Cluster Computing (2019) 22:223–239
https://doi.org/10.1007/s10586-018-2847-y

Resource allocation in the cloud for video-on-demand applications using multiple cloud service providers

P. Muthi Reddy[1] · Ansaf Ahmed[1] · S. H. Manjula[1] · K. R. Venugopal[1]

Received: 15 March 2018 / Revised: 18 July 2018 / Accepted: 29 August 2018 / Published online: 26 September 2018

Abstract
Video-on-demand (VoD) applications have become extensively used nowadays. YouTube is one of the most extensively used VoD application. These applications are used for various purposes like entertainment, education, media, etc., of all age groups. Earlier, these applications were supported by private data centers and application servers. Sufficient infrastructure had to be bought and maintained, to support the demand even during unexpected peak times. This approach caused huge loss of resources when the demand is normal as a large portion of the resources remained idle. To overcome this, VoD application providers moved to the cloud, to host their video content's. This approach reduced the wastage of resources and the maintenance cost of the VoD application provider. The problem is to determine the number of resources to handle the demand while maintaining QoS for every instance. We have designed two algorithms in this paper, namely the multiple cloud resource allocation (MCRA) algorithm and the hybrid MCRA algorithm. Most of the cloud service providers (CSPs) basically provide two types of resource allocation schemes: (i) the reservation scheme and (ii) the on-demand scheme. The reservation scheme provides time-based tariff prices, where the discount is provided for the resources depending on their quantity and reservation time. This scheme is used in the MCRA algorithm to reduce the *cost* of the VoD application provider. In Hybrid MCRA algorithm both the reservation scheme and on-demand scheme are implemented, to overcome the drawbacks of the MCRA algorithm which are *under-subscription* and *over-subscription*. We have analyzed both the algorithms in terms of cost and allocation of resources. These algorithms can help allocate resources in of cloud for VoD applications in a cost-effective way and at the same time not compromise on the QoS of the video content.

Keywords Cloud service provider · Hybrid · Resource allocation · Reservation scheme · Time-discount tariffs · Video-on-demand

1 Introduction

Video-on-demand (VoD) applications are already popular. People use them for one or more purposes like entertainment, learning, and education, etc. Earlier, VoD applications were hosted on the application servers at the application provider's private data centers. The application providers had to maintain at any point of time. The sufficient resources at their data centers to handle the demand for the video content provided by the VoD application. The resources were sufficient enough to sustain the QoS of the video content. The demand for any video content provided by the VoD application is very unpredictable. The demand may be stable for some period of time while it may be varying and unexpected for some other period of time.

The drawback of private data centers is that the frequency scaling of resources either up or down, depending on the changes in the demand for the video content is not possible. A large amount of investments are done on the procurement of hardware and other infrastructure for hosting the VoD application and video contents. The maximum demand that the application can expect at any point in time is taken into consideration for the purchase of infrastructure resources. The varying and unpredictable demand of the video content cannot make full utilization of the resources available to private data centers

✉ P. Muthi Reddy
muthireddy2016@gmail.com

1 Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, Karnataka 560001, India

and hence causes wastage of resources. This indirectly affects the environment as a lot of power is used to keep the servers up and running [1]. Complete utilization of the resources may happen sometimes, but mostly a part of the resources remain idle. QoS of the video content may be compromised with available resources at the data center when the demand exceeds the maximum expected level, though it may not usually occur [2]. These drawbacks do not make private data centers a viable option for hosting VoD applications, and hence require a more scalable and dynamic resource allocation paradigm, which is the cloud.

One of the notable features in cloud computing is the scalability of resources. Resources can be scaled up or down depending on the requirement. There are many remarkable advantages of cloud computing. A company which provides cloud services is simply called a *cloud service provider* (CSP). The CSP is responsible for maintaining all the servers, providing power supply, upgrading the systems, performing timely service of the equipment, cooling of the systems, etc. The server systems are physically secured as well as software protection from viruses, malware and other forms of attacks is provided by the CSP. The resources are remotely accessible from anywhere at any time. The resources are accessible virtually as their physical location such as memory location is hidden due to security reasons. The resources are leased and they are charged on the basis of utilization. There are different service models in the cloud such as platform as a service, infrastructure as a service and software as a service. Infrastructure as a service model provides hardware resources such as a processor, hard disk space, bandwidth, etc. These benefits of cloud computing makes it a feasible solution to address the problems of private data centers. The VoD application provider need not make huge investments in the purchase of hardware and other infrastructure. The application provider need not worry about the wastage of resources thereby promoting the revolution in green computing. Automatic scaling of resources can opt where the resources are scaled up or down automatically if some conditions are met, for example, if the demand increases above some threshold level then increase the resources by some amount [3].

The cloud resources can be dynamically allocated depending on the demand [4]. We can detect the workload patterns and allocate the resources accordingly. Cloud supports the resource allocation for complex applications which require different resources for their execution [5]. Cloud computing is favorable for hosting a VoD application, as it offers many advantages over procuring own infrastructure [6]. Some of the advantages are scalability of resources, flexibility, latest technology, hardware and software maintenance done by CSP and high Internet connectivity.

Allocation of resources in the cloud is easy whereas allocating *appropriate* number of resources is a big challenge. The VoD application provider wants to allocate resources in such a way that the cost is minimized while QoS of the video content is not compromised. There are two schemes provided by CSPs for resource allocation—(i) the reservation scheme and (ii) the on-demand scheme. In the reservation scheme, resources are reserved prior to their consumption. Time-discount tariff prices are offered on the reservation scheme where the prices are discounted depending on the number of resources and reservation time. CSPs provide such type of offerings to attract buyers and to make them purchase more resources for longer duration's of time to make efficient and profitable utilization of their infrastructure [7]. Such type of discount prices is not offered in the on-demand scheme. In the on-demand scheme, resources are allocated at the time of consumption and follow the pay-as-you-go model. The on-demand resources are more expensive than the reservation resources. Hence, the VoD application provider would like to reserve as many resources as possible and limit the allocation of resources using the on-demand scheme. The VoD application provider has to predict the demand for a video content. There are many techniques for the prediction of demand that have been designed [8, 9]. The reservation of resources has to be done on this prediction. The demand for a video content follows a log-normal distribution, as shown in [10]. Hence, we have used a lognormal distribution for the predicted demand in our analysis. The energy efficiency of video content is computed by using the dynamic data virtualization (DDA) algorithm [11].

Another aspect of CSPs which can be exploited is that different CSPs have different tariff prices. Cloud tariffs are given in tabular form as shown in Table 1. Two algorithms

Table 1 An example of a tariff table presented by a CSP

w (in seconds)	Alloc. (in Mbps)	Tariff (in $ per unit time)
1	1	10
1	2	19.4
1	3	28.518
1	4	37.363
1	5	45.943
2	1	9.7
2	2	18.818
2	3	27.663
2	4	36.243
2	5	44.565

for the allocation of resources have been designed in [6], Prediction-based resource allocation (PBRA) algorithm and hybrid algorithm. PBRA algorithm has used the reservation scheme whereas the Hybrid algorithm has used both the reservation and the on-demand schemes. But the drawback of these algorithms is that they are based on single CSP. They do not take into consideration the different tariff prices provided by different CSPs, which can be further exploited to reduce the cost. In this paper, we have designed two algorithms which overcome the drawback of the PBRA and the hybrid algorithms. We call them the *multiple cloud resource allocation (MCRA)* algorithm and *hybrid MCRA* algorithm. Both the algorithms involve multiple CSPs thereby overcoming the drawback of their predecessor algorithms [6].

The different tariff prices offered by different CSPs can be exploited by the VoD application provider to minimize the cost of resource allocation. In CSP, that charges the least can be chosen for resource allocation. This idea is incorporated in our proposed algorithms. The MCRA algorithm implements only the reservation scheme whereas the Hybrid MCRA algorithm implements both the reservation and the on-demand schemes. The hybrid MCRA algorithm overcomes some of the drawbacks of the MCRA algorithm namely, *under-subscription* (fever resources are allocated compared to demand) and *over-subscription* (more resources are allocated compared to demand). This paper takes into consideration security aspects related to allocation of resources. The allocation cannot be performed unless the user has access to the system. The system maintains a table of users that can login to the system. Any unauthorized users cannot login to the system. The system validates the user every-time when the user tries to login to the system.

1.1 Motivation

The PBRA and the hybrid algorithms [6], have implemented both the reservation and the on-demand schemes provided by CSP to reduce the cost of resource allocation. *PBRA* and *hybrid* algorithms have taken into consideration only one CSP. In the distributed system the different CSPs have variant tariff rates which can be exploited still further to reduce the cost. This idea has been implemented in our MCRA and hybrid MCRA algorithms.

1.2 Contributions

Two algorithms have been designed in this paper to perform resource allocation in of the cloud—the *MCRA* and the *hybrid MCRA* algorithms. The MCRA algorithm has implemented the reservation scheme while the hybrid MCRA has implemented both the reservation and the on-demand schemes, and hence the name *hybrid*. These algorithms have provided a way to perform allocation in the cloud such that sufficient resources are allocated to maintain the QoS of the video content while at the same time reduce the cost on the VoD application provider. The different tariff rates provided by different CSPs are exploited to further reduce the cost.

1.2.1 Organization

The paper is structured as follows. In Sect. 2 we discuss the related work. The problem definition is presented in Sect. 3. The system model is introduced in Sect. 4. Section 5 demonstrates the MCRA algorithm while the hybrid MCRA algorithm is presented in Sect. 6. The analysis of the algorithms is shown in Sect. 7. Finally, the conclusions are drawn in Sect. 8.

2 Related work

The prediction of user access demands for video streaming application and resource utilization has been studying in literature. An auction-based online technique is proposed for the supplying of virtual machines (VMs) [12]. The allotment of VMs and their cost evaluation in different clouds is taken into consideration with many different categories of resources. The designed technique does not make any guesses regarding the future need for VMs, hence resembling the settings in the real cloud. The users are given incentives, to encourage them to provide their true requests. Cloud computing provides the flexibility to users by providing the facility to obtain virtual machine resources on-the-fly and hence supports the pay-as-you-go scheme of CSPs. A technique has been designed to reduce the utilization of energy by effectively allotting virtual machine resources to physical machines [13]. The virtual machine resources allocated in a decentralized multi-agent manner. CSPs provide instances with varying configurations. A group of instances with a specific configuration are usually bought in order to maintain the applications' performance predictability. The specific configuration of the instances has to be chosen such that the required performance of the applications is achieved. This has been incorporated by using the mixed-integer programming paradigm to perform resource allocation [14]. The mechanism handles mispredictions and supports automatic scaling of resources.

One of the critical requirements for storage systems in the cloud is to support for deadline guaranteed services compliant to service level agreements (SLAs) for its services. A new mechanism for SLAs has been proposed [15], which allowing users to state a part of their requests which require to complete within the given time limit. To make maximum benefit out of the cloud resources, it is necessary to allot and schedule them in such a way that the QoS needs of users are met. These QoS requirements were specified in the SLAs. A mechanism is designed in [16], for the allotment of resources and their scheduling using constraint programming. The mechanism can handle efficiently MapReduce jobs represented by SLAs. One of the important features of cloud computing is that it provides *flexibility* for users to increase or decrease the number of resources depending on the requirement. Cloud resources can be used *efficiently* by multiplexing resources. A mechanism has been proposed [17], using virtualization technology that assigns resources on-the-fly depending on the needs of the user. This assists green computing by keeping the number of servers that are running to a minimum.

In cloud computing, all the devices and equipment's are in the hands of the cloud providers. This makes it possible for the cloud providers to assign incoming requests to appropriate machines and allot cloud resources on the fly as virtual machines. Reducing the required time to complete the incoming requests is very critical to cloud providers as reducing this can benefit the cloud providers in many ways. A technique has been proposed [18], to allot the resources in an optimal way where the assignment of requests to the machines will be determine prior to allotment. The scientific work-flow of applications have started moving to cloud computing for their deployment. These applications are real-time applications which have time limits for their execution. Therefore it is very important to make sure that the resources are reliable, as a number of machines are deploying in the cloud. A mechanism is designed [19], by enhancing the traditional primary backup technique to include the cloud features. The cloud has been used to host many applications that require a lot of bandwidth. But there are some drawbacks with the cloud which include failure of devices, overload of resources especially bandwidth and transmission over long distances that bring down the QoS with regard to availability of data, provision of resources and local access to resources. A mechanism has been designed [20], with the goal of making the maximum use of resources in the cloud. The applications that require a lot of bandwidth can attain the required QoS specified in SLA within the deadline and with low costs [21].

The efficiency of data centers' can be further increased with the use of *Dockers*. But the current techniques for using Dockers are not quite efficient. An application oriented docker container (AODC)-based model has been developed [22] for resource allocation which reduces the cost of deploying applications in data centers' and provides automatic scaling of resources as the demand of cloud applications changes. Auction based allocation in the cloud for resource bundling on-the-fly and provisioning of VMs was still immature. In [23], an auction-based technique had been developed considering the bundling of resources on-the-fly and the servers distributed geographically across different areas. The cost incurred in running the servers is taken into consideration. Furthermore, the welfare of the cloud users, as well as the maximization of the profit to a cloud provider, is focused. Many businesses today have started using the cloud for their work. Therefore, it is crucial to achieving their requirements which require a mechanism to categorize and increase the performance of cloud services. A mechanism has been designed [24], to handle the service of requests of customers. Cloud providers pay back the customers when the performance of the services go below a particular threshold. This is not feasible and is unsatisfactory for the customers as a drop in performance may affect their work adversely and the cloud provider pays back to the customers only in the next period. In [25], a technique has been designed to allocate VMs in which there are set of classes of cloud users and various spectrum's of resource allocations.

Cloud provides various types of resources which were provided to users in the form of services. Users access these services for their tasks which indirectly utilize the resources to carry out the service. The cloud services may be requiring a lot of I/O related operations or processing power depending on the type of service offered. Resources at the cloud can perform different types of tasks like processing of data, rendering of high-resolution complex graphics, etc. A mechanism had designed [26], to perform resource allocation in the cloud to maintain the performance metrics between the cloud provider and the user. Virtualization technology in the cloud is used to offer scalability of resources to users. The VMs with different configurations can be grouped together to provide specialized services and thereby reduce the wastage of resources. A technique has been designed [27], for resource allocation in the cloud using the concept of the uncertainty principle and the formation of the coalition. Some types of applications were composed of large number of discrete tasks. These applications require a high amount of computing power and heterogeneous resources which can be provided with supercomputers, clusters, grids, and cloud computing. We have to take into consideration the factors that the allocation of resources among users must be fair, the utilization of the system should be maximized. The response time for the user should be decreased. A mechanism has been proposed [28], for resource allocation in multiple users and multiple application environment with heterogeneous resources and computing platforms.

Cloud computing has been used by business enterprises to carry out their business processes. This helps the business enterprises to reduce their costs on resource procurement and also increases their performance in terms of delivering business services. A large number of mechanisms for resource allocation have been designed but very few have been developed for verifying the resource allocation in the cloud [29]. A framework was proposed [30], that allows ensuring proper allocation of resources in the cloud, from the perspective of business processes.

Faiz et al. [31] proposed the camera identification process using conditional probability features and Apache Hadoop to overcome the manipulation of the digital images. Shamshirband et al. [32] designed intrusion detection and prevention systems (IDPS) to date, diverse soft computing machine learning techniques. This solves wireless environment intrusion recognition issues in the cloud computing. Elliptic curve cryptography (ECC) used for multi-user message broadcasting and to avoid attackers in the WSNs [33]. It improves privacy and user untracking in WSNs. These methods could not find the minimum cost of the distributed cloud service providers.

Tajiki et al. [34] proposed a state-of-the-art traffic engineering model for SDN-MPLS network. The proposed design is more efficient compared to the traditional MPLS networks. Al-Janabi et al. [35] developed a compression algorithm which can achieve highest compression percentage and high quality for video compression. The developed approach is more effective compared to its previous counterparts.

3 Problem definition

The PBRA and the hybrid algorithm [6], have utilized the reservation and the on-demand schemes of resource allocation provided by the CSPs. These algorithms have taken into consideration only a single CSP. By taking multiple CSPs, we can exploit the different tariff rates provided by different CSPs.

3.1 Problem statement

Our problem statement can be stated as “Only one CSP is used for resource allocation using PBRA and hybrid algorithms [6]”. In distributed systems, more CSPs exist, which can be exploited to find the minimum cost CSP. In the existing system, the price is more for the allocation of resources. This is the main drawback of the existing system.

3.2 Objective

The objective of the proposed MCRA and the hybrid MCRA algorithms is to facilitate VoD application providers to allocate resources in the cloud with minimum cost while maintaining QoS of the video content. To accomplish this, the different tariff rates of different CSPs are exploited to further reduce the cost.

4 System model

The system that we have designed for resource allocation in of the cloud for VoD applications are composed of the following elements as illustrated in Fig. 1.

- *Demand prediction* component, that forecasts the demand for resources for every video over a given future span of time. The prediction of demand can be performed by various methods [8, 9].
- *Cloud broker* component, allocates the suitable number of resources and reserves the time for which the resources are allocated in the cloud. The broker performs the resource allocation depending on the demand prediction it receives. The broker interacts with the different CSPs to decide which one to select for resource allocation. It makes use of the MCRA or the hybrid MCRA algorithm for resource allocation in the cloud. Both the demand prediction component and the broker are situated in the VoD application provider’s location.
- *Cloud provider* is responsible for providing the resources to the VoD application provider for hosting the VoD application. Cloud provider directly delivers the video content to the viewers.

We consider multiple CSPs for the allocation of resources. Different CSPs provide different tariff prices for resource allocation. This can be exploited to select the CSP [36] which gives the minimum cost for resource allocation. CSPs presents their tariff prices in tabular form, as shown in Table 1. The MCRA or hybrid MCRA algorithm iterates through the set of CSPs to find the one which yields a minimum cost. For each CSP, the algorithm performs resource allocation such that the number of resources allocates and the time for which the resources are allocated, costs the minimum. The MCRA algorithm makes use of the reservation scheme while the hybrid MCRA algorithm makes use of both the reservation and the on-demand schemes. These schemes are provided by most of the CSPs. In the reservation plan, the CSP offers discount prices based on the number of resources reserved and the time for which the resources are reserved. Figure 3 shows a typical example of the comparison of the tariff functions of three different CSPs, for 3 units of resources. From this comparison, we observe that CSP ‘A’have tariff prices lesser than the tariff prices of CSPs ‘B’and ‘C’. Hence, CSP ‘A’can be a potential selection for resource allocation.

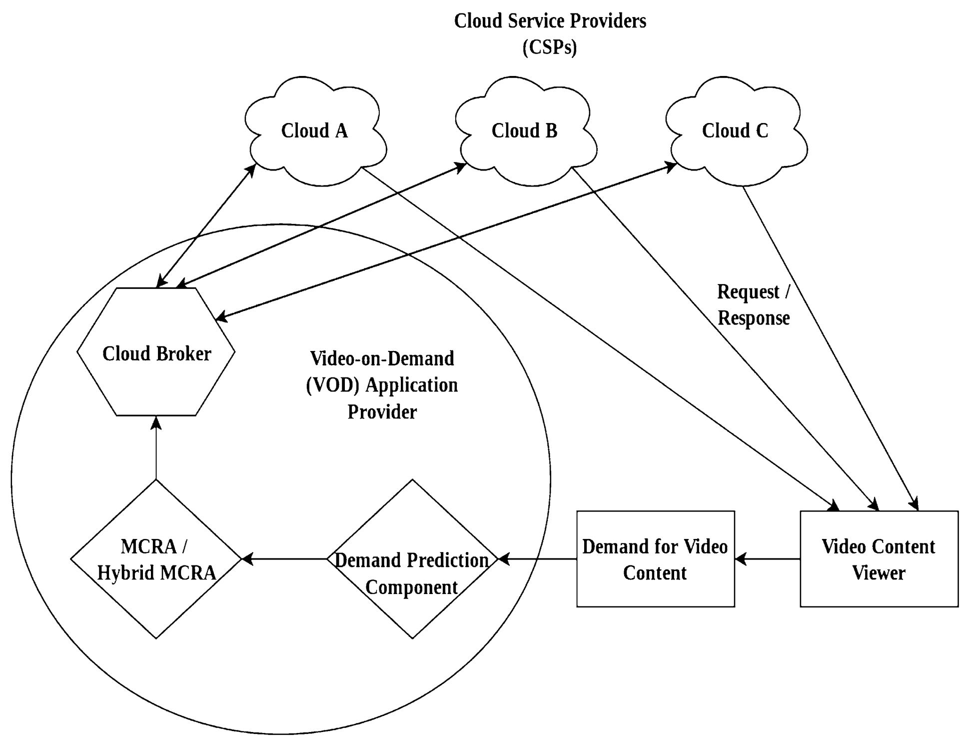

Fig. 1 An overview of the system model

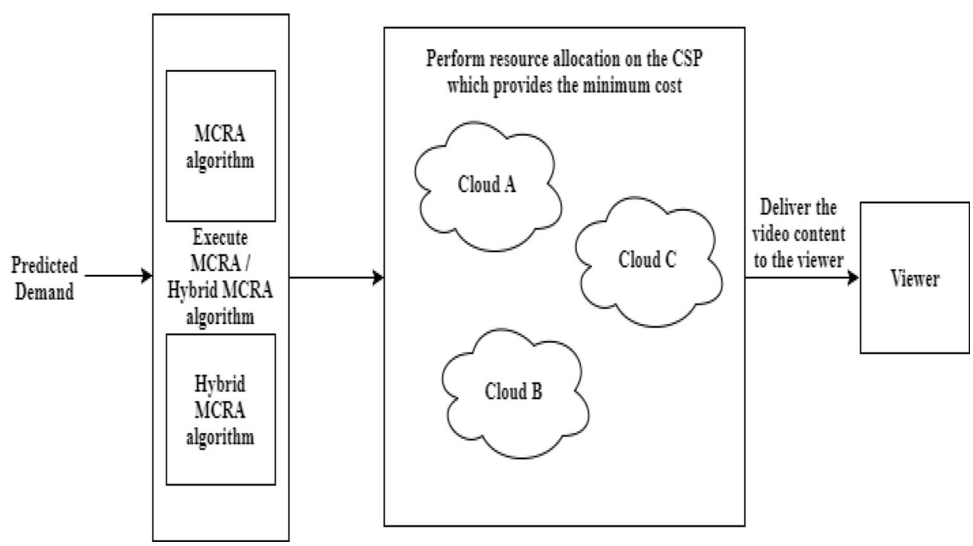

Fig. 2 The detailed design of the system model

A detailed design of the system model is given in Fig. 2. The input to the system predicts the demand for any future period of time P. Either the MCRA or Hybrid MCRA algorithm can be executed for resource allocation. The choice of algorithm to be executed solely depends on the VoD application provider. Once the algorithm is executed, the allocation of resources is performed on the CSP which gives a minimum cost. Finally, the viewer receives the video content directly from the cloud. The execution of the algorithms happens at the site of the VoD application provider. Only the allocation of the resources is passed to the CSP. The process of determination of the number of resources and the reservation time is hidden from the CSP (Fig. 3).

The resource allocation is performed with some probability η. At any point in time, the probability that the demand of the video content is less than or equal to the allocated resources must always be less than the probability η. The value of η has to be carefully decided such that it will not be very high or very low. A very high value implies that the allocated resources have a high probability of sustaining the QoS of the video content while it may

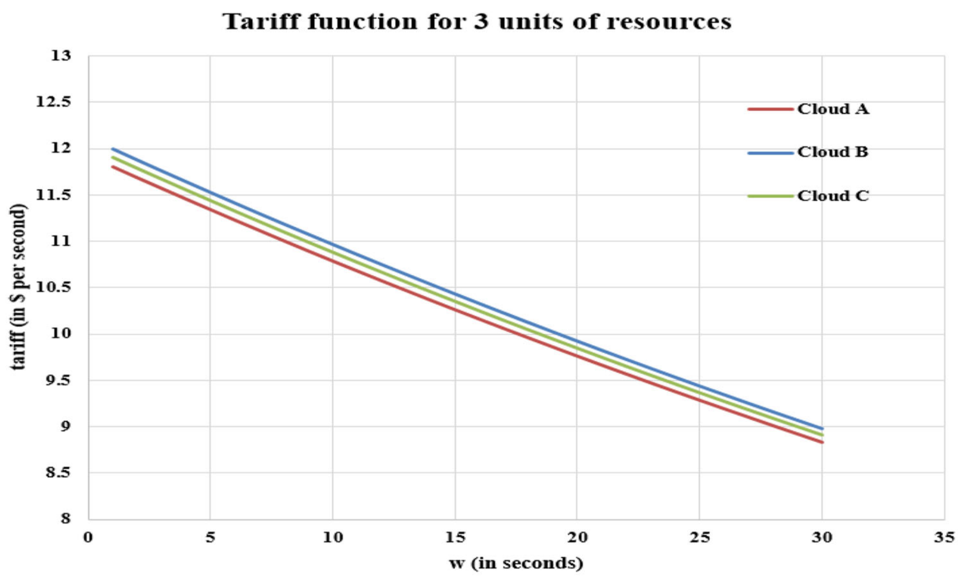

Fig. 3 An example of tariff function for three different CSPs

lead to over-subscription which leads to wastage of resources. On the other hand, a very low value implies that the allocated resources have a low probability of sustaining the QoS of the video content while it may lead to under-subscription which leads to degradation of QoS. Hence, an optimal value of η has to be selected, such that sufficient resources are allocated to handle the demand without degradation in the QoS. The demand at any instant of time t is denoted by *Demand(t)* and the allocation is denoted by *Alc(t)*.

$$Probability(Demand(t) \leq Alc(t)) \geq \eta. \quad (1)$$

The resource allocation at any instant of time has to satisfy the constraint in Eq. (1).

5 Multiple cloud resource allocation (MCRA) algorithm

The MCRA algorithm uses only the reservation scheme provided by CSPs for resource allocation. The demand for a video content follows a lognormal distribution. We assume an example of predicted demand for a future period of time $P = 30$ is shown in Fig. 4. The predicted demand can be increased or decreased depending on the requirement of the user. This has to be changed by the Administrator of the system. When a video content becomes available on the VoD application, the demand increases slowly. This is because of the lower number of viewers aware of the video content. As these viewers share the video content across to other viewers the demand increases at time 10. During the time period from 11 to 17 the demand is constant. Therefore the curve is linear. Following the time 17 the demand increases steadily up to timestamp 21. When the lifespan of the video content expires the demand slowly diminishes and becomes stable (time period between 20 and 30). The lifespan of a video content expires when all the potential viewers have viewed the content.

5.1 Selection of CSP

The selection of CSP is based on the availability of free resources on the CSP and the total cost of resource allocation for the period for which allocation is performed. Every CSP has two parameters *flag* and *fullyAllocated*. The *flag* parameter identifies whether the CSP is involved in the allocation or not whereas the *fullyAllocated* parameter identifies if the CSP has any free resources or not. These two parameters help in the process of resource allocation.

The process of selection of CSP is listed in Algorithm 1. For the combination of CSPs for which the condition satisfies the MCRA or Hybrid MCRA algorithm is executed. The type of algorithm executed depends on the choice of the VoD application provider. Each of the algorithms has their merits and demerits which will be discussed later in the paper.

Consider three CSPs A, B and C. Each of these CSP have parameters *flag* and *fullyAllocated*. The CSP which has *flag* is set to *true* then *fullyAllocated* is set to *false* takes part in the resource allocation. The selection of the allocation algorithm depends on the choice of the user. The user is provided with the option to select either the MCRA or hybrid MCRA algorithm.

The algorithm can be scaled to any number of CSPs. Consider three CSPs A, B and C. A has 20 units of resources, B has 50 units of resources and C has 30 units of resources. Imagine, CSP A is selected for allocation when

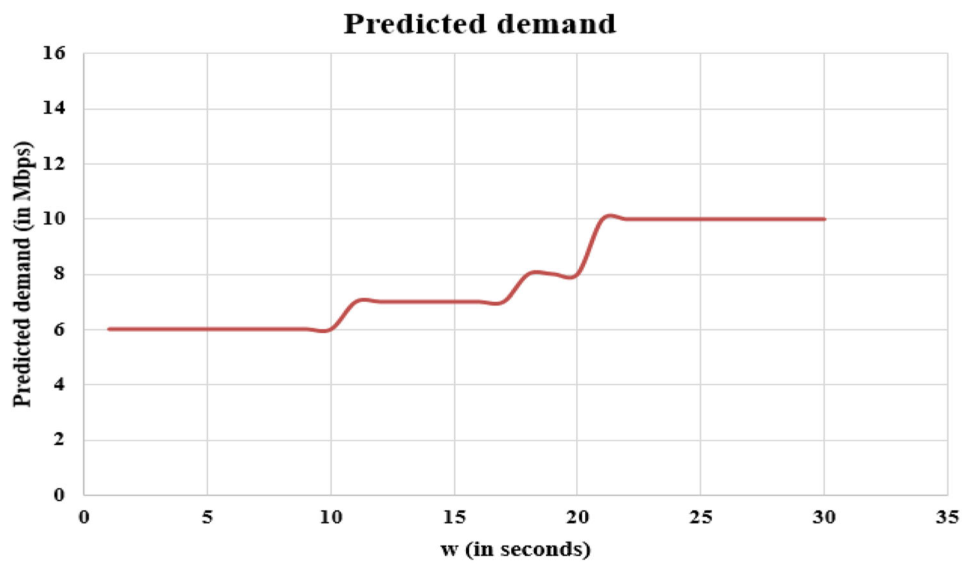

Fig. 4 An example of the predicted demand for a future period of time $P = 30$

the algorithm is executed. All the resources on CSP A are exhausted and there are no more resources for allocation. During the next allocation session, CSPs B and C are also exhausted. In this case, the algorithm does not have any CSP to perform the allocation. We need to add to the list of CSPs involved in the selection process. This supports for scalability of the algorithm. This can be extended to n number of CSPs based on the requirement. The time complexity of Algorithm 1 is $\theta(n^3)$.

Algorithm 1 Selection of CSP for allocation of resources

Consider n number of CSPs C_i, C_j, ... , C_n which can be used for resource allocation
selectedCSPs is an array of selected CSPs which can be involved in the resource allocation
For each CSP,
Define:
flag that identifies whether the CSP is involved in the resource allocation or not,
fullyAllocated that determines if there are unallocated resources available on the CSP
for each CSP C_i **do**
  if $C_i.flag$ **and** $!C_i.fullyAllocated$ **then**
    $selectedCSPs \leftarrow selectedCSPs + C_i$
  end if
end for
if *selectedCSPs* **then**
  Run MCRA or Hybrid MCRA algorithm with the CSPs in *selectedCSPs*
else
  Resource allocation cannot be performed as there are either no CSPs which can be involved in the resource allocation or none of the involved CSPs have unallocated resources available with them
end if

5.2 Design of the algorithm

The following assumptions have been taken into consideration in our analysis:

- The CSP allocates the resources immediately when it receives the request from the VoD application provider without any delay.
- The CSP is responsible for the delivery of video content to the viewers situated at different geographical locations at the guaranteed data-rate.
- The VoD application provider is charged for the resources to be allocated, at the time of making the request for the resources. The VoD application provider cannot modify or change a request already submitted to the CSP.
- The CSP can allocate the only a discrete number of resources in the cloud. There is a minimum reservation time for which the resources can be allocated.
- The size of the video content depends on the CSP selected for allocation. For example, the iCloud service provider allows 1000 videos per hour and 10,000 videos per day from a single user.
- A user can allocate resources only on a single CSP during a particular allocation session when the algorithm is executed to determine the CSP with the minimum cost. In other words, a user cannot allocate across different CSPs during a particular allocation session.

Our Algorithm 2 is based on time slots. The allocation is performed for a future period of time P. The resource allocation algorithm begins with the first time slot and proceeds incrementally with the next time slot until the future period of time P. For each time slot, the number of resources to be reserved and the time for which the reservation is to be made is decided. The allocation which gives the minimum cost for the particular number of resources and reservation time is chosen. This process is repeated for each CSP, and the one which gives the minimum cost is chosen. Each time slot is also called as a *window*. The size of each window is denoted as *win*. The window size is the duration of the time slot. The allocation of resources is performed with some probability η.

The number of resources reserved in window j is denoted by Alc_j and the size of the window is denoted by win_j. The Cost of resource allocation is denoted by C, which is the product of the tariff function $tar(win_j, Alc_j)$ and the window size win_j, as given in Eq. (2).

$$C(win_j, Alc_j) = tar(win_j, Alc_j) * win_j, \tag{2}$$

where $tar(win_j, Alc_j)$ is a function of both the number of resources reserved and the time for which the resources are reserved, i.e., Alc_j and win_j. *tar* is measured in $ per unit time. The resource allocation has to be performed such that the value of C is minimized.

The objective of the algorithm can be mathematically formulated as, reduce the cost $C(win_j, Alc_j)\ \forall j$ of resource allocation subject to the constraint in Eq. (3),

$$Probability(Demand(t) \leq Alc(t)) \geq \eta,\ \forall t \in P. \tag{3}$$

The formula which is used in our resource allocation algorithm in window j is given by Eq. (4).

$$\int_0^{Alcj} \frac{1}{x.\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}(\frac{ln(x)-\mu_{max}}{\sigma})^2} dx = \eta, \tag{4}$$

where σ is the variance of the lognormal distribution and μ_{max} is the maximum predicted demand during that particular window j.

The CSP allows only discrete number of resources to be reserved. A minimum duration of the window win_{mn} is enforced by the CSP which serves as the duration baseline for reservation. The CSP cannot reserve any resources for duration below win_{mn}. Therefore, the algorithm proceeds incrementally, beginning with the first iteration whose window size is equal to win_{mn}. The algorithm computes the cost rate for this iteration and proceeds to the next iteration. In every iteration, the window size will be the product of win_{mn} and a positive integer i. To make the process simple, we have used a test window whose size is denoted as win_i which is equated to the size of the particular window in every iteration. The cost rate is computed as $X_i = tar(win_i, Alc_i)$, where i is the iteration index, and Alc_i is calculated using Eq. (4) for *Alc*. Once the algorithm is executed for the entire future period of time P, the iteration k which gives the minimum cost rate X_F is found. The corresponding values of window size win_k and number of resources reserved Alc_k are assigned to win_j and Alc_j respectively. These are the final computed values for window *j*. The MCRA algorithm is listed in Algorithm 2. The time complexity of algorithm 2 is $\theta(n^2)$.

Algorithm 2 Multiple Cloud Resource Allocation Algorithm.

```
The resource allocation is performed for a future period
  of time P
Define:
win_i which is the size of the test window for every itera-
  tion i
win_mn which is the minimum duration for any window,
  enforced by the CSP below which duration no resources
  can be reserved
j which denotes the window j
To calculate win and Alc for every window j, do
win_i ← 0
i ← 1
while win_i ≤ P do
    win_i = win_i + win_mn
    Evaluate μ_max_i
    Evaluate Alc_i using Eq. (4) for Alc
    X_i = tar(win_i, Alc_i)
    i ← i + 1
end while
X_F = argmn(X_i ∀i)
Get the iteration index k that corresponds to X_F
win_j ← win_k
Alc_j ← Alc_k
j ← j + 1
```

5.3 An example of MCRA algorithm

Consider the predicted demand for a video content shown in Fig. 4. The corresponding reservation of resources using the MCRA algorithm is shown in Fig. 5 and the values are listed in Table 2. The resource reservation is performed for the entire duration of future period of time $P = 30$. We have chosen $win_{mn} = 1$. The steps listed in the MCRA algorithm are executed. Out of all values of X_i, the minimum value is found, and the corresponding values of window size and number of resources reserved are assigned to w_j and Alc_j respectively. In the table, we see that the iteration 20 gives the minimum value of X_i. Hence, $win_j = 20$ and $Alc_j = 5$.

In the Fig. 5, we observe that the resources reserved are less compared to the predicted demand for the entire duration of the future period of time P. This is a drawback of the MCRA algorithm. There are two drawbacks to the MCRA algorithm - under-subscription and over-subscription. Under-subscription occurs when the number of resources reserved is less than the predicted demand whereas over-subscription occurs when the number of

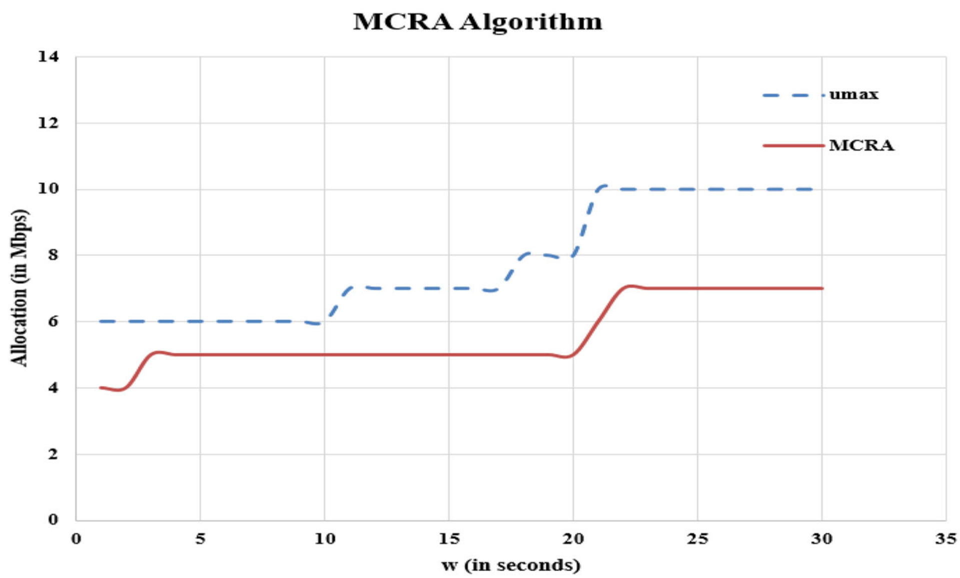

Fig. 5 An example of MCRA algorithm

Table 2 An example of MCRA algorithm executed for a future period of time P = 30

Iteration (i)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
w_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
μ_{max}	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7
Alc_i	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5
X_i	12.80	12.67	13.53	13.39	13.26	13.13	13.00	12.87	12.74	12.61	12.49	12.36	12.24	12.12	12.00
Iteration (i)	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
w_i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
μ_{max}	7	7	8	8	8	10	10	10	10	10	10	10	10	10	10
Alc_i	5	5	5	5	5	6	7	7	7	7	7	7	7	7	7
X_i	11.88	11.76	11.64	11.53	11.41	12.12	12.81	12.68	12.55	12.43	12.30	12.18	12.06	11.94	11.82

resources reserved is more than the predicted demand. Under-subscription causes a degradation in the QoS of the video content whereas over-subscription causes unnecessary wastage of resources. The number of resources reserved depends on the probability η discussed in Sect. 4. Increasing the value of η increases the number of resources reserved and hence reduces the problem of under-subscription. But on the other hand, a high value of η may increase the chances of over-subscription. Hence, an optimal value of η has to be chosen such that a balance is achieved between the predicted demand and the number of resources reserved. The QoS of the video content is maintained and the wastage of resources is mitigated. Although we do not solve the problems of under-subscription and over-subscription completely, we negotiate them to some extent. These problems are solved in the Hybrid MCRA algorithm where we use both the reservations and the on-demand schemes provided by CSPs. This algorithm will be discussed in Sect. 6.

6 Hybrid MCRA algorithm

The hybrid MCRA algorithm is an extension of the MCRA algorithm. In the MCRA algorithm, we use only the reservation scheme offered by the CSPs. But in the Hybrid MCRA algorithm, we make use of both the reservation and on-demand schemes provided by CSPs. In the reservation scheme, time-discount tariff prices are offered. This type of discount prices are not offered in the on-demand scheme and the prices of resources are higher compared to the reservation scheme. This algorithm overcomes the drawbacks of MCRA algorithm, by using the benefits of the on-demand scheme. In the on-demand scheme, resources can be allocated on-the-fly. The algorithm reserves resources for the predicted demand such that the number of resources reserved is less than or equal to the demand. When the number of resources reserved is less than the predicted demand, the algorithm makes use of an on-demand scheme and allocates resources on-the-fly to make up for the remaining resources. In this way, the problems of under-subscription and over-subscription are completely solved.

6.1 Design of the algorithm

The total cost of resource allocation in the Hybrid MCRA algorithm will be a sum of the cost for resources reserved through reservation scheme and the cost of the resources allocated through the on-demand scheme. The formula for the total cost is shown in Eq. (5).

$$Cost_{total} = \sum_{j} (Cost_{RSV_j} + Cost_{OD_j}), \tag{5}$$

where $Cost_{total}$ is the total cost for resource allocation, $Cost_{RSV_j}$ is the cost of reserving resources using the reservation scheme, and $Cost_{OD_j}$ is the cost of allocating resources using the on-demand scheme. Alc_{RSV_j} is the number of resources reserved in window j using the reservation scheme and Alc_{OD_j} is the number of resources allocated in window j using the on-demand scheme. $tar(win_{RSV_j}, Alc_{RSV_j})$ is the tariff for the reservation scheme and $tar(Alc_{OD_j})$ is the tariff for the on-demand scheme. The tariff of the reservation scheme depends on both the window size and the number of resources reserved while the tariff of the on-demand scheme depends only on the window size. This is because of the time-discount tariff prices are offered only in the reservation scheme and not in the on-demand scheme.

In hybrid MCRA algorithm, there are three possible cases which may arise in the allocation of resources.

(i) The on-demand scheme is used only when the number of resources reserved using the reservation scheme is less than the predicted demand.

(ii) If the number of resources reserved using the reservation scheme is equal to the predicted demand, then there is no need to use the on-demand scheme. The resources reserved are enough to handle the predicted demand. In this case, the total cost will be the cost of using the reservation scheme only, and the cost of using the on-demand scheme will be zero.

(iii) When the resources reserved using the reservation scheme is less compare to the predicted demand, the on-demand scheme is used to fulfill the remaining demand. In this case, the total cost will include both the components of cost, i.e., the cost using reservation scheme and the cost using on-demand scheme.

Algorithm 3 Hybrid MCRA Algorithm.

```
Define:
Let S which is the set of all η values for which the algorithm will be executed
For every window j, do
for every value η in the set S, do
    i ← 1,
    Execute Algorithm 2 to evaluate the values of win_j and Alc_RSV_j using the reservation scheme,
    Calculate Alc_OD_j = μ_max − Alc_RSV_j,
    Calculate X_i = tar(w_RSV_j, Alc_RSV_j) + tar(Alc_OD_j),
    i ← i + 1,
end for
Y_F = argmn(X_i ∀i),
Get the iteration index k that corresponds to Y_F,
η ← η_k,
Alc_RSV_j ← Alc_RSV_k,
Alc_OD_j ← Alc_OD_k
```

The number of resources reserved using the reservation scheme depends on the probability η. As stated previously, higher the value of η higher in the number of resources reserved. But, a high value may cause wastage of resources. Hence, an optimal value has to be chosen. To facilitate this, we have used a set S of η values. For each value of η, Algorithm 2 is executed and the optimal values of window size and number of resources to be reserved are calculated. If the quantity of reserved resources is less than the predicted demand, then the remaining resources is allocated using the on-demand scheme. This can be easily calculated by $Alc_{OD_j} = \mu_{max} - Alc_{RSV_j}$. The total cost rate for allocation using the Hybrid MCRA algorithm is given by $X_i = tar(win_{RSV_j}, Alc_{RSV_j}) + tar(Alc_{OD_j})$. The η value is evaluated which gives the minimum cost, and the corresponding values of Alc_{RSV_j} and Alc_{OD_j} are reserved and allocated using the reservation scheme and on-demand scheme respectively. The Hybrid MCRA algorithm is listed in Algorithm 2. The time complexity of Algorithm 2 is $\theta(nj^2)$.

6.2 An example of hybrid MCRA algorithm

Consider the predicted demand for a video content shown in Fig. 4. The corresponding reservation of resources using the Hybrid MCRA algorithm is shown in Fig. 6 and the values are listed in Table 3. We have taken the set S = {0.75, 0.8, 0.9, 0.95}. For each value of S Algorithm 2 is executed to find optimal values of window size win_j and quantity of reserved resources Alc_{RSV_j} for each window j. If the reserved resources is less than the predicted demand, the remaining resources will be allocate using the on-demand scheme as Alc_{OD_j} = μ_{max} - Alc_{RSV_j}. Hence, the total

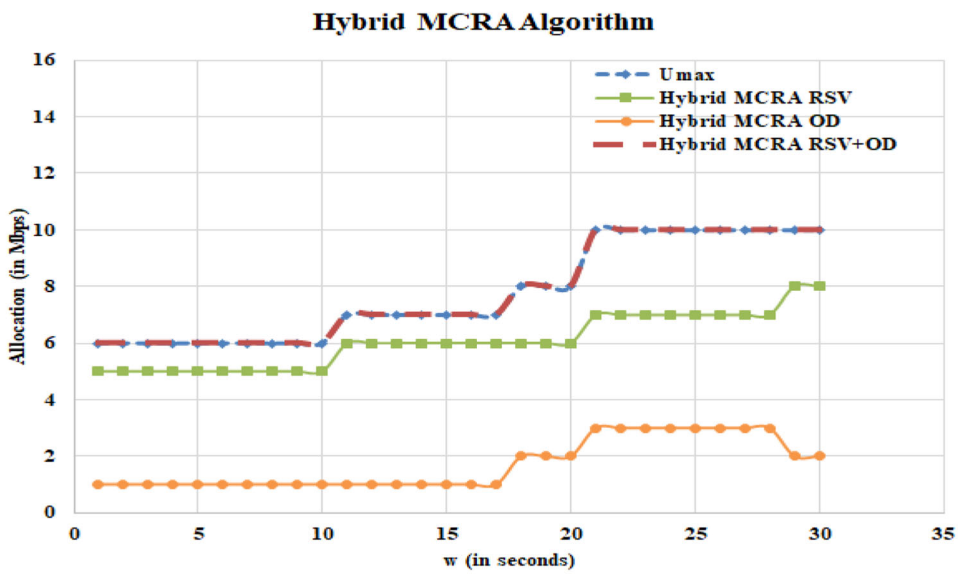

Fig. 6 An example of hybrid MCRA algorithm

Table 3 An example of hybrid MCRA algorithm executed for a future period of time P = 30

Iteration (i)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
w_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
μ_{max}	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7
Alc_{RSV_i}	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6
Alc_{OD_i}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$Alc_{(RSV+OD)_i}$	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7
X_i	23.60	23.46	23.33	23.19	23.06	22.93	22.80	22.67	22.54	22.41	23.19	23.06	22.93	22.79	22.67
Iteration (i)	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
w_i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
μ_{max}	7	7	8	8	8	10	10	10	10	10	10	10	10	10	10
Alc_{RSV_i}	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8
Alc_{OD_i}	1	1	2	2	2	3	3	3	3	3	3	3	3	2	2
$Alc_{(RSV+OD)_i}$	7	7	8	8	8	10	10	10	10	10	10	10	10	10	10
X_i	22.54	22.41	32.08	31.96	31.84	42.33	42.21	42.08	41.95	41.83	41.70	41.58	41.46	32.39	32.16

allocated resources is the sum of resources reserved using reservation scheme Alc_{RSV_j} and resources allocated using on-demand scheme Alc_{OD_j}. The cost rate X_i is calculated for each value of S in each window j. Out of all the values of X_i, the minimum value is found, and the corresponding values of window size and allocated resources are assigned to win_j, and Alc_{RSV_j} and Alc_{OD_j}. In the table, we see that the iteration 10 gives the minimum value of X_i. Hence $win_j = 10, Alc_{RSV_j} = 5$ and $Alc_{OD_j} = 1$. $Alc_{(RSV+OD)_i} = 6$, which is equal to the maximum demand μ_{max} in that window. The hybrid MCRA algorithm facilitates the selection of an optimal value of η. This is done easily by selecting the value η from the set S which gives the minimum cost rate X_i. We can clearly justify that the Hybrid MCRA algorithm overcomes the drawbacks of under-subscription and over-subscription that occurred with the MCRA algorithm. The different advantages of both the reservation scheme and the on-demand scheme are utilized to maintain the QoS of video content and on the other hand reduce wastage of resources.

7 Analysis of the algorithms and results

In this section, we analyze the MCRA and the hybrid MCRA algorithms in terms of allocation of resources and cost. We have compared the two algorithms with their

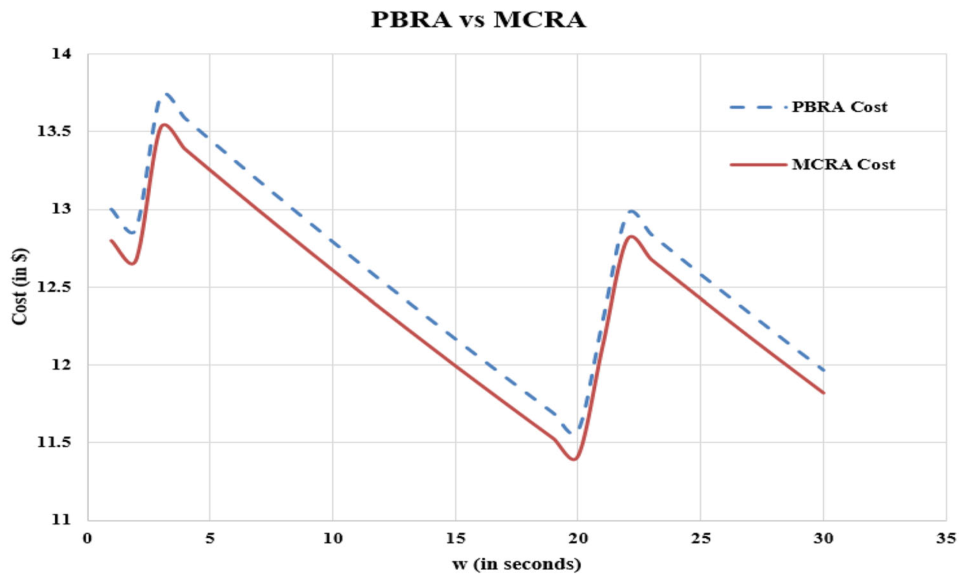

Fig. 7 Cost comparison of PBRA and MCRA algorithms

single cloud counterparts the PBRA and hybrid PBRA algorithms [6], respectively. The MCRA and hybrid MCRA algorithms are the enhanced versions which are designed to include multiple CSPs and have taken into consideration the availability of resources with the CSPs.

7.1 Cost comparison of PBRA and MCRA algorithms

A comparison of PBRA and MCRA algorithms in terms of cost is shown in Fig. 7. We observe that the MCRA algorithm is better compared to the PBRA algorithm in terms of cost. This is because, in the MCRA algorithm, we choose the CSP which gives the minimum cost. This is the lack of PBRA algorithm, as the user does not have control over the selection of CSP. Once the CSP is selected, it is fixed and cannot be changed automatically. If the CSP has to be changed, it has to be done manually. But in the case of the MCRA algorithm, we have a set of CSPs. The MCRA algorithm determines the CSP which provides the minimum cost for resource allocation. The two peak values (time 3 and 22 in Fig. 7) in the PBRA and MCRA curves indicate the change in the number of resources. For a given number of resources, the price gradually decreases as the duration of reservation time increases. This resembles in the Fig. 7, for time between 4 to 19, and 23 to 30.

7.2 Cost comparison of hybrid PBRA and hybrid MCRA algorithms

The hybrid PBRA and hybrid MCRA algorithms are compared in terms of cost as illustrated in Fig. 8. From the figure, we observe that the Hybrid MCRA algorithm is better compared to the hybrid PBRA algorithm although the difference is very small. The hybrid PBRA and hybrid MCRA graphs are quite invariant between times 0 and 17. This shows that the number of resources is relatively stable for this duration of time, as is the demand. The increase in the cost after time 17 and 20 show that the number of resources has increased depending on the demand. The demand is stable again for the time between 22 and 27 and hence the invariance in the curves. The fall in the curves between time 28 and 29 is because of the reduction in the number of resources allocated using the on-demand scheme. The number of resources reserved using the reservation scheme is increased during this time. This can be seen in Table 3 for iteration 29. Furthermore, the figure gives us an insight into the cost patterns of the reservation and the on-demand schemes. The prices of the on-demand resources are quite expensive compared to that of the reservation resources. Hence, the reservation scheme has to be utilized to the maximum extent while the on-demand scheme has to be limited to peak times only.

7.3 Allocation comparison of MCRA and hybrid MCRA algorithms

A comparison of the MCRA and the hybrid MCRA algorithms in terms of allocation of resources has been depicted in Fig. 9. The figure clearly implies that the hybrid MCRA algorithm is better than the MCRA algorithm in terms of allocation of resources, as its curve is closer to the predicted demand curve. The predicted demand is shown in Fig. 4 has been taken into consideration for this comparison.

The MCRA algorithm has allocated less number of resources compared to the hybrid MCRA algorithm. When compared to the predicted demand, we observe that the

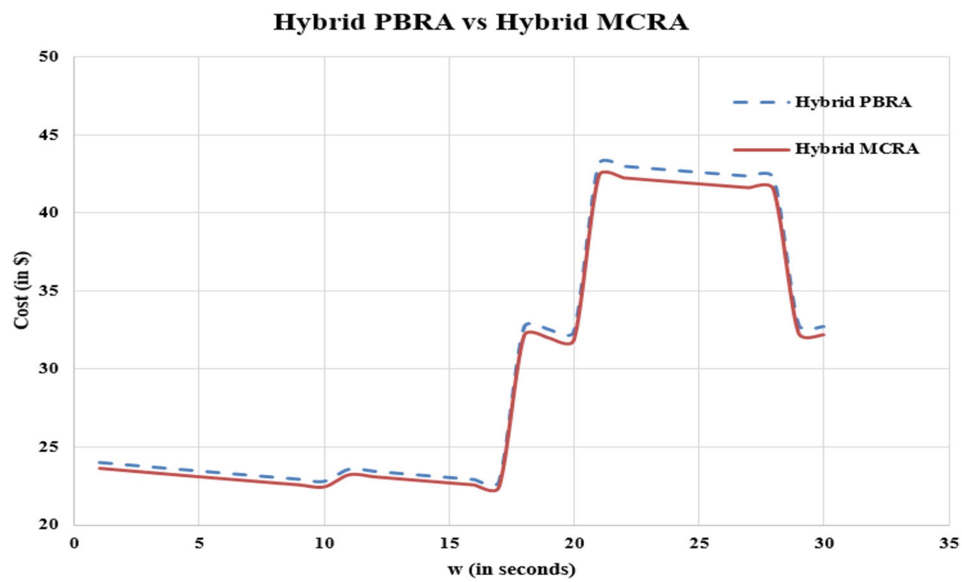

Fig. 8 Cost comparison of hybrid PBRA and hybrid MCRA algorithms

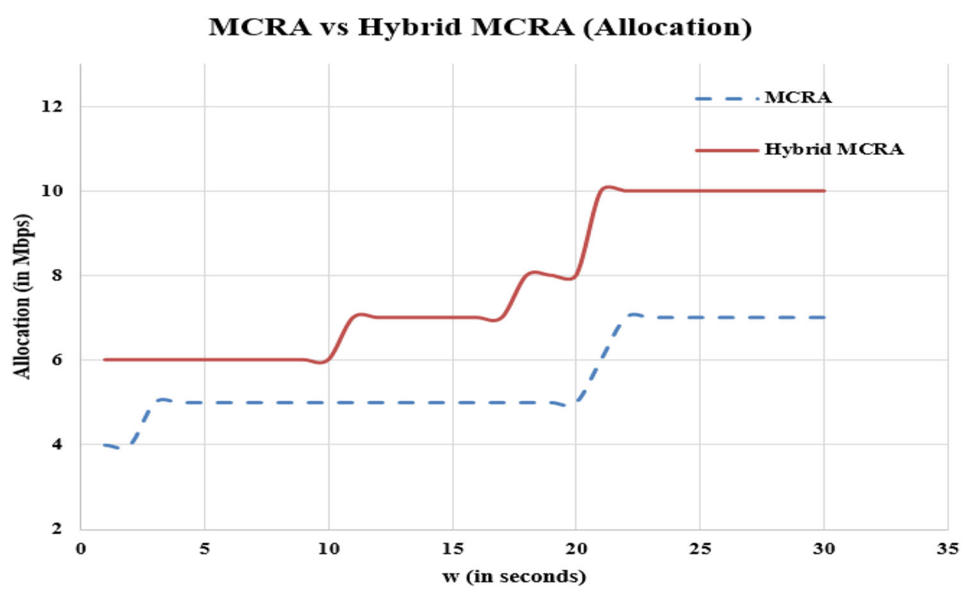

Fig. 9 Allocation comparison of MCRA and hybrid MCRA algorithms

curve of the Hybrid MCRA algorithm overlaps the curve of the predicted demand. This demonstrates the QoS aware characteristic of the Hybrid MCRA algorithm. The cost minimization property is exhibited by allocating resources sufficient enough to handle the demand. The curve of the MCRA algorithm on the other hand is below the predicted demand curve. This exposes the drawback of the MCRA algorithm, the under-subscription problem, where the number of resources reserved is insufficient to fulfill the demand. The MCRA algorithm faces another problem of over-subscription, where the number of resources reserved is much more than the predicted demand. This causes a wastage of resources.

7.4 Cost comparison of MCRA and hybrid MCRA algorithms

The MCRA and the hybrid MCRA algorithms have been compared in terms of cost as illustrated in Fig. 10. We see that the MCRA and the hybrid MCRA graphs are quite different from one another. This is because in the MCRA algorithm we are using only the reservation

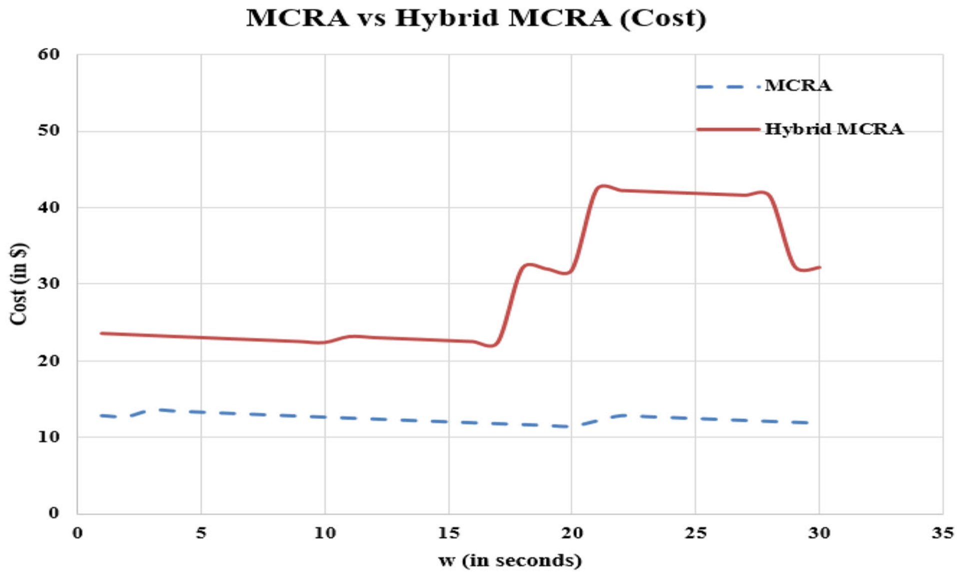

Fig. 10 Cost comparison of MCRA and hybrid MCRA algorithms

scheme whereas in the hybrid MCRA algorithm we are using both the reservation and the on-demand schemes.

The MCRA curve is quite invariant for the entire duration of resource allocation. The small increments at time 3 and 22, followed by a gradual decrease, represent the change in the quantity of reserved resources at these times. The gradual decrease demonstrates the time-discount tariff prices of the reservation scheme, resources reserved for longer duration's have discounted prices. On the other hand, the hybrid MCRA curve is quite varying. The increased values at times 10, 18 and 21 seconds indicate the change in the number of resources allocated using the on-demand scheme. The fall in the curve at times 16, 20, 28 and 30 s indicate the change in the number of resources of both the schemes, as seen in Table 3 for iteration 29. In this iteration, the number of resources reserved using the reservation scheme has gone down, while the number of resources allocated using the on-demand scheme has gone up.

8 Conclusions

In the existing system, the PBRA and hybrid PBRA algorithms [6] are used a single CSP for resource allocation. The disadvantage of using one CSP is high cost. In distributed systems more CSPs are available. These algorithms do not involve more than one CSP for the resource allocation in cloud. To avoid these drawbacks, we have designed two algorithms—the MCRA and Hybrid MCRA algorithms, for allocation of resources in of cloud for VoD applications. The two basic schemes provided by most CSPs—(i) the reservation scheme and (ii) the on-demand scheme have been utilized to reduce the cost of resource allocation and at the same time maintain the QoS of the video content. We have incorporated multiple CSPs in our design. This further reduces the *cost* as different CSPs provide different tariff rates which can be exploited to select the CSP provides the minimum cost. The analysis of the algorithms shows that they maintain a balance between the cost and the QoS of the video content. The cost is minimized while the QoS is not compromised. The availability of resources with the CSP has been taken into consideration while the resources are allocated.

This can be further improvised by considering the geographical aspects of the data centers and application servers. The allocation of resources can be done on the basis of the distance between the viewer of the video content and the physical location of the servers. This would improve the latency of the video content and make the delivery process more efficient.

References

1. Beloglazov, A., Abawajy, J., Buyya, R.: Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener. Comput. Syst. **28**(5), 755–768 (2012)
2. Liu, Y., Guo, Y., Liang, C.: A survey on peer-to-peer video streaming systems. Peer-to-Peer Netw. Appl. **1**(1), 18–28 (2008)
3. Muthi Reddy, P., Manjula, S.H., Venugopal, K.R.: Secure data sharing in cloud computing: a comprehensive review. Int. J. Comput. (IJC) **25**(1), 80–115 (2017)
4. Iqbal, W., Dailey, M.N., Carrera, D.: Unsupervised learning of dynamic resource provisioning policies for cloud-hosted multitier web applications. IEEE Syst. J. **10**(4), 1435–1446 (2016)
5. Marinescu, D.C., Paya, A., Morrison, J.P.: A cloud reservation system for big data applications. IEEE Trans. Parallel Distrib. Syst. **28**(3), 606–618 (2017)

6. Alasaad, A., Shafiee, K., Behairy, H.M., Leung, V.C.: Innovative schemes for resource allocation in the cloud for media streaming applications. IEEE Trans. Parallel Distrib. Syst. **26**(4), 1021–1033 (2015)
7. Armstrong, M., Vickers, J.: Competitive non-linear pricing and bundling. Rev. Econ. Studies **77**(1), 30–60 (2010)
8. Peichang, S., Huaimin, W., Gang, Y., Fengshun, L., Tianzuo, W.: Prediction-based federated management of multi-scale resources in cloud. AISS: Adv. Inf. Sci. Serv. Sci. **4**(6), 324–334 (2012)
9. Gürsun, G., Crovella, M., and Matta, I.: Describing and forecasting video access patterns. In: INFOCOM, 2011 Proceedings IEEE. IEEE, pp. 16–20 (2011)
10. Niu, D., Xu, H., Li, B., and Zhao, S.: Quality-assured cloud bandwidth auto-scaling for video-on-demand applications. In: INFOCOM, 2012 Proceedings IEEE. IEEE, pp. 460–468 (2012)
11. Muthireddy, P., Manjula, S.H., and Venugopal, K.R.: Energy optimization for virtual machines scheduling in cloud data centers. In: IEEEFORUM, Proceedings of IEEEForum International Conference on Computer Science, Industrial Electronics (ICCSIE), pp. 26–30 (2018)
12. Mashayekhy, L., Nejad, M.M., Grosu, D., Vasilakos, A.V.: An online mechanism for resource allocation and pricing in clouds. IEEE Trans. Comput. **65**(4), 1172–1184 (2016)
13. Wang, W., Jiang, Y., Wu, W.: Multiagent-based resource allocation for energy minimization in cloud computing systems. IEEE Trans. Syst. Man Cybern.: Syst. **47**(2), 205–220 (2017)
14. Adam, O.Y., Lee, Y.C., Zomaya, A.Y.: Constructing performance-predictable clusters with performance-varying resources of clouds. IEEE Trans. Comput. **65**(9), 2709–2724 (2016)
15. Liu, G., Shen, H., Wang, H.: Deadline guaranteed service for multi-tenant cloud storage. IEEE Trans. Parallel Distrib. Syst. **27**(10), 2851–2865 (2016)
16. Lim, N., Majumdar, S., Ashwood-Smith, P.: Mrcp-rm: a technique for resource allocation and scheduling of mapreduce jobs with deadlines. IEEE Trans. Parallel Distrib. Syst. **28**(5), 1375–1389 (2017)
17. Xiao, Z., Song, W., Chen, Q.: Dynamic resource allocation using virtual machines for cloud computing environment. IEEE Trans. Parallel Distrib. Syst. **24**(6), 1107–1117 (2013)
18. Shi, L., Zhang, Z., Robertazzi, T.: Energy-aware scheduling of embarrassingly parallel jobs and resource allocation in cloud. IEEE Trans. Parallel Distrib. Syst. **28**(6), 1607–1620 (2017)
19. Zhu, X., Wang, J., Guo, H., Zhu, D., Yang, L.T., Liu, L.: Fault-tolerant scheduling for real-time scientific workflows with elastic resource provisioning in virtualized clouds. IEEE Trans. Parallel Distrib. Syst. **27**(12), 3501–3517 (2016)
20. Zhao, Y., Jiang, H., Zhou, K., Huang, Z., Huang, P.: Dream-(l) g: a distributed grouping-based algorithm for resource assignment for bandwidth-intensive applications in the cloud. IEEE Trans. Parallel Distrib. Syst. **27**(12), 3469–3484 (2016)
21. Muthi Reddy, P., Rekha Rangappa Dasar, Tanuja, R., Manjula, S.H., and Venugopal, K.R.: Forward secrecy in authentic and anonymous cloud with time optimization. In: Proceedings of the IEEE Fifteenth International Conference on Wireless and Optical Communications Networks (WOCN 2018) **7**(2): 10036–10043 (2018). ISBN:978-1-5386-4798-1,
22. Guan, X., Wan, X., Choi, B.-Y., Song, S., Zhu, J.: Application oriented dynamic resource allocation for data centers using docker containers. IEEE Commun. Lett. **21**(3), 504–507 (2017)
23. Zhang, X., Huang, Z., Wu, C., Li, Z., Lau, E.: Online auctions in iaas clouds: welfare and profit maximization with server costs. ACM SIGMETRICS Perform. Eval. Rev. **43**(1), 3–15 (2015)
24. Wang, Z., Hayat, M.M., Ghani, N., Shaban, K.B.: Optimizing cloud-service performance: efficient resource provisioning via optimal workload allocation. IEEE Trans. Parallel Distrib. Syst. **28**(6), 1689–1702 (2017)
25. Simão, J., Veiga, L.: Partial utility-driven scheduling for flexible sla and pricing arbitration in clouds. IEEE Trans. Cloud Comput. **4**(4), 467–480 (2016)
26. Wang, H., Kang, Z., Wang, L.: Performance-aware cloud resource allocation via fitness-enabled auction. IEEE Trans. Parallel Distrib. Syst. **27**(4), 1160–1173 (2016)
27. Pillai, P.S., Rao, S.: Resource allocation in cloud computing using the uncertainty principle of game theory. IEEE Syst. J. **10**(2), 637–648 (2016)
28. Hwang, E., Kim, S., Yoo, T.-K., Kim, J.-S., Hwang, S., Choi, Y.-R.: Resource allocation policies for loosely coupled applications in heterogeneous computing systems. IEEE Trans. Parallel Distrib. Syst. **27**(8), 2349–2362 (2016)
29. Muthi Reddy, P., Rekha Rangappa Dasar, Tanuja, R., Manjula, S.H., Venugopal, K.R.: Coirs: cost optimized identity based ring signature with forward secrecy in cloud computing. Int. J. Comput. Sci. Inf. Secur. (IJCSIS) **16**(3), 71–79 (2018)
30. Graiet, M., Mammar, A., Boubaker, S., Gaaloul, W.: Towards correct cloud resource allocation in business processes. IEEE Trans. Serv. Comput. **10**(1), 23–36 (2017)
31. Faiz, M., Anuar, N.B., Wahab, A.W.A., Shamshirband, S., Chronopoulos, A.T.: Source camera identification: a distributed computing approach using hadoop. J. Cloud Comput. **6**(1), 18 (2017)
32. Shamshirband, S., Anuar, N.B., Kiah, M.L.M., Patel, A.: An appraisal and design of a multi-agent system based cooperative wireless intrusion detection computational intelligence technique. Eng. Appl. Artif. Intell. **26**(9), 2105–2127 (2013)
33. Bashirpour, H., Bashirpour, S., Shamshirband, S., Chronopoulos, A.T.: An improved digital signature protocol to multi-user broadcast authentication based on elliptic curve cryptography in wireless sensor networks (wsns). Math. Comput. Appl. **23**(2), 17 (2018)
34. Tajiki, M.M., Akbari, B., Mokari, N., and Chiaraviglio, L.: Sdn-based resource allocation in mpls networks: a hybrid approach. arXiv preprint arXiv:1803.11486 (2018)
35. Al-Janabi, S., and Al-Shourbaji, I.: A smart and effective method for digital video compression. In: 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT). IEEE, pp. 532–538 (2016)
36. Murthi Reddy, P., Manjula, S.H., and Venugopal, K.R.: Sdspg: secured data sharing with privacy as a group in cloud computing. Int. J. Curr. Adv. Res. **7**(2), 10036–10043 (2018)

P. Muthi Reddy is a full time Research Scholar in the Department of Computer Science and Engineering from University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India. He was awarded Master of Technology in Computer Science and Engineering from Davangere University, Davangere. He obtained his Bachelor of Engineering degree in Computer Science and Engineering from BMS College of Engineering, Visvesvaraya Technological University. He was received the Diploma in Computer Science and Engineering from Govt. Polytechnic, Board of Technical Examinations, Karnataka. His research interests are in the field of Data Security, Data Sharing and Internet of Things (IoT) in the Cloud Computing.

Ansaf Ahmed received the B.E. degree in Computer Science and Engineering from the Visvesvaraya Technological University, Belagavi, India in 2015. He is currently pursuing his M.E. in Web Technologies in University Visvesvaraya College of Engineering, Bengaluru. His areas of interests include Cloud Computing, Semantic Web, and the Internet of Things.

S. H. Manjula is currently Associate Professor, Department of Computer Science and Engineering, UVCE, Bangalore University, Bengaluru. She has obtained B.E., M.Tech., Ph.D. in Computer Science and Engineering, Chennai. Her research interests are in the field of Wireless Sensor Networks and Data mining.

K. R. Venugopal is currently the Principal, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru. He obtained his Bachelor of Engineering from University Visvesvaraya College of Engineering. He received his Masters degree in Computer Science and Automation from Indian Institute of Science Bengaluru. He was awarded Ph.D. in Economics from Bangalore University and Ph.D. in Computer Science from Indian Institute of Technology, Madras. He has a distinguished academic career and has degrees in Electronics, Economics, Law, Business Finance, Public Relations, Communications, Industrial Relations, Computer Science and Journalism. He has authored and edited 64 books on Computer Science and Economics, which include Petrodollar and the World Economy, C Aptitude, Mastering C, Micro-processor Programming, Mastering C++ and Digital Circuits and Systems etc., He has filed 101 patents. During his three decades of service at UVCE he has over 640 research papers to his credit. His research interests include Computer Networks, Wireless Sensor Networks, Parallel and Distributed Systems, Digital Signal Processing and Data Mining. He is a Fellow of IEEE, ACM and ISTE.