Workflow Scheduling in Cloud by Hybridization of Particle Swarm Optimization (PSO) with Grey Wolf Optimization (GWO)

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering
in
Software Engineering

Submitted By
Damini Arora
(Roll No. 801631005)

Under the supervision of:
Dr. (Mrs) Seema Bawa
Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY
PATIALA – 147004

July 2018

Certificate

I hereby certify that the work which is being presented in the thesis entitled, *"Workflow Scheduling in Cloud by Hybridization ofParticle Swarm Optimization (PSO) with Grey Wolf Optimization (GWO)"*, in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Software Engineering*submitted in Computer Science and Engineering Department of ThaparInstitute of Engineering and Technology, Patiala, is an authentic record of my own work carried out under the supervision of *Dr. SeemaBawa* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

(DaminiArora)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. SeemaBawa)
Professor,CSED
TIET, Patiala

Acknowledgement

I would like to express my deep sense of gratitude to my supervisor Dr. Seema Bawa, Professor, Computer Science & Engineering Department for her invaluable guidance, support and encouragement during this thesis work. She provided me all the resources and guidance throughout thesis work.

I express my gratitude to Dr. Maninder Singh, Head, Computer Science & Engineering Department for providing us adequate environment, facility for carrying thesis work.

I would also like to thank all staff members and my colleagues who were always there at the need of hour and provided with all help and facilities, which I required, for the completion of my thesis work.

I would also like to thanks to all my friends for helping me in the hour of need and providing me all the help and support for completion of my thesis.

Last but not the least I am highly grateful to all my family members for their inspiration and ever encouraging moral support, which enables me to purse my studies.

Damini Arora

Abstract

Cloud Computing is used commonly in almost every business or research field. It is young but familiar technology which enables the client to use its services without being bothered to know how the services run and leave this job to providers. Cloud providers offer three main services namely Infrastucture as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Users can choose appropriate services which suit their requirements. The user has to pay the amounts of resources used in particular duration of time accordingly. It is pay-as-you-go model. It has opened up many new opportunities for researchers as well as business organizations. It is flexible as the users can scale up or scale down resources depending on their requirement.

Various scientific workflows depend on the static configuration of virtual machines, which is not a real condition. The workflow scheduling creates an issue which resists the efficiency of parameters like makespan and cost in the cloud environments. This optimization depends on the Random distribution for local users like Virtual Machine (VM) or global users (Datacentre) which sometimes take more time therefore increase the overall cost.

To overcome this problem we propose our work on Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO) for effective scheduling. This work is based on the optimization of Total Execution Time (TET) and Total Execution Cost (TEC). The results of the proposed approach are found to be effective when compared with existing methods. The hybridization of both the optimization technique is done. The results were compared with the existing BAT algorithm while parsing the workflow into task. These tasks were mapped onto virtual machine. The results obtained are found better in terms of execution time and cost.

Table of Content

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cloud Computing Evolution

Cloud computing is a distributed design which brings together server resources on an acceptable stage in order to provide on request figuring resources and administrative data. A cloud specialist organization (CSP's) provides the different stages to their customers to use the services and make the web administrative control. This service is similar to a broadband band connection offered by the service provider for the internet connection. Cloud computing provides the services through the internet; these service belongs to hardware and software both. Cloud computing concept is growing high day by day due to its service pay per usage concept. When cloud provides the service in the form of platform it is called as Platform as a service (PaaS) model. When cloud provides hardware to the user it is called as Infrastructure as a service (IaaS) model. When cloud is provided to the software services it is also called as Software as s service.

1.2 Cloud Service Models

Into three service models cloud computing can be circulated: IaaS, PaaS and SaaS. Depending upon their particular needs an organization may come by gathering of any these service model [3].In Fig. 1.3 these services are appeared.

i. Software as a Service (SaaS): Over the web the SaaS delineates some cloud organization from where purchasers can get programming applications. For both individuals and affiliations the applications can be uplift in "the cloud" and can be used for an expansive assortment of assignments. By techniques for any web engaged contraption the Twitter, Facebook and Flickr are all examples of SaaS, with users prepared to get to the organizations. Rather than securing it the programming as Service clients, subscribes, and regardless to the things, if all else fails on a month to start. Rather than on singular PCs the applications are acquired and used in online with the help of records to save as a bit of the cloud.

ii. Platform as a Service (PaaS): Using instruments given by the provider in the PaaS are licensing the clients to make programming applications. Customers can subscribe to the

Paas organizations can contain preconfigured fragments; while discarding those that don't they can combine the parts that meet their necessities.

iii. Infrastructure as a Service (IaaS): In the IaaS clouds the cloud customers clearly use IT bases (dealing with, structures, stockpiling, and other essential planning assets) is given. Virtualization is comprehensively put to use as a touch of IaaS cloud reviewing a legitimate focus to sort out/break down physical resources in an exceptionally assigned way to deal with meeting or contracting resource request from customers. The elementary course of action of virtualization is to provide free VM which are detached from both the secured mechanical assembly and specific VMs. This framework is totally not the same as of the multi-residency show up, which intends to change design of application programming so that different cases can be kept for running on a particular application.

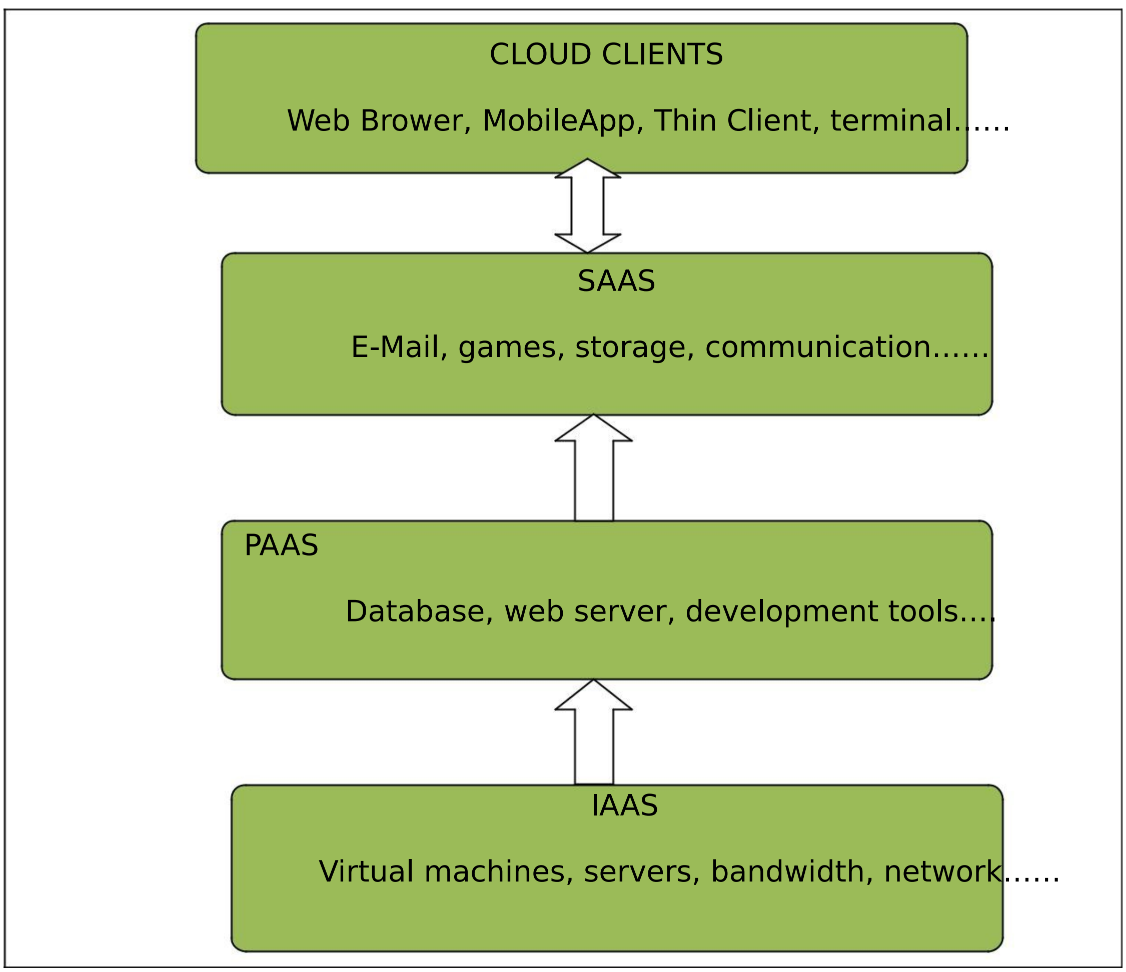

Figure 1.1 Cloud service Model [34]

1.3 Cloud Deployment Model

i. As depicted in Fig.1.4 subject on necessities the going with four sending models have been perceived, each with particular qualities that reinforce the prerequisites of the organizations and customers of the clouds specifically ways [3].

ii. Private cloud (PC) the cloud foundation is worked in a particular union, and facilitated by the associated cloud or untouchable notwithstanding whether it is found prelude or off the reason. The inspiration to build a private cloud inside a federation has two or three perspectives. Regardless, to strengthen and streamline the implementation of existing in-house assets. The second, security concerns includes information protection and belief in way makes PC a likelihood for a couple of affiliations. Third, information exchange passed on from near to IT foundation to a PC is still rather stunning. The fourth, affiliations reliably require full control over mission-basic exercises that stay back of the firewalls.

iii. Community cloud a few affiliations usually make and have a relative cloud base and in like way diagrams, necessities, qualities, and concerns. The cloud bunch shapes into a level of gainful and free change.

iv. Public cloud is the key to sort current Cloud managing sending model. People considering all things cloud is utilized by the normal masses cloud customers and the cloud association supplier has the all time duty as to open cloud with its own particular system, respect, and great position, cost charges and charging model. Particular comprehended cloud affiliations are open clouds including Amazon EC2, S3 and Force.com.

v. Hybrid cloud is a base cloud is a mix of no under two clouds (private, storing up, or open) that stay striking parts however are bound together by systematized or select development that pulls in information and application transportability (e.g., cloud influencing for weight changing between cloud). Affiliations utilize the cream cloud appear with a specific phenomenal attentiveness to push their advantages for build up their inside points of confinement by margining out edges business limits onto the cloud while directing spotlight rehearses on-premises through private cloud.

1.4 Cloud Applications

i. **Development and Testing:** Cloud plays an effective role as it is used for test and development. It saves the cost of setting up environment by setting up physically which includes the manpower and time. The installation and configuration of the software also take more time and this problem is also solved by using clod resources.

ii. **Big Data Analytics:** Cloud is using the concept of big data and provides theeffective data extraction of the business value. It provides the effective data to the retailers and suppliers by extracting the buying patterns of the consumers. The buying patterns of the consumers show their liking and disliking of the consumers to the product.

iii. **File storage:** Cloud offers the facility of data storage, retrieval and access from any web-enabled interface. The user can access data anytime, anywhere with high speed and availability. The large organization stores their data on cloud and only pay for the storage of data and they do not worry about the daily maintenance of the storage system.

iv. **Disaster Recovery:** Cloud provides the effective data recovery in case of disaster at very effective cost. Data recovery by traditional method is very expensive and slow.

v. **Backup:** Backing up data is always a complex and time consuming process. The backup includes the tapes and drives to collect the data manually and then dispatching them for backup. Cloud provides the data backup automatically and no need to worry if the data is deleted. By using cloud it is easy to recover the data.

1.5 Benefits of Cloud Computing

Cloud computing have many essential or unique characteristics as shown in Fig. 1.2 which helps to provide qualitative services. Some of the characteristics are as follows [2]

i. **On-demand self-service** This is self-advantage notification to the organization given by appropriated registering merchants that enables the course of action of cloud assets on demand for at whatever stage they are required for. In such services, the customer finds the opportunity to access online control facilities which is basic benefit from cloud.

ii. **Broad network access** Cloud computing isolates computing capability from their consumers, with the goal that they don't need to keep up the capacities themselves. A consequence of this is that computing capabilities are found somewhere else, and that must be aquired over a network.

iii. **Resource pooling** Resource pooling is an Information Technology term used as a bit of distributed computing conditions which depict a situation in which suppliers serve diverse clients, clients or "inhabitants" with impermanent and flexible federation. These organizations can be usual to suit every client's needs with no developments being clear to the client or end client. Occurrences of advantages combine stockpiling, arranging, memory, and framework data trade restrain.

iv. **Fast elasticity** It is described as the ability to modify resources both all over as required. To the buyer, the cloud has every one of the reserves of being immense, and the purchaser can purchase to such an extent or as pitiful enlisting power as they need.

v. **Measured service** Cloud systems therefore control also, redesign asset are used by utilizing a metered limit at some altitude of direction sensible to the sort of affiliation (e.g., dealing with, stockpiling, information transmission, and dynamic customer accounts).

vi. Stockpiling, information transmission, and dynamic customer accounts).

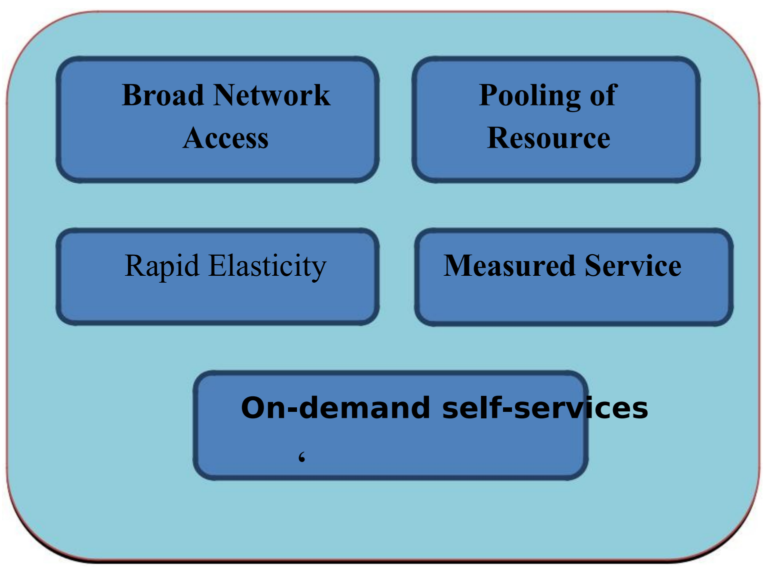

Figure.1.2 Unique characteristics of cloud computing [35]

1.6 Research Motivation

The distributed cloud computing condition gives numerous administrations to the vast scale logical work processes. Right off the bat, assets can be flexibly provisioned and de-provisioned on-request. It gives the capacity to the work process framework to utilize the assets as per the assignment and its compose artfully and should be prepared at a given time. Along these lines, the accessible resource pool can be scaled out and in the work process prerequisites change. This is a helpful element for logical work processes as normal topological structures, for example, information appropriation and conglomeration [30] prompt noteworthy changes in the parallelism of the work process after some time. This prompts circumstances in which altering the quantity of assets being utilized is very attractive with a specific end goal to build execution and guarantee the accessible assets are proficiently used. Another advantage of sending work processes in cloud computing situations gets from the way that they are for the most part inheritance applications that contain heterogeneous programming segments. Virtualization takes into consideration the execution condition of these parts to be effectively redone. The working framework, programming bundles, registry structures, and info information records, among others, would all be able to be customized for a particular segment and put away as a VM picture. This picture would then be able to be effectively used to send VMs equipped for executing the product segment they were intended for. Another favorable position of utilizing VM pictures for the sending of work process errands is the way that they empower logical approval by supporting analysis reproducibility. Pictures can be put away and redeployed at whatever point an investigation should be imitated as they empower the production of the same correct condition utilized as a part of past analyses.

1.7 Organization of Thesis

Chapter 2- this chapter describes the literature review related to the scheduling process its advantages and challenges faced in different techniques.

Chapter-3 describes the problem statement and its formulation and also the objectives of this research.

Chapter-4 describes the proposed methodology of the problem formulation and algorithm used for it.

Chapter-5 represents the results after the implementation and also shows the comparison with existing system.

Chapter-6 represents the conclusion of the study and future scope of this work.

Chapter 2

Literature Review

To impart a high quality of service to the client is the important factor for service provider so that the client can have maximum profit from them. Different types of parameters are used to decide the quality of service and these parameters are clearly mentioned in a document called Service Level Agreement (SLA). The SLA document consists of the information related to the service name, contract duration, desired outcome, and requirements of the service interface, responsibilities, and Glossary. The service provider and client agreed upon the SLA for the services and also perform a lot of computation and data analysis. For this reason, the scheduling becomes more complex and enhances the overhead of the applications and leads to increases the overall cost of the computation. If the scheduling process is not proper it takes a long time to complete a task and execution time is increased. For developing an effective and efficient scheduling technique it is necessary to study the challenges and needs of the scheduling. In this chapter challenges in scheduling, different algorithms used in scheduling and their positive and negative impacts on services are discussed in it.

2.1 Challenges Faced by Scientific Applications

Qi Zhang et al. [1] have discussed the various cloud computing technologies and commercial products in detail. The commercial products have been compared on parameters like cloud provider, computing classes, target application, computation, auto scaling and storage. The research challenges presented in this work are service provisioning automation, server consolidation, virtual machine migration, energy management and data security.

Christian Vecchiola et al. [2] have given a comparison of computing solutions such as Amazon EC2, Google App Engine, and Microsoft Azure. The comparison is done on the basis of parameters like type of service, value added provider, if PaaS, has capability to work on third party IaaS platform, Virtualization, deployment Model and interface for user access. The aneka architecture, deployment model and application model are discussed in detail. Programming models like task model, map reduce model, thread model, parameter sweep model, workflow have been compared on the basis of execution services, applications and execution unit.

Yong Zhao et al. [3] have given many opportunities that the cloud has brought in, such as better utilization of resources, improved responsiveness thereby improving user experience, enabling a generation of collaborative scientific workflows and reducing the cost in challenges and opportunities in running scientific workflows on cloud. The challenges faced by the applications are architectural challenges, service challenges for integration tools, high-end computing support language-conversion challenge, challenge in compute intensive applications, challenge for data management, service management challenge.

Jens Sonke et al. [4] a scientific application is executed on Future Grid, Amazon EC2 cloud and NERSC's Magellan in this paper. The result of this paper have been compared and analyzed to comprehend various challenges that came across during the process. In this work Pegasus workflow management system has been used to execute a scientific application which was used to process data from the Kepler project by NASA to find out planets similar to the earth.

Suraj Pandey et al. [5] has worked on reducing the computing cost of the application by using Particle Swarm Optimization algorithm which is basically a meta-heuristic algorithm used for scheduling. PSO is used for calculating the fitness function. The overall cost calculated is the cost of execution and the transfer cost of data. This algorithm ensures the cost of the highest task is reduced by heuristic scheduling. This algorithm helps to schedule the resources and mapping.

Alexandru Iosup et al. [6] have explained the differences between the actual scope field of cloud and the requirements of scientific applications. It evaluates the cloud and check the capability of a cloud to run the applications efficiently. The evaluation is done by quantifying the number of users that require scientific computing services followed by evaluating four cloud services mostly used for scientific applications.

Simon Ostermann et al. [7] have discussed about various features of cloud computing which help ease the execution of scientific applications. It evaluates these features by different workloads like SJSI, MJSI and SJMI on Amazon EC2 cloud platform. Different types of benchmarks like Lmbench, Bonnie, and HPC are used to access the performance of EC2 cloud for scientific applications.

EwaDeelman et al. [8] have discussed dependency of cost on execution models in this work. In this work, the cost is calculated as a function of number of processors. The cost of executing montage workflow has been estimated by running simulation using GridSim tool. Three montage workflows have been Remote I/O, regular and dynamic cleanup. Three montage workflow that are executed are following montage degree 1, degree 2 and degree 4. The cost of running each of the data management models have been compared graphically. To maintain the trade-off between the number of processors and reduction in execution time.

Christina Hoffa et al. [9] worked on four different workflows and four different environments in order to compare the performances. The tools like WMS, DAGman. GridFTP, condoe and GRAM are explained in detail. The different montage degree workflows are executed on local machines and cluster and multiple virtual clusters.

Scott Callaghan et al. [10] have discussed and tackled the problems relating to managing a workflow in this work. In this work, the probabilistic seismic hazard analysis is performed for calculation estimates. It requires ground motions caused by past earthquakes as input. The ground motions are calculated by CyberShake 3D ground motion simulations with analytic wave propagation model.

Table 2.1 Challenges Faced By Scientific Application in Cloud Computing

Author's Name	**Year**	**Scientific Applications**	**Challenges**
Qi Zhang et al. [1]	2010	GoogleApp Engine	Automation of service provisioning, Migration of energy management and server consolidation.
Christian Vecchiola et al. [2]	2009	FMRI	Prediction, scheduling, Pricing and Computation Time.
Yong Zhao et al. [3]	2011	Amazon Map Reduce Work flows	Workflow scheduling, computation and management.
Jens Sonke et al. [4]	2011	Work flow represented by DAG	Scheduling, Computation
SurajPandey et al. [5]	2010	Bio-informatics workflow	Scheduling and Cost of execution
AlexandruIosup et al. [6]	2011	Grid workload and parallel workload archives	Cost of execution, storage and Response time

Simon Ostermann et al. [7]	2009	SJSI,SJMI,MJSI Workload	Computation Time, Response Time, Throughput
EwaDeelman et al. [8]	2008	Montage 1 Degree	Computation Cost, Execution Time
Christina Hoffa et al. [9]	2008	Montage	Scheduling overhead and Computation time
Scott Callaghan et al. [10]	2010	CyberShake Overflow	Scaling up of resources and execution time

2.2 Existing Scheduling Models

Alkhanak et al. [11] presented a cost optimization approach for scientific workflow scheduling in cloud computing. The proposed approach employs the four meta-heuristic algorithms which are based on the population. It helps in reducing cost and time of the service providers. The execution time and cost are reduced as compared to existing approaches.

Anubhav et al. [12] introduced a gravitational search algorithm for workflow scheduling in the cloud environment. The optimizations in workflow reduce the cost and makespan. In this process, two algorithms are hybridized GSA and HEFT for workflow scheduling. The performance evaluation is done on the basis of two metrics that are budgetary cost ratio and schedule length ratio. The validation of result is also tested by ANOVA test and it shows that the proposed approach outperforms.

Sagnika et al. [13] proposed BAT algorithm for workflow scheduling in cloud computing which helps to handle the large size of data. The scheduling process decides that which task is executed first and which is last according to their requirement of the resources. It manages the resources according to the task size and execution time. The result of the proposed algorithm is compared with particle swarm optimization algorithm and Cat swarm optimization algorithm. The convergence of the proposed algorithm is better than the existing algorithms.

Vinothina et al. [14] proposed Ant Colony Optimization algorithm (ACO) for workflow scheduling in cloud computing. This model is presented for heterogeneous distributed systems. The service level agreements are used to check the quality of service of the service providers. The problem of workflow scheduling is solved by using parameters cost, makespan and resource utilization. The ACO algorithms reduce the cost and makespan and enhance the resource utilization.

Liu, Li et al. [15] proposed the genetic algorithm for workflow scheduling in cloud computing with deadline-constrained. The crossover and mutation probability is adjusted by using convolution approach. This approach prevents from the prematurity and enhances the convergence. The proposed approach is compared with existing algorithms on the simulator at 4 different workflows. The results show that the total execution cost is reduced in this approach.

Garg et al. [16] formulated the scheduling problem in cloud by using the Genetic Algorithm. The proposed work is done to reduce the computation time and execution cost of the task. This work is done on the cloudSim simulator and it maximizes the resource utilization. The performance evaluation is done on the different parameters and performs well.

Rimal et al. [17] Cloud supports the Multi-tenancy feature and provides the scalability and other benefits to the other users. Resource management is an important task in the multi-tenant cloud computing which is done by using the scheduling process. In this work cloud based workflow scheduling policy is proposed for efficient computing in cloud. This policy reduces the overall workflow completion time, cost of execution and properly utilize the resources. The result of the proposed work is compared with existing approaches and algorithms. The simulation result of the proposed approach shows more effective results than the existing approaches.

Casas, Israel et al. [18] proposed a scheduling approach called Balanced and file Reuse-Replication (BFRR) scheduling. This approach is used to schedule the scientific application workflows. It splits the workflows into sub-workflows which help in proper utilization via parallelization process. This approach provides the facility if data reuse and replication which helps in optimization of data and transfer it at run time. The optimization process is based on execution time and monetary cost of workflows.

Kaur et al. [19] the proposed work is done on the Infrastructure as a service platform of the computer for scheduling and resource provisioning. The scheduling process is done by using the Shuffled Frog Leaping Algorithm (ASFLA). The performance evaluation is done by comparing the result or proposed algorithm with PSO (Particle Swarm optimization). The experiment is performed on different workflows by using Java Simulator and it gives outcome at low cost and completes the task on deadline.

Bölöni et al. [20] proposed the concept of computation scheduling which is used for prediction of computation cost and financial cost. It also predicts the benefit of the output and it is called as value of information. This work is based on the analysis process of real-estate investment opportunities. The scheduling algorithm used in this work is called as volume based scheduling algorithm.

Manojit Ghose et al. [21] have given the energy efficient scheduling approach in cloud environment. In this work six different scheduling strategies are proposed for a collection of scientific workflows. The performance evaluation of the scheduling approaches is compared with existing policies and presented the average energy reduction of 70%.

Li Yibin et al. [22] introduced the concept of Dynamic Voltage Scaling for maintaining the power and lowering the supply of voltage and frequency of processor. In this work, Energy aware dynamic task scheduling algorithm is used to reduce the energy consumption. This algorithm reduces the more energy consumption as compare to parallelism and critical path scheduling algorithm.

Quang-Hung et al. [23] proposed an approach for energy saving virtual machine scheduling in cloud computing with fixed interval constraints. This approach reduces the busy time and total energy consumption by the resources. The experiment is performed on the parallel workload models. The simulation result shows that the total energy consumption is less than the existing model.

Abrishami et al. [24] proposed a deadline constrained workflow scheduling algorithm for IaaS. In this paper, the author proposed partial critical path algorithm with IaaS environment and this is called as PCP with deadline distribution. Both the algorithm supports the polynomial time complexity which is good for the scheduling in large workflows. This algorithm also reduces the execution cost and response on the deadline. The simulation result of the proposed methodology is also compared with other methods and gives better results.

Netjinda et al. [25] focused on optimizing the value of buying infrastructure-as-a-service cloud competencies to attain clinical work goes with the flow execution in the unique closing dates. Authors considered the quantity of purchased times, example types, buying alternate sources, and venture scheduling as constraints in an optimization technique. Particle swarm optimization augmented with a variable community seeks approach turned into used to discover the superior solution. Results display promising performance from the views of the total fee and fitness convergence when in comparison with other trendy algorithms.

Verma et al. [26] recommended that the users put up their workflows alongside a few QoS constraints like closing date, budget, and consider, reliability and so on. For computation, Authors considered the two constraints: closing date and finances and recommend cut-off date and finances Due date and Budget Distribution based cost-Time Optimization (DBD-CTO) work process scheduling set of rules that points of confinement execution regard while get together time diagram for giving over outcomes and separate the direct of the estimation.

Xu et al. [27] recommended a various Quality of services compelled scheduling strategy of multi-work processes (MQMW). The procedure can plan different work processes which are begun whenever and the QoS prerequisites are considered and ready to build the planning achievement rate essentially.

Rahman et al. [28] proposed an Adaptive Hybrid Heuristic for purchaser limited insights examination work process planning for hybrid Cloud surroundings through organizing the dynamic method for heuristic based techniques and furthermore work process degree change helpfulness of meta-heuristic based frameworks. The sufficiency of the proposed framework was sketched out by strategy for an extensive case take a gander at in assessment to introduce strategies.

Mao et al. [29] offered a procedure whereby the major figuring elements are virtual machines (VMs) of various sizes/costs, employments are exact as work processes, clients indicate execution prerequisites by method for allotting (delicate) time points of confinement to occupations, and the reason for existing is to ensure all employments are finished inside their due dates at negligible money related charge. Creators finish their objective by utilizing progressively dispensing/deal locating VMs and scheduling duties at the most extreme esteem green cases. They assessed approach in four delegate cloud workload styles and show charge money related reserve funds from 9.8% to 40.four% contrasted with various methodologies.

Nancharaiah et al. [30] displayed hybrid routing algorithm, Ant Colony Optimization (ACO) algorithm and Particle Swarm Optimization (PSO) is utilized to enhance the different measurements in MANET routing. The ACO algorithm utilizes portable specialists as ants to distinguish the most possible and best way in a system. Likewise ACO algorithm finds ways between two hubs in a system and gives contribution to the PSO strategy. The PSO finds the best answer for a particle's position and speed and limits cost, power and end to end delay. This hybrid routing shrewd algorithm has an enhanced execution when contrasted and basic ACO algorithm as far as delay, power, consumption and communication cost.

Alkhanak et al. [31] researched and investigated different cost mindful difficulties of WFS in cloud computing, for example, Quality of administration, execution, framework usefulness and framework engineering. In this they additionally talked about different WFS cost mindful methodologies from the accessible pools of options. Different WFS challenges influencing particular WFS execution cost has additionally been thought about.

Zhan et al. [32] proposed progressed PSO based venture scheduling set of rules in Cloud Computing which can lessen the assignment common jogging time and raises the supply of sources which complements the convergence price and improves the efficiency. They investigated that that is due to the fact in every generation worldwide fast confluence of simulated annealing algorithm is applied to amalgamate particle swarm optimization algorithm.

Bilgaiyan et al. [33] focused that in computing environment there is a large amount of data that is processed every second, so there is a time where scheduling plays a vital role which helps to manage the cost and make span. They analyzed down the different swarm optimization algorithm that recommended that scheduling guideline expect to mean to lessen the measure of information exchange with slightest cost and guarantee balanced dispersion of errands according to handling capacity.

Table 2.2: Existing Scheduling Model.

Author's Name	**Year of Publication**	**Algorithm/ Technology Used**	**Summary**
Alkhanak et al. [11]	2018	Meta-Heuristic Algorithms	The proposed approach employs the four meta-heuristic algorithms which are based on the population. The approach helps in reducing cost and time of the service providers. The execution cost and time are reduced as compared to baseline approaches.
Anubhav, et al. [12]	2018	GSA based hybrid algorithm	The optimizations in workflow reduce the cost and makespan. In this process, two algorithms are hybridized GSA and HEFT for workflow scheduling. The performance evaluation is done on the basis of two metrics that are monetary cost ratio and schedule length ratio.
Sagnika et al. [13]	2018	Bat Algorithm	The scheduling process decides that which task is executed first and which is last according to their requirement of the resources. It manages the resources according to the task size and execution time.
Vinothina et al. [14]	2018	ACO algorithm	Ant Colony Optimization algorithm is presented for heterogeneous distributed systems. The service level agreements are used to check the quality of service of the service providers. The problem of workflow scheduling is solved by using parameters cost, makespan and resource utilization.

Liu, Li et al. [15]	2017	Genetic Algorithm	This approach prevents from the prematurity and enhances the convergence. The proposed approach is compared with existing algorithms on the simulator at 4 different workflows. The results show that the total execution cost is reduced in this approach.
Garg et al. [16]	2017	Genetic Algorithm	The proposed work is done to reduce the computation time and execution cost of the task. This work is done on the cloudSim simulator and it maximizes the resource utilization.
Rimal et al. [17]	2017	Workflow Scheduling Policy	In this work cloud based workflow scheduling policy is proposed for efficient computing in cloud. This policy reduces the overall workflow completion time, cost of execution and properly utilize the resources.
Casas Israel et al. [18]	2017	Balanced and file Reuse-Replication scheduling. (BFRR)	BFRR scheduling is used to schedule the scientific application workflows. It splits the workflows into sub-workflows which help in proper utilization via parallelization process. This approach provides the facility if data reuse and replication which helps in optimization of data and transfer it at run time.
Kaur et al. [19]	2017	Shuffled Frog Leaping Algorithm	The scheduling process is done by using the Shuffled Frog Leaping Algorithm (ASFLA). The performance evaluation is done by comparing the result or proposed algorithm with PSO (Particle Swarm optimization).
Bölöni et al. [20]	2017	Volume based scheduling algorithm.	This work is based on the analysis process of real-estate investment

			opportunities. The scheduling algorithm used in this work is called as volume based scheduling algorithm.
Ghose, Manojit et al. [21]	2017	Energy Efficient Scheduling	Energy efficient scheduling approach in cloud environment. In this work six different scheduling strategies are proposed for a collection of scientific workflows.
Li, Yibin et al. [22]	2017	Dynamic Task Scheduling	In this work, Energy aware dynamic task scheduling algorithm is used to reduce the energy consumption. This algorithm reduces the more energy consumption as compare to parallelism and critical path scheduling algorithm.
Quang-Hung et al. [23]	2017	Energy-Saving Virtual Machine Scheduling	This approach reduces the busy time and total energy consumption by the resources. The experiment is performed on the parallel workload models.
Abrishami et al. [24]	2012	Hybrid Heuristic Algorithm	In this paper, the author proposed partial critical path algorithm with IaaS environment and this is called as PCP with deadline distribution. Both the algorithm supports the polynomial time complexity which is good for the scheduling in large workflows. This algorithm also reduces the execution cost and response on the deadline.
Netjinda et al. [25]	2011	NIST Cloud Model	Authors considered the quantity of purchased times, example types, buying options, and venture scheduling as constraints in an optimization technique. PSO augmented with a variable community seeks approach turned into used to

			discover the superior solution.
Verma et al. [26]	2012	Date and Budget Distribution based cost-Time Optimization	For computation, Authors considered the two constraints: closing date and finances and recommend cut-off date and finances Due date and Budget Distribution based cost-Time Optimization (DBD-CTO) work process scheduling set of rules that points of confinement execution regard while get together time diagram for giving over outcomes and separate the direct of the estimation.
Xu et al. [27]	2011	Multi-work processes	The procedure can plan different work processes which are begun whenever and the QoS prerequisites are considered and ready to build the planning achievement rate essentially.
Rahman et al. [28]	2014	Hybrid Heuristic Algorithm for scheduling	Proposed an Adaptive Hybrid Heuristic for purchaser limited insights examination work process planning for hybrid Cloud surroundings through organizing the dynamic method for heuristic based techniques and furthermore work process degree change helpfulness of meta-heuristic based frameworks.
Mao et al. [29]	2011	Hybrid Heuristic Algorithm	Creators finish their objective by utilizing progressively dispensing/deal locating VMs and scheduling duties at the most extreme esteem green cases.
Nancharaiah et al. [30]	2013	ACO and PSO Algorithm	This hybrid routing shrewd algorithm has an enhanced execution when contrasted and basic ACO algorithm as far as delay, power, consumption, and communication cost.
Alkhanak et al. [31]	2015	Hybrid Particle Swarm Optimization Algorithm	In this they additionally talked about different WFS cost mindful

			methodologies from the accessible pools of options. Different WFS challenges influencing particular WFS execution cost has additionally been thought about.
Zhan et al. [32]	2013	Particle Swarm Optimization Algorithm	They investigated that that is due to the fact in every generation worldwide fast confluence of simulated annealing algorithm is applied to amalgamate particle swarm optimization algorithm.
Bilgaiyan et al. [33]	2009	PSO Algorithm	They analyzed down the different swarm optimization algorithm that recommended that scheduling guideline expect to mean to lessen the measure of information exchange with slightest cost and guarantee balanced dispersion of errands according to handling capacity.

2.3 Existing Predication Model

BAT Algorithm

Bat algorithm is a meta-heuristic optimization algorithm which is used to obtain the optimal solution. It is basically based on echolocation behaviour of the bats with varying pulse rates of emissions and loudness. The working of this algorithm is depending on the velocity and position of bat which vary according to the frequency, wavelength and loudness. Following are the steps that are performed in the Bat algorithms.

Initialization: Firstly generation counter t is set to be 1, p is the population of NP bats which is initialize randomly. Each bat gives a possible solution of the given problem. Here,

A: it defines the loudness

Q: is the frequency

V: are the initial velocities.

s: is the pulse rate

F: is the weight factor.

Step 1: Evaluate the quality f for each bat in P determined by f(x).

Step 2: while the termination criteria are not satisfied or t <MaxGeneration**do**

Sort the population of bats P from the best to worst by order of quality f for each bat; **for** i = 1: NP (all Bats) **do**

4 = [NP * rand] Select uniform randomly 1≠s2≠s3≠ i

= −1 + (- x*) * Q

= −1 +

if (rand > r**) then**

t

=

= 1 + F (2- 3)

end if

Evaluate the fitness for the offspring , , 4

Select the offspring with the best fitness among the off springs

, , 4

if (rand < A) **then**

4 = ;

enf if

end for i

t = t+1 ;

Step 3:end while

Step 4: Post- processing the results and visualization;

End.

2.4 Research Questions

i. Dynamic and Fluctuating workloads- The major challenge to elasticity in cloud computing is that the workloads are unpredictable. The workload fluctuation may occur in planned or unplanned way. In case of planned fluctuations in the workload, the situation can be predicted in advance so that resource could be allocated smoothly and in time.

ii. Ensuring efficient resource utilization- The resources must be allocated instantly whenever needed, although it's unplanned demand. This is called Auto-scaling in cloud computing. The incoming workload must be allocated to resources (Virtual Machines) such that the cloud service provider has to ensure efficient resource utilization. This requires optimal scheduling techniques to allocate the tasks to available machines.

iii. Heterogeneous physical nodes in cloud datacenters-the tasks are allocated (or scheduled) across available nodes which are widely distributed at different location and vary in computational power, architecture, memory and even the network performance. Different tasks perform differently at different nodes [3].

iv. Increased scheduling granularity than traditional scheduling- the size of scheduling problem has increased from simple task(process) scheduling in traditional computing systems with small data transfers to intensive VM resource scheduling and VM migrations in cloud computing environment [4]

This chapter discusses various approaches used of scheduling of workflows. The issues related to these approaches is also discussed and mostly related to the resource scheduling and allocation. To overcome these issues we discussed new approach in the next chapter with problem statement and scope of the work also.

Chapter 3

Research Gaps and Problem Statement

The above literature survey consists of various challenges faced by scientific applications as well as solution to tackle these challenges. Various scheduling approaches have been discussed. The research that has been done in this field of resource scheduling has been included. This chapter contains the research gap analysis, problem formulation and objectives of the research has been explained.

3.1 Gap Analysis

1. In the cloud, evolution service along with collaboration framework ought to be considered to help conniving and unlikelihood environments.

2. The workflow scheduling issue which streamlines both makes span and cost as a Multi-Objective Optimization Problem (MOP) for the Cloud environments.

3. On the off chance that workload prediction procedure can be utilized, all the more fine-grained and accurate spending allocation strategies can be adopted. This can help us to enhance the mechanism's performance.

Table 3.1 Analysis of various Reliable Workflow Scheduling Algorithms

Sr. No.	Algorithm Name	Technique	Parameters	Advantage	Disadvantage
1	BAT [13]	Scheduling	Time and cost	Distance base optimization	Reduce generality
2	ACO [14]	Scheduling	Time and cost	Effective utilization	Reduce the accuracy
3	GWO [29]	Scheduling	Time and cost	Improved bias performance	Increase time
4	SO [31]	Scheduling	Time and cost	Local and global optimization	Increase cost
5	Energy-Aware fault	Rescheduling	Fault tolerant, Energy	i. No resource wastage due	Delay in execution in

	tolerant Scheduling			to replication. ii. Optimizes energy consumption.	case of fault.

3.2 Problem Formulation

During the literature survey, many papers studied related to workflow scheduling depends on the static configuration virtual machine, which is not a real condition. Many authors worked on optimization depends on the Random distribution of task uses local (VM) or global (Datacentre), which sometimes conflicts like time and cost, which take more time for convergence.

In previous work, deadline constrained is depended on workflow dependency and optimization of the virtual machine is the context of overflow and underflow. It was found that parsing of the workflow level wise was not any distribution.

3.3 Objectives

Following are the objectives of thesis:

i. To study and analyze various existing workflow scheduling algorithm for makespan optimization and reliable workflow execution.
ii. To propose and design an efficient workflow scheduling approach for cost and time efficient workflow execution in cloud.
iii. To implement and validate the proposed algorithm in WorkflowSim environment on a scientific application.

Chapter 4

Proposed PSO-GWO Technique

The solution of the problem in this thesis has been proposed in pervious chapter. The proposed algorithm is compared with the BAT algorithm by giving inputs of workflow. After that these inputs are parsed into tasks. These tasks were mapped on VM. We are using Pareto Distribution instead of random initialization. The hybridization of proposed PSO-GWO technique helps to get effective results with the help of workflow simulator when compared with BAT algorithm.

4.1 Design of Proposed Technique

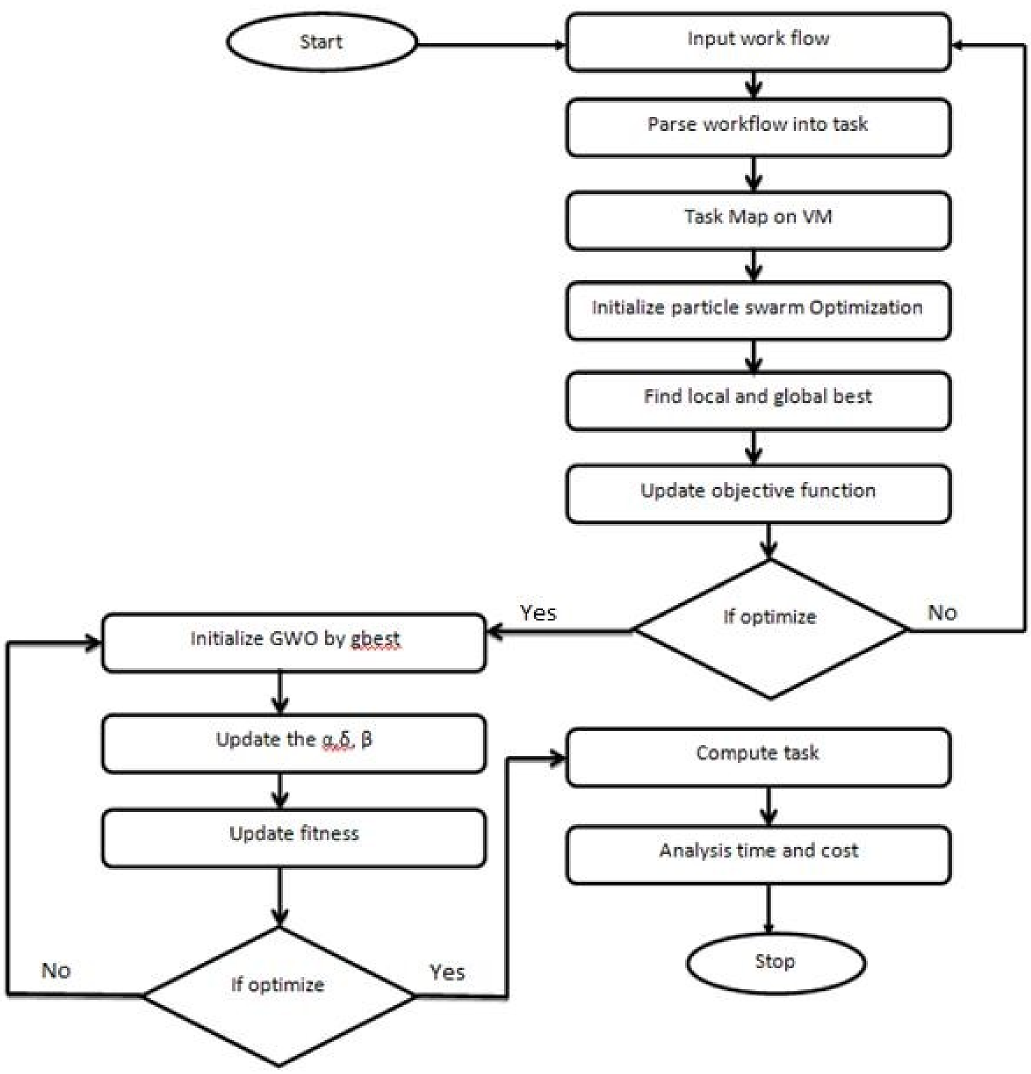

Figure 4.1: Flowchart of proposed PSO-GWO technique

4.2 Proposed PSO-GWO Technique

PSO stands for Particle Swarm Optimization which is a stochastic optimization algorithm based on the behavior of birds. It works similar to the genetic algorithm. In PSO is initialized with a group of arbitrary particles. In every iteration, each particle is improved by the two "best" values. The first best solution shows the fitness of the particles and this called as pbest. The second best value is followed by the optimizer is the best value. This value is known as global best (gbest).When a particle takes part of the population as its topological neighbors therefore the best value is a local best and is called lbest.

GWO

Grey Wolf Optimization algorithm is a bio-inspired algorithm which is based on the directorship and hunting behavior of the wolves in the pack. The grey wolves prefer to live in the pack which is a family of approximate 5-12 wolves. In the pack each member has social dominant and consisting according to four different levels. The below given figure shows the social hierarchy of the wolves which plays and important role in hunting.

Figure 4.2 GWO Hierarchy [36]

i. The wolves at the first level are called as alpha wolves (α) and these are leaders of the hierarchy. The wolves at this level are the guides of the hunting process in which other wolves search, track and chase and working like a team. The decision making is the main task that is executed by the alpha wolves and the order by alpha wolves is chased by all pack members.

ii. The 2nd level wolves are called as Beta (β). These wolves are called as subordinate and advisors of the alpha nodes. The advice by the Beta wolves helps in decision making. The beta wolves transmit the alpha's command to the whole pack and give the feedback to the alpha.

iii. The 3rdlevel wolves are called Delta (δ) wolves and called as scouts. The scout wolves in this level are accountable for observing frontier and domain. Sentinel wolves are answerable for securing the pack and caretakers are in-charge-of caring the liverish and wounded pack.

iv. The last and 4th level of the hierarchy is called Omega (ω). They are also called as scapegoats and they have to present to all the other commanding wolves. These wolves follow the other three wolves.

Proposed PSO-GWO technique The proposed technique is compared with the BAT algorithm by providing inputs of workflow. After that these inputs were parsed into tasks. These taskes were mapped on Virtual Machine (VM). With the help of Particle Swarm Optimization (PSO) we are able to find the best global and local outcomes from where we can update the objective function. Grey Wolf Optimization (GWO) will help to migrate the best located task to optimize execution time and cost. As we are using Pareto Distribution instead of random initialization. The hybridization of PSO with GWO helps to get effective results with the help of workflow simulator when compared with BAT algorithm.

4.3 Proposed PSO-GWO Algorithm Used

ALGORITHM USED PSO_GWO

Step 1: Input the mammographic images.

Step 2: Apply Gray Scale on the images.

Step 3: Edge detection by using the Prewitt Filter.

Step 4: For optimization input in the PSO model.

Step 5: Apply the loop in PSO model. for each particle n in S do

Step6 :   for each dimension d in D do

Step7:   //initialize each particle's position and velocity

Step8: $y_{p,q}$= $_{min}$)

Step9:   =   /3 ,   /3)

(l+1) =   (l) +   (l)) + 2 (G- (l))

New velocity

(l+1) =   (l) +   (l+1)

Where

p denotes the particle index

l denotes discrete time index

z_p denotes velocity of n^{th} particle

y_p denotes place of n^{th} particle

p_n denotes best position found by n^{th} particle (personal best)

J denotes best position found by swarm(global best, best of personal bests)

$J_{(1,2)i}$- random number on the interval[0,1]applied to the n^{th} particle

Step12:   =

// update global best position

Step13: if   ) <

Step 14:

Step15: end if

end for

Input the optimized output into GWO.

Step16:Initialize GWO   (i=1, 2, …n)
Initialize x, X, and Y

Step 1 :Calculate fitness function for every search agent
best search agent
second best search agent
Third best search agent

While (T<Max iterations)
For (in every pack)
Update current position of wolf by eq. (1)
Update x, X and Y
Calculate the fitness function for all search agents Update , , and
End for
For best pack insert migration ( )
Evaluate fitness function for new individuals selection of best pack New random individuals for migration
End if
End while

In this chapter we proposed PSO-GWO technique which is discussed in detail and also explain step by step the working of proposed PSO-GWO technique. The BAT algorithm is used in this work for comparison of results with it. PSO and GWO both are used for the optimization locally and globally. Both the algorithm is based on the biological behavior of swarms and Bats.

Chapter 5

Implementation and Experimental Results

In the previous chapter we had discussed the proposed algorithm. In this chapter we are going to discuss the experimental setup along with the results.

5.1 Experimental Setup

There are many scientific workflow simulators which are designed to simulate the workflows which are running on multi-core cluster system. Scientific workflow is a paradigm for presenting complex computations.

Scientific workflow application usually has complicated data dependence and process dependence. It also has very high requirement on computing, storage and network resources. Meanwhile, the high-performance cloud computing platform is more and more intricate. So the performance of scientific workflow under the intricate computing environment is more and more arduous to predict. For more accurateness simulation framework for scientific workflow is required, which can simulate the features of workflow, the wide range of resources, the resource scheduling algorithm and system performance.

There are many simulators used in cloud namely CloudSim, CloudAnalyst, Green Cloud, iCanCloud, Integrated Emulation and Simulation (EMUSIM), Workflowsimulator, GroundSim, Data Centre Simulator (DCSim). Above all CloudSim is an expandable simulation toolkit that permits modeling and simulation of cloud computing systems and application provisioning environments. It aims to quantify and compare the performance of different application's scheduling and allocation policies. However, CloudSim can only describe individual task, it cannot be used to simulate the workflow. WorkflowSim extends the CloudSim by providing a layer of workflow management. Based on Directed Acyclic Graph, a workflow model was established that can simulate the process dependency of tasks. However, the process dependency and data dependency exists in the scientific workflow at the same time. WorkflowSim can simulate the data dependency of the workflow. In other words, it can simulate the data communication among the tasks. Therefore we are using workflowsimulator in our work..

Table 5.1: Comparison of Cloud Computing Simulators [37]

Simulator	Programming Language	Underlying Platform	Software/Hardware
CloudSim	Java	GridSim	Software
WorkflowSim	Java	CloudSim	Software
CloudAnalyst	Java	CloudSim	Software
GreenCloud	C++, OTcl	Ns2	Software
NetworkCloudSim	Java	CloudSim	Software
EMUSIM	Java	AEF, CloudSim	Software
SPECI	Java	SimKit	Software
GroudSim	Java	-	Software
DCSim	Java	-	Software

Table 1 shows the analysis and difference between the cloud computing simulators which are based on underlying platform, developing language, software or hardware. Most of these simulators are based on software and are developed using Java.

5.2 WorkflowSim

WorkflowSim has been used to simulate the application for implementing the optimized PSO_GWO algorithm. WorkflowSim is a workflow simulator. It extends the features of cloudsim as it provides workflow support for simulation. It is developed by Wei Wei Chen, a University of Southern California student pursing Phd. Workflows are modeled here in the form of Directed Acyclic Graph (DAG). It models the failure node, WMS stack delay in various levels and implements various dynamic and static schedulers of workflow and algorithms of task clustering. Real executions form the source of parameters used in the workflows.

Workflow is formed of various components:

i. **Workflow Mapper:** It is responsible for creating a list of user activities called as tasks and assigning this list to the execution site. It coordinates with the workflow generator and imports xml format DAG files and metadata from there.

ii. **Workflow Engine:** It releases free tasks to clustering engine. It also makes sure that dependencies are kept intact while releasing them. This means that only that task is released whose parent task has been completed.

iii. **Clustering Engine:** Its main task is to get various tasks together to form a job. While merging the tasks, it also makes sure that the scheduling overhead is reduced. It also merges failed tasks which are sent back by workflow scheduler.

iv. **Workflow Scheduler:** It considers the user criteria and matches jobs with the worker node. Simulation accuracy is greatly improved by workflow scheduler.

5.3 Scientific Workflows Used

i. **SIPHT:** It is used to automate the search for untranslatable RNA for bacterial replicons in the NCBI database.

ii. **CYBERSHAKE:** This workflow is used by the California to distinguish the earthquake hazards in a division.

iii. **Ligo:** It is used to originate and inspect gravitational waveforms from data composed during coalescing of consize binary system.

iv. **Montage:** It is created by NASA/IPAC together input image to fabricate custom mosaics of the sky.

v. **Genome:** It is generated by USC Epigenome center and used to automate diverse operation in genome sequence processing.

5.4 Results

Table 5.2: Comparison table of BAT Algorithm and proposed PSO-GWO Technique using SIPHT

RESULTS OF BAT ALGORITHM AND PROPOSED PSO-GWO TECHNIQUE USING –SIPHT				
Ensemble	**BAT Algorithm**		**Proposed PSO_GWO Technique**	
Size	**TET (ms)**	**TEC (Rs)**	**TET (ms)**	**TEC (Rs)**
2	23.06	5620.855	5.59	4550.063
4	41.16	10009.35	12.26	8603.69
6	38.05	13456.19	14.13	12125.43
8	52.98	13523.39	34.83	14265.19
10	87.61	17540.09	29.77	16958.17
12	62.5	20456.6	32.34	19965.58
14	52.64	21665.64	33.06	22432.37
16	62.21	25785.99	55.59	30112.12
18	93.3	33482.56	75.36	33230.33
20	72.04	34112.99	62.55	33012.34

Table 5.2: depicts the comparison of proposed and existing optimization algorithm using scientific workflow namely SIPHT. It shows the proposed technique outperforms as compared to BAT algorithm because it has better efficiency rates in terms of total estimated cost and time.

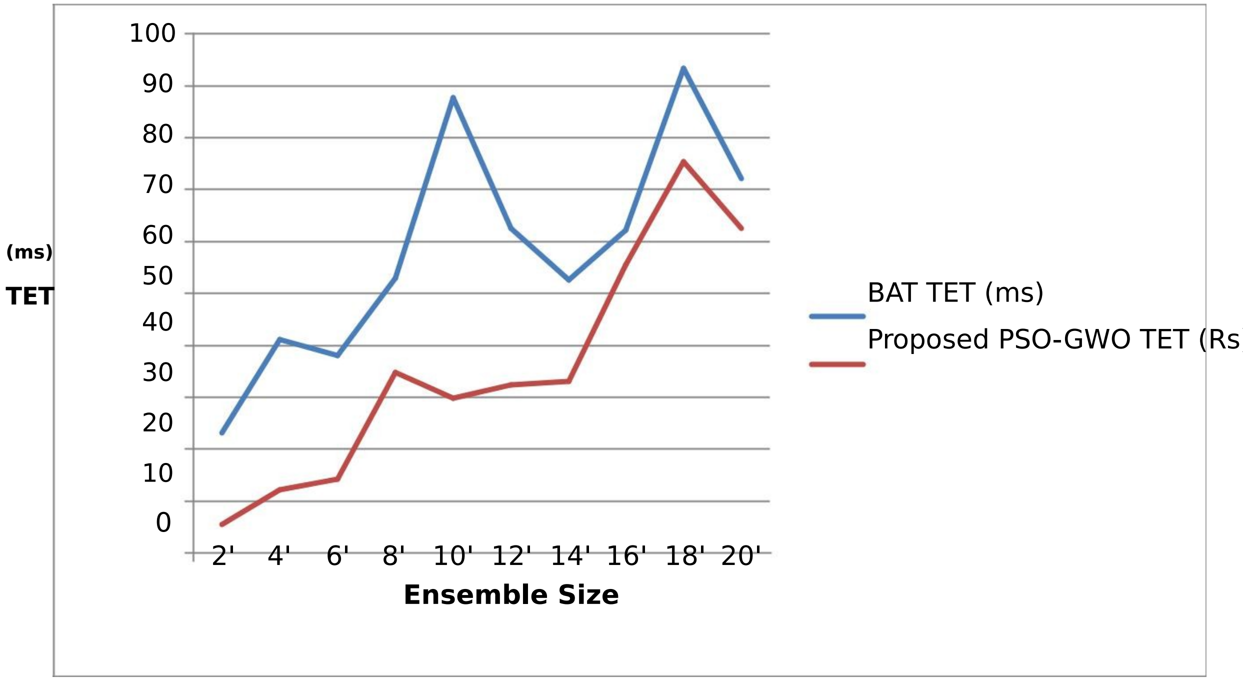

Figure 5.1: Comparison graph of TET of BAT Algorithm and proposed PSO-GWO technique using SIPHT

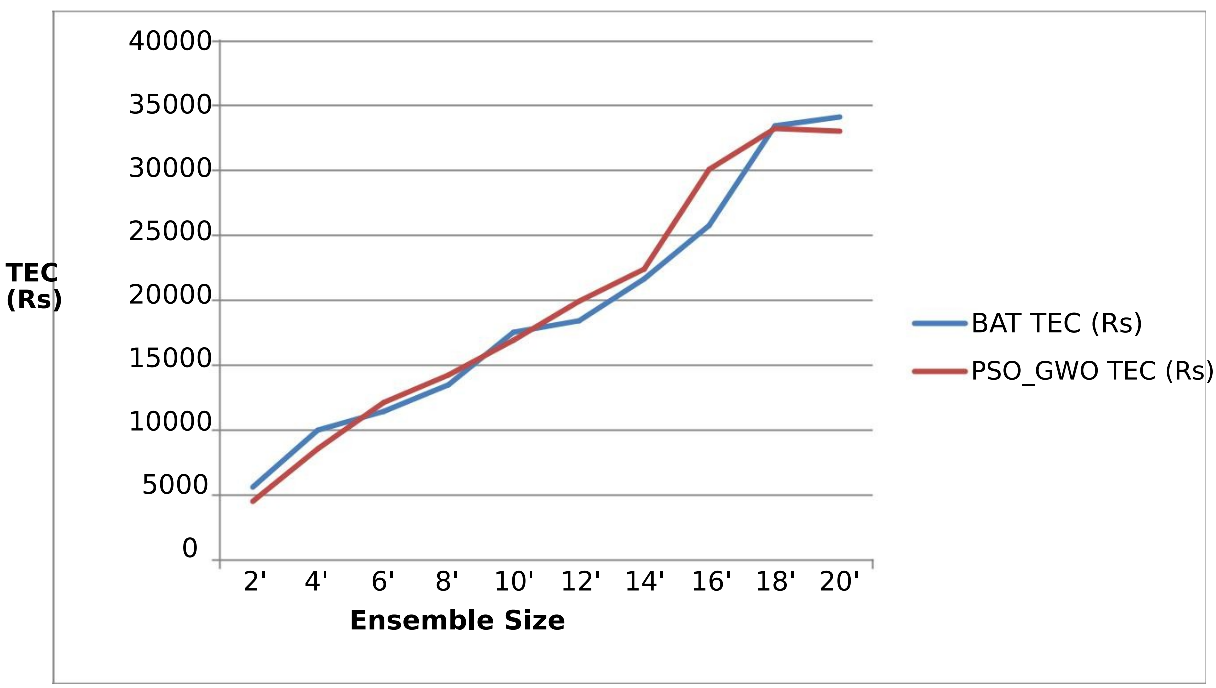

Figure 5.2: Comparison graph of TEC of BAT Algorithm and proposed PSO-GWO Technique using SIPHT

Table 5.3: Comparison table of BAT Algorithm and proposed PSO-GWO Technique using MONTAGE

RESULTS OF BAT ALGORITHM AND PROPOSED PSO_GWO TECHNIQUE USING –MONTAGE				
Ensemble Size	**BAT Algorithm**		**Proposed PSO_GWO Technique**	
	TET (ms)	**TEC (Rs)**	**TET (ms)**	**TEC (Rs)**
2	0.0	0	0.0	0.0
4	15.09	447.3718	1.73	299.958
6	17.69	628.3048	1.82	356.9556
8	20.95	730.9887	4	496.4608
10	25.47	1093.207	4.8	496.4608
12	27.89	1017.366	6.84	1147.339
14	31.15	1329.824	7.61	1131.746
16	32.2	1288.644	9.52	1480.906
18	35.07	1538.399	8.66	1278.957
20	36.3	1418.877	11.4	1745.794

Table 5.3: depicts the comparison of proposed and existing optimization algorithm using scientific workflow namely Montage. It shows the proposed technique outperforms as compared to BAT algorithm because it has better efficiency rates in terms of total estimated cost and time.

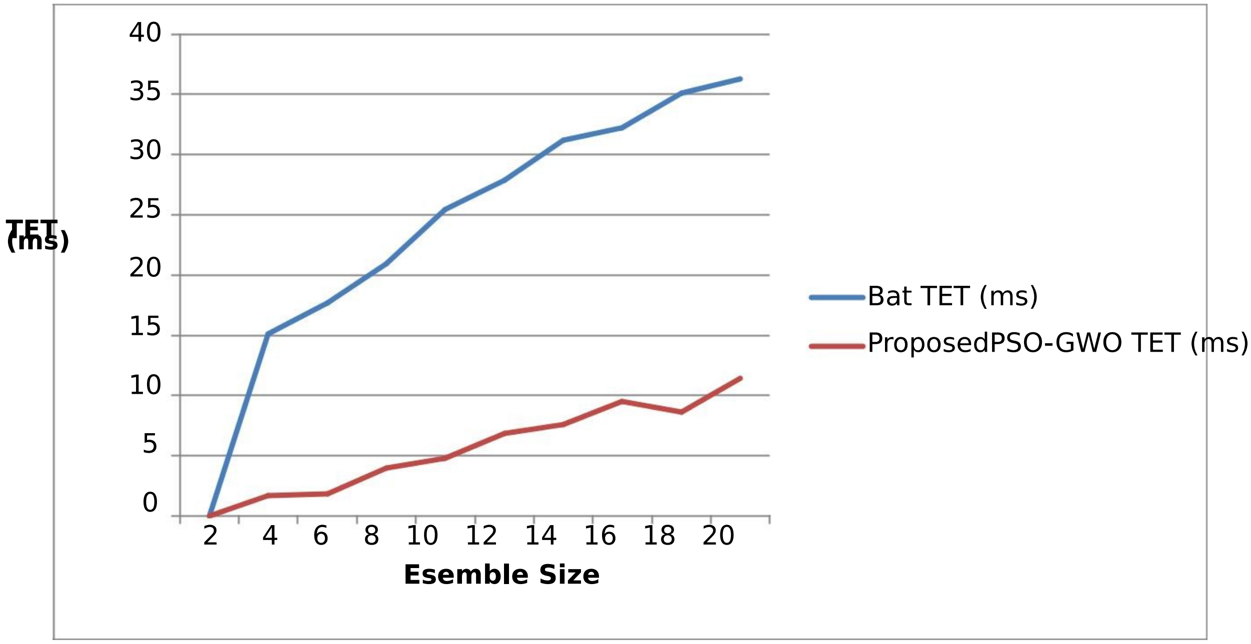

Figure 5.3: Comparison graph of TET of BAT Algorithm and proposed PSO-GWO Technique using MONTAGE

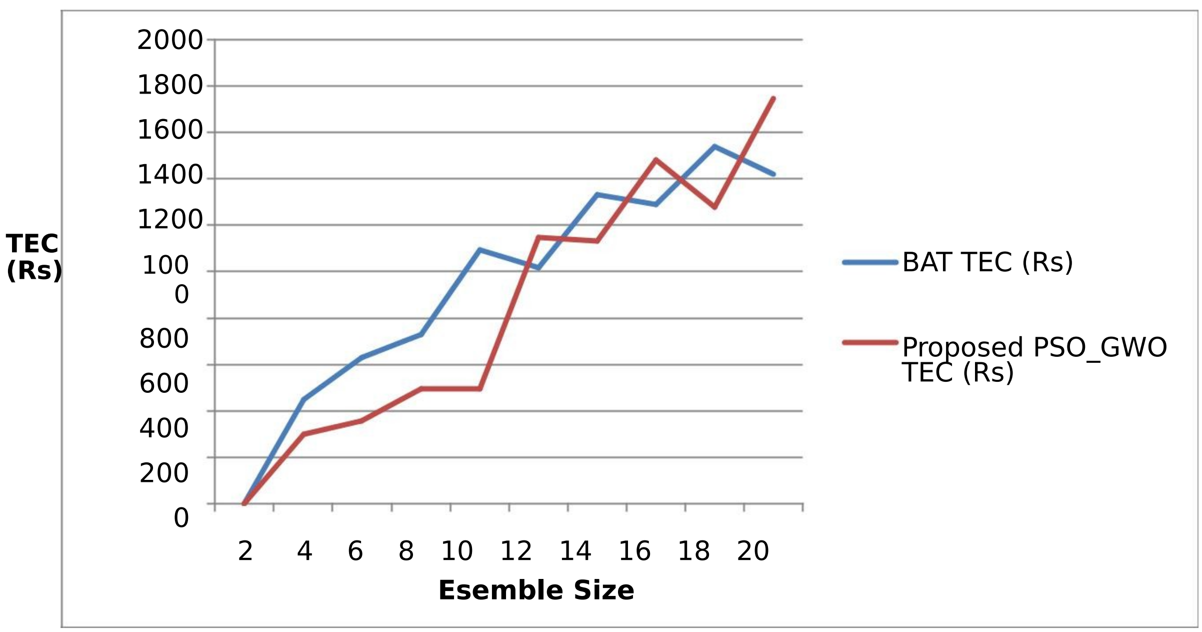

Figure 5.4: Comparison graph of TEC of BAT Algorithm and proposed PSO-GWO Technique using MONTAGE

Table 5.4: Comparison table of BAT Algorithm and proposed PSO-GWO Technique using CYBERSHAKE

RESULTS OF BAT ALGORITHM AND PROPOSED PSO-GWO TECHNIQUE USING CYBERSHAKE				
Ensemble size	**BAT Algorithm**		**Proposed PSO-GWO Technique**	
	TET (ms)	**TEC (Rs)**	**TET (ms)**	**TEC (Rs)**
2	0	0	0	0
4	18.57	822.6585	2.4	687.9981
6	24.25	1136.987	3.36	950.981
8	26.05	1545.962	6.92	1518.366
10	29.85	1883.18	4.50	1689.733
12	30.78	1729.39	8.08	1724.117
14	36.80	2530.267	10.96	2463.681
16	34.34	2106.646	9.25	2731.877
18	37.77	2311.253	11.58	1989.96
20	39.50	2541.908	11.42	1982.903

Table 5.4: depicts the comparison of proposed and existing optimization algorithm using scientific workflow namely Cybershake. It shows the proposed technique outperforms as compared to BAT algorithm because it has better efficiency rates in terms of total estimated cost and time.

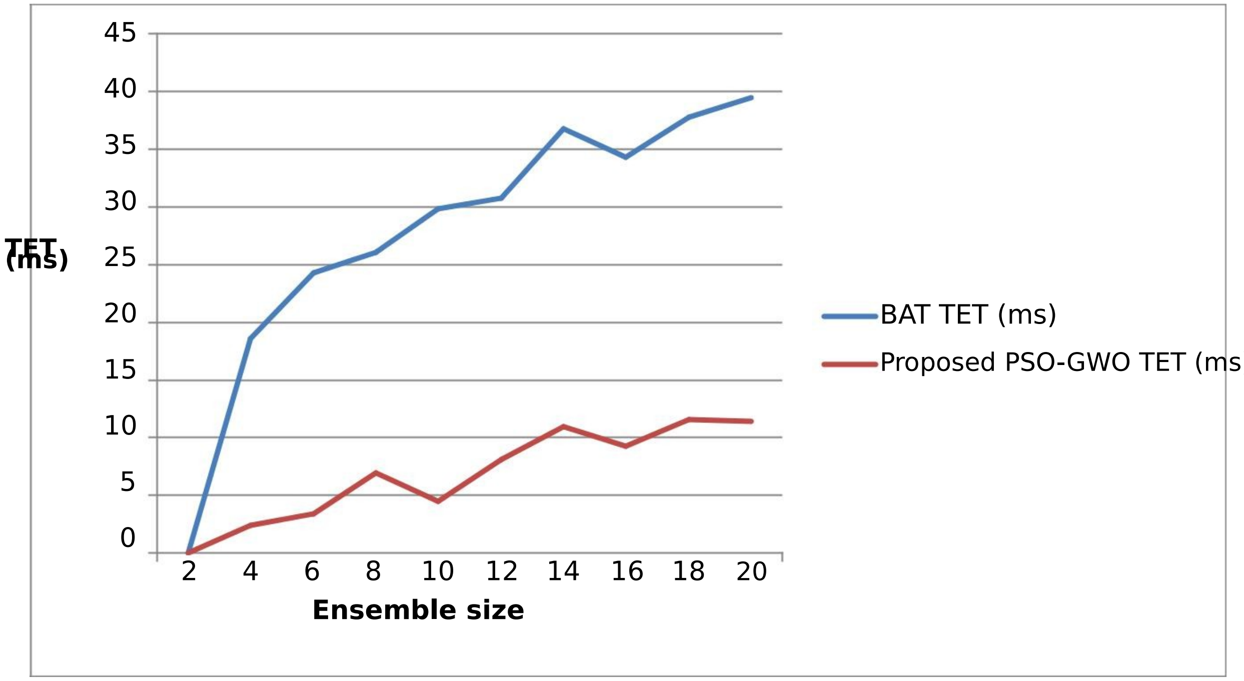

Figure 5.5: Comparison graph of TET of BAT Algorithm and proposed PSO-GWO Technique using CYBERSHAKE

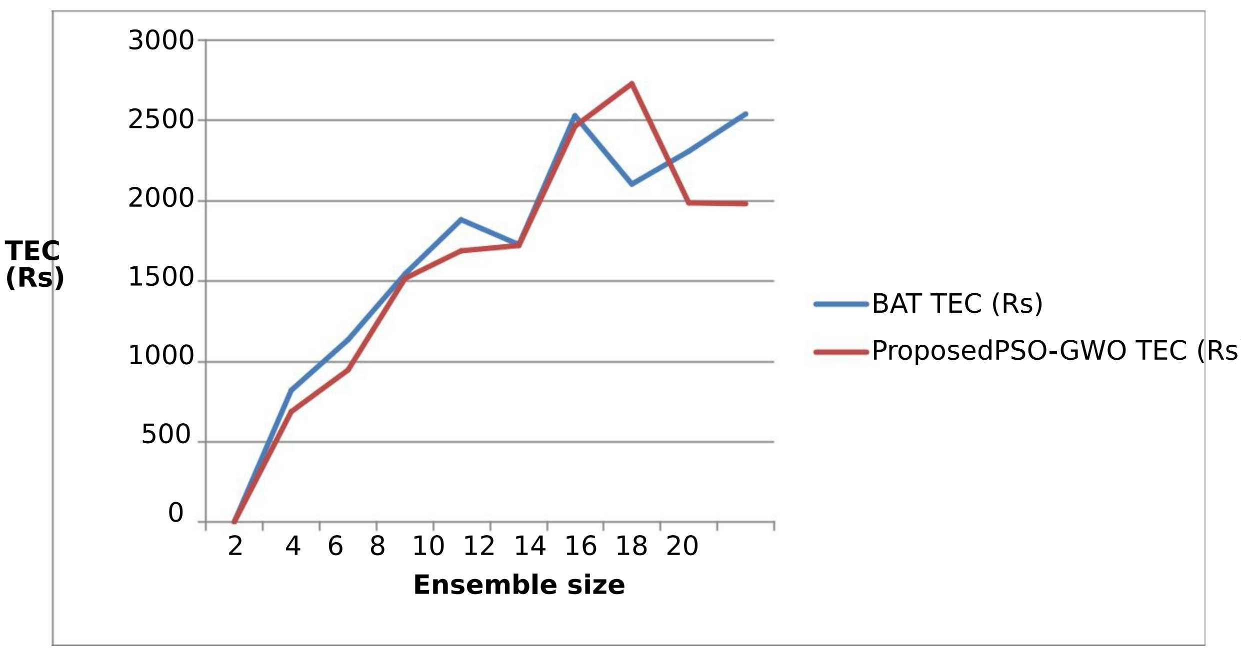

Figure5.6: Comparison graph of TEC of BAT Algorithm and proposed PSO-GWO Technique using CYBERSHAKE

Table 5.5: Comparison table of BAT Algorithm and proposed PSO-GWO Technique using LIGO

RESULTS OF BAT ALGORITHM AND PROPOSED PSO-GWO TECHNIQUE USING –LIGO				
Ensemble Size	**BATAlgorithm**		**Proposed PSO-GWO Technique**	
	TET (ms)	**TEC (Rs)**	**TET (ms)**	**TEC (Rs)**
2	0.0	0.0	0.0	0.0
4	44.91	2762.633	6.61	1467.243
6	65.89	4191.678	24.09	4022.422
8	35.66	4883.178	45.34	5699.963
10	68.82	5304.774	15.98	4708.369
12	102.39	7331.103	26.65	6271.538
14	58.96	7729.967	33.7	7550.524
16	151.28	10608.93	49.58	9948.208
18	87.77	11384.08	71.95	11188.03
20	153.22	13094.86	90.9	12582.79

Table 5.5: depicts the comparison of proposed and existing optimization algorithm using scientific workflow namely LIGO. It shows the proposed technique outperforms as compared to BAT algorithm because it has better efficiency rates in terms of total estimated cost and time.

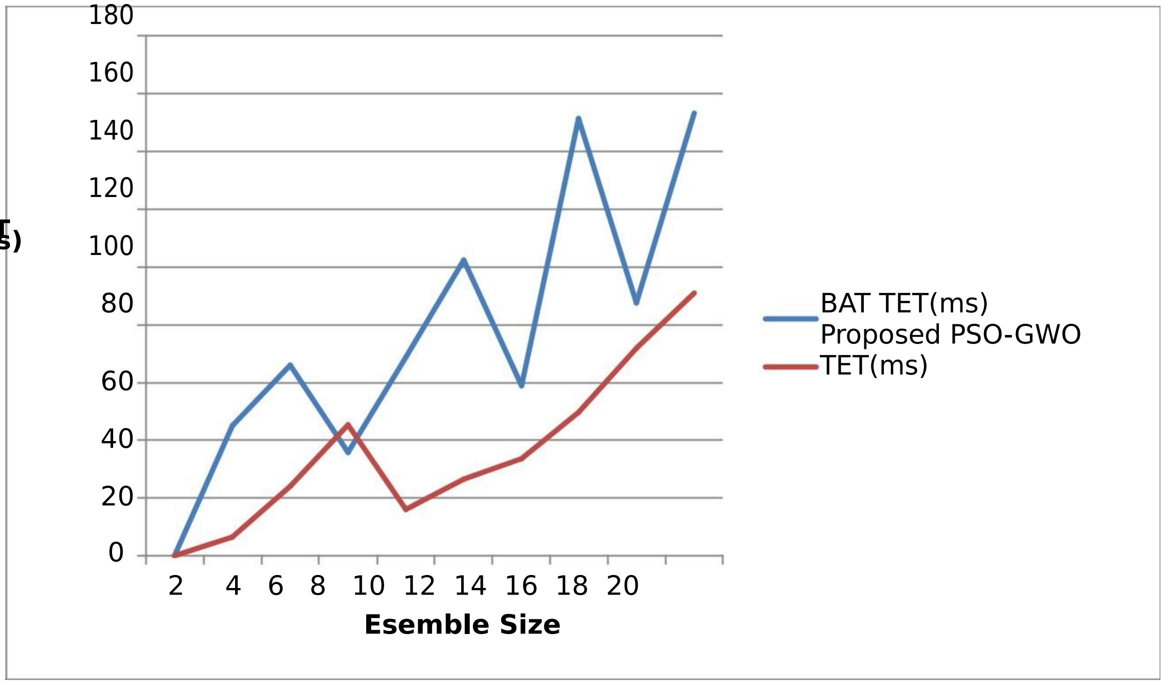

Figure 5.7: Comparison graph of TET of BAT Algorithm and proposed PSO-GWO Technique using LIGO

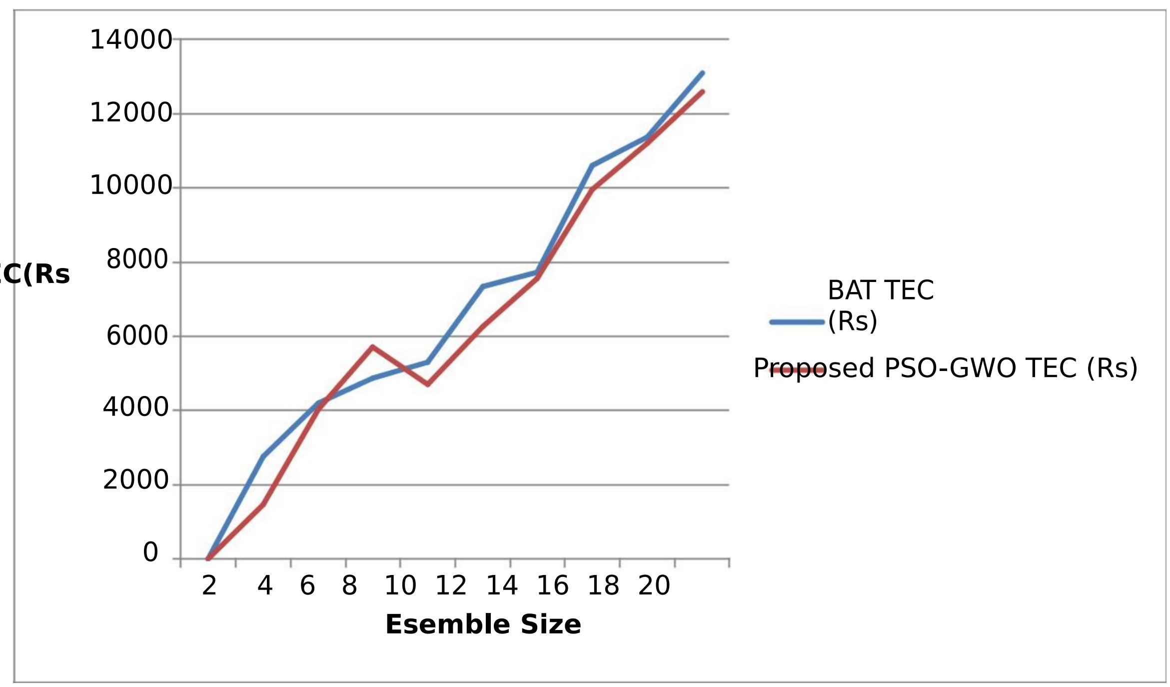

Figure 5.8: Comparison graph of TEC of BAT Algorithm and proposed PSO-GWO Technique using LIGO

Table 5.6: Comparison table of BAT Algorithm and Proposed PSO-GWO Technique using GENOME

RESULTS OF BAT ALGORITHM AND PROPOSED PSO-GWO TECHNIQUE USING –GENOME				
Ensemble Size	**BATAlgorithm**		**Proposed PSO-GWO Technique**	
	TET (ms)	**TEC (Rs)**	**TET (ms)**	**TEC (Rs)**
2	**12.23**	**8825.232**	23.2	9915.513
4	137.78	7549.552	73.24	27030.74
6	201.57	42511.93	133.83	42569.68
8	370.15	28155.45	315.65	41656.28
10	284.46	58872.21	745.24	86234.89
12	365.5	77034.56	546.78	88012.68
14	423.76	66004.98	317.76	93654.41
16	486.7	82720.56	401.95	106579.9
18	530.13	113448.6	791.04	123916.8
20	419.1	119935.5	491.21	139920

Table 5.6: depicts the comparison of proposed and existing optimization algorithm using scientific workflow namely Genome. It shows the proposed technique outperforms as compared to BAT algorithm because it has better efficiency rates in terms of total estimated cost and time.

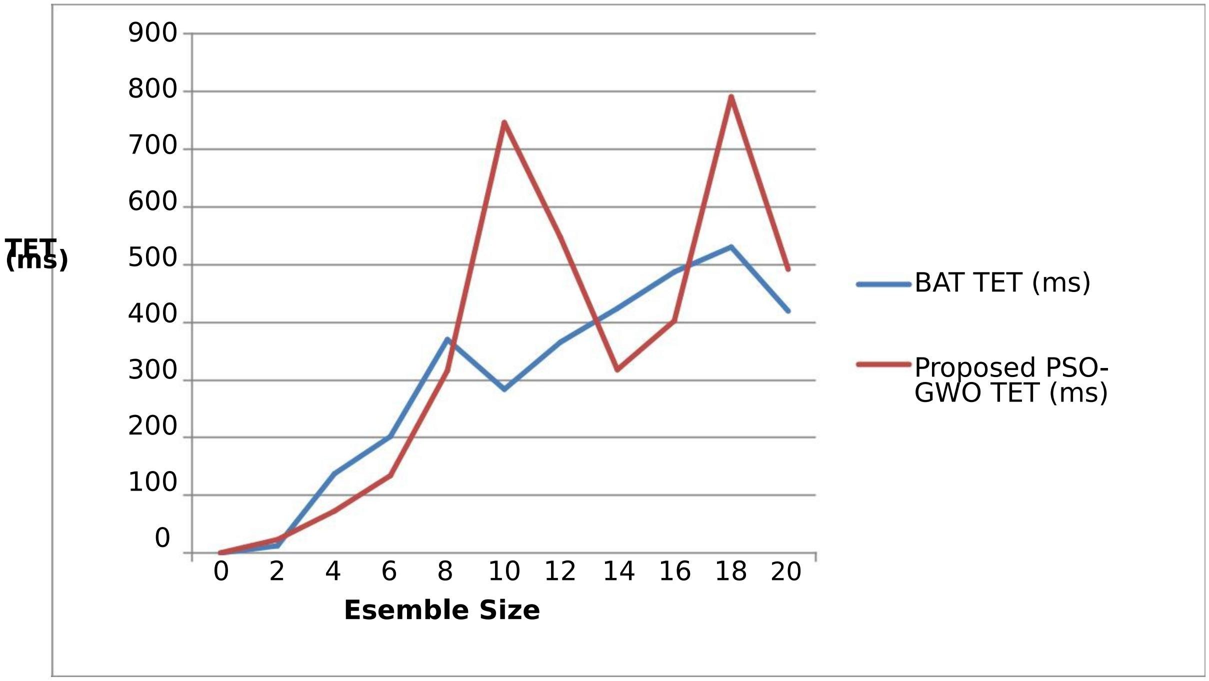

Figure 5.9: Comparison graph of TET of BAT Algorithm and proposed PSO-GWO Technique using GENOME

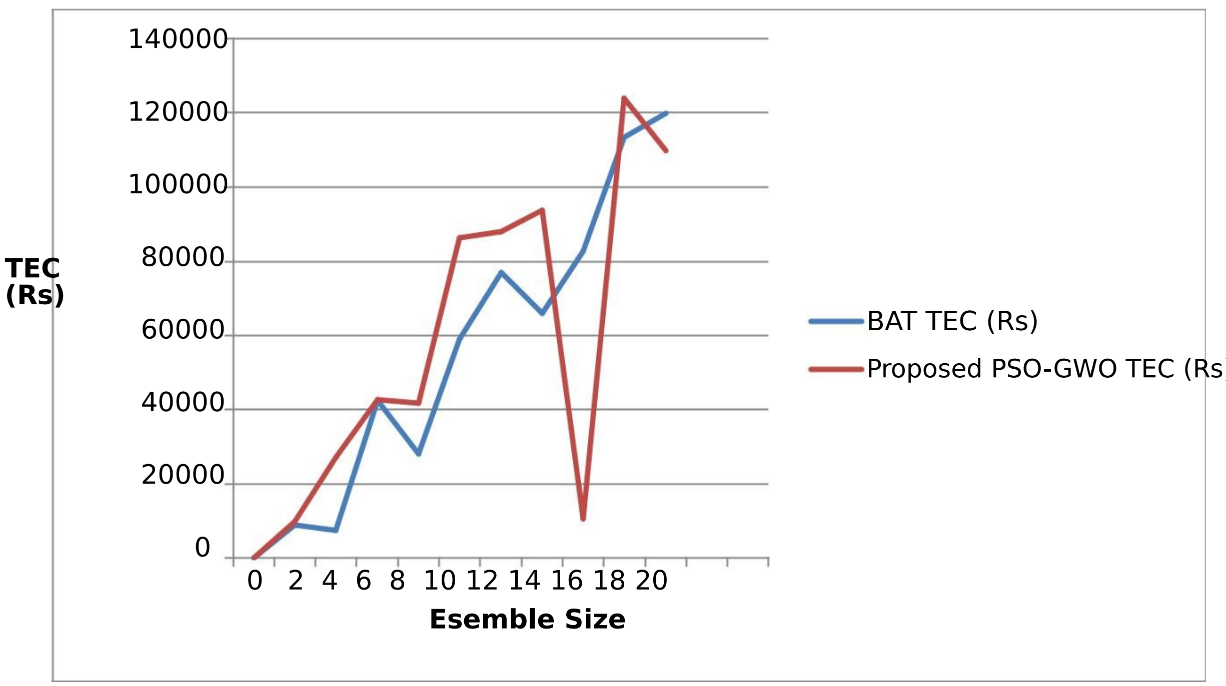

Figure 5.10: Comparison graph of TEC of BAT Algorithm and proposed PSO-GWO Technique using GENOME

5.5 Result Analysis:

In above given graphs and tables, represented a comparative analysis of TET and TEC parameters on the foundation of Bio inspired optimization (BAT) and Particle Swarm optimization (PSO) with Grey Wolf Optimization. In experiment, we used workflow planning in cloud domain with the utilization of different type of scientific workflow. In our analysis, total cost and execution time are improved by optimization but optimization also dependent on initializing factors. In the proposed approach, we use Pareto distribution instead of random initialization. If random distributions are used, more time will be taken to converge and sometime enforces the convergence by iteration but enforcing of convergence will increase the computation and execution time therefore does not meet the deadline condition. So, task initialization is an important task as defined in this paper. Another thing represented in these graphs and tables is that PSO_GWO performs better in comparison to BAT for reduction of cost and time because of the random crossover.

Chapter 6

Conclusion and Future Work

In this work, we proposed the mechanism for execution of the sensible forms on the IaaS clouds. The main issue in the cloud computing while decreasing makespan is execution cost. This issue is solved by using Hybridizarion of PSO with GWO. The tests were directed by mimicking four surely understood work processes (Cybershake, Ligo, Genome, Montage) on Workflowsim, which demonstrates that our answer has a general more beneficial execution than other existing algorithms. The worthy results are achieved because Particle Swarm Optimization (PSO) play important role in global optimization and GWO optimize locally and we have merge the two algorithms by taking the best out of them. With the proposed approach in most of the work processes we can bring down cost efficient schedule then additionally decreasing the time delay.

Future Scope

In future, we might need to investigate peculiar choices for your choice of the preliminary resource pool as it is influenced by the execution of the figure. Various optimization approaches and algorithms would also be examines for their execution. Another task for future would be analyzing the data exchange between the servers to show that virtual machines may use in various areas. Also we try to include response time and quality factors At long last, we need to actualize our approach in a work process motor with the target so it may be used for sending applications, things considered, and conditions.

References

[1] Zhang, Qi, Lu Cheng and RaoufBoutaba. "Cloud computing: state-of-the-art and research challenges." *Journal of internet services and applications* 1.1 (2010): 7-18.

[2] Vecchiola, Christian, SurajPandey and RajkumarBuyya. "High-performance cloud computing: A view of scientific applications." *Pervasive Systems, Algorithms, and Networks (ISPAN), 2009 10th International Symposium on*. IEEE, 2009.

[3] Zhao, Yong et al. "Opportunities and challenges in running scientific workflows on the cloud." *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2011 International Conference on*. IEEE, 2011.

[4] Vöckler, Jens-Sönke et al. "Experiences using cloud computing for a scientific workflow application." *Proceedings of the 2nd international workshop on Scientific cloud computing*. ACM, 2011.

[5] Pandey, Suraj et al. "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments." *Advanced information networking and applications (AINA), 2010 24th IEEE international conference on*. IEEE, 2010.

[6] Iosup, Alexandru et al. "Performance analysis of cloud computing services for many-tasks scientific computing." *IEEE Transactions on Parallel and Distributed systems* 22.6 (2011): 931-945.

[7] Ostermann, Simon et al. "A performance analysis of EC2 cloud computing services for scientific computing." *International Conference on Cloud Computing*. Springer, Berlin, Heidelberg, 2009.

[8] Deelman, Ewa et al. "The cost of doing science on the cloud: the montage example." *High Performance Computing, Networking, Storage and Analysis, 2008. SC 2008. International Conference for*. IEEE, 2008.

[9] Deelman, Ewa et al. "The cost of doing science on the cloud: the montage example." *High Performance Computing, Networking, Storage and Analysis, 2008. SC 2008. International Conference for*. IEEE, 2008.

[10] Callaghan, Scott et al. "Scaling up workflow-based applications." *Journal of Computer and System Sciences* 76.6 , 2010: 428-446.

[11] Alkhanak, EhabNabiel and Sai Peck Lee. "A hyper-heuristic cost optimisation approach for Scientific Workflow Scheduling in cloud computing." *Future Generation Computer Systems*(2018).

[12] Choudhary, Anubhav et al. "A GSA based hybrid algorithm for bi-objective workflow scheduling in cloud computing." *Future Generation Computer Systems* (2018).

[13] Sagnika, Santwana, Saurabh Bilgaiyan and Bhabani Shankar Prasad Mishra. "Workflow Scheduling in Cloud Computing Environment Using Bat Algorithm." *Proceedings of First International Conference on Smart System, Innovations and Computing*. Springer, Singapore, 2018.

[14] Vinothina and R. Sridaran. "An Approach for Workflow Scheduling in Cloud Using ACO." *Big Data Analytics*. Springer, Singapore, 2018. 525-531.

[15] Liu, Li et al. "Deadline-constrained coevolutionary genetic algorithm for scientific workflow scheduling in cloud computing." *Concurrency and Computation: Practice and Experience* 29.5 (2017).

[16] ErJyoti Garg, and Gurjit Bhathal. "Research Paper on Genetic Based Workflow Scheduling Algorithm in Cloud Computing." *International Journal of Advanced Research in Computer Science* 8.5 (2017).

[17] Rimal, Prasad Bhaskar, and Martin Maier. "Workflow scheduling in multi-tenant cloud computing environments." *IEEE Transactions on Parallel and Distributed Systems* 28.1 (2017): 290-304.

[18] Casas, Israel et al. "A balanced scheduler with data reuse and replication for scientific workflows in cloud computing systems." *Future Generation Computer Systems* 74 (2017): 168-178.

[19] Parmeet Kaur and Shikha Mehta. "Resource provisioning and work flow scheduling in clouds using augmented Shuffled Frog Leaping Algorithm." *Journal of Parallel and Distributed Computing* 101 (2017): 41-50.

[20] Bölöni, Ladislau and DamlaTurgut. "Value of information based scheduling of cloud computing resources." *Future Generation Computer Systems* 71 (2017): 212-220.

[21] Ghose, Manojit et al. "Energy Efficient Scheduling of Scientific Workflows in Cloud Environment." *High Performance Computing and Communications; IEEE 15th*

International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2017 IEEE 19th International Conference on. IEEE, 2017.

[22] Li, Yibin et al. "Energy optimization with dynamic task scheduling mobile cloud computing." *IEEE Systems Journal*11.1 (2017): 96-105.

[23] Quang-Hung, Nguyen, Nguyen Thanh Son, and Nam Thoai. "Energy-Saving Virtual Machine Scheduling in Cloud Computing with Fixed Interval Constraints." *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXXI*. Springer, Berlin, Heidelberg, 2017. 124-145.

[24] Ghanam, Yaser, J. Ferreira and F. Maurer, "Emerging issues and Challenges in Cloud Computing - A hybrid approach," in *Journal of Software Engineering and Applications*, vol. 5, no.11A, pp. 923- 937, Nov. 2012.

[25] Mell, Peter and T. Grance ,"The NIST definition of cloud computing." Oct. 25, 2011.

[26] Malawski, Maciej, G.Juve, E. Deelman and J. Nabrzyski, "Cost-and deadline-constrained provisioning for scientific workflow ensembles in Iaas clouds." in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, pp. 22, 2012.

[27] Mao, Ming, and M. Humphrey, "Auto-Scaling to Minimize Cost and Meet Application Deadlines in Cloud Workflows" in *International Conference High Performance Computing, Networking, Storage and Analysis (SC),* pp. 1-12, Nov. 2011.

[28] Rodriguez Sossa and Rajkumar Buyya, "Deadline based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds." In *IEEE Transactions* on *Cloud Computing,* vol. 2, no. 2, pp.222-235, June 2014.

[29] Rahman, Mustafizur, Xiaorong Li and Henry Palit, "Hybrid Heuristic for Scheduling Data Analytics Workflow Applications in Hybrid Cloud Environment" in *IEEE International Parallel & Distributed Processing Symposium,* pp. 966-974, May2011.

[30] Sajid, Mohammad and Z. Raza, "Cloud Computing: Issues and Challenges," in *International conference on Cloud, Big Data and Trust*, pp. 35 – 41, Nov. 2013.

[31] M. Sridhar and G.R.M Babu ,“ Hybrid Particle Swarm Optimization scheduling for cloud computing” in *IEEE International Advance Computing Conference (IACC)*, pp. 1196- 1200 , June 2015.

[32] S.Xavier and S.P.J Lovesum, “ A Survey of Various Workflow Scheduling Algorithms” in Cloud Enviroment” in *International Journal of Scientific and Research Publications,* vol. 3, no.2, Feb. 2013.

[33] S.Tilak and D. Patil,“ A Survey of Various Algorithms in Cloud Environment” in *International* Chen, Wei-Neng and Jun Zhang, “An Ant Colony Optimization Approach to a GridWorkflow Scheduling Problem With Various QoS Requirements” in *IEEE Transactions on systems, man and cybernatics-Part C: Applications and reviews,* vol. 39, no. 1, Jan. 2009.

[34] Service and Deployment.[online]. http://www.ques10.com/p/3971/enlist-and-explain-various-service-model-and-dep-1/ Accessed on 30 May,2018

[35] Characteristics of cloud computing.[online]https://www.researchgate.net/figure/Essential-characteristics-of-cloud-computing-16_fig1_310477656. Accessed on 29 May,2018

[36] I. A. Hameed, R. T. Bye & O. L. Osen “Grey wolf optimizer (GWO) for automated offshore crane design.” In *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on* (pp. 1-6). IEEE, December, 2016.

[37] Wei Zhao, Yong Peng, Feng Xie, Zhonghua Dai, “Modeling and Simulation of Cloud Computing: A Review” , *2012 IEEE Asia Pacific Cloud Computing Confreence,* IEEE, 2012

List of Publications

[1] Damini Arora, Seema Bawa. "Workflows Scheduling in Cloud by Hybridization of Particle Swarm Optimization (PSO) with Grey Wolf Optimization (GWO)." *International Conference on Advanced Computing and Intelligent Engineering* 2018. [Communicated]

Thesis File

ORIGINALITY REPORT

7%	4%	6%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Cleveland State University Student Paper	1%
2	www.slideshare.net Internet Source	1%
3	Ghafarian, Toktam, and Bahman Javadi. "Cloud-aware data intensive workflow scheduling on volunteer computing systems", Future Generation Computer Systems, 2015. Publication	<1%
4	Submitted to Jawaharlal Nehru Technological University Student Paper	<1%
5	Khatibi Bardsiri, Vahid, Dayang Norhayati Abang Jawawi, Siti Zaiton Mohd Hashim, and Elham Khatibi. "A flexible method to estimate the software development effort based on the classification of projects and localization of comparisons", Empirical Software Engineering, 2014. Publication	<1%

6	Ehab Nabiel Alkhanak, Sai Peck Lee. "A hyper-heuristic cost optimisation approach for Scientific Workflow Scheduling in cloud computing", Future Generation Computer Systems, 2018 Publication	<1%
7	Wang, Gaige, and Lihong Guo. "A Novel Hybrid Bat Algorithm with Harmony Search for Global Numerical Optimization", Journal of Applied Mathematics, 2013. Publication	<1%
8	www.nema.org Internet Source	<1%
9	Li, Yibin, Min Chen, Wenyun Dai, and Meikang Qiu. "Energy Optimization With Dynamic Task Scheduling Mobile Cloud Computing", IEEE Systems Journal, 2015. Publication	<1%
10	Shadi Torabi, Faramarz Safi-Esfahani. "A dynamic task scheduling framework based on chicken swarm and improved raven roosting optimization methods in cloud computing", The Journal of Supercomputing, 2018 Publication	<1%