ARTIFICIAL BEE COLONY BASED OPTIMIZATION FOR DATA SHARING IN INTERNET OF THINGS

ANEES ASGHAR

NATIONAL UNIVERSITY OF MODERN LANGUAGES

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with overall exam performance, and recommend the thesis to the Faculty of Engineering and Computer Sciences.

THESIS TITLE: ARTIFICIAL BEE COLONY BASED OPTIMIZATION FOR DATA SHARING IN INTERNET OF THINGS

Anees Asghar	22MS/CS/S19
Submitted By:	Registration #:
Master in Computer Science	Computer Science
Title of the Degree	Name of Discipline

__________	Signature: __________
Name of External Examiner	
__________	Signature: __________
Name of Internal Examiner	
Dr. Ata Ullah	Signature: __________
Name of Research Supervisor	
__________	Signature: __________
Name of Co-Supervisor	
Dr. Sajjad Haider	Signature: __________
Name of HoD (CS)	
Dr. Basit Shahzad	Signature: __________
Name of Dean (FE&CS)	
Prof. Dr. Muhammad Safeer	Signature: __________
Name of Pro-rector Academics	

July 5th, 2021

"I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of the Degree of Master of Science in (*Computer Science*)"

Signature	:	______________________
Name	:	Assoc. Prof. Dr. Ata Ullah
Date	:	July 5th, 2021

ARTIFICIAL BEE COLONY BASED OPTIMIZATION FOR DATA SHARING IN INTERNET OF THINGS

ANEES ASGHAR

A thesis submitted in fulfillment of the
requirements for the award of the degree of
Master of (Computer Science)

Department of Computer Sciences
National University of Modern Languages

July 2021

DECLARATION

I declare that this thesis entitled "*Artificial Bee Colony Based Optimization for Data Sharing in Internet of Things*" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	________________________
Name	:	Anees Asghar
Date	:	July 5th, 2021

This thesis work is dedicated to my parents and my teachers throughout my education career who have not only loved me unconditionally but whose good examples have taught me to work hard for the things that I aspire to achieve.

ACKNOWLEDGEMENT

First of all, I wish to express my gratitude and deep appreciation to Almighty Allah, who made this study possible and successful. This study would not be accomplished unless the honest espousal that was extended from several sources for which I would like to express my sincere thankfulness and gratitude. Yet, there were significant contributors for my attained success and I cannot forget their input, especially my research supervisors, Assoc. Prof. Dr. Ata Ullah, who did not leave any stone unturned to guide me during my research journey.

I shall also acknowledge the extended assistance from the administrations of Department of Computer Sciences who supported me all through my research experience and simplified the challenges I faced. For all whom I did not mention but I shall not neglect their significant contribution, thanks for everything.

ABSTRACT

Internet of Things (IoT) comprises of complicated and dynamical aggregation of smart units that normally need decentralized command for data sharing across the networks. The most popular swarm intelligent techniques artificial bee colony (ABC) is inspired from collective actions of honey bees that can be used for solving problems during clustering in large scale data of IoT. The main problem is that each food source is compared with every other food source in neighborhood to determine the best global food source. It requires unnecessary comparisons to compare the pair of poor quality food sources as well. It results in consuming more utilization time, slow convergence speed and increased delay. This work presents an enhanced ABC (E-ABC) based optimization for data collection and replication mechanism. E-ABC improves the previous ABC algorithm by reducing the unnecessary comparisons. E-ABC compares the best source with the available sources which will excludes the comparison of poor resources. The proposed E-ABC algorithm was applied on replica selection to prove its supremacy as compared to counterparts in terms of convergence speed, data availability and response time. Results show the supremacy of proposed E-ABC over previous algorithms. The proposed algorithm provides 65% better response time from DCR2S and 20% better than MOABC when number of cloudlets are 1000. The file availability probability for E-ABC is noticed as 85% when total cost is 20. Additionally, some open research challenges are highlighted on the basis of literature which will help the researchers to find the research gap with respect to IoT and ABC.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABC	-	Artificial Bee Colony
qABC	-	Quick Artificial Bee Colony
iqABC	-	Improved Quick Artificial bee colony
IoT	-	Internet of Things
SI	-	Swarm Intelligence
MOABC	-	Multi Objective Artificial Bee Colony

LIST OF SYMBOLS

Ø	-	Phi symbol, represents random number
li	-	Lower bound
ui	-	Upper bound
$C_{jk}(st)$	-	Sending time of cloudlet
$C_{jk}(rt)$	-	Receiving time cloudlet
$\boldsymbol{x_{best}}$	-	Global best food source
$\boldsymbol{i_{best}}$	-	Index of the best food source
$\boldsymbol{v_{best}}$	-	New food source

CHAPTER 1

INTRODUCTION

1.1 Overview

Internet of things (IoT) is an arrangement of interrelated physical objects that are capable of exchanging the data over a network at an extraordinary rate. IoT empowers these physical objects to sense, hear, identify, think, communicate and exchange the data [1]. IoT is getting growing interest as it finds out a reliable solution for every problem. The use of IoT is equally crucial for the industrial systems, as IoT is a rising technology and provides the encouraging convenience to build up the industrial systems [2]. The domain and application of IoT are shown in Figure 1.1

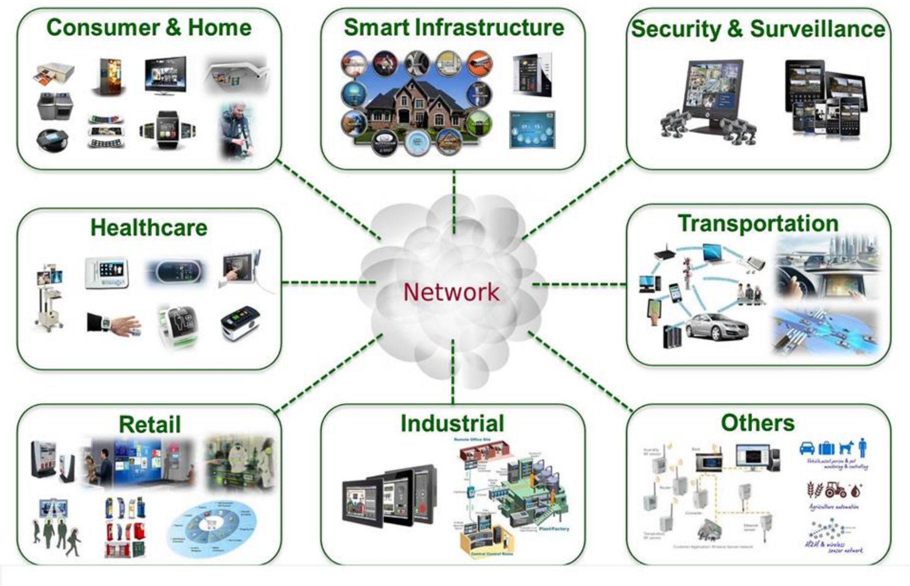

Figure 1.1: IoT Architecture and Applications

IoT can be utilized in numerous applications. These application areas of IoT can be classified into following fields but not bounded to: i) Healthcare field, ii) Smart Environment (smart cities, smart homes, smart hospitals, etc.) field, iii) Logistics and transportation field, iv) Personal and social field. These applications can face the issues such as limited storage, and limited processing. To tackle these issues swarm intelligence can be used [3].

Cloud servers refers to the on-demand accessibility of the services such as storing and processing the data. Cloud computing is widely utilized in numerous fields such as wireless networks, IoT, and big data etc. The job of cloud with respect to IoT is to function as a feature of a coordinated effort. The cloud servers are utilized in order to store the IoT data. As cloud is a central server therefore it is responsible to provide the accessibility of the services whenever required. Cloud computing reduces the cost of IT systems, and provides efficient, scalable, and flexible accessibility to the resources anywhere, anytime and is capable to transmit the huge data bundles/packets created by IoT via internet. Considering the role of Cloud Computing in IoT, it works for storing IoT data and for collaboration as well. The Cloud is an established server which carries computer resources that can be attained whenever required. In short cloud environment is a method that is utilized for transferring the data packages developed by internet of things via the internet.

Swarm intelligence (SI) is a technique inspired from the behavior of social insects' colonies or animal groups in nature such as ant and bee colonies, etc. An extensive scope of intelligent/smart machines in IoT exchange a tremendous amount of sensing data that should be proceeded intelligently to decrease the data redundancy that consequently decreases the transmission cost. SI-based algorithms are motivated from some natural events therefore such algorithms are widely known as nature-inspired algorithms (NIAs). These algorithms follow the smart behavior of social insects such as ant and honeybees, etc. several interactions among social colonies is the main concern of swarm based optimization algorithms. These algorithms have capability to solve the notably complex problems including complex real world optimized challenges, complex engineering problems, and complex multi-objective reliability optimization problems etc. [4].

D.Karaboga et al. [5] presented a swarm intelligence based optimization technique called Artificial Bee Colony (ABC) in 2005. ABC is applicable in Continuous optimal problems, data clustering, artificial neural systems, medical model categorization, travel salesman problem, data replication, optimizing the data aggregation problems, image processing [4]. ABC algorithm is rated as most frequently used in searching for optimum solutions. It is because of its uniqueness in problem-solving technique in which the solution of the problem arises from the intelligent behavior of honeybee insects. In ABC, basically, there are three types of bees: Employed, Onlooker and Scout Bees. Employed bee attached with the particular source of food for exploitation purpose. The onlooker looks out the dance of employed bees from their hives to pick the solution and the scout bees explore the sources at random as shown in Figure 1.2.

Trust worthy data collection is a crucial problem to help emergency applications within IoT enabled systems. The devices dispatch the crucial information to the base station for taking a decision. Moreover, base station dispatchs several commands toward the devices which are under its control [2], [6]. Transmitting essential information across the network needed a reliable backbone. The most frequently used method for data collection in IoT enabled systems is constructing a spanning tree across the available devices [7]–[11].

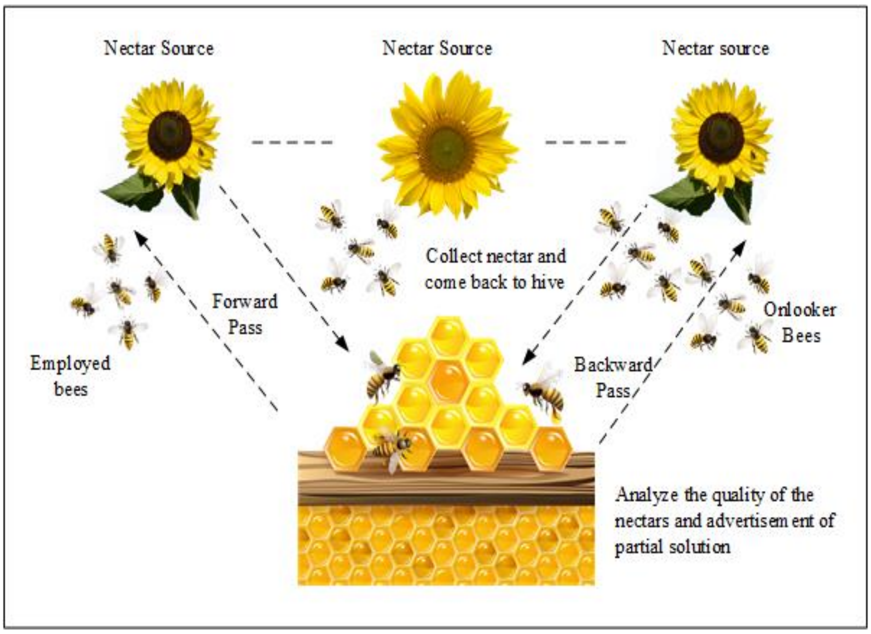

Figure 1.2: Intelligent behavior of bees

In this electronic generation, it is a truly challenging task to save, proceed and examine the gigantic and quickly growing data without proper tools or expertise. Data mining is a procedure to extract the relevant information and expertise by considering and analyzing such data through analytical algorithms. Data Clustering or cluster analysis is a kind of indirect data mining, since the purpose is to discover the link between each and every variable in contrast to direct data mining, in which several variables are highlighted as targets. ABC has proven to be effectively utilized in broad range of applications for example neural networking, sensor networking, image processing, knowledge discovery or data mining, industrial, mechanical and civil engineering, electronics and communication engineering. ABC have been effectively applied in data clustering problems and attain encouraging results which leads to the finer prediction and data analysis. The survey that presents the extensive understanding about the modified versions of ABC along with its applications in addressing the data clustering issues is discussed in [12].

1.2 Research Motivation

The goal of IoT is to fostering a paradigm whereby each and everything in our environment is turned into an intelligent object with sensing, storing, interacting and actuating capacity and is connected all the time. It is a field of research in which digitized as well as physical units (i.e. humans, objects, machines) are connected with each other through the internet, hence allowing for an entire new group of applications and services due to the fact that IoT enabled networks are complicated and have dynamical character. Flexibility and versatility makes SI an effective designed framework for the algorithm which solves the growing number of complicated problems, such that IoT enabled networks. Therefore, swarm intelligence establishes a cause of encouragement to the IoT enabled devices which can be designed as swarm of simple system or can incorporate swarm intelligence enabled techniques that attains the number of global objectives.

1.2.1 Application Areas

ABC is one of the most frequently studied algorithm, which is continuously encouraging the researchers to apply in solving various real world problems. ABC is a trending topic for the application areas including data clustering, aggregated data optimization, optimizing data replication, optimizing network attacks, improving the convolutional ANN, training ANN, travel time estimation, etc. These key features of the swarm intelligence and ABC motivated the researchers to work in this field. Most specifically ABC is the one which is grabbing the interests of the researchers in the recent times due to its efficient, optimized solutions and vast application areas.

1.3 Problems Statement

One of the modified variants of ABC algorithm is used in MOABC [13] for the placement of replications. In reality the artificial employees and the artificial onlookers chooses their nutrition sources in different way but this problem is not considered in the original ABC. Instead of this it chooses the nectar sources using same criteria or formula for both the artificial employed as well as onlooker bees phase. This problem is resolved in a modified version of ABC algorithm named qABC algorithm[14]. But in qABC during the best global food source selection, each food source is selected to compare with every other food source in neighborhood to determine the best solution. In this way, extensive number of comparisons are involved that consumes more utilization time, slow convergence speed and increases delay. Moreover, after consuming a lot of time the fitness evaluation function does not help to improve the quality of global best food source selection. In addition to this, its diversity can be further enhanced because it considers the best possible solution and ignore the rest of solutions. Another factor that causes slow convergence speed is, only one parameter is changed in solution search equation of parent solution.

1.4 Research Questions

In this section number of research questions are elaborated. These research questions are based on the research objectives that are discussed above.

i. What are the mechanisms to enhance the search schemas of ABC algorithm?
ii. What are the mechanisms to optimize the replication process using ABC?
iii. How to manage the balance between global and local search criteria in ABC?

1.5 Aim of the research

The aim of this work is to enhance the availability of data within low cost and achieve the better efficiency. It also aims to enhance the performance of servers as ABC provides optimal solution and makes data access faster. Data replication using ABC algorithm decreases the response time because data can be retrieved quickly. Files are saved at various locations which helps the users to select the files quickly. This process improves the response time of cloudlets.

1.6 Research Objectives

The research objective of the proposed work are listed below;

i. The objective of the proposed work is to improve the search schema of standard ABC and implement E-ABC algorithm in data replication scenario.
ii. To enhance the performance of ABC algorithm in terms of convergence speed and response time.
iii. To efficiently balance the global and local search criteria of ABC in terms of convergence rate.

1.7 Scope of Research

This research aims to improve the performance of the standard ABC algorithm in terms of convergence speed, and implement it in to the replication scenario to improve data availability and reduce delay. But the proposed E-ABC algorithm did not consider the storage cost.

1.8 Thesis Contributions

The major contributions of the proposed work includes:

i. This work explored the literature on ABC based schemes and identified a valid problem.
ii. This thesis presents the modified ABC based algorithm that improves the replication identification. It utilizes the modified ABC algorithm so that a user can achieve shorter and cost effective route in order to access or place the replicas.
iii. Finally, E-ABC is applied and tested by utilizing C# to investigate the performance where E-ABC is compared with MOABC and DCR2S.

1.9 Thesis Organization

The rest of work is structured as:

Chapter 2 presents a literature review for the data replication strategies. Chapter 3 portrays the system design and proposed structure, followed by chapter 4 that narrates the E-ABC algorithm in depth. Chapter 5 portrays the findings, formulations and simulations; moreover, it provides the details about the proposed algorithm's performance and its comparison with other algorithms. Finally, consequences and future research direction are put forward in Chapter 6.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In this Section, a number of bee colony based schemes are explored that focus on the optimization of different processes where datasets are also used. In contrast to other SI based algorithms, ABC is selected due to less number of parameters, good optimization and robustness. The information exchange among the individuals is an effective way to deal with the issues in the original ABC including slow convergence speeds and limited communication between the bees. The literature is mainly divided in two categories: 1) Different versions of ABC algorithm; 2) Implementation of ABC algorithm in various fields.

2.2 Preliminaries - Standard ABC algorithm

At initial period of the work, the bees search for the food sources. When they successfully find the food source, they store its position in their mind before leaving the hive. These bees start dancing around the food source in the dancing zone. In the first phase, food source position is initialized randomly whereas the calculations for the food source position are obtained as $x_{mi} = l_i + rand\,(0, 1) * (u_i - l_i)$. In the equation, x_{mi} represents n number of possible solutions for the problem with i dimensions. Moreover, l_i is bottom limit of x_{mi} and u_i represents the upper limit of the parameter x_{mi}. In the second stage, employee bees seek for better food source in their neighborhood. This new food source is identified as $v_{mi} = x_{mi} + \emptyset_{mi} (x_{mi}- x_{kj})$ where X_k is a food source which is chosen randomly and i represents the dimension selected randomly while $\emptyset_{mi}$ is a certain number chosen in a random way within a range [-1,1]. The function $fit(Xm)$ represents fitness of solution, calculated using equation (2.1) where $abs(f(Xm))$ represents absolute objective function value of the function $fit(Xm)$.

The onlookers use the shared information of employees to choose food source probabilistically. Probability value $Pm = \frac{fit(Xm)}{\sum_{m=1}^{SN} fit(Xm)}$ where SN is size of population [14].

$$fit(Xm) = \begin{cases} \frac{1}{1+f(Xm)} & if\ Xm \geq 0 \\ 1 + abs\big(f(Xm)\big) & if\ Xm < 0 \end{cases} \qquad (2.1)$$

2.3 Different versions of ABC algorithm

A number of schemes treat the data as a single unit and apply ABC algorithm to extract the parameters for decision making. In this Section, we explore the different versions of ABC algorithm. In [5], ABC is adopted for food foraging attitude of honey bees. ABC provides fast, robust, and flexible solutions for optimization problems, but it still requires improvements for further optimization. Later on different schemes are introduced to enhance the efficiency of the original ABC [15]. The summary of existing survey papers that focuses on ABC based schemes are elaborated in table 2.1. The researchers use different techniques to tackle the shortcomings of the original ABC algorithm. Parameter adjustment, modification of search equation and hybridization of the different algorithms are the well-known techniques adopted by the researchers to achieve the goal of global optimization.

Table 2.1: Summary of Existing Surveys focused on ABC based Schemes

Focused Topic of ABC	**Description**
A comprehensive survey: ABC algorithm and applications	Dervis Karaboga et al. [15] presented a survey that thoroughly discussed the advancements in ABC algorithm and its applications. Furthermore, different areas of researchers' interest are discussed comprehensively.
A systematic review on ABC optimization technique	Dishti Agarwal et al.[16] Described a solution to the different optimization problems using the Artificial bee colony. The basic idea of the author was to analyze the work done by utilizing ABC for optimization in different areas and to discover the advantages and disadvantages of the approach along with their applications.
A survey of ABC algorithm	Ying Liu et al. [17] surveyed, and comprehensively presents the basic biological principles and mathematical operators of ABC algorithm. And analyzes the recent research on ABC with aspects of algorithm enhancements and engineering applications.
ABC Algorithm: A Survey	Sangeeta Sharma et al. [18] made a detailed survey on ABC algorithm. The basic purpose was to investigate the performance of ABC with various size of population.
Review on ABC Algorithms and Applications to Data Clustering	Ajit Kumar et al. [12] presented a concise survey on ABC algorithm along with variants and discussed the numerous applications of ABC in clustering.
A review on ABC algorithm	Balwant kumar et al. [19] provided a literature survey which concluded that a substantial part of research is focusing on enhancing the ABC algorithm to tackle the wide range of issues.
Overview of Artificial bee colony and its Applications	Fahad S. Abu-Mouti et al. [20] presented a concise overview of literature regarding ABC algorithm. Moreover, provides a comprehensive survey on the meta-heuristic optimization algorithms.

2.3.1 Search Equation based Schemes

In the first two critical stages of the ABC algorithm, the search formula contributes significantly by determining the diversity and convergence of the algorithm. By simulating the performance of swarm populations, the SI algorithms were designed and analyzed. But there are some differences in the behavior of real bee foraging and original ABC algorithm. Therefore, by modifying the search formula some scholars improved the original ABC. A new model of ABC for onlooker bees titled qABC overcomes the local search proficiency problem that occurs in ABC. The qABC provides a new description for onlooker bees. In qABC, way of choosing the sources of onlooker bees is changed from the employee bees. The qABC introduces a new description for attitude of onlooker bees which is defined in equation (2.2) where the ideal solution among the neighbor source of Xm and itself Nm is represented by the $X_{Nm,i}^{best}$. By considering mean Euclidean distance among Xm and other possible solutions for the numerical problems, the neighborhood of the solution Xm is defined. Therefore, the Euclidean distance among Xm and Xj is represented by $d(m,j)$ and mean Euclidean distance is represented by mdm, which is computed using Equation (2.3).

$$\boldsymbol{U_{Nm,i}^{best}} = \boldsymbol{X_{Nm,i}^{best}} + \emptyset_{m,d} \, (\boldsymbol{X_{Nm,i}^{best}} - \boldsymbol{X_{k,i}}) \tag{2.2}$$

$$\boldsymbol{md_m} = \frac{\sum_{j=1}^{SN} d(m,j)}{SN-1} \tag{2.3}$$

The best solution among S solutions is calculated using Equation (2.4) where r represents “neighborhood radius”. To determine the neighbors ofXm, $if\ d(m,j) \leq r \times \mathrm{md}_m$, then X_j is neighbor of Xm. To evaluate the performance, qABC is compared with native ABC and different variants of ABC. Outcomes demonstrates, qABC method provides encouraging consequences in terms of convergence speed and complexity [14].

$$\boldsymbol{fit(X_{Nm}^{best}) = max(fit(X_{Nm}^{1}), fit(X_{Nm}^{2}), \ldots, fit(X_{Nm}^{S}))} \tag{2.4}$$

An optimization technique named multi-strategy ensemble ABC (MEABC) associates advantages of different nectar search methods. Every honey bee produces offspring in accordance to the distribution approach of food source during search. Results indicates the MEABC provides superior performance as compared to evolutionary algorithms including original ABC, GABC and MABC with respect to convergence rate and precision [21]. A rank-based adaptive ABC (ARABC) was used for global numerical optimization. In this scheme, honey bees pick the solution according to the rank. The sites that has higher attributes has maximum chance of being chosen. The proposed approach strengthens the exploitation problem of primary ABC algorithm without disturbing the performance of exploration ability of ABC. The experimental evidence proves that ARABC gives improved performance with respect to convergence speed without affecting the population diversity [22]. A depth-first search framework improves exploitation capability to give accurate solutions. It achieves better convergence speed of different versions of ABC algorithm including original ABC, CABC, and GABC. Furthermore, the Depth DFS system is combined with two search schemas in order to enhance the efficiency of ABC in the form of reliable and accurate solutions with faster convergence speed [23]. A framework for fireworks explosion based on ABC (FW-ABC) upgrades the effectiveness of different ABC algorithms and solves the complicated optimization problems. For this purpose, two search stages are introduced in the proposed framework. The first stage is named as bee search stage while the second one is fireworks explosion search stage. Experimental findings show that FW-ABC improves the efficiency of previous ABC algorithms [24]. Data driven ABC (DDABC) scheme addresses the modeling and impersonations of the enhanced ABC with data-driven optimization (DDABC). The scheme was to enhance the convergence rate as well as exploration and exploitation capability of artificial bee colony. To accelerate the convergence rate, the looking through procedure is derived by directional guided data. By comparing the presented DDABC approach with certain existing schemes, the experimental findings show that DDABC achieves improvements in the performance by enhancing the convergence rate of ABC algorithm [25].

2.3.2 Hybrid Algorithms based Schemes

Hybrid algorithms merge the advantages of two or more algorithms to attain improved performance of standard ABC algorithms. These algorithms achieve better convergence speed, algorithm accuracy and global search proficiencies. Hybrid algorithms are further categorized into three categories as follows. i) Original ABC Combines with other optimization algorithms where a modified ABC algorithm inspired from the Particle Swarm Optimization (PSO) called Velocity-based ABC (VABC) is proposed. VABC utilizes exploring capability of ABC to identify the different sites. It also examines the best solution the exploitation ability of PSO. It prevents the presented algorithm from falling into the local optima[26]. Li et al, [27] present a Bacterial Foraging Optimization algorithm to strengthen the efficiency of ABC in terms of convergence rate and precision. ii) Original ABC combines with mathematics to present an adaptive hybrid ABC(AHABC) algorithm [28] which merged Nelder-Mead simplex search method with ABC in dynamic manner to solve the parameter approximation of q-Weibull distribution. It overcomes slow convergence rate of ABC. iii) Original ABC combines with optimization strategy where Sun *et al*. presented a parallel strategy which combines SI, ABC, local search operator and Nelder-Mead process to attain the convergence performance of the overall identification procedure [29].

2.3.3 Parameter Adjustment based ABC Schemes

The performance of ABC algorithm can be optimized by selecting the parameter properly. Some researchers have modified the parameters in past few years. For example, Zhou *et al.* modified the selection parameter of ABC to achieve the goal of multi objective problem; and to avoid the local maxima, modified the local selection formula of the employed bee phase [30]. Li *et al.* describes a self-adaptive approach to address the constrained optimization problems. For each population based on viable rule, the employee bee severs as a global search engine. The onlooker bee colony search space transformed the constrained optimization problems into a multi objective problem. A self-adaptive modified rate is proposed to improve the convergence rate [31].

2.4 The schemes that implemented ABC algorithm in various fields

This section briefly describes the literature review of the existing schemes that implemented different versions of ABC algorithm in different fields.

2.4.1 Service Optimization, IoT and Cloud based Schemes

Service optimization problems (SOPs) can be solved using the ABC optimization techniques. Xiaofei Xu et al. presented a scheme named service domain-oriented ABC (S-ABC) that aims to strongly dominant on solving the SOPs. S-ABC is utilized to achieve the enhanced effective and efficient solutions for the SOPs. Experimental results demonstrate the excellence of S-ABC over various existing approaches in terms of solving SOPs [32]. A cross-modified ABC algorithm (CMABC) optimizes the service composition along with the accuracy and appropriate time. To bring out the optimal solution with eye catching accuracy and acceptable time, a service model is established. In CMABC, 6 service sets are built for simulation. Experimental results authenticate the superiority of CMABC when it is compared with different algorithms [33].A Reliable spanning tree construction algorithm in IoT (RST) uses ABC algorithm to generate proper trees for achieving high-throughput for IoT scenario. The primary superiority of using this technique is that it generates near-optimal trees. These trees are sorted according to their preferences. The simulation results indicate that RST for IoT improves the existing approaches in terms of reliability and energy consumption [34]. Hadoop-based ABC (HABC) aims to select the features from the big data with high performance. To achieve this goal, a system architecture is presented in which features are selected using ABC, while for removing the noise from the Hadoop ecosystem the kalman filter is used. Furthermore, to neglect the unnecessary data and for data aggregation, and to analyze the data an exhaustive four-tier architecture is presented. The results verify that the presented system is more efficient, scalable, and accurate [35]. An IoT based hybrid ABC algorithm with an effective schedule change (HABCA-EST) aims to resolve the scheduling issue and guarantees the rapid enhancement within fitness function of every source of food, because the presented approach fully utilizes the enormous information between the scheduled devices more efficiently and effectively. HABCA-EST combines the efficient schedule transformation operations with ABC to obtain the optimal search solution in minimum running time. Results show that HABCA-EST achieves encouraging performance when compared with the existing dominating approaches [36]. Reem E. Mohamed et al. presented a collaborative distributed antenna routing protocol for WSN named CDA. The presented approach belongs to the proactive category of routing protocols for ad-hoc networks. The purpose of presented protocol was to achieve the fair load balance with respect to energy efficiency. The presented technique is based on Degree Constrained Tree (DCT) and utilizes the optimal node degree to enhance the lifetime of the network. The purpose of developing CDA was to monitor the data periodically in WSNs apps. In CDA, the gathered (aggregated) data packet that is transmitted to the base station PTB is sent towards the distributed antenna element DAE along with minimal energy overhead. The input is accepted by a protocol for the purpose of topology maintenance (TM). This input can be OHCR or OHA, though literature proved that OHCR gives better results than OHA in terms of energy efficiency. Virtual base station VBS choses n-1sensor nodes for distributed antenna method. When the energy level of distributed antenna system arrives to the level of the threshold energy of VBS then it serves as a central processing unit that is liable to gather the data of every mote (sensor nodes) in the region of interest ROI incorporated in the data packet. The empirical findings show that the demonstrated work provides encouraging results as it doubles the lifetime of a network. Seyed Farhad Aghili et al. presented a technique for the IoT systems named Proactive Authentication and Key Agreement Protocol for Internet of Things PAKIT. The idea of presented scheme was to provide an energy efficient authentication protocol for the IoT systems. In PAKIT, there are four entities. Among these, GW represents the gateway and its responsibility is to execute the initial phase of presented scheme and offer the private or secret shares, which are requested by the other entities. Registering a user to the gateway and providing a securing login system is a second entity of the presented model. Sensor node SN is a third entity of the PAKIT, which takes the responsibility to gather the environmental details and send the collected details towards the appropriate users. Final entity of the presented scheme is cluster head CH that is responsible to produce a session key for user, SN, and GW at any time. The presented approach is proactive since the shared value of every single entity is up to dated. Experimental results demonstrates the efficiency of PAKIT over other state of art algorithms. Maha Bouaziz et al. introduced a mobility support protocol for IoM systems on the basis of Extended Kalman Filter and low-power standards for routing protocols and named it as EKF-MRPL. The basic idea of presented scheme was to minimize the signaling overhead and enhance the power consumption capabilities. In addition to this, present a prediction system regarding Extended Kalman Filter in order to forecast its non-linear trajectory. EKF-MRPL predicts the movement direction of the process and provides a persistent continuous connectivity. Such movement is predicted on the basis of a well-known filter called extended kalman filter that is responsible for modeling the non-linier movement trajectory of mobile node. Experimental evaluations provide encouraging results when the presented schemes is contrasted with different schemes. EKF-MRPL decreases signaling cost, reduces power consumption and accelerate the delivery rate. Elham Zamani et al. presented a scheme that merges the idea of AODV-ABL with AR-AODV protocols to reduce overhead called M-AODV. In first phase, the same idea of AR-AODV is utilized according to which pick the alternative path for data packet or node. In a case when alternative route is not available or the alternative path elapsed then move forward to the next step, in accordance to idea of ABL. According to ABL, utilize the substitute paths nearby at one hop distance (ABR). In that situation, transmitting a control message is prevented as far as possible. In order to update the real time table the concept of ODV-BFABL protocol is utilized. Simulated outcomes indicates that presented scheme surpass the rest of schemes for the purpose of reducing the control overhead. M. Anand et al. presented a scheme to provide optimal and energy efficient solutions to enhance the quality of battery power in mobile ad-hoc networks. For this purpose better quality AODV protocol is utilized and named the scheme as Intelligent Routing AODV protocol (IRAODV). In IRAODV, every single node figure out the distance of all other nodes in first step. In the second step, every single node evaluates the distance with every other node for packet transmission. In third step, nodes that are in the similar region and can accept the packet are discovered. In next phase one node is selected for transmitting the data package and deactivating the transferring of remaining data packets. Just the similar procedure is repeated in step 6 to receive the acknowledgement. The whole process is reiterated until all packets are transmitted.

2.4.2 Data sharing based schemes

In recent times numerous routing schemes have been presented in the networking world. This section briefly describes the literature review of the existing schemes that uses different techniques for the data gathering and data sharing. After concise study of the literature we categorized the schemes into three groups. The first category includes routing algorithms amongst IoT enabled systems. These algorithms do not utilize the tree structure while the second category includes the schemes that uses a tree structure. The tree based structure established a backbone tree for data transferring towards the main station in IoT enabled devices. In the third section, the data sharing and data replication based schemes are discussed.

2.4.2. Routing algorithms without Tree-based structure

The group of schemes which do not utilizes a tree based structure examined several measures such as reliability of the data and energy efficiency. Reliable routing of data within the IoT domain was considered in[37]–[40]. These schemes considered various criterions for example, mobility of the devices and reliability of the links for choosing the best path for transmission of data. Huang et al. [37] presented a scheme which examined the mobility of the devices and identified that with the passage of time IoT enabled devices have certain mobility patterns. So considering the history of each device, mobility of such device can be predicted. Ali et al. [38] presented a stochastic routing algorithm for IoT enabled device. In the presented approach, the authors formed a network as an absorbing Markov chain to figure out the delivery ratio and delay. Consequently, the probability of the transmission of every single link is calculated on the basis of local information. The drawback of presented approach is that it has chances of selecting improper routes for data transmission. Gia et al. [39] presented a scheme to ensure reliability by providing backup paths for the nodes. if the main routes failed, these routes are utilized for the data transmission in such case. Goins et al. [41] presented an algorithm for a cross-layer solution to attain the efficient data delivery. In the presented work the MAC layer is supposed as TDMA, in which for the scheduling purpose a randomized technique is utilized. The division of time into several frames is made, to achieve this goal. The IoT enabled systems competed to obtain the channel in every single frame. This procedure reduces the collision significantly. In addition to this, there is no need of rescheduling in a case if the routing paths are changed. A bio-inspired approach is applied at the routing layer of algorithm to send the packets over the base stations depending on its availability. Number of pheromones are dropped by the devices on the data transmission routes. Such attitude results in pheromone trails that have the maximum gatherings near the base stations. Every single device forward the packets toward its neighbor with the most pheromone in order to perform data transmission. After performing number of iterations the packets are forwarded towards the closest base station, as the device which is closer to the base station has more pheromone. However, the mobility of the device is not handled in the presented algorithm because of its slow convergence speed. Rani et al. [42] presented an energy-efficient multitier data gathering framework. In energy-efficient data gathering framework the nodes in each tier can collect the data from the lower tier. At first tier the nodes are responsible to monitored the surroundings and transfer the sensed data to the next tier. Such process remains as far as each and every sensed data meet the topmost tier. The uppermost tier includes numerous base stations. The second tier nodes serve as a cluster head considering the objective of reducing the power expenditures and aggregation of the sensed data. The presented work designed the issue of determination of nodes at each tier as the problem of optimization and demonstrated its NP-hardness. After that they present an efficient heuristic technique to handle such problem. Debroy et al. [43] examined the device-to-device routing problem in cognitive IoT networks. In order to properly control the simultaneous device-to-device communication and an evolutionary game-based route construction technique is presented. In order to enhance the rate of end-to-end data transferring, the availability of the channel is examined. Additionally, the presented approach adjusted the transmission power in every single node to maximize the transmission rate whereas the interference is maintained at the admissible level. Zhong et al. [44] examined cognitive IoT networks as well. Opportunistic routing for data transmission is exploited where forwarder nodes are chosen dynamically on the basis of number of criteria per packet.

2.4.3 Schemes using Tree-based structure

The tree based structure established a backbone tree for data transmission to the base station in IoT enabled systems [7], [8], [11]. Qiu et al. presented an algorithm [8], which construct the spanning tree in few rounds. In initial phase, the spanning tree considers only the base station. The base station transmits the control packets to all of its neighbors. These nodes choose the base station as their parent. In the second phase, IoT enabled devices repeat the same process and announced their presence in the tree by transmitting the control packets. The devices which are non-tree devices accept such packets and choose their parents to enter into the tree. This process runs as long as the tree in completed. There is a possibility that a device that have non-tree structure receives more than one control packets. In that case, the most proper device is picked as a parent. The appropriateness of a parent can be measured by calculating the number of children, residual energy, and hop count distance from the base station. More specifically, the one having maximum residual energy and minimum children will be preferred and will be placed closer to the base station. Li et al. [11] presented a scheme that utilizes the spanning tree for collection of the data. The presented approach supposed that the tree is provided in advance and the concern was to ensure the appropriate data transmission schedule so that the delayed requirement is kept. In number of applications, it is enough to circulate the messages to a particular subset of nodes. In such a case, a multicast tree is built between the base station and the indicated nodes. Samad NAJJAR et al. [34] presented a Reliable spanning tree construction algorithm in IoT (RST-IoT) uses ABC algorithm to generate proper trees for achieving high-throughput. The primary superiority of using this technique is that it generates near-optimal trees. These trees are sorted according to their preferences. The simulation results indicate that RST-IoT improves the existing approaches in terms of reliability and energy consumption.

2.5 Data replication based schemes

Rasheed saleem et al. [13]presented an artificial bee colony based multi-objective optimization technique for the selection and placement of the data replications and named it as multi-objective optimization with ABC (MOABC). ABC is utilized to place the replica at best position with respect to shortest distance and within the minimum cost. Furthermore, MOABC utilized knapsack approach to afford the low cost and attain the load balancing via data centers. To calculate the optimal sequence for the data replications the data centers uses ABC algorithm. The presented approach is implemented through CloudSim to examine the efficiency, availability of data and to test the optimality of the data replication. The simulation results show that MOABC achieves efficient results and is superior as compared to other algorithms. Azami et al. [46] presented an artificial bee colony based data replication in cloud computing. The presented approach acknowledged hierarchical topology consists of three levels. In the first level cluster of devices connect through higher bandwidth, in second level local area networks are connected with relatively lower bandwidth as compared to the first level. In the final level the regions of each LANs are connected together via higher bandwidth. In the presented algorithm if new food locations demonstrate superior quality, the bees stay in the new location and trial index will be incremented one unit. The site is selected on the basis of considerable number of request for the file. Experimental results show that presented scheme successfully reduced the mean job time.

2.6 ABC Schemes using Data Clustering

In this Section, we explore the unsupervised techniques that aim to segregate large data into small homogenous clusters. These clusters gather the data in this fashion that, items within the same cluster should be similiar to each other as much as possible. While the items belonging to different clusters should be dissimilar [47]. Data clustering can be defined as finding k groups on the basis of similarity measure from the set of n inputs. The similarity and dissimilarity between these groups or clusters is done in a way that within the same sub-group or cluster the similarity should be high while objects in different clusters should have higher dissimilarity [48]. The researchers depict that a cluster should be internally homogenous while externally the clusters should be separable. There are various kind of data clustering algorithms that are used to classify objects in a specific cluster on the basis of inter class homogeneity or intra class dissimilarity. The comprehensive surveys related to clustering can be found in [34–38].

2.6.1 ABC with K-Means Algorithm based Schemes

A simplified data clustering algorithm known as k-means is an unsupervised learning that is known to solve the common clustering problems. It uses a very simplified and easy approach to segregate the considerable quantity of dataset into small non-overlapping k subgroup or clusters. After selecting the cluster centers, the distance is calculated to associate it with the nearest data points in the cluster. K-mean algorithm is robust and efficient in this scenario [54].

Zeeshan et al. presented a scheme Global ABC Search Algorithm (GABCS) for data clustering. In GABCS, the searched equations of original ABC algorithm are rectified because of the fact that in search space, bees utilize their previous knowledge of quantity of food and position to accommodate their actions. Data clustering is done with the help of three widely known datasets: Iris, wine and thyroid. These data repositories are approached from the UCI database. Experimental findings are then compared with some widely known algorithms for example ABC, GA, ACO, K-NM-PSO, TS and GA. GABCS gives encouraging performance compared to its counterparts [55].

A Modified bee colony optimization (MBCO) scheme highlighted the forgiveness features of bees and to give the equal opportunity to the both trustworthy and untrustworthy bees. An approach in which selection is made on the basis of probability is introduced in the proposed work to assign the only unassigned test points in each iteration. MBCO and K-means algorithm are hybridized and two methods namely MKCLUST and KMCLUST are proposed to enhance the efficiency of MBCO and hence achieving the globally optimized and diverse results. The results testify that MBCO performance is better than other existing algorithms and the proposed MBCO can be utilized effectively in data clustering [56]. A hybrid scheme enhances the performance of k-means algorithm, named Enhanced ABC and K-means (EABCK). It employs a mutation operation and global solution which is modelled using equation (5) where x_{ij} represents a current food source while v_{ij} is a new source of food by onlooker bees and employed bees, x_{k1j} and x_{k2j} are randomly selected food sources. The gbest represents the best solution that one is updated by k-means clustering after every single iteration to perform data clustering. The symbol R represents the random value. The performance of clustering is evaluated using eleven different datasets. The experiment-based results show that EABCK outperforms existing ABC and data clustering algorithms [57].

$$v_{ij} = R(0,1) \cdot (x_{ij} - x_{k1j}) + R(0,1) \cdot (gbest_j - x_{k2j}) \quad \textbf{(2.5)}$$

A two-step ABC presented clustering to achieve efficient and effective results. For this purpose, enhancement is done in all three phases of canonical ABC algorithm. K-means algorithm is included in the initial phase. In the second phase, also known as onlooker bees phase, a new search equation is added to accelerate the convergence speed. While in the third and final phase, also known as scout bees phase, the Hooke and Jeeves-based searching procedure is used and the results are claimed to be encouraging [58].

2.6.2 ABC with Fuzzy C-Means based Schemes

Fuzzy C-Means (FCM) is a well-known and commonly used clustering technique which is based on a classic technique named as c-means. FCM algorithm was originated by Dunn in 1973 [59] and afterwards, in 1981, it was enhanced by Bezdek [60]. Fuzzy data clustering is an approach in which a single data item refers to more than one group or cluster. FCM is frequently used for the processing of images or pattern recognition [51]. FCM is based upon the concept of K-means, so FCM also partitions the data into various groups known as clusters. It is a “soft” computing technique which attempts to divide the large data into small c fuzzy clusters. These clusters contain various objects which are assigned to a cluster on the basis of a degree of belief. This degree of belief allows an object to be associated with more than one cluster [61]. Yunbin et al. present an FCM clustering based improved ABC scheme with new rank fitness selection. It is a strategy to tackle the limitations of original FCM clustering algorithm. In FCM, there is a problem that it can be easily trapped into the native optimum, considering the delicate selection procedure of primary or initial cluster centers. For such purpose, an enhanced ABC algorithm is combined with the procedure of rank fitness decision. The proposed algorithm utilizes the efficient behavior of fuzzy C-Means clustering and optimal searching ability of ABC algorithm which results in a robust and efficient data clustering. empirical findings demonstrates the proposed RABC-FCM method surpasses other conventional algorithms in terms of efficiency [62]. A hybridized clustering analysis on the basis of an enhanced ABC and FCM clustering titled IABCFCM utilizes the characteristics of both ABC and FCM algorithms. FCM clustering faces a problem of being stuck into the local optimum. The IABCFCM technique overcomes this problem. Six data packages of the UCI datasets are used for performing the experiments. The scheme escapes from the local optima and can handle the datasets in a better way [63]. A technique for MRI brain segmentation is built upon the hybridizing of FCM and ABC algorithm and is named as FCMABC. To achieve effective results on the basis of actual and simulated magnetic resonance imaging of brain, FCMABC sets the number of clusters in an appropriate way. Results show the effectiveness of FCMABC as the proposed system can segment the MRI brain image automatically and can overcome the deficiencies of the FCM clustering [64]. Brain tumor segmentation approach in MRI images uses ABC and FCM to optimize the process. Segmentation quality is enhanced by using robust and efficient fitness function of ABC algorithm. In the similar vein, FCM algorithm is used for clustering. It identifies brain tumor spots in the segmented image [65]. An optimal algorithm utilizes FCM operator in scout bee phase of ABC. The experimental results determine significance of the proposed approach over other algorithms as it provides quality solutions [66].

2.6.3 Cluster-Head Selection based IoT, WSNs Schemes

Shamim et al. presented an ABC based cluster head (CH) selection method for IoT named ABC-DC. Aim of the presented work is to enhance the limited life span of IoT, and to tackle the transmission delay as well as load balancing issues. The presented framework is made up of two fundamental phases. First phase is responsible to choose the best possible optimized CH with the help of ABC. While the main goal of second stage is to make the clusters of different devices on the basis of Euclidean distance. Experimental results prove the efficiency of presented approach in terms of residual energy, lifespan and transmission delay[67]. ABC is hybridized with GSA algorithm aims to enhance the energy efficiency, transmission delay and load balancing among the IoT devices. CH is selected with the help of ABC algorithm. The outcomes are compared with standard ABC, GA, PSO, and GSO. Results show that presented approach provides better results in terms of energy efficiency and transmission delay for IoT devices[68]. Zongshan wang et al. presented an energy efficient framework that utilizes an enhanced version of ABC algorithm to choose the cluster heads. Numerous factors such as energy, CH location, CH density and many more parameters are initiated in the IABC to resolve the clustering difficulties. In addition to this an improved ACO based energy effective algorithm for routing is presented. The presented model gives encouraging results as compared to its counterparts[69]. An efficient hybrid ACO and ABC algorithm named HACO-ABC-CHS is presented to choose the cluster heads efficiently. The presented approach tackles the draw backs of ACO and ABC algorithm. Results verify the superiority of presented technique. HACO-ABC-CHS enhances the throughput and residual energy of the network[70].

2.7 Comparative Analysis

In this Section, we analyze existing schemes in literature to explore the optimization ability of modified ABC based schemes as explored in Table 2.2. In most of the schemes, performance is judged by comparing algorithms by utilizing several UCI datasets for the purpose of performance evaluation [55] [62]. It has been observed that the standard ABC algorithm itself has an excellent capability of exploration, but it is not efficient in exploitation capability. ABC has three control parameters i.e. Colony size CS which is composed of employee bees and onlooker bees, limit to determine the trial run for the source of food position to be abandoned and maximum cycle number MCN. The limit value can be calculated easily after determining the value of CS. So technically ABC approach holds simply two parameters to adjust and hence is simple to implement. These fewer parameters make the ABC algorithm highly flexible as bee can be added and removed without reinitializing the algorithm. Therefore, small parameters are considered as strength of the ABC algorithm. Furthermore, some researchers modify the solution search equation of ABC to achieve better efficiency. In schemes [16] [19]–[21], the ABC algorithm gives accurate solutions with fast convergence speed when search equation of original ABC is modified. Convergence is a precisely defined mathematical term which is essentially determined as a sequence of elements that eventually approaches to a single value called limit. Convergence itself is not an algorithm instead it is a value that an algorithm manipulates or iterate. In the term “fast convergence speed”, the word “fast” defines the convergence time of the algorithm which shows, how fast ABC converges toward the good quality solutions. While robustness refers to the ability of tolerating errors which means a function that can handle erroneous input and output argument, considered as robust. Modifying the crossover operator and upgrading the search equation of the traditional ABC provides significant results but still requires improvements to solve the complex problems [22], [24]. Another technique in which researcher’s showed the interest is hybrid technique in which researchers combine the benefits of different algorithms for enhancing the performance of canonical ABC. Hybrid algorithms shows the significance over other algorithms with respect to accuracy of the algorithm and convergence rate. Recently, another trending approach is hybridization of ABC algorithm with different data clustering algorithm. The approach shows appropriate performance in terms of efficiency, accuracy, robustness, and reliability [55]–[57], [62], [63]. These schemes also consume less time for intended operations for ABC schemes. Existing literature describes that ABC algorithms achieve encouraging results when applied in data clustering scenario. In [21]–[31], the simulation scenarios using MATLAB are explored where multiple iterations are performed to achieve better results. These schemes explore scheduling, spaning tree based process tree, efficient feature selection, service optimization problems and optimized service composition and selection as well.

Table 2.2: Summary of Existing ABC Based Schemes

Scheme	Basic Idea	Mechanism	Advantages	Limitations
Modified versions of ABC				
ABC[5]	A new algorithm on the basis of the behavior of honey bees to attain the optimal solution.	Standard ABC algorithm is introduced to simulate the behavior of real honey bee swarms.	Simple, robust, flexible, and small parameters	Exploitation capability of ABC is not good[71].
qABC [14]	Enhances the performance of original ABC.	Modified the search equation by adding a new schema for the onlookers phase to select the best possible solution.	Faster convergence speed with accurate solutions.	In each step, it needs more CPU time in contrast to original ABC[72]
MEABC [21]	To overcome the exploitation problem of original ABC.	Modified the search equation. Number of solution search strategy exists in the whole searching process competing for generating offspring.	Faster convergence speed and gives accurate solutions.	To work well on combinatorial optimization problem MEABC required more enhancements[21].
ARABC [22]	Overcome exploitation problem of ABC by selecting a food source on the basis of ranking.	Modified the search equations based on ranking. Ranking is assigned on fitness of each food source.	Enhances the slow convergence of ABC.	Unable to resolve complicated non-detachable, multi-purpose optimization issues[22].
DFSABC_elite[23]	Enhances the exploitation problem and convergence speed using depth-first search (DFS) method.	Search schemas incorporates the information of elite solution in employee bee phase and second equation for onlooker bee phase to perform current best solution.	Faster convergence speed, and gives reliable and accurate solutions.	DFS is embedded over limited algorithms and can be extended to other algorithms to handle practical optimization problem[23].
FW-ABC [24]	To strengthen the exploitation ability of ABC using fireworks explosion based framework	Firework explosion search is implemented after three regular search phases including work of employee, onlooker, and scouts in ABC algorithm.	Overcomes the exploitation problem and gives encouraging results.	Not able to deal with complex engineering optimization problem[24].
DDABC[25]	To upgrade the convergence rate, exploitation efficiency of artificial bee colony.	Search mechanism chooses local data of each onlooker is generated from multiple searching trials.	DDABC is an accurate, stable and time saving approach.	Convergence rate can be further improved.
S-ABC [32]	Efficient and effective optimal	Enhances the optimization strategies of ABC to	Efficient, and effective.	No sub-criterion for

	solutions for service optimization problems (SOPs).	presents a service-oriented ABC.		SOPs. It is applied in smart healthcare and smart home services[32].
CMABC[33]	To get the optimized solution for the service composition and for the selection purpose with maximum accuracy and in sufficient time.	Set up a framework for web service instantiation. Utilizes the cross operation of GA.	Enhanced the convergence speed, accuracy and the stability.	Not able to provide optimal solution for the task nodes[33].
RST [34]	Efficient and reliable spanning tree construction algorithm, used to generate the proper trees.	Uses ABC to generate proper trees. Appropriateness of tree is measured by using hop count and residual energy.	Improves the reliability of the data gathering for emergency applications.	WSN are not capable to send data to sink directly due to long distance[73].
HABC [35]	To select the features efficiently and appropriately using ABC algorithm by neglecting the unnecessary data.	Selects the features using ABC. Kalman filter is used to remove noise from Hadoop ecosystem. MapReduce with ABC is used for processing efficiency.	HABC selects the features efficiently. In addition to this, HABC is a scalable approach.	Not able to analyze the data efficiently[74].
HAB [27]	Improves the efficiency of ABC	Bacterial foraging method is introduced in employee and onlooker's stage of ABC.	Enhances convergence speed and accuracy	Six well known functions are further extendable [27]
HABCA[29]	Parallel computations enhance convergence efficiency	Hybrid ABC with local search operator (LS), Nelder–Mead simplex, and a search space reduction method.	Powerful, robust and efficient.	Accurate estimation of parameters is a real concern[75].
HABCA-EST[36]	Overcomes the scheduling problems during ABC algorithm based calculations.	Fitness value used to initialize sources. EST of SDs for system scheduling of more crucial items, and entire usage of repetitive information.	Uses small population size. Efficiently initializes the population.	Only considered the target-coverage problem[36].
VABC [26]	To attain the optimal solution for continuous numerical functions.	Added a new search equation inspired from PSO search strategy in onlooker bee phase	Enhanced search tasks of high-dimensional optimization problem.	Not able to figure out the multi-purpose optimization issues[26].
MOABC[30]	To achieve the optimal solution for hydrothermal devices with respect	MOABC employs the procedure of dynamic parameter monitoring.	Enhances the algorithm's search capability.	MOABC must be

	to short term scheduling.			further enhanced to segregate acquired non-dominant set[30].
SACABC[31]	Achieve optimal solution for the constrained optimization problem	Introduced a self-adaptive modification rate method for parameter adjustment.	Simple, reliable, robust and efficient.	Benchmark test function require additional computations [31]
HABC [76]	It increases optimization ability of original ABC.	Crossover operator of GA is applied in ABC then employed for clustering.	It is better in convergence rate and accuracy.	HABC stuck into local minimum on several function[76].
GABCS [55]	Improve search equations of ABC for better performance.	Improves the search schemas of ABC and applied the modified ABC in K-Means clustering.	Saves time	Experiments are performed on only three datasets[77].
MBCO [56]	It presents a probability based selection approach.	Presents a probability based selection technique to allocate only remaining unallocated data points to clusters.	Provides better quality solutions when compared with other algorithms.	Needs more enhancement with respect to multi-objective optimization[56].
EABCK [57]	Improves k-means by upgrading ABC for solving data clustering.	A new mutation operation is added in original ABC and in each iteration best solution is updated using K-Means.	Efficient and robust.	Could not solve dynamic clustering problem[57].
Two-step ABC [58]	To achieve efficient, robust and reliable solution for optimization problem.	Initial food sources are selected with K-Means and an enhanced solution search schema inspired from PSO is used in onlooker's stage of ABC.	Reliable, efficient and robust.	Two-step ABC needs more enhancements to attain better solution for complex problems[58].
RABC_FCM [62]	Strengthen selection probability criteria for every individual with enhanced condition.	Introduces a Rank fitness selection in ABC and combined it with FCM.	Robust and efficient.	Selection probability criteria can be further enhanced[62]
IABCFCM [63]	To present a hybrid algorithm that utilizes the properties of both the ABC and FCM algorithm properly.	Replace roulette wheel selection with variable tournament selection method in onlooker bee phase. And modified Hooke and Jeeves search method is used in scout bee phase. Finally, hybridizes it with Fuzzy C-Means.	Improved the convergence speed and robust approach for clustering.	Until lacking of previous knowledge about the number of clusters, unable to achieve automatic clustering[63].

FCMABC [64]	Hybrid approach for MRI brain images segmentation.	Segmentation is done using hybridization of Fuzzy C-Means with modified ABC.	Effective and robust performance.	Finding the intensity of tumour[64].
ABCFCM [65]	To improve tumour part of the segmented image so that it can provide the intensity of tumour.	Segmentation is done using ABC and clustering is done using FCM.	Detects the tumour from image with intensity.	Visual segmentation results are provided but numerical results are not provided[65].
AB-FF [66]	To achieve the significance results.	Utilizes fuzzy C-Means operator at the scout bee phase of ABC.	Accurate and time saving.	Only three different datasets are used[66]
ABC-DC[67]	To achieve energy efficient CH based IoT system.	Uses ABC to select cluster heads.	Energy efficient, increases lifetime of IoT.	Need more enhancement in lifespan of IoT[67].
HABC-CHS[68]	To attain energy efficiency and reduce transmission delay.	Uses ABC to select cluster heads.	Minimizes the transmission delay, energy consumption.	Fixed network[68].
EE-IABC[69]	To balance the energy consumption.	Utilizes IABC to choose the cluster heads.	Extend lifespan of network.	Fixed network[70].

2.8 Open Research Challenges

ABC algorithms are applicable in variety of innovative application, but still it needs attention of active researchers to propose dependable solutions for open challenges. We have identified a number of open research challenges that should be examined by the researchers to propose novel ABC based efficient schemes.

2.8.1 High Dimension, discrete multi-objective problem

While dealing with high dimensional multi-objective problem, the original ABC still faces lots of problems. The traditional multi-objective optimization algorithm can efficiently solve the optimization problem of two to three targets, but still it has high-dimensional multi-objective problem when object is more than three. Furthermore, most of the enhanced ABC algorithms are used to convert the discrete variables into continuous variables to settle the discrete multi-purpose challenges. To find a specialized optimization algorithm that can directly deal with the discrete variables is not an easy task [17]. Considering these key points, the high dimensional discrete multi-objective problem is quite difficult but not impossible. Therefore, discrete multi-objective problem is still researcher's topic of interest. Researchers should work on this problem to propose a suitable solution to handle this problem.

2.8.2 Exploration-Exploitation problem in ABC algorithm

ABC is very effective at exploration, but not good for exploitation problem. There are some other SI based algorithms that are good at exploitation. Therefore, some scholars combined the advantages of ABC with advantages of some other algorithms to overcome the problem. In [26], ABC is combined with PSO to solve the Exploration-Exploitation problem. However, there are still many problems such as how to find the determining factors of the best Exploration-Exploitation trade-offs, trade-off between performance and calculation cost of solutions, and many more [5]. Regarding these vital issues and their possible solutions proposed by the researchers, it is clear that researchers should work more on the exploration-exploitation problem to enhance the performance of ABC algorithm. Novel solutions are expected to mitigate these problems efficiently.

2.8.3 Managing huge number of smart devices

Managing a wide range of devices in real-world systems is a very difficult task, as most of the researchers utilize the design variables at smaller scale usually fewer than hundred variables, and hence not able to handle the vast range of devices. To manage large scale problems or to manage tremendous range of devices for the real-world applications, researchers should require designed variables at much larger scale i.e. more than thousand variables. Rank based ABC scheme ensures the optimized solution for the numerical problems. ARABC achieves success with respect to convergence rate. It also provides possible solution for the exploitation problem but still it needs some enhancements, since ARABC is not able to solve complex and many more real-world problems [22]. FW-ABC provides encouraging results, but still it is not able provide the optimized solution for the complex engineering problems [24]. Therefore, it could be a challenging task for the researchers in future to produce the promising findings regarding managing the large amount of smart devices and to design an algorithm which is scalable for the complex large-scale problems.

2.8.4 Combinatorial problems

Combinatorial problem / optimization is an approach that achieves optimal solution for the discrete space problems. There are many complex problems that should be solved optimally, among these complicated optimization problems combinatorial problem is the one that is usually considered to be a very difficult one to solve. Although, there are usually no good ways to deal with most combinatorial problems, but still ABC algorithm is a good approach in most cases. A huge number of problems can be considered under the combinatorial problems, most popular among them is traveling salesmen problem (TSP). To achieve the optimal solution for the TSP, many researchers have utilized the ABC algorithm with TSP and attain the encouraging results [78]–[81]. While scheduling problem, shortest path tree, minimum spanning tree, quadratic assignment problem, vehicle routing problem and many more combinatorial optimization problems from various fields and from various applications could be the future interest of the researchers. Researchers should work on these problems to propose the encouraging solutions.

2.8.5 Web Storage and Retrieval

Many researchers use ABC algorithm in IoT to attain the goal of optimum solution in an efficient and appropriate way [32], [33], [35]. In IoT, the most difficult and challenging task is to transfer the correct data at right time. IoT demands intelligent communication among the various devices and put all the data of everything on the web. Therefore, intelligent computation is necessary at the device as well as at cloud based servers. During this whole phenomenon, the power and as well as the bandwidth are consumed during communication with the server. By considering the points mentioned above, it can be concluded that an important research area in the IoT is web storage and data retrieval. It would be an interesting and challenging task for the researchers in the future to attain the best optimal solution for the ABC based IOT applications with appropriate storage requirements. Furthermore, in IoT data moves multiple times which create complexity, therefore updating the data and data storage with in an appropriate time period is a challenging task.

2.8.6 Ensure full time availability of smart devices in IoT

Another challenging issue in IoT is to ensure the availability of the various devices for the maximum time to communicate with each other. Full time connectivity among the devices ensure the better performance with cost effective solutions. In IoT, a device or network of devices which ensures the availability for the maximum time is considered to be reliable and scalable. Here the availability of the devices means, a device should be available whenever a user wants to access it. It is a challenging whenever a user replicates a device with the new one that should be intelligent enough. It should easily access the historical data of the previous devices without any delay [24][82]. It is a daring task for the researchers to find out a reliable solution that ensures the full time availability and scalability among the various smart devices in IoT. Furthermore, it is quite challenging to find out the resources especially the full time availability of smart devices to achieve the best performance within the budget constraints.

2.8.7 Limitations of task optimization

ABC is an optimization approach based on SI; specifically motivated from the foraging actions of the honey bees. The ABC algorithm was proposed to attain the best optimal solution for the numerical problems. Many researchers modified the original ABC either by modifying the search equations or by adjusting the parameters of original ABC algorithm [16], [19], [21] and [22]. These algorithms provide optimal solution but face various limitations such as not able to solve complex problems or requiring more CPU time than original ABC algorithm etc. To gain the best possible optimal solution which has ability to solve the complex real world problems with accuracy and efficiency is still a challenging issue for the researchers. Furthermore, the unique features and various advantages of the ABC algorithm open the doors for the researchers to utilize the best features of ABC by applying them in different real world problems. For example, data clustering is one of the most powerful and widely used approach which is continuously impressing the researchers. Data clustering is used in various applications and it gives encouraging results. However, these techniques face various problems such as slow convergence, easily trapping into the local optima, and initially depending on the center etc. considering these shortcomings researchers utilizes the features of ABC algorithm by combining them with the different data clustering techniques [55], [56], [83]. These hybrid algorithms are successful to achieve the optimal solution, but in most cases experiments are performed on the small datasets so, in future it would be a challenging task for the researchers to propose reliable solutions to handle these issues on large datasets. To achieve the global best solution for the dynamic clustering problems with efficiency, accuracy, robustness and effectiveness is still a challenging issue for the researchers. So, in future researcher should work in these areas to propose a reliable solution for such issues.

2.8.8 Efficiency of tasks for large scale data in IoT

IoT is one of those topics which grabs the researchers' interest and intentions in recent days due to its significant role in the internet world. The IoT systems are capable to generate the huge amount of data but this huge amount of data must be processed effectively and efficiently. Literature guarantees that ABC algorithm enhances the performance of IoT based systems in terms of efficiency [32], [33], [35], [36]. These schemes authenticate the superiority of ABC algorithm, as it provides efficient solutions for the tasks of large-scale data in IoT. So, it would be a great choice for the researcher to apply ABC in different areas such as healthcare-related IoT applications to attain the efficiency of the tasks for the large-scale data in IoT and to generate the optimal solution for such problems.

2.9 Summary

In this chapter initially standard ABC. Numerous ABC based survey papers are discussed to identify the research gap. After this numerous different variant of ABC algorithm are discussed. After this implementation of artificial bee colony algorithm in different fields such as data clustering, internet of things, data sharing and data replication etc. is presented. A comparative analysis for different versions of ABC algorithm and its implementation in various is presented. And table for performance analysis is presented. Finally, open research challenges are explored that will open the doors for researchers to work in different fields. And fulfill the research gaps.

CHAPTER 3

METHODOLOGY

3.1 Overview

This chapter describes the system design and proposed structure for the identified problem. In this chapter a proposed model is illustrated where different types of bees are explored for various activities, a number of tasks are explored among the bees for finding the food sources. Next, we present the proposed solution to select the global best food source with better quality as compared to counterparts.

3.2 Problem Formalization & System Model

This section presents an enhanced optimization modal for replica selection and placement. Different behavior of employed bees and onlooker bees is modified in order to achieve the better convergence speed.

3.2.1 System Model

There are three types of bees: Employed, Onlooker and Scout Bees. One-half are employees and reaming are onlooker bees. First employed bees explores the sites. When they found a good enough food sources, then they visit that source, collect the nectar, and come back to their hive and start dancing. Such movement demonstrates the foundation of food source and its specific direction. Onlooker bees are dependent on the employed bees because as onlooker bees watch the employed bees dancing then they choose the food source accordingly and go out for looking the food sources near to those sources found by the employed bees as shown in Figure 3.1. Scouts bees are independent of other two bees (i.e., onlooker bees and employed bees). Scout bees search the environment randomly to find out the new food sources. We utilize ABC to find out the shortest path and to reduce the cost for accessing and placing the data replica across the nodes. The primary steps of the ABC algorithm are briefly discussed in the proposed solution section.

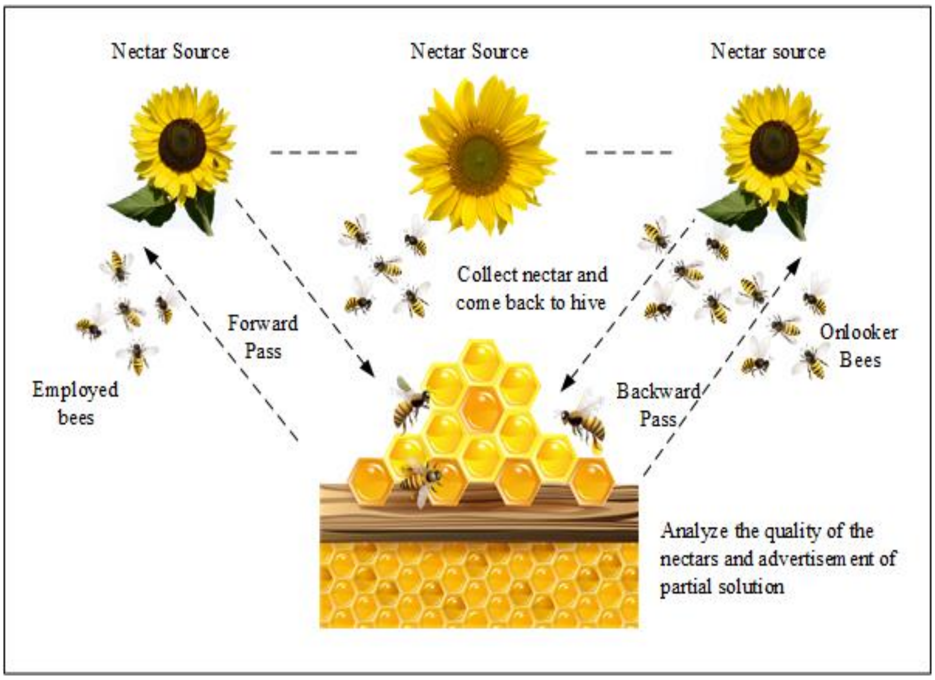

Figure 3.1: Intelligent Behavior of Honey Bees

Moreover, in this section, an enhanced model for data sharing among various nodes and placement of the replica in cloud environment is described. In literature numerous researchers [13], [84], [85] utilizes the same structure. The replica accessing and appropriate placement in cloud via nodes is presented in the proposed system. Therefore, we rely on the existing schemes in order to achieve the better results in terms of selecting the shortest path among the DCs within minimum cost and utilized the heterogeneous system to place the replica optimally. Statistical distribution is utilized to achieve the optimal results for replicas placement. Each DC comprises of various stages, with the end goal that the VM contrasts from DCs, etc. Simultaneously, it varies as far as RAM, CPU, PE, etc. The data centers vary with respect to the expenses and the number of accessible data replicas. In addition to this it is capable of approaching the replicas and placing them optimally within affordable cost and shortest distance. In the proposed system the employed bees are responsible for accessing and placing the data centers properly within minimum cost and simultaneously all the data centers are associated hierarchically as well as circularly at each level. The users are at the external level of the system and they are able to send tasks to data replication to obtain the best placed accessibility in the shortest time, distance and lowest cost through data centers. Placing the replicas on the sites that are closer to the users within reasonable cost through data centers is challenging task therefore artificial intelligence based technologies are utilized to attain the goal of optimal replica placement in data centers. The proposed model utilizes one of the enhanced version of ABC algorithm that reduces the number of comparisons and attain the goal of high file availability, and reduces response time.

3.2.2 Convergence Rate/Speed

Convergence is a precisely defined mathematical term which is essentially determined as a sequence of elements that eventually approaches to a single value called limit. Convergence itself is not an algorithm instead it is a value that an algorithm manipulates or iterate. In this work, the term “fast convergence speed” is used, herein the word “fast” defines the convergence time of the algorithm which shows, how fast ABC converges toward the good quality solutions. Here the value of limit is calculated using following equation 3.1:

$$Limit = \frac{CS \times D}{2} \tag{3.1}$$

Where CS represents colony size and D represents dimension[14].

3.2.3 File Availability

A device or network of devices which ensures the availability for the maximum time is considered to be reliable and scalable. Here the availability of the devices means, a device should be available whenever a user wants to access it. In other words availability refers to the capability of system to provide access to all the services at any moment. While if a system fails to provide services or any error, faults occurs, the device will be considered unavailable or unreliable. Therefore, ensuring the high data file availability is very important in cloud computing. Data file availability and unavailability can be calculated using the equation 3.2-3.5 as follows:

$$pro(bap_j)_{high_{DC}} > pro(bap_j)_{mid_{DC}} > pro(bap_j)_{low_{DC}} \quad \textbf{(3.2)}$$

$$pro(fla_k) = \begin{cases} (1 - \prod_{i=1}^{bnr_k}(1 - pro(bap_j)_i))^{nb_k} & for\ case\ 1 \\ \prod_{i=1}^{nb_k}(1 - \prod_{i=1}^{bnr_k}(1 - pro(bap_j)_i)) & for\ case 2 \end{cases} \quad \textbf{(3.3)}$$

Where case 1 represents a structure at which all blocks of data files are located together at DC. While case 2 represents a scenario at which blocks are located separately on various DC's.

Since $\boldsymbol{pro(fla_k)'}$=**1**-$\boldsymbol{pro(fla_k)'}$ therefore**,**

$$pro(fla_k)' = \begin{cases} 1 - (1 - \prod_{i=1}^{bnr_k}(1 - pro(bap_j)_i))^{nb_k} & for\ case\ 1 \\ 1 - \prod_{i=1}^{nb_k}(1 - \prod_{i=1}^{bnr_k}(1 - pro(bap_j)_i)) & for\ case 2 \end{cases} \quad \textbf{(3.4)}$$

$$high_{DC} = 0.9 > mid_{DC} = 0.6 > low_{DC} = 0.3 \quad \textbf{(3.5)}$$

Using these equations the probability of data file availability and unavailability is calculated[13], [84]. Table 3.1 shows the list of notations and their description used in above equations.

TABLE 3.1 List of Notations

Notation	Description
b	Blocks
$\mathbf{nb_k}$	Number of blocks
$\boldsymbol{pro(fla_k)}$	File availability probability
$\boldsymbol{pro(fla_k)'}$	File unavailability probability
$\boldsymbol{pro(bap_i)}$	Block availability probability
$\boldsymbol{bnr_k}$	Number of replica for data file $\boldsymbol{df_k}$

3.2.4 Delay/Response Time

Response time is a time that a device takes to react to a service. Moving data via the shortest path leads to decrease the response time of a network. The average response time is calculated using equation 3.6:

$$ART = \frac{\sum_{j=1}^{m}\sum_{k=1}^{m_j}(C_{jk}(rt)-C_{jk}(st))}{\sum_{j=1}^{m}M_j} \tag{3.6}$$

Where $(C_{jk}(st)$ represents the sending time and $C_{jk}(rt)$ represents receiving time of cloudlet k in user j. Moreover, M_j represents the no. of cloudlet for user j[86].

3.2.5 Design Objectives

The main objective of the proposed work is to construct an optimal system for the cloud environments. And achieve the better performance in terms of convergence speed and file availability.

3.3 Summary

In this chapter a system design and proposed structure for the identified problem is discussed in detailed. The intelligent behavior of honey bees is discussed and a solution for an optimal path with cost effective features is described. For which the proposed algorithm will be modelled and discussed in detailed in the next chapter.

CHAPTER 4

PROPOSED ENHANCED ABC (E-ABC) BASED OPTIMIZATION FOR DATA COLLECTION AND REPLICATION

4.1 Overview

In this chapter, we present an enhanced artificial bee colony (E-ABC) based optimization for data collection and replication schemes. In this section the proposed solution is illustrated to select the global best food source with better quality as compared to counterparts. The proposed model elaborates the access of data replication and its placement via nodes in cloud. To access and place the replicas by means of shortest route ABC algorithm is utilized. ABC algorithm provides optimal solution in lower cost and achieve load balancing through data centers (DCs). To access replica and to place data centers at the ideal position is critical according to guidance bees as these bees considered the most cost effective path. To achieve significant data availability within least cost we will use ABC based techniques that will perform optimal replica placement in DCs. In the proposed mechanisms there are four search schemas in the workflow of the ABC algorithm based on exploration-exploitation capability of honey bees to enhance the convergence speed of the algorithm. In the initial phase food sources are initialized randomly. Thereafter the obtained food sources are assessed and values of their objective function are measured. Next, the main three phases including employee bee phase, onlooker bee phase and replication position optimization phase are explored as shown in Figure 4.1. The main difference between the proposed scheme and the previous schemes is that, an enhanced definition is utilized for the employed and onlooker bees in the proposed scheme.

Furthermore, a comparison of different algorithms to each other as well as to meta-heuristic algorithms can be presented through the standards of number of fitness evaluations.

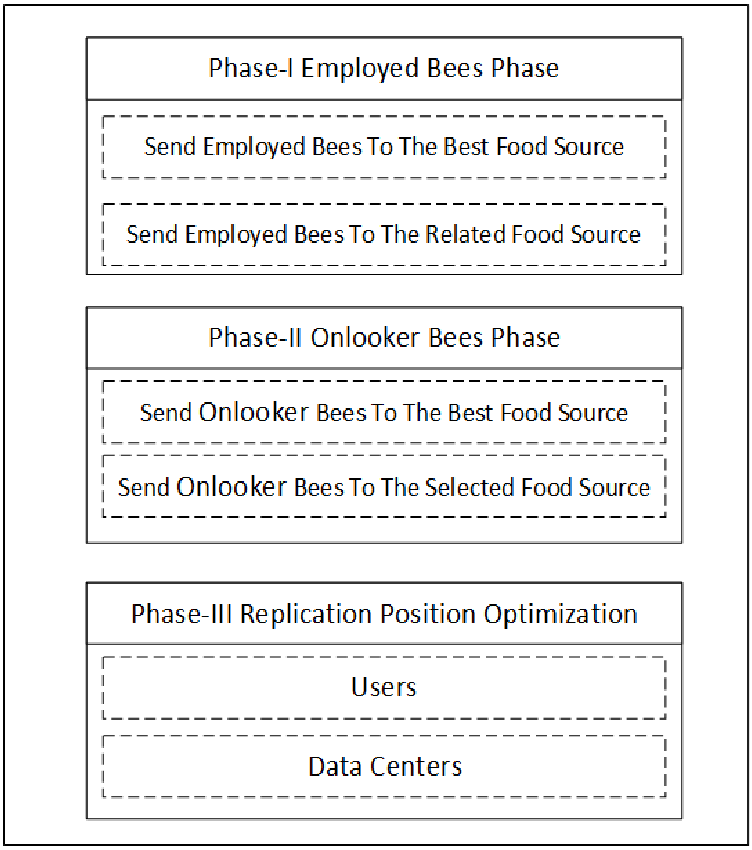

Figure 4.1: Main modules of proposed solution

Table 4.1 elaborates the different notations along with their description that are used in the algorithm.

TABLE 4.1 List of Notations

Notation	Description
SN	Population size
CS	Colony Size
x_{best}	Global finest food site
i_{best}	Index of the finest food site

$\boldsymbol{v_{best}}$	New food source
$\boldsymbol{MR}$	Modification Rate
$\boldsymbol{rand}$	Random value
p	Probability
$\boldsymbol{l_i}$	Lower bound
$\boldsymbol{u_i}$	Upper bound
$\boldsymbol{fit}$	Fitness value
$\boldsymbol{MCN}$	Maximum cycle number

4.2.1 Initialization Phase

The controlled parameters are initialized in this section which include colony size CS, maximum number of cycles MaxNum and limit l. In the initial phase, the food sources are initialized randomly by utilizing the equation (4.1) within a given range.

$$x_{m,i} = l_i + \text{rand } (0, 1) * (u_i - l_i) \qquad \textbf{(4.1)}$$

Where i=1,2,3,....,sn, m=1,2,3,….,d. Here sn represents number of food sources while d represents number of optimization parameters. l_i Represents lower bound of $x_{m,i}$ and u_i represents upper bound of $x_{m,i}$. Moreover rand represents a random value in the range (0, 1).

4.2.2 Employed Bee Phase

In standard ABC algorithm the frequency of perturbation is fixed therefore in order to generate a new candidate solution ***vi***, only one parameter of the parent solution $\boldsymbol{x_i}$ is changed which causes in slow convergence speed. Therefore, a control parameter named modification rate is introduced in the employed and onlooker bee phase to fix this issue. So by using this modification rate MR there is an evenly assigned random number ($0 \leq R_{i,j} \leq 1$). And the value of $\boldsymbol{v_{i,j}}$ is calculated according to modification rate as in equation 4.2.

$$v_{i,j} = \begin{cases} x_{i,j} + \varphi\left(x_{i,j} - x_{k,j}\right) & if R_{i,j} < MR \\ x_{i,j} & otherwise \end{cases} \qquad \textbf{(4.2)}$$

Where MR represents modification rate that picks the value between 0 and 1 and k∈{1,2,3,……,SN} is a randomly selected index which should be different from i. Moreover, after consuming a lot of time due to extensive number of comparisons the fitness evaluation function does not help to improve the quality of global best food source selection. In addition to this, its diversity can be further enhanced because it consider the best possible solution and ignore the rest of solutions. Therefore, such undesired consequences which eliminates the chances of selection of the other solutions results in poor performance. And sometimes standard ABC gives better results than the qABC. To tackle these issues a new parameter named finestLimit is added into the ABC which will help in changing the flow of work from best limit employ and onlooker to the new phase with the help of equation 4.2.

In the employed bee phase initially SN food sources are randomly generated in the population. Next the totalEval are compared with the maxEvl. If totalEvl are less than or equal to maxEvl then compare a counter trialInEBP that keeps the record of how often the best food represented as $X_{best,}$ fails to improve the employee bee phase with the finestLimit value, in order to take a decision that for each and every evaluation either the best limited employee bee phase will be utilized or not. If the value of the finestLimit is greater than or equal to the value of trialInEBP than the i^{th} employee bee produce a candidate solution around X_{best} by utilizing the equation **4.3**:

$$v_{best,j} = x_{best,j} + \emptyset\,(x_{best,j} - x_{i,j}) \qquad \textbf{(4.3)}$$

In this equation x is food source used by ith employee foragers while v_{best} is a candidate solution having jth parameter. The parameters of x_{best} and v_{best} are similar except the jth parameter. After producing new source the fitness of best food source and the fitness of newly generated source is compared. If new solution is better than the previously memorized best solution than the x_{best} will be replaced with v_{best}. Fitness of solution is calculated using equation 4.4:

$$fit(X_{best}) = \begin{cases} \frac{1}{1+f(X_{best})} & if\ X_{best} \geq 0 \\ 1 + abs\big(f(X_{best})\big) & if\ X_{best} < 0 \end{cases} \quad \textbf{(4.4)}$$

In this case the value of trials (i_{Best}) will be 0 because employed bee phase is updated. Else it will be incremented and trialInEBP will also be incremented.

And if the value of finestLimit is smaller than the value of trailsInEBP than *ith* employee bee produces a candidate solution by using equation 4.2. By using equation 4.2 employed bee will be sent to the related food source. The algorithm for employed bee phase is presented in table 4.2.

TABLE 4.2: Algorithm 1 (Employed Bee Phase)

1. **for** i= 1,2,3,…,sn **do**
2.   **if** totalEvl ≤ maxEvl **then**
3.   **if** trialInEBP ≤ finestLimit **then**
4. i_{best}← get finest food site index
5. x_{best} ← finest food site
6.   Generate v_{best} using equation 4.3;
7. v_{best}← new food source
8.   **if** fit (x_{best}) ≥ fit(v_{best}) **then**
9.   Replace x_{best} $with$ v_{best}
10.   trial(i_{best}) ←0
11.   **else**
12.   trial(i_{best}) ←trial(i_{best})+1

trialInEBP←trialInEBP+1

end if

else

v_i← new food source using Equation 4.2.

if fit(x_i)≥fit(v_i) **then**

Replace x_i *with* v_i

trial(i) ← 0

else

trial(i) ← trial(i)+1

end if

end if

totalEvl ← totalEvl + 1

end if

end for

4.2.3 Onlooker Bee Phase

In the second stage, initially SN food sources are randomly generated in the population. And the probability value is calculated using Equation 4.5.

$$P(x_i) = \frac{fit(x_i)}{\sum_{j=1}^{SN} fit(x_j)} \quad \textbf{(4.5)}$$

Where SN represents the size of population. Next the enhancements made in EBP are also applied into the OBP according to which, totalEval are compared with the maxEvl and totalBees compared with SN. If totalBees are less than are equal to SN and totalEval are less than are equal to maxEval then go for the next condition according to which check if rand (0,1) is less than or equal to the p($x_{recentBee}$) if true then value of the finestLimit will be compared with the value of trialsInOBP. If value of trialsInOBP is greater than or equal to the value of finestLimit then onlooker bee will be sent to the best food source and it will produce a candidate solution around x_{best} by utilizing the equation 4.6:

$$v_{best,j} = x_{best,j} + \emptyset\,(x_{best,j} - x_{s,j}) \qquad \textbf{(4.6)}$$

In this equation x_s is food source chosen by onlooker bee having *jth* parameter while v_{best} is a candidate solution having *jth* parameter $v_{best,j}$. The parameters of x_{best} and v_{best} are similar except the jth parameter. If the value of finestLimit is smaller than the value of trailsInOBP than to produce a candidate solution equation 4.2 will be used by onlooker bees. Algorithm 2 that elaborates the onlooker bee phase is presented in table 4.3.

TABLE 4.3: Algorithm 2 (Onlooker Bee Phase)

for i= 1,2,3,…,sn **do**
    p(x_i) ←prob. Val. cal. utilizing Equation 4.5.
end for
totalBee ←1
recentBee ←1
while totalBee ≤ sn and totalEvl ≤ maxEvl **do**
        if ran(0,1) ≤ p($x_{recentBee}$) **then**
            if trialInOBP ≤ finestLimit **then**
              i_{best}← get finest food site index
              x_{best} ← finest food site.
              v_{best}← Generate new food site v_{best} with Equation 4.6;
                if fit (x_{best})≥fit(v_{best}) **then**
                  x_{best}←v_{best}
                  trials(i_{best}) ←0

else

trial(i_{best}) ←trial(i_{best})+1

trialInOBP←trialInOBP+1

end if

else

$v_{recentBee}$←a new solution by utilizing Equation 4.2

if fit($x_{recentBee}$) ≥ fit($v_{recentBee}$) **then**

$x_{recentBee}$←v_i

trial(recentBee) ←1

else

trial(recentBee) ← +1

end if

end if

totalEvl← totalEvl+1

totalBee← totalBee+1

end if

recentBee ← +1

if recentBee ≥ sn **then**

recentBee ←1

end if

end while

4.2.4 Replication Position Optimization

Users are at exterior level of system and they are capable of sending tasks to data replication for the accessibility of best placement with respect to shortest time, distance and least cost. The accessibility and the placement of the data centers will be done via nodes in cloud. Therefore, to obtain the finest accessibility to pick the nodes having least cost and minimum route between the data centers. To attain optimal data replication placement, we will use heterogeneous method. Across all the levels, each data center is connected hierarchically as well as circularly simultaneously. Thus, accessibility and placement of the data center at proper location is very crucial for the bees as they are responsible to examine the most cost effective path. The replicas are placed at data centers with respect to the user's task. In order to distribute the data replica, geometric distribution and Zipf are used. Zipf is utilized to randomly distribute the replica file and place the file among data centers that are closer to the users as in equation 4.7.

$$p(f_i) = \frac{1}{i^{\alpha}} \tag{4.7}$$

Where i=1…n and α represents factor data replica distribution $0 \leq \alpha < 1$. While second one is geometric distribution that depicts the random allocation to place the replica files optimally with different parameters. It can be computed using $p(\mathrm{i}) = (1-p)^{i-1}.p$. Where i=1…n and p represents file replica access, $0 \leq p < 1$.

4.3 Summary

In this chapter an improved version of artificial bee colony is presented. And then the proposed model is implanted in cloud environment for the purpose of sharing and placing the data optimally in addition to this with in cost effective manner and shortest distance. The algorithm is discussed in detailed in this chapter.

CHAPTER 5

RESULTS AND ANALYSIS

5.1 Overview

This chapter presents the performance analysis of the proposed data replication based enhanced E-ABC and provides the details about the configuration and experimental results. The simulation parameters and their values are shown in the table. While the experiments are performed and validated using c#. Various graphs are drawn to illustrate the convergence rate of proposed algorithm. To check the cost of replicas, probability of data availability, and data transmission various scenarios can be considered to draw the graph for replication cost, data availability and data transmission. An important issue in IoT systems is the reliability of data sharing. So, we can draw the graph for average reliability to illustrate the reliability of the considered algorithms. The reliability can be measured by considering the number of failed devices, number of nodes and average energy consumed.

Experimental results are discussed in this section for the proposed algorithm. And the efficiency of the E-ABC scheme is contrasted with the original ABC algorithm and also compared with various variants of ABC algorithm. Eight different benchmark functions are utilized in the experiments. The experiments are further divided into two categories. First group evaluates the performance of proposed algorithm with respect to the finestLimit parameter on the various functions presented in CEC2015[87]. While in the second category, a comparison is drawn for the proposed E-ABC to the standard ABC algorithm, qABC and iqABC algorithm. Table

5.1 describes the test functions along with their formulas that are used for the experiments.

Table 5.1. Test Functions and Formulations

Function	**Formulation**
Rastrigin	$f(x) = \sum_{i=1}^{D}(x_i^2 - 10\cos(2\pi x_i) + 10)$
Dixen Price	$f(x) = (x_1 - 1)^2 + \sum_{i=2}^{D}(i(2x_i^2 - x_{i-1})^2)$
Ackley	$f(x) = 20 + e - 20\,exp - (0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D}x_i^2}) - \exp(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi x_i))$
Griewank	$f(x) = \frac{1}{4000}\sum_{i=1}^{D}(x_i^2) - (\prod_{i=1}^{D}\cos(\frac{x_i}{\sqrt{i}})) + 1$

5.2 Performance evaluation with respect to finestLimit parameter

In this section the performance of the finestLimit parameter is evaluated regarding its different values. In Table 5.2, the names of different functions which are utilized in order to analyze the quality of eventual solution of these complicated functions are elaborated.

Table 5.2. Benchmark Functions

No.	**Function**
$f1$	Rotated Bent Cigar function
$f4$	Shifted and Rotated Schwefel's function
$f6$	Shifted and Rotated HappyCat function
$f7$	Shifted and Rotated HGBat function

While the $f1, f2$ are unimodal and remaining functions are multimodal functions. The colony size CS is 20, the limit is determined as (CS*D)/2, the peak value for fitness evaluations are selected five hundred and fifteen hundred. The dimensions chosen for these fitness values are selected as ten and thirty. And the total of twenty autonomous moves or runs are carried out[87].

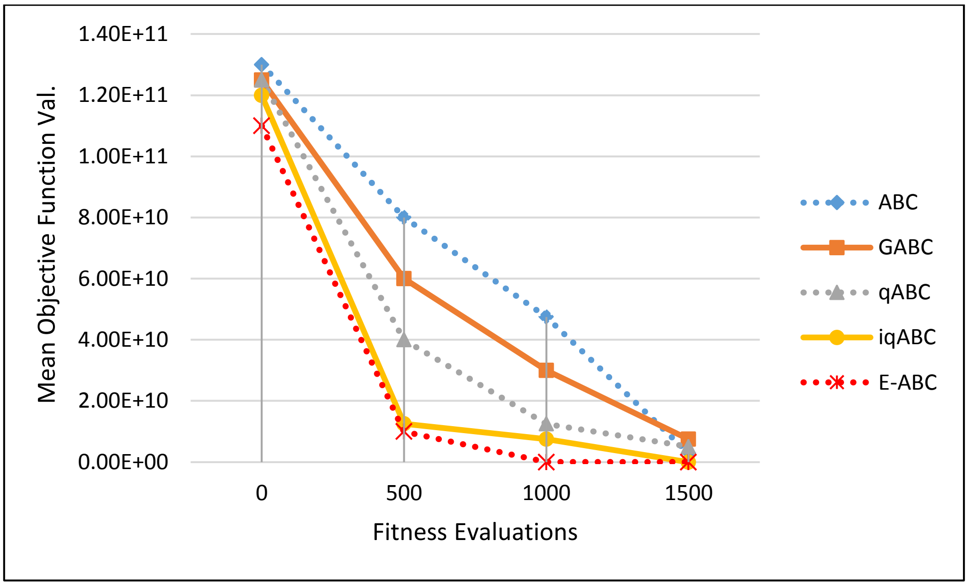

Figure 5.1: Convergence Graph for $f1$

Figure 5.1 shows the graph of convergence rate for the function $f1$. The graph shows the details of convergence speed of standard ABC, GABC, qABC, iqABC, and for the proposed algorithm E-ABC regarding the average finest value of objective function and fitness evaluation. For an example when the fitness evaluation value is 500 the mean objective function values for ABC, GABC, qABC, iqABC and E-ABC are noted as 8.00E+03, 7.5E+03, 5.0E+3, 3.0E+3, and 2.50E+3 respectively. After exploring the graphs in detail it is notified that the proposed algorithm provides 45.83%, 41.67%, 20.83%, 4.17% better results as compared to standard ABC, GABC, qABC, iqABC respectively.

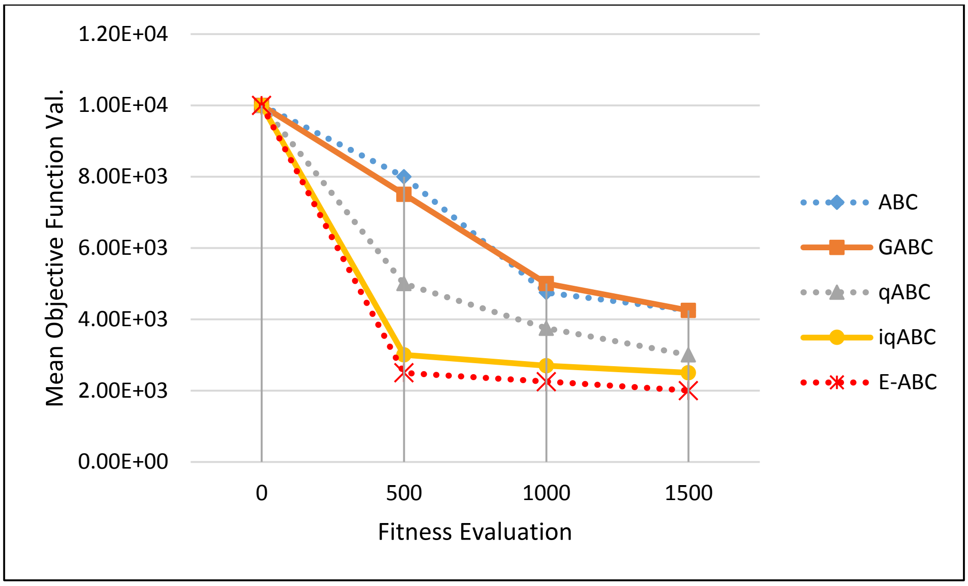

Figure 5.2: Convergence Graph for $f4$

Figure 5.2 elaborates the convergence rate for the function $f4$. The details about the convergence speed of standard ABC, GABC, qABC, iqABC, and for the proposed E-ABC are explored which shows the superiority of the proposed algorithm with respect to convergence speed. The graph explore the performance of algorithms considering the average finest value for y-axis function and fitness evaluation. Graph shows detailed comparison among the standard ABC, GABC, qABC, iqABC in terms of convergence speed. Results show that the proposed algorithm gives noticeably better results as compared to the other algorithms.

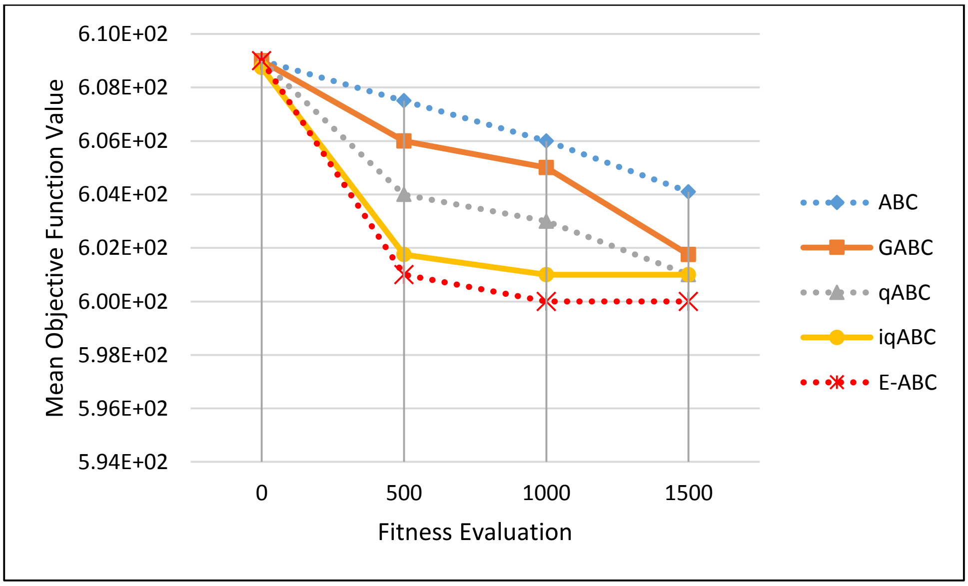

Figure 5.3 convergence graph for $f6$

The graph for the function $f6$ which shows the details of convergence speed of different versions of ABC and the proposed E-ABC is shown in the Figure 5.3. From the figure the detailed picture of convergence rate of standard ABC, GABC, qABC, iqABC, and the proposed algorithm E-ABC is illustrated. The convergence of the algorithms is measured with respect to the average finest value expressed on y-axis with the fitness evaluation narrated on x-axis. And it is seen that when a comparison is made for proposed E-ABC with standard ABC, GABC, qABC, iqABC, it provides much better performance considering the convergence rate.

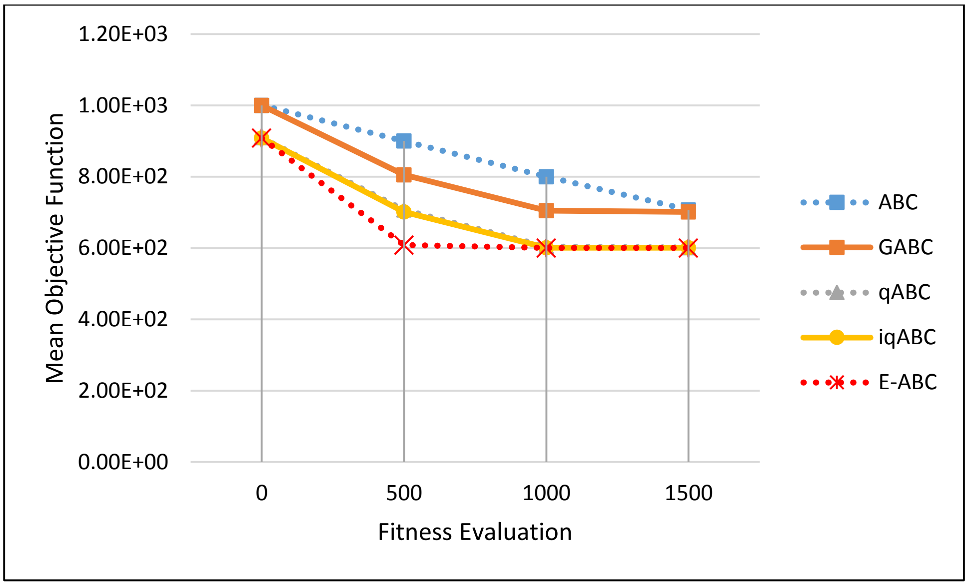

Figure 5.4: Convergence Graph for $f7$

The graph for the convergence speed of the function $f7$ is elaborated in Figure 5.4. The graph shows the detailed picture of the convergence rate of standard ABC, GABC, qABC, iqABC, and for the proposed algorithm E-ABC according to the mean best value of objective function and fitness value. After examining the graphs in detailed, it is noted that the proposed algorithm gives remarkably better results when compared with standard ABC, GABC, qABC, iqABC in terms of convergence speed.

5.3 Comparisons of algorithms on benchmark functions

In this section the experiments are performed on the various classical benchmark functions. The details about these benchmark functions including their names, along with formulae, global minimum values and their characteristics are listed are considered. From literature it is to be noted that all these classical benchmark functions are frequently utilized functions in order to investigate the efficiency of various well known algorithms. In the table C represents characteristic of function while U and M represents that the function is unimodal or multimodal. Additionally, the symbols S and N represents that function is separable or nonseparable. For these problems the maximum fitness evaluations were picked 500,00. And total 30 independent runs were carried out with various seeds and noted the mean value of the functions.

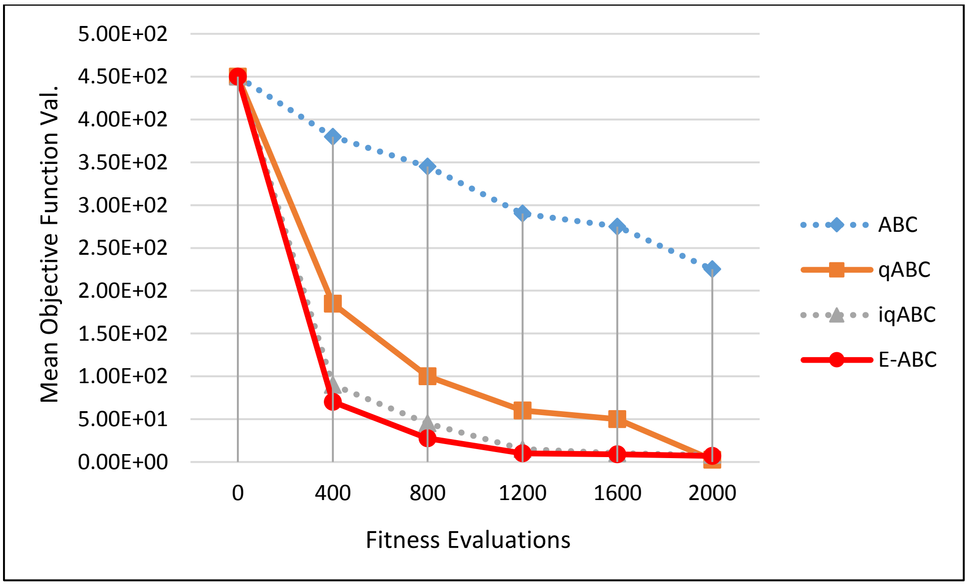

Figure 5.5: convergence graph for Rastrigin function

The performance of proposed algorithm is tested on the Rastrigin function as shown in the Figure 5.5. While the average objective values of finest possible outcomes found were noted and compared with the standard ABC, qABC, and iqABC. The results have proven the efficiency of E-ABC algorithm.

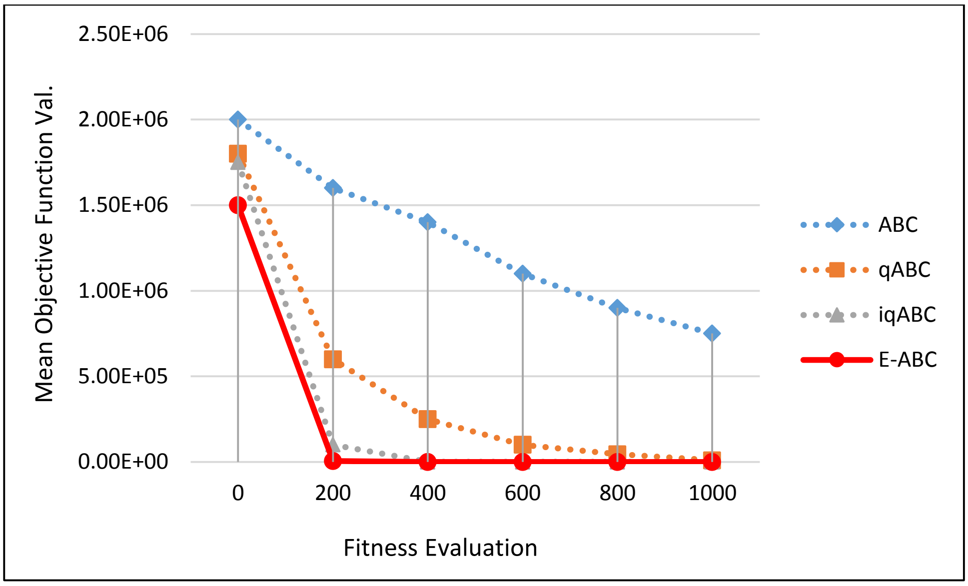

Figure 5.6: convergence graph for Dixen Price function

Figure 5.6 explains the convergence rate of standard ABC, qABC, iqABC, and of the proposed M-ABC algorithm for the Dixen Priece function. We analyze the performance of proposed algorithm as well as the related algorithms by observing the convergence speed with mean function value of best objective solution and the fitness evaluations. Results show that proposed algorithm totally outperform the standard ABC and qABC while iqABC produces encouraging results but still the M-ABC gives better results.

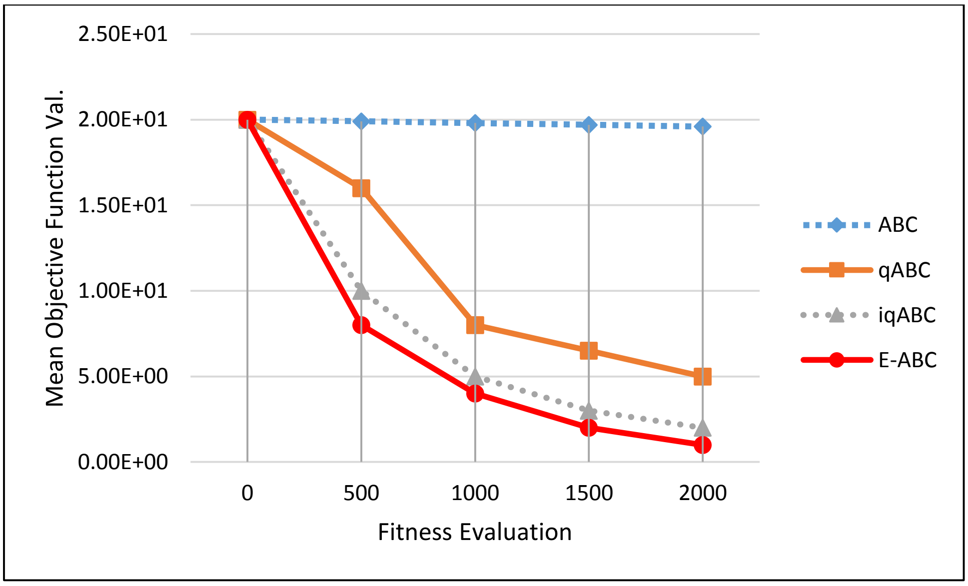

Figure 5.7: convergence graph for Ackley

Next we applied the proposed algorithm for the Ackley function to check the performance of E-ABC. And compared its performance with the other algorithms. The graph has shown the detailed view of convergence speed of the original ABC algorithm, and its modified variants qABC, and iqABC. The graph is presented with respect to the mean objective value of the function and fitness evaluation. And it is clearly notified that the proposed algorithm outperform other algorithms and provides more robust results.

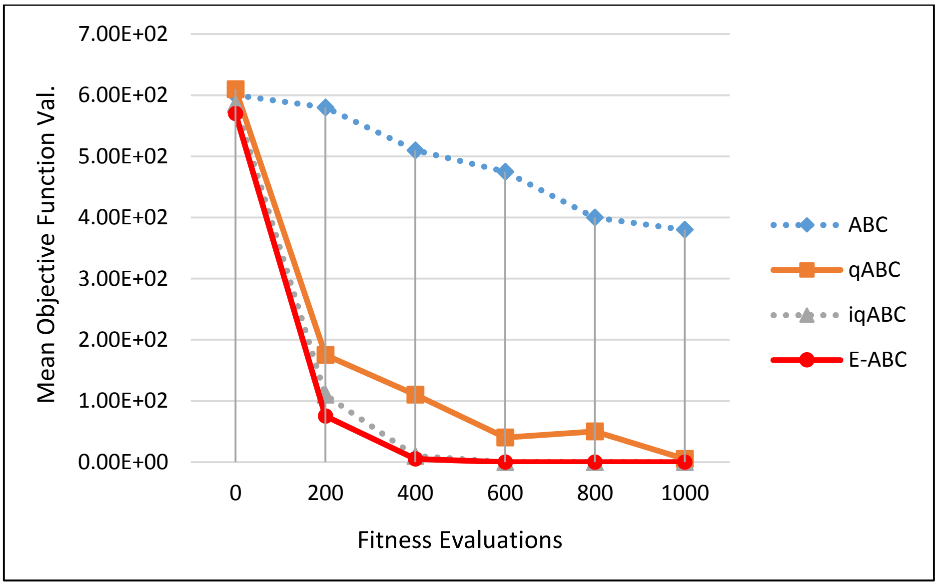

Figure 5.8: convergence graph for Griewank

In Figure 5.8 we determine the convergence rate for the Griewank function. The results have shown that our proposed model provides robust and effective results as compared to the other algorithms. Graph shows the behavior of convergence rate according to the mean function value and fitness evaluation. As shown in the Figure 5.8 the convergence rate of the proposed algorithm is better than other schemes.

5.4 Experiment for data file availability

Figure 5.9 demonstrates the cost effect on the availability of replicas. The number of cloudlets is set as 50 and cost to reach the best replica varies from 10 to 50.

We compute the accessibility of replica with every cost value ranging from 10 to 50. On the basis of results as shown in figure 5.9 it is noted that file availability increases with increase in the cost and it's because of increase in number of replicas. For example when cost is 20 the probability of file availability is 85%.

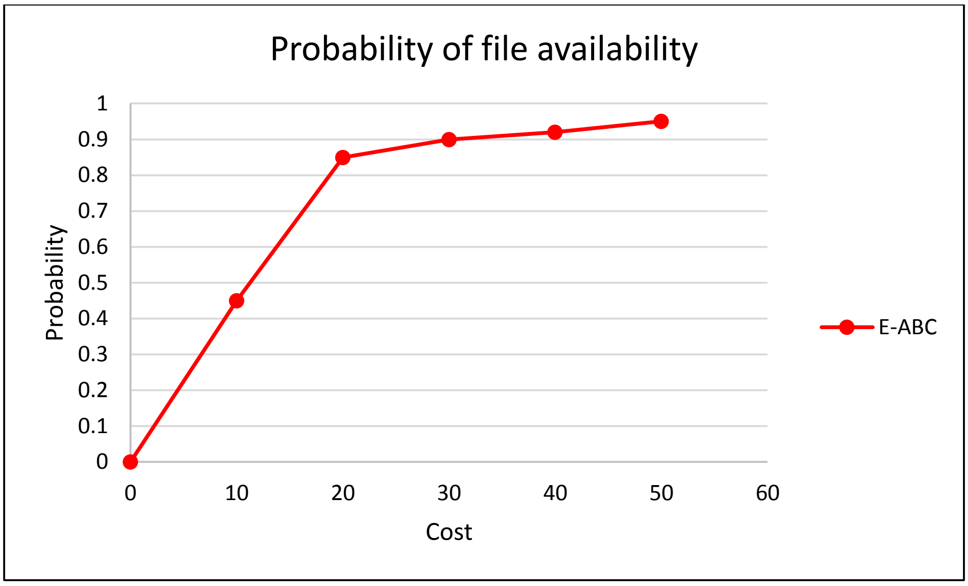

Figure 5.9: Probability of file Availability

5.5 Experiment for Response Time

Figure 5.10 determines the response time for increase in the number of cloudlets. And performance of the proposed algorithm is analyzed with MOABC and DCR2S. In this context, we noticed the response time for DCR2S, MOABC, and E-ABC when number of cloudlets are 1000, the average response time increases and noted as 38 seconds, 19 seconds and 11 seconds for DCR2S, MOABC, and E-ABC respectively. A significant decrease in the response time for E-ABC is noted as 67.5% and 20% for DCR2S and MOABC respectively.

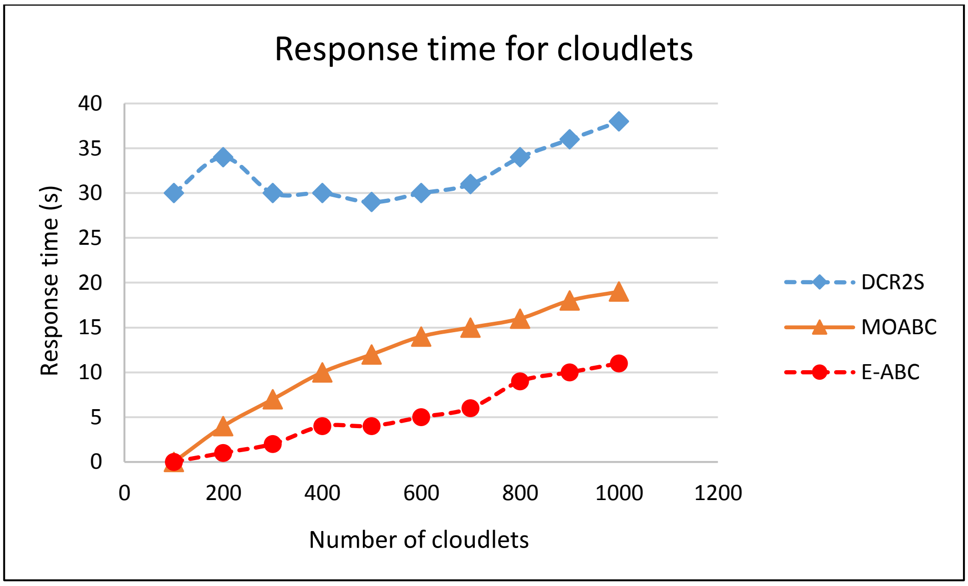

Figure 5.10: Average Response Time

5.6 Summary

One of the main reason of proposed algorithm is to enhance the convergence rate without effecting the quality of the final solutions. In this chapter performance of the proposed model is examined regarding convergence speed, data availability, and response time. For this purpose experiments are conducted for finestLimit parameter and comparison of the proposed algorithm with other algorithms on set of various classical benchmark functions. Experimental results prove the efficiency of E-ABC.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Overview

The primary focus of this thesis was to increase the availability of the data within low cost and achieve the better efficiency than previous schemes. To achieve this purpose an enhanced version of artificial bee colony algorithm is introduced. In the proposed algorithm the attitude of employee bees and onlookers are modified. The experimental findings prove the efficiency of proposed algorithm over different state of art algorithms. Finally, the proposed algorithm is implemented in data sharing and data replication scenario. The results show the superiority of proposed algorithm over counter parts in terms of availability, and response time.

6.2 Summary of Contributions

In this paper, an enhanced version of ABC algorithm is proposed. The search schemas are improved in employed bee phase and onlooker bee phase. Convergence speed of the proposed algorithm is examined on various functions. Additionally, performance of E-ABC is evaluated with respect to finestLimit parameter. Proposed algorithm is implemented in data replication scenario in order to achieve the better response time and file availability. Results shows that E-ABC provide better results when compared with counterparts.

6.3 Future Work

In future, this work can be extended to explore the use of ABC schemes for clustering data with software-defined networking to analyze data and take decisions to forward it or not. In addition to this the proposed framework is capable to solve the continuous, binary and combnitorial problems. So it can be utilized in many fields.

REFERENCES

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015, doi: 10.1109/COMST.2015.2444095.

[2] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Trans. Ind. Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014, doi: 10.1109/TII.2014.2300753.

[3] B. Dhanalaxmi and G. A. Naidu, "A survey on design and analysis of robust IoT architecture," *IEEE Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2017 - Proc.*, no. Icimia, pp. 375–378, 2017, doi: 10.1109/ICIMIA.2017.7975639.

[4] O. Zedadra, A. Guerrieri, N. Jouandeau, H. Seridi, and G. Fortino, *Swarm Intelligence and IoT-Based Smart cities*. Springer International Publishing, 2019.

[5] D. KARABOGA, "An idea based on honey bee swarm for numerical optimization," vol. 48, no. 9, pp. 1–16, 2005.

[6] Y. S. Chen and J. S. Hong, "Intelligent Device-to-Device Communication in the Internet of Things," *IEEE Syst. J.*, vol. 7, no. 1, pp. 77–91, 2013.

[7] N. Meghanathan, "Link selection strategies based on network analysis to determine stable and energy-efficient data gathering trees for mobile sensor networks," *Ad Hoc Networks*, vol. 62, pp. 50–75, 2017, doi: 10.1016/j.adhoc.2017.04.007.

[8] T. Qiu, X. Liu, L. Feng, Y. Zhou, and K. Zheng, "An Efficient Tree-Based Self-Organizing Protocol for Internet of Things," *IEEE Access*, vol. 4, no. c, pp. 3535–3546, 2016, doi: 10.1109/ACCESS.2016.2578298.

[9] J. Tang, Z. Zhou, J. Niu, and Q. Wang, "An energy efficient hierarchical clustering index tree for facilitating time-correlated region queries in the Internet of Things," *J. Netw. Comput. Appl.*, vol. 40, no. 1, pp. 1–11, 2014, doi:

10.1016/j.jnca.2013.07.009.

[10] Z. Zhou, J. Tang, L. J. Zhang, K. Ning, and Q. Wang, "EGF-tree: An energy-efficient index tree for facilitating multi-region query aggregation in the internet of things," *Pers. Ubiquitous Comput.*, vol. 18, no. 4, pp. 951–966, 2014, doi: 10.1007/s00779-013-0710-y.

[11] Z. Li, W. Zhang, D. Qiao, and Y. Peng, "Lifetime balanced data aggregation for the internet of things," *Comput. Electr. Eng.*, vol. 58, pp. 244–264, 2017, doi: 10.1016/j.compeleceng.2016.09.025.

[12] A. Kumar, D. Kumar, and S. K. Jarial, "A Review on Artificial Bee Colony Algorithms and Their Applications to Data Clustering," *Cybern. Inf. Technol.*, vol. 17, no. 3, pp. 3–28, 2017, doi: 10.1515/cait-2017-0027.

[13] R. Salem, M. A. Salam, and A. A. Mohamed, "An Artificial Bee Colony Algorithm for Data Replication Optimization in Cloud Environments," *IEEE Access*, vol. 8, pp. 51841–51852, 2020, doi: 10.1109/ACCESS.2019.2957436.

[14] D. Karaboga and B. Gorkemli, "A quick artificial bee colony (qABC) algorithm and its performance on optimization problems," *Appl. Soft Comput. J.*, vol. 23, pp. 227–238, 2014, doi: 10.1016/j.asoc.2014.06.035.

[15] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: Artificial bee colony (ABC) algorithm and applications," *Artif. Intell. Rev.*, vol. 42, no. 1, pp. 21–57, 2014, doi: 10.1007/s10462-012-9328-0.

[16] D. Agarwal, A. Gupta, and P. K. Singh, "A systematic review on artificial bee colony optimization technique," *Int. J. Control Theory Appl.*, vol. 9, no. Specialissue11, pp. 5487–5500, 2016.

[17] G. yaung Ying Liu, Lianho Ma, "a survey of artificial bee colony," pp. 1510–1515, 2017.

[18] S. Sharma and P. Bhambu, "Artificial Bee Colony Algorithm: A Survey," *Int. J. Comput. Appl.*, vol. 149, no. 4, pp. 11–19, 2016, doi: 10.5120/ijca2016911384.

[19] B. K. Verma and D. Kumar, "A review on Artificial Bee Colony algorithm," *Int. J. Eng. Technol.*, vol. 2, no. 3, p. 175, 2013, doi: 10.14419/ijet.v2i3.1030.

[20] F. S. Abu-Mouti and M. E. El-Hawary, "Overview of Artificial Bee Colony (ABC) algorithm and its applications," *SysCon 2012 - 2012 IEEE Int. Syst. Conf. Proc.*, pp. 590–595, 2012, doi: 10.1109/SysCon.2012.6189539.

[21] H. Wang, Z. Wu, S. Rahnamayan, H. Sun, Y. Liu, and J. S. Pan, "Multi-strategy

ensemble artificial bee colony algorithm," *Inf. Sci. (Ny).*, vol. 279, pp. 587–603, 2014, doi: 10.1016/j.ins.2014.04.013.

[22] L. Cui *et al.*, "A ranking-based adaptive artificial bee colony algorithm for global numerical optimization," *Inf. Sci. (Ny).*, vol. 417, pp. 169–185, 2017, doi: 10.1016/j.ins.2017.07.011.

[23] L. Cui *et al.*, "A novel artificial bee colony algorithm with depth-first search framework and elite-guided search equation," *Inf. Sci. (Ny).*, vol. 367–368, pp. 1012–1044, 2016, doi: 10.1016/j.ins.2016.07.022.

[24] X. Chen, X. Wei, G. Yang, and W. Du, "Fireworks explosion based artificial bee colony for numerical optimization," *Knowledge-Based Syst.*, no. xxxx, p. 105002, 2019, doi: 10.1016/j.knosys.2019.105002.

[25] M. Zhang, Y. T. Tan, J. H. Zhu, Y. N. Chen, and H. M. Liu, "Modeling and simulation of improved artificial bee colony algorithm with data-driven optimization," *Simul. Model. Pract. Theory*, vol. 93, no. June, pp. 305–321, 2019, doi: 10.1016/j.simpat.2018.06.004.

[26] N. Imanian, M. E. Shiri, and P. Moradi, "Velocity based artificial bee colony algorithm for high dimensional continuous optimization problems," *Eng. Appl. Artif. Intell.*, vol. 36, pp. 148–163, 2014, doi: 10.1016/j.engappai.2014.07.012.

[27] L. Li, F. F. Zhang, C. Liu, and B. Niu, "A hybrid Artificial Bee Colony algorithm with bacterial foraging optimization," *2015 IEEE Int. Conf. Cyber Technol. Autom. Control Intell. Syst. IEEE-CYBER 2015*, pp. 127–132, 2015, doi: 10.1109/CYBER.2015.7287922.

[28] M. Xu, E. L. Droguett, I. D. Lins, and M. das Chagas Moura, "On the q-Weibull distribution for reliability applications: An adaptive hybrid artificial bee colony algorithm for parameter estimation," *Reliab. Eng. Syst. Saf.*, vol. 158, no. September 2016, pp. 93–105, 2017, doi: 10.1016/j.ress.2016.10.012.

[29] and R. B. D. Hao Sun*, "Simultaneous identification of structural parameters and dynamic input with incomplete output-only measurements Hao," *Struct. Control Heal. Monit.*, no. May 2013, p. n/a-n/a, 2013, doi: 10.1002/stc.

[30] J. Zhou, X. Liao, S. Ouyang, R. Zhang, and Y. Zhang, "Multi-objective artificial bee colony algorithm for short-term scheduling of hydrothermal system," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 542–553, 2014, doi: 10.1016/j.ijepes.2013.10.013.

[31] X. Li and M. Yin, "Self-adaptive constrained artificial bee colony for

constrained numerical optimization," *Neural Comput. Appl.*, vol. 24, no. 3–4, pp. 723–734, 2014, doi: 10.1007/s00521-012-1285-7.

[32] X. Xu, Z. Liu, Z. Wang, Q. Z. Sheng, J. Yu, and X. Wang, "S-ABC: A paradigm of service domain-oriented artificial bee colony algorithms for service selection and composition," *Futur. Gener. Comput. Syst.*, vol. 68, pp. 304–319, 2017, doi: 10.1016/j.future.2016.09.008.

[33] L. Huo and Z. Wang, "Service composition instantiation based on cross-modified artificial Bee Colony algorithm," *China Commun.*, vol. 13, no. 10, pp. 233–244, 2016, doi: 10.1109/CC.2016.7733047.

[34] S. Najjar-Ghabel, S. Yousefi, and L. Farzinvash, "Reliable data gathering in the Internet of Things using artificial bee colony," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 26, no. 4, pp. 1710–1723, 2018, doi: 10.3906/elk-1801-100.

[35] A. Ahmad *et al.*, "Toward modeling and optimization of features selection in Big Data based social Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 82, pp. 715–726, 2018, doi: 10.1016/j.future.2017.09.028.

[36] Z. Muhammad, N. Saxena, I. M. Qureshi, and C. W. Ahn, "Hybrid Artificial Bee Colony Algorithm for an Energy Efficient Internet of Things based on Wireless Sensor Network," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 34, no. sup1, pp. 39–51, 2017, doi: 10.1080/02564602.2017.1391136.

[37] J. Huang, S. Wang, X. Cheng, and J. Bi, "Big data routing in D2D communications with cognitive radio capability," *IEEE Wirel. Commun.*, vol. 23, no. 4, pp. 45–51, 2016, doi: 10.1109/MWC.2016.7553025.

[38] Z. Ali, Z. H. Abbas, and F. Y. Li, "A stochastic routing algorithm for distributed IoT with unreliable wireless links," *IEEE Veh. Technol. Conf.*, vol. 2016-July, 2016, doi: 10.1109/VTCSpring.2016.7504110.

[39] T. N. Gia, A. M. Rahmani, T. Westerlund, P. Liljeberg, and H. Tenhunen, "Fault tolerant and scalable IoT-based architecture for health monitoring," *SAS 2015 - 2015 IEEE Sensors Appl. Symp. Proc.*, 2015, doi: 10.1109/SAS.2015.7133626.

[40] S. Rani, S. H. Ahmed, R. Talwar, J. Malhotra, and H. Song, "IoMT: A Reliable Cross Layer Protocol for Internet of Multimedia Things," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 832–839, 2017, doi: 10.1109/JIOT.2017.2671460.

[41] J. B. Goins, A. G. Roy, K. Mayaram, and T. S. Fiez, "JetNet: A proposed protocol for reliable packet delivery in low-power IoT applications," *2016 IEEE*

3rd World Forum Internet Things, WF-IoT 2016, pp. 48–53, 2017, doi: 10.1109/WF-IoT.2016.7845421.

[42] S. Rani, R. Talwar, J. Malhotra, S. H. Ahmed, M. Sarkar, and H. Song, “A novel scheme for an energy efficient internet of things based on wireless sensor networks,” *Sensors (Switzerland)*, vol. 15, no. 11, pp. 28603–28626, 2015, doi: 10.3390/s151128603.

[43] S. Debroy, P. Samanta, A. Bashir, and M. Chatterjee, “SpEED-IoT: Spectrum aware energy efficient routing for device-to-device IoT communication,” *Futur. Gener. Comput. Syst.*, vol. 93, pp. 833–848, 2019, doi: 10.1016/j.future.2018.01.002.

[44] X. Zhong, R. Lu, L. Li, and S. Zhang, “ETOR: Energy and trust aware opportunistic routing in cognitive radio social internet of things,” *2017 IEEE Glob. Commun. Conf. GLOBECOM 2017 - Proc.*, vol. 2018-Janua, pp. 1–6, 2017, doi: 10.1109/GLOCOM.2017.8254213.

[45] M. S. Pan and S. W. Yang, “A lightweight and distributed geographic multicast routing protocol for IoT applications,” *Comput. Networks*, vol. 112, pp. 95–107, 2017, doi: 10.1016/j.comnet.2016.11.006.

[46] M. H. Olyaei, H. Jalali, A. Olyaei, and A. Noori, *Fundamental Research in Electrical Engineering*, vol. 480. Springer Singapore, 2019.

[47] M. N. M. A.K Jain, “data clustering: a review,” vol. 3, no. September, 1981.

[48] A. K. Jain, “Data clustering: 50 years beyond K-means,” *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, 2010, doi: 10.1016/j.patrec.2009.09.011.

[49] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 1999, doi: 10.1145/331499.331504.

[50] K. C. Wong, “A short survey on data clustering algorithms,” *Proc. - 2015 2nd Int. Conf. Soft Comput. Mach. Intell. ISCMI 2015*, pp. 64–68, 2016, doi: 10.1109/ISCMI.2015.10.

[51] S. Bano and M. N. A. Khan, “A survey of data clustering methods,” *Int. J. Adv. Sci. Technol.*, vol. 113, no. December, pp. 133–142, 2018, doi: 10.14257/ijast.2018.113.14.

[52] J. Yu, H. Huang, and S. Tian, “Cluster validity and stability of clustering algorithms,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3138, no. 3, pp. 957–965, 2004, doi: 10.1007/978-3-540-27868-9_105.

[53] D. Karaboga and C. Ozturk, “A novel clustering approach: Artificial Bee Colony (ABC) algorithm,” *Appl. Soft Comput. J.*, vol. 11, no. 1, pp. 652–657, 2011, doi: 10.1016/j.asoc.2009.12.025.

[54] J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A K-Means Clustering Algorithm,” *Appl. Stat.*, vol. 28, no. 1, p. 100, 1979, doi: 10.2307/2346830.

[55] Z. Danish, H. Shah, N. Tairan, R. Gazali, and A. Badshah, “Global artificial bee colony search algorithm for data clustering,” *Int. J. Swarm Intell. Res.*, vol. 10, no. 2, pp. 48–59, 2019, doi: 10.4018/IJSIR.2019040104.

[56] P. Das, D. K. Das, and S. Dey, “A modified Bee Colony Optimization (MBCO) and its hybridization with k-means for an application to data clustering,” *Appl. Soft Comput. J.*, vol. 70, pp. 590–603, 2018, doi: 10.1016/j.asoc.2018.05.045.

[57] D. C. Tran, Z. Wu, Z. Wang, and C. Deng, “A novel hybrid data clustering algorithm based on artificial bee colony algorithm and K-means,” *Chinese J. Electron.*, vol. 24, no. 4, pp. 694–701, 2015, doi: 10.1049/cje.2015.10.006.

[58] Y. kumar and G. Sahoo, “A two-step artificial bee colony algorithm for clustering,” *Neural Comput. Appl.*, vol. 28, no. 3, pp. 537–551, 2017, doi: 10.1007/s00521-015-2095-5.

[59] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *J. Cybern.*, vol. 3, no. 3, pp. 32–57, 1973, doi: 10.1080/01969727308546046.

[60] J. C. Bezdek and J. C. Bezdek, “Objective Function Clustering,” *Pattern Recognit. with Fuzzy Object. Funct. Algorithms*, pp. 43–93, 1981, doi: 10.1007/978-1-4757-0450-1_3.

[61] A. Fahad *et al.*, “A survey of clustering algorithms for big data: Taxonomy and empirical analysis,” *IEEE Trans. Emerg. Top. Comput.*, vol. 2, no. 3, pp. 267–279, 2014, doi: 10.1109/TETC.2014.2330519.

[62] H. Yunbin, X. Yupeng, W. Jing, and L. Song, “The fuzzy clustering combined an improved artificial bee colony algorithm with new rank fitness selection.”

[63] A. Kumar, D. Kumar, and S. K. Jarial, “A hybrid clustering method based on improved artificial bee colony and fuzzy C-means algorithm,” *Int. J. Artif. Intell.*, vol. 15, no. 2, pp. 40–60, 2017.

[64] Mutasem k. alsamadi, “MRI brain Segmentation using a hybrid ABC and FCM.Pdf.” .

[65] N. Menon and R. Ramakrishnan, “Brain Tumor Segmentation in MRI images

using unsupervised Artificial Bee Colony algorithm and FCM clustering," *2015 Int. Conf. Commun. Signal Process. ICCSP 2015*, pp. 6–9, 2015, doi: 10.1109/ICCSP.2015.7322635.

[66] M. Krishnamoorthi and A. M. Natarajan, "Artificial bee colony algorithm integrated with fuzzy C-mean operator for data clustering," *J. Comput. Sci.*, vol. 9, no. 4, pp. 404–412, 2013, doi: 10.3844/jcssp.2013.404.412.

[67] S. Yousefi, F. Derakhshan, H. S. Aghdasi, and H. Karimipour, "An energy-efficient artificial bee colony-based clustering in the internet of things," *Comput. Electr. Eng.*, vol. 86, 2020, doi: 10.1016/j.compeleceng.2020.106733.

[68] M. Praveen Kumar Reddy and M. Rajasekhara Babu, "Energy Efficient Cluster Head Selection for Internet of Things," *New Rev. Inf. Netw.*, vol. 22, no. 1, pp. 54–70, 2017, doi: 10.1080/13614576.2017.1297734.

[69] Z. Wang, H. DIng, B. Li, L. Bao, and Z. Yang, "An Energy Efficient Routing Protocol Based on Improved Artificial Bee Colony Algorithm for Wireless Sensor Networks," *IEEE Access*, vol. 8, pp. 133577–133596, 2020, doi: 10.1109/ACCESS.2020.3010313.

[70] S. Janakiraman, "A hybrid ant colony and artificial bee colony optimization algorithm-based cluster head selection for IoT," *Procedia Comput. Sci.*, vol. 143, pp. 360–366, 2018, doi: 10.1016/j.procs.2018.10.407.

[71] H. Wang *et al.*, "Artificial bee colony algorithm based on knowledge fusion," *Complex Intell. Syst.*, vol. 7, no. 3, pp. 1139–1152, 2021, doi: 10.1007/s40747-020-00171-2.

[72] B. Görkemli and Z. Al-Dulaimi, "On the performance of quick artificial bee colony algorithm for dynamic deployment of wireless sensor networks," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 27, no. 6, pp. 4038–4054, 2019, doi: 10.3906/ELK-1902-189.

[73] F. Derakhshan and S. Yousefi, "A review on the applications of multiagent systems in wireless sensor networks," *Int. J. Distrib. Sens. Networks*, vol. 15, no. 5, 2019, doi: 10.1177/1550147719850767.

[74] A. Kobusińska, C. Leung, C. H. Hsu, S. Raghavendra, and V. Chang, "Emerging trends, issues and challenges in Internet of Things, Big Data and cloud computing," *Futur. Gener. Comput. Syst.*, vol. 87, pp. 416–419, 2018, doi: 10.1016/j.future.2018.05.021.

[75] X. Jia, S. Nadarajah, and B. Guo, "Inference on q-Weibull parameters," *Stat.*

Pap., vol. 61, no. 2, pp. 575–593, 2020, doi: 10.1007/s00362-017-0951-3.

[76] X. Yan, Y. Zhu, W. Zou, and L. Wang, "A new approach for data clustering using hybrid artificial bee colony algorithm," *Neurocomputing*, vol. 97, pp. 241–250, 2012, doi: 10.1016/j.neucom.2012.04.025.

[77] M. S. Beg and A. A. Waoo, "A Comprehensive Study of Artificial Bee Colony (ABC) Algorithms and Its A Comprehensive Study of Artificial Bee Colony (ABC) Algorithms and Its Applications," no. August, 2020.

[78] S. S. Choong, L. P. Wong, and C. P. Lim, "An artificial bee colony algorithm with a Modified Choice Function for the traveling salesman problem," *Swarm Evol. Comput.*, vol. 44, pp. 622–635, 2019, doi: 10.1016/j.swevo.2018.08.004.

[79] D. Karaboga and B. Gorkemli, "Solving Traveling Salesman Problem by Using Combinatorial Artificial Bee Colony Algorithms," *Int. J. Artif. Intell. Tools*, vol. 28, no. 01, p. 1950004, 2019, doi: 10.1142/s0218213019500040.

[80] W. Li, W. Li, Y. Yang, H. Liao, J. Li, and X. Zheng, "Artificial bee colony algorithm for traveling salesman problem," *Adv. Mater. Res.*, vol. 314–316, pp. 2191–2196, 2011, doi: 10.4028/www.scientific.net/AMR.314-316.2191.

[81] B. Gorkemli, "Quick Combinatorial Artificial Bee Colony -qCABC- Optimization Algorithm for TSP," vol. 3536, no. 1, pp. 97–101, 2013.

[82] M. M. Dhanvijay and S. C. Patil, "Internet of Things: A survey of enabling technologies in healthcare and its applications," *Comput. Networks*, vol. 153, no. June, pp. 113–131, 2019, doi: 10.1016/j.comnet.2019.03.006.

[83] A. Kumar, D. Kumar, and S. K. Jarial, "A novel hybrid K-means and artificial bee colony algorithm approach for data clustering," *Decis. Sci. Lett.*, vol. 7, pp. 65–76, 2017, doi: 10.5267/j.dsl.2017.4.003.

[84] N. K. Gill and S. Singh, "A dynamic, cost-aware, optimized data replication strategy for heterogeneous cloud data centers," *Futur. Gener. Comput. Syst.*, vol. 65, pp. 10–32, 2016, doi: 10.1016/j.future.2016.05.016.

[85] D. W. Sun, G. R. Chang, S. Gao, L. Z. Jin, and X. W. Wang, "Modeling a dynamic data replication strategy to increase system availability in cloud computing environments," *J. Comput. Sci. Technol.*, vol. 27, no. 2, pp. 256–272, 2012, doi: 10.1007/s11390-012-1221-4.

[86] N. Mansouri and M. M. Javidi, "A new Prefetching-aware Data Replication to decrease access latency in cloud environment," *J. Syst. Softw.*, vol. 144, pp. 197–215, 2018, doi: 10.1016/j.jss.2018.05.027.

[87] Q. Chen, B. Liu, Q. Zhang, J. J. Liang, P. N. Suganthan, and Q. B. Y., "Problem Definitions and Evaluation Criteria for CEC 2015 Special Session on Bound Constrained Single-Objective Computationally Expensive Numerical Optimization," *2015 IEEE Congr. Evol. Comput.*, pp. 84–88, 2015.