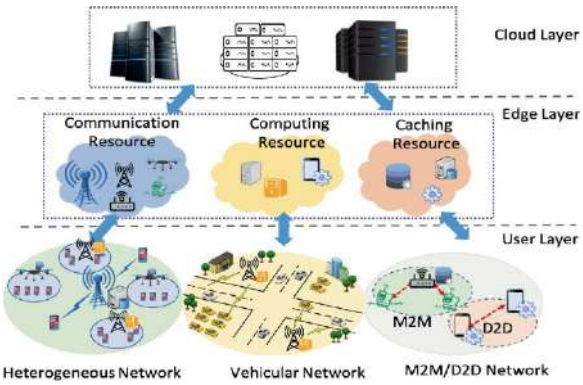

Fig.1: Example figure

To address this challenge, a novel mobile cloud computing paradigm, called mobile-edge cloud computing, has been proposed [4]. As illustrated in Figure 1, mobile-edge cloud computing can provide cloud-computing capabilities at the edge of pervasive radio access networks (e.g., 3G, 4G, WiMax, femtocells) in close proximity to mobile users. In this case, the need for fast interactive response can be met by fast and low-latency connection (e.g., via fiber transmission) to resource-rich cloud computing infrastructures (called mobileedge clouds) deployed by network operators (e.g., AT&T and T-Mobile) within the network edge and backhaul/core networks. By endowing ubiquitous radio access networks with powerful computing capabilities, mobile-edge cloud computing is envisioned to provide pervasive and agile computation augmenting services for mobile users at anytime and anywhere. In this paper, we aim at devising an efficient mobile-edge cloud computing framework that can provide rich flexibility in meeting different mobile users' demands. To this end, in this paper we propose a comprehensive framework consisting of a resource-efficient computation offloading mechanism for the users and a joint communication and computation (JCC) resource allocation mechanism for the network operator. Specifically, we first address the resource-efficient computation offloading problem, in order to reduce a user's resource occupancy. We then study the admission control problem and design the pricing scheme for JCC resource allocation, by jointly taking into account the objective of system-wide performance optimization as well as the practical constraints such as computational complexity for practical implementation and truthfulness for preventing manipulation.

2. LITERATURE REVIEW

Cloud computing for mobile users: Can offloading computation save energy.

The cloud heralds a new era of computing where application services are provided through the Internet. Cloud computing can enhance the computing capability of mobile systems, but is it the ultimate solution for extending such systems' battery lifetimes.

Cloud-vision: Real-time face recognition using a mobile-cloudlet cloud acceleration architecture.

Face recognition applications for airport security and surveillance can benefit from the collaborative coupling of mobile and cloud computing as they become widely available today. This paper discusses our work with the design and implementation of face recognition applications using our mobile-cloudlet-cloud architecture named MOCHA and its initial performance results. The challenge lies with how to perform task partitioning from mobile devices to cloud and distribute compute load among cloud servers (cloudlet) to minimize the response time given diverse communication latencies and server compute powers. Our preliminary simulation results show that optimal task partitioning algorithms significantly affect response time with heterogeneous latencies and compute powers. Motivated by these results, we design, implement, and validate the basic functionalities of MOCHA as a proof-of-concept, and develop algorithms that minimize the overall response time for face recognition. Our experimental results demonstrate that high-powered cloudlets are technically feasible and indeed help reduce overall processing time when face recognition applications run on mobile devices using the cloud as the backend servers.

Exploiting social ties for cooperative d2d communications: A mobile social networking case.
Thanks to the convergence of pervasive mobile communications and fast-growing online social networking, mobile social networking is penetrating into our everyday life. Aiming to develop a systematic understanding of mobile social networks, in this paper we exploit social ties in human social networks to enhance cooperative device-to-device (D2D) communications. Specifically, as handheld devices are carried by human beings, we leverage two key social phenomena, namely social trust and social reciprocity, to promote efficient cooperation among devices. With this insight, we develop a coalitional game-theoretic framework to devise social-tie-based cooperation strategies for D2D communications. We also develop a network-assisted relay selection mechanism to implement the coalitional game solution, and show that the mechanism is immune to group deviations, individually rational, truthful, and computationally efficient. We evaluate the performance of the mechanism by using real social data traces. Simulation results corroborate that the proposed mechanism can achieve significant performance gain over the case without D2D cooperation.

Mobile-edge computing – introductory technical white paper.
This white paper is authored by the founders of the Mobile-edge Computing (MEC) industry initiative. The objectives of this paper are to introduce the concept of Mobile-edge Computing and the related key market drivers, and to discuss the business, consumer and technical value/benefits that this technology offers. The paper discusses the enablers, the requirements and challenges for Mobile-edge Computing as well as the objectives of the MEC initiative. This white paper presents the high-level architectural blueprint of Mobile-edge Computing which, together with the scope of work, will form the basis for the first release of the work in the initiative. In addition, it highlights the relationships between and the interfaces with other industry efforts. The authors invite the various players in the value chain to actively participate in the work of the initiative.

3. METHODOLOGY

Existing studies in wireless resource auction as an initial thrust and to enable tractable analysis, in this paper we consider a static setting such that users are stationary.
The more general case that users may dynamically depart and leave the mobile-edge cloud system is very challenging and will be addressed in a future work.
We proposed a comprehensive framework consisting of a resource-efficient computation offloading mechanism for the users and a joint communication and computation (JCC) resource allocation mechanism for the network operator. We first solved the resource-efficient computation offloading problem for each individual user, and derive the optimal communication and computation resource demanding profile that minimizes the resource occupancy and mean while satisfies the delay constraint

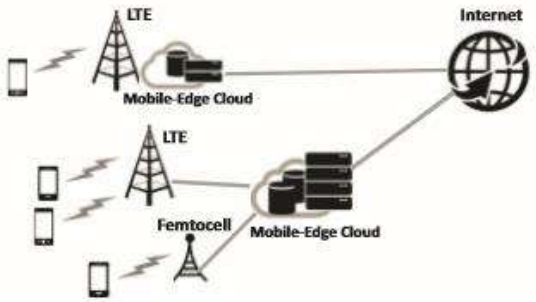

Fig.2: System architecture

4. IMPLEMENTATION

The major modules of the project are

1. USER
2. ADMIN

MODULES DESCRIPTION:

USER

In this application the user should register with the application and the user can search for the friends and send request to the friends and also check friends list, view all videos and he can recommend videos to the friends and also check the recommended videos.

ADMIN

Admin can login directly with the application, and the admin can perform the following operation such as view Users, upload Dataset, recommended Video, view Dataset, view All Videos View In Grpah , view All Videos Likes In Graph.

The above operations are done by the admin

5. EXPERIMENTAL RESULTS

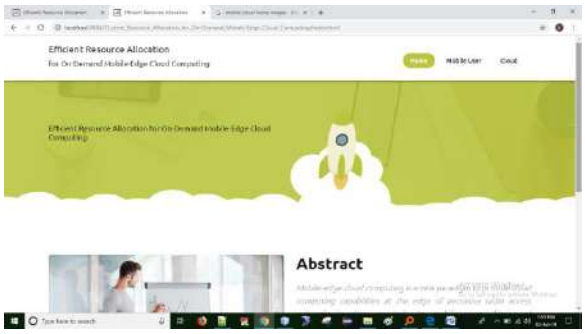

Fig.3: Mobile user login screen

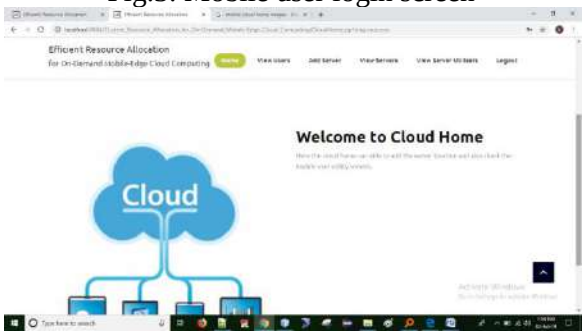

Fig.4: Clod home

Fig.5: Add server here

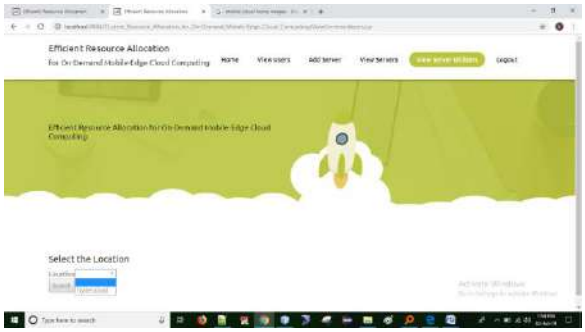

Fig.6: location selection

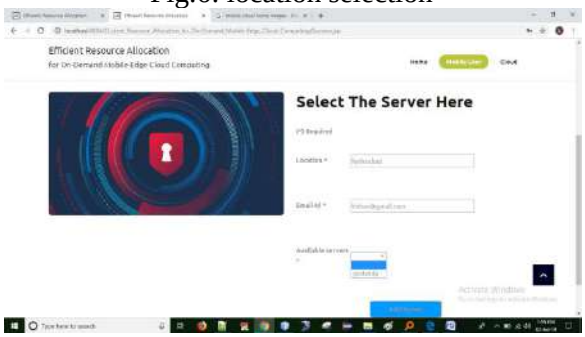

Fig.7: Select server here

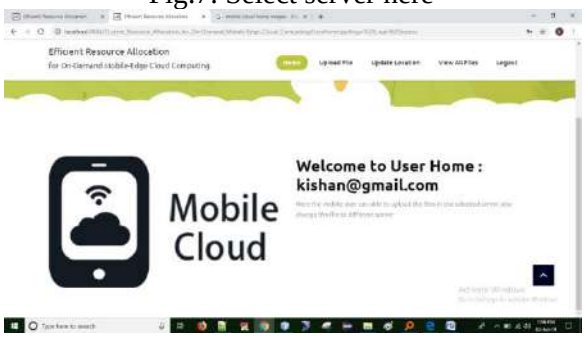

Fig.8: Upload file

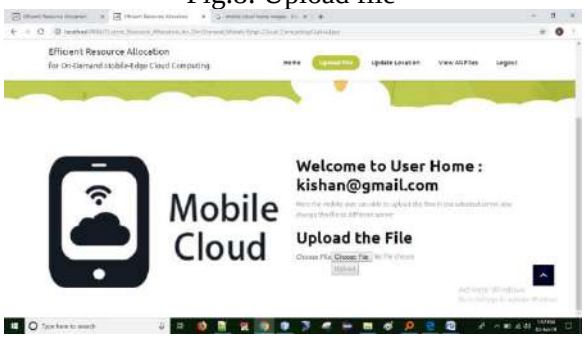

Fig.9: Update location

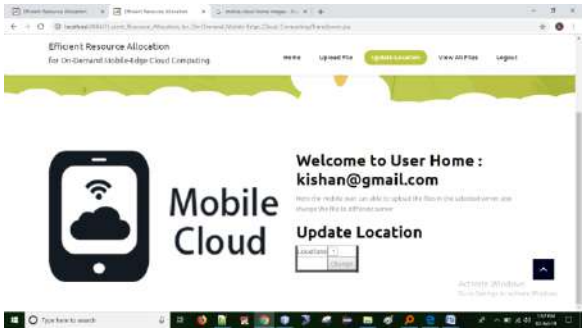

Fig.10: view all files

6. CONCLUSION

Aiming at provisioning flexible mobile-edge cloud service, in this paper we proposed a comprehensive framework consisting of a resource-efficient computation offloading mechanism for the users and a joint communication and computation (JCC) resource allocation mechanism for the network operator. We first solved the resource-efficient computation offloading problem for each individual user, and derive the optimal communication and computation resource demanding profile that minimizes there source occupancy and mean while satisfies the delay constraint. We tackled the admission control problem for JCC resource allocation, and developed an efficient approximation solution of a low complexity. We also addressed the truthful pricing problem by resorting to the powerful tool of critical value approach. Extensive performance evaluation demonstrates that the proposed mechanisms can achieve superior performance for on-demand mobile-edge cloud computing.

For the future work, we are going to consider the more general case that users may dynamically depart and leave the mobile-edge cloud system. We will take into account users' mobility patterns and devise efficient online resource allocation algorithms to cope with such system dynamics.

REFERENCES

[1] K. Kumar and Y. Lu, "Cloud computing for mobile users: Can offloading computation save energy?" IEEE Computer, vol. 43, no. 4, pp. 51–56, 2010.

[2] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon, and W. Heinzelman, "Cloud-vision: Real-time face recognition using a mobile-cloudletcloud acceleration architecture," in IEEE Symposium on Computers and Communications (ISCC), 2012.

[3] X. Chen, B. Proulx, X. Gong, and J. Zhang, "Exploiting social ties for cooperative d2d communications: A mobile social networking case," IEEE/ACM Transactions on Networking, vol. 23, no. 5, pp. 1471–1484, 2015.

[4] European Telecommunications Standards Institute, "Mobile-edge computing – introductory technical white paper," September 2014.

[5] Y. Wen, W. Zhang, and H. Luo, "Energy-optimal mobile application execution: Taming resource-poor mobile devices with cloud clones," in IEEE INFOCOM, 2012.

[6] L. Yang, J. Cao, Y. Yuan, T. Li, A. Han, and A. Chan, "A framework for partitioning and execution of data stream applications in mobile cloud computing," ACM SIGMETRICS Performance Evaluation Review, vol. 40, no. 4, pp. 23–32, 2013.

[7] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," IEEE/ACM Transactions on Networking, vol. 24, no. 5, pp. 2795–2808, 2016.

[8] Y. Zhang, H. Liu, L. Jiao, and X. Fu, "To offload or not to offload: an efficient code partition algorithm for mobile cloud computing," in IEEE 1st International Conference on Cloud Networking (CLOUDNET), 2012.

[9] Y.-H. Kao and B. Krishnamachari, "Optimizing mobile computational offloading with delay constraints," in IEEE GLOBECOM. IEEE, 2014, pp. 2289–2294.

[10] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," in IEEETransactionsonSignalandInformationProcessingoverNetworks. IEEE, 2015.
[11] X. Chen, Q. Shi, L. Yang, and J. Xu, "Thriftyedge: Resource-efficient edge computing for intelligent iot applications," IEEE Network, vol. 32, no. 1, pp. 61–65, 2018.
[12] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," IEEE Transactions on Communications, vol. 50, no. 2, pp. 291–303, 2002.
[13] C. Stein, T. Cormen, R. Rivest, and C. Leiserson, Introduction to algorithms. MIT press, 2009.
[14] A. Caprara and M. Monaci, "On the two-dimensional knapsack problem," Operations Research Letters, vol. 32, no. 1, pp. 5–14, 2004. [15] P. R. Milgrom, Putting auction theory to work. Cambridge University Press, 2004.

