

Cost Effective Resource Allocation in C-RAN with Mobile Cloud Computing

M. Jame Rishi

M.Tech, Department of CSE
Christ College of Engineering & Technology
Puducherry, India

Abstract:

To facilitate mobile users service and increase their revenue mobile operators uses Cloud radio access network(C-RAN) and mobile cloud computing(MCC). The aim of this paper is to reduce mobile operator's cost and mobile users task time constraints. Such that we assume the operation happens like the mobile cloud completes the task first and then pass the results back through the C-RAN to users. Considering constraints like user's data rate requirements, maximum transmission power and fronthaul links capacity joint cost-effective resource allocation is proposed between C-RAN and MCC. We solve the optimization problem for resource allocation and admission control of users. Taking the non-linear constraint and solving the problem through big-M method framework. We solve joint cost effective by non-convex optimization with the constraints of task executing time, transmitting power, computation capacity and fronthaul data rates. Using weighted minimum mean square(WMMSE) the non-convex problem is changed to convex problem.

Keywords: C-RAN, Cost efficient, Data rates, Mobile Cloud Computing, Resource-Allocation.

1.Introduction

Cloud computing is a mean of hardware and software accessed through online instead of using our desktop or our company's network where that service will be provided by another company in a complete seamless manner[1]. The location of the hardware and software is not needed such that it is placed in cloud which is represented by the internet. The buzzword cloud computing means different things to different people[5]. Some says its another method to describe information technology "outsourcing" and some says its used to mean any computing service through internet or similar network and some say it as bought-in computer service that is placed outside our firewall[12].

Resource allocation is a process of allocating available resources in a economic, efficient, effective way. Resource allocation process schedules available resources and activities required by other activities which considers both resource availability and project time[3]. Both resource provisioning and allocation solves the problem by providing response through service

provider who manage the resources by individual resource request. Resource allocation strategy (RAS) is a strategy which allocates and utilize inadequate resources within the cloud platform which meets the needs of cloud application. This requires the resource amount and type of resources which each application needs to finish their task. From cloud provider perspective its impractical to predict the dynamic nature user,their demands and applications demands. But for cloud users the tasks should be completed in a minimum of cost. We need a resource allocation system which is suitable for cloud environment with limited resources, resource heterogeneity, environmental neccessities, locality restrictions and dynamic nature of resource demands. Cloud resources are made of virtual resources[4]. Through virtualization and provisioning the physical resources are connected to multiple computers. The virtualized resources are made with set of parameteres which guides the processing, memory and disk needs. Cloud provisioning is a process of mapping virtualized resources to physical resources. The hardware and

software resources are allocated in only demand basis in a cloud environment[13].

It is very lucky that one new promising network infrastructure has discovered recently is the cloud radio access network(C-RAN). This discovery will be soon receive a lot of attention in academia and industry[6]. C-RAN is said to be 4C network which is cloud computing based, centralized, clean and collaborative radio access network. C-RAN divides traditional base station into three parts they are remote radio heads (RRH), base band unit (BBU) and a fiber transport (or fronthaul) link which connects RRH to BBU cloud pool[2]. The link should be of high bandwidth, high speed, low latency level. Industries enjoys more the benefits of C-RAN.

Like C-RAN another impressive technique has been discovered recently, i.e., mobile cloud computing(MCC) has attracted huge of number of fields. This attraction is due to the integration of efficient cloud computing with mobile environment, which facilitates the mobile user with high the computing demands and low the computing resource can balance task to the powerful cloud platforms. Investigation is happening that if the offloading operation can save energy and increase the battery life of user equipments[14]. The reference provides a theoretical framework of cost effective mobile cloud computing under stochastic wireless channel and for obtaining efficient computation offloading for MCC the reference also proposes a game theoretical approach. We can also learn cost efficiency oriented traffic offloading in wireless networks. Cloud computing is integrated with vehicular networks has been deployed by which the vehicle can share each other with computation resources, storage resources and bandwidth resources.

2.Related Works

In [2],C-RAN is a impressive and promising network architecture in the mobile communications now and in future and one practical obstacle is that its large scale implementation. This implementation is the important requirement for low latency fronthaul and high capacity connecting the remote radio heads to the base band units in the C-RAN whereas the remote radio heads are distributed and the base band units are centralised.

For the C-RAN improved scalability we should consider the fronthaul loading which is important in the signal deduction and it is very attractive to reduce the fronthaul loading in C-RAN. In this technique we take uplink C-RAN systems and we model a distributed fronthaul compression scheme at the distributed RRH and a joint recovery algorithm in the base band unit by applying the techniques of compressive sensing(CS).

CS problem in C-RAN system differs from conventional distributed CS that requires to adopt the underlying effect of multi-access fading for the point to point recovery of the transmitted data from the users. We calculate the performance of the proposed point to point recovery algorithm and we show that the sum measurement matrix in C-RAN system which has both the distributed fronthaul compression and multi-access fading which can still compensate the limited isometric property with higher probability. According to these research we design a tradeoff results between fronthaul loading and uplink capacity in C-RAN system.

In [7], today's world smartphone have been exploded to higher popularity which is becoming more sophisticated and portable day by day. So the smartphone developers are developing complex applications that needs large amount of computational power and cost. In this technique ThinkAir is architecture which reduces the work for developers which can move the smartphone applications to the cloud. ThinkAir enhanced the concept of smartphone virtualization in cloud and offers method level computation offloading. Mainly concentrates on the elasticity and scalability of the cloud and exploits mobile cloud computing power by executing parallelizing method using multiple virtual machine images. We deploy ThinkAir and calculate it using range of benchmarks. Where the benchmarks ranges from simple micro benchmarks to complex applications. We test the results with N-queen problem, face detection application and virus scan application in two orders of magnitude where we can find the reduction in execution time and cost consumption. This technique also facilitates parallelizable application which can use multiple virtual machines for seamless and on demand basis execution so that we can attain greater reduction on execution time and cost consumption. At the final phase we use a

memory-hungry image combiner tool to show that the application can dynamically request virtual machines for their computational requirements with their computational power.

In [10], mobile cloud computing is discovered as a promising method to attach computation capabilities of mobile phones for augmenting resource-hungry mobile applications. In this technique we design a game theoretic method for obtaining efficient computation offloading for mobile cloud computing. We calculate the decentralized computation offloading decision making problem by taking samples from mobile phone users which appears to be decentralized computation offloading game. We testify game structural property and demonstrate that game always meet a Nash equilibrium. Next we will model a decentralized computation offloading method that can attain a game Nash equilibrium and centralized optimal solution is obtained with quantifying along its efficiency ratio. Numerical values shows that the proposed method can receive efficient computation offloading performance and it will scale as the system size increases.

In [8], this technique gives a theoretical architecture of energy-optimal mobile cloud computing under stochastic wireless channel. Aim of the technique is that to conserve power of mobile phones, by executing mobile applications in mobile phones in optimal manner (such that mobile execution) or cloud offloading (such that cloud execution). In the old case we can reconfigure the CPU frequency sequentially; or in the new case we can dynamically change the data transmission rate to the cloud, which corresponds to the stochastic channel condition. We calculate both allocation problems as constrained optimization problems, and we attain optimal allocation policies from closed form solutions. Testing for the energy- optimal execution strategy of applications with smaller output value (say for example., CloudAV).

In [9], we evaluate a threshold policy which gives the data consumption rate, and finally a ratio is obtained which lies between the data size(L) and the delay constraint (T). This ratio is compared with threshold which is based on both the power consumption model and the wireless link model. The final output is of numerical form will shows

that a sufficient amount of power can be saved for the mobile phone by offloading mobile applications optimally to the cloud in some cases. The obtained theoretical architecture and numerical values will enhance the system implementation part of mobile cloud computing under stochastic wireless channel.

In [11], we are in the internet generation, to advance transport safety, release traffic congestion, make free from air pollution, and make sophisticate the driving all components in intelligent transportation systems will be attached. But this idea will be a open challenge from the vehicles connected to the huge amount of collected and stored traffic-related data. In this technique , we propose to combine cloud computing into vehicular networks so that the vehicles can share resources for computation, for storage and for bandwidth.

The designed model has a vehicular cloud, a roadside cloud, and a central cloud. Further we learn cloud resource allocation and virtual machine migration for efficient resource allotment in this cloud depending vehicular network. To allocate cloud resources optimally a game-theoretical method is also designed. Because of vehicle mobility virtual machine migration is solved by resource reservation scheme.

3. Proposed Work and Models

In this paper we proposed a technique which can improve the system's performance and reduce the cost. The technique namely cost effective resource allocation which is run through cloud radio access network in the mobile cloud computing. C-RAN has moved to the next generation as a potential candidate to save the increasing mobile traffic.

Mobile cloud computing gives a prospective solution to the reduced resource mobile users for the execution of computation intensive task to advance both the performance and cost efficiency. We are very lucky that a new promising architecture has been developed recently that is cloud radio access network and it will soon attract large amount of users in industry and academia.

C-RAN divides the traditional base stations into three parts namely Remote Radio Heads, Base Band Unit and Fibre Transport Link. Intensive network computational tasks such as pre-coding

matrix calculation, baseband signal processing, channel state information which will be transferred from baseband unit to the cloud. This cloud is made up of huge amount of software which is defined by the virtual machines with the features like Dynamic configuration, Scalability, Sharing, Re-allocation on demand basis. On the other side remote radio heads will play a soft relay which can compress and transmit the received signals of the base band unit and transmit further to the user equipments.

In this scenario the remote radio heads with restricted functions can only include A/D and D/A conversion, amplification, frequency conversion which can easily share based on the network requirement in the most impressing feature of C-RAN is the separation of base band unit and remote radio heads and the integration with different base band units which correspondingly increases significant performance which can be gained from the efficient interference cancellation and management which can also increase the network capacity and reduce the cost.

3.1 Advantages

- We transform the cost effective optimization into non-convex optimization, which is Non-deterministic Problem.
- We change the cost reduction along with data rate maximisation problems which will reduce the total cost.
- Mean square problem is avoided to solve the weighted minimum mean square error.
- We can view the cost effective resource allocation between C-RAN and Mobile cloud computing which is used in beam forming vector design.

3.2 Module Description

3.2.1 Mobile Clone

Normally, when the mobile users come across the computational intensive or high energy required tasks, they sometimes do not want to offload those tasks into the mobile cloud, as transmitting those program data to the cloud still costs some energy. In some cases, it is even better to execute those tasks locally if transmission overhead is too high. Therefore, it is better to have the mobile user's computational tasks and some of the corresponding data in the mobile cloud first. To deal with this concerns, we propose to have

mobile clones which are co-located with the BBU in the cloud pool. The mobile clone will have the user task information and data on board. Mobile clone can be implemented by the cloud-based virtual machine which holds the same software stack, such as operating system, middleware, applications, as the mobile user.

3.2.2 Network Model

After the mobile clone finishes the execution of the task, the results will return to i, mobile user through C-RAN. The received signal at the user equipment under the complex baseband equivalent channel can be return as

$$y_i = \sum h_{ij} v_{ij} x_i + \sum h_{ij} v_{kj} x_k + \sigma_i$$

$$i=1,2,\dots,n$$

where x_i denotes the transmission data for the i th UE with $E\{|x_i|^2\} = 1$, $C \subseteq L$ is the set of serving RRHs, $H_{ij} \in C^{K \times I}$ denotes the channel vector from RRH j to UE i , while σ_i denotes the white Gaussian noise which is assumed to be distributed as CN $(0, \sigma_i^2)$. Denote $V_{ij} \in C^{K \times I}$ as the transmitting beamforming vector from RRH j to UE i .

3.2.3 Fronthaul Constraints

The fronthaul link can carry the task results from the mobile clone to the UE through C-RAN. One can see that the number of non-zeros elements of the transmitting beamforming vector also indicates the number of data symbol streams, carried by the fronthaul link from BBU to RRH j for the i -th mobile user. Assume that each fronthaul link is only capable of carrying at most $C_{j,max}$ signals for user equipments.

3.2.4 QoS Requirement

The QoS can be given as the constraints of the whole time cost for completing the required task and returning the results back to the mobile user. We define the total time spent in executing and transmitting the task results to UE. Also, the whole energy cost in executing this task and transmitting the results back to i -th UE can be given as

$$T_i = T_i^{Tr} + T_i^c n_i$$

$$i=1,2,\dots,n$$

where $n_i \geq 0$ is a weight to trade off between the energy consumptions in the mobile cloud and the C-RAN, and it can be also explained as the inefficiency coefficient of the power amplifier at RRH.

3.2.5 Joint Optimization Solution

In this section, we will solve the energy minimization optimization and resource allocation jointly between the mobile cloud and mobile network. The objective is to minimize the total energy consumption in mobile cloud for executing the task and in C-RAN for transmitting the processing results back to the mobile user. We assume that the task has to be completed in the total time constraint (QoS) of the given task, including the executing time plus the transmitting time.

4. Dataflow Diagram

The DFD is additionally called as air pocket outline. It is a basic graphical formalism that can be utilized to speak to a framework as far as info information to the framework, different handling did on this information, and the yield information is created by this framework.

The information stream chart (DFD) is a standout amongst the most vital demonstrating devices. It is utilized to demonstrate the framework segments. These parts are the framework procedure; the information utilized by the procedure, an outer element that communicates with the framework and the data streams in the framework.

DFD indicates how the data travels through the framework and how it is adjusted by a progression of changes. It is a graphical strategy that portrays data stream and the changes that are connected as information moves from contribution to yield.

DFD is otherwise called bubble diagram. A DFD might be utilized to speak to a framework at any level of deliberation. DFD might be divided into levels that speak to expanding data stream and practical detail.

A data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system, modelling its process aspects. For a clear understand of the outline of the system DFD is used, and then the content will be elaborated later. Another use of DFD is that it can showcase the data processing i.e., structure design.

The DFD states what are all will be the input information and the output information of the system, and gives the location of the stored data and shows the movement of the data through the system.

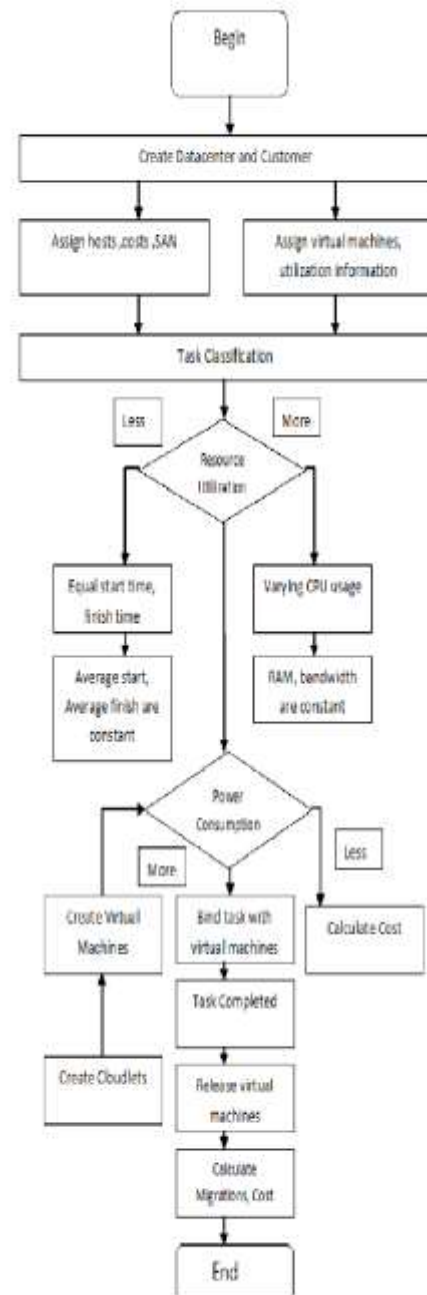


Figure 1: Data Flow Chart

5. Experimental Results

As per the simulation results, we proposed framework to define the utilized resource, power consumed, cloudlets used and execution time analysis.

5.1 Resource Allocation

It is used to the ratio the amount of overall usage of CPU, RAM and bandwidth where RAM and bandwidth remains equal. The relations between the total energy consumption and different QoS or time constraints are examined under different D_i with total CPU cycles $F_i = 1500$. One can see that with the increase of the time constraints, the energy consumption decreases, as expected. Also, with the increase of the data size, the energy increases, but the gap between them is small, due to the tradeoff factor we set.

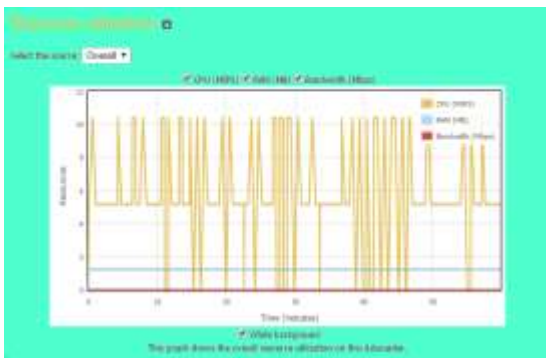


Figure 2: Resource Utilization

5.2 Power Consumption

It is used to the increase or decrease the total amount of power consumed by the data provider and customer based on their usage. Similar to Fig. 1, Fig. 2 shows that the whole energy consumption of mobile cloud and C-RAN decreases either with the increase of the time constraints or with the decrease of the CPU cycles required by each task.

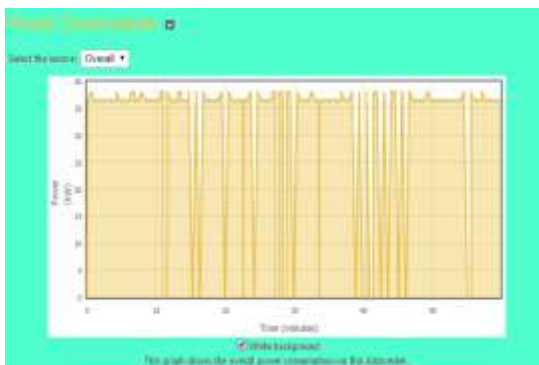


Figure 3: Power Consumption

5.3 Cloudlets

It is used to the display the number of cloudlets used and their performance individually. In Fig. 2 and Fig. 3, we compare the proposed joint energy minimization optimization with the

separate energy minimization solutions, which has been used in some works etc. For the separate energy minimization, we set two time constraints as $T_i^{Tr} \leq T_{i,max}^{Tr}$ and $T_i^C \leq T_{i,max}^C$, where $T_{i,max}^{Tr} + T_{i,max}^C = T_{i,max}$. $T_{i,max} = 0.1s$ is set in both Fig. 2 and Fig. 3 while $D_i = 1000$ and $F_i = 1500$ are set in Fig. 2 and Fig. 3, respectively

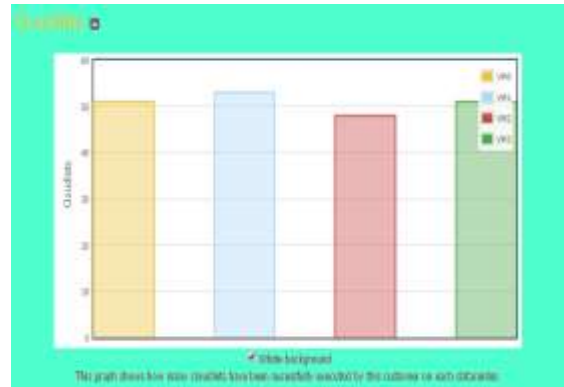


Figure 4: Cloudlets

5.4 Execution Time

It is used to the ratio the amount of time taken for initializing ,finishing and for average start and average finish. In Fig. 4, we assume that one additional user has been added in C-RAN system in Fig. 2 and other parameters are set the same as in Fig. 3. One can see that our proposed optimization method has nearly the same performance gain as in Fig. 3. As expected, more power is used for all the solutions in Fig. 4 than Fig. 3. Also, we have checked our our solution for different number of antennas and similar performance gain can be achieved.

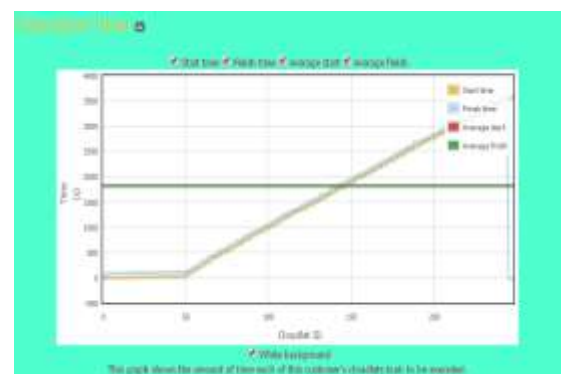


Figure 5: Execution Time

6. Conclusion

In this paper the C-RAN framework along with mobile cloud computing is proposed to gain the

advantage of two cloud based techniques. We assume separate task is to be executed in each user equipment in the mobile cloud network and we design this task with two functions i.e., the maximum number of CPU cycles needed to complete the task and the total memory needed to transmit the result back to the user equipment through C-RAN. We reduce the total cost of mobile cloud and mobile network by addressing this problem into optimization problem when considering quality of service(QoS) i.e., we take time constraint into consideration. Adding to this we also have fronthaul constraints in C-RAN to gain the RRH clusters. The proposed system gives numerical results which can clearly show the cost reduction and resource allocation solution which can improve the system performance and reduce cost. Future work of this technique will be based on whole data transmission process including uplink transmission and downlink transmission. The uplink transmission sends data from user equipment to RRH whereas the downlink transmission send the data from RRH to another RRH.

7. References

1. J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
2. X. Rao and V. Lau, "Distributed fronthaul compression and joint signal recovery in Cloud-RAN," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1056–1065, February 2015.
3. Rajasekar, B., and S. K. Manigandan. "An Efficient Resource Allocation Strategies in Cloud Computing." *International Journal of Innovative Research in Computer and Communication Engineering* 3.2 (2015): 1239-1244.
4. Mangla, Neeraj, and Jaspreet Kaur. "Resource Allocation in Cloud Computing."
5. J. Wu, "Green wireless communications: from concept to reality," *IEEE Wireless Communications*, vol. 19, no. 4, pp. 4–5, August 2012.
6. C. M. R. Institute., "C-RAN white paper: The road towards green Ran. [online]," (June 2014), Available: <http://labs.chinamobile.com/cran>.
7. S. Kosta, A. Aucinas, P. Hui, R. Mortier, and X. Zhang, "Thinkair: Dynamic resource allocation and parallel execution in the cloud for mobile code offloading," in *2012 IEEE Proceedings INFOCOM*, March 2012, pp. 945–953.
8. K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?" *Computer*, vol. 43, no. 4, pp. 51–56, April 2010.
9. W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4569–4581, September 2013.
10. X. Chen, "Decentralized computation offloading game for mobile cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 974–983, April 2015.
11. X. Chen, J. Wu, Y. Cai, H. Zhang, and T. Chen, "Energy-efficiency oriented traffic offloading in wireless networks: A brief survey and a learning approach for heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 4, pp. 627–640, April 2015.
12. Armbrust, Michael, et al. "A view of cloud computing." *Communications of the ACM* 53.4 (2010): 50-58.
13. J. Tang, W. P. Tay, and T. Quek, "Cross-layer resource allocation in cloud radio access network," in *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, December 2014, pp. 158–162.
15. L. Yang, J. Cao, S. Tang, T. Li, and A. Chan, "A framework for partitioning and execution of data stream applications in mobile cloud computing," in *2012 IEEE 5th International Conference on Cloud Computing (CLOUD)*, June 2012, pp. 794–802

