Securing the data using ABC Algorithm and Secure Multi-Party Computation Protocol In Cloud

S. Artheeswari[#1], RM. Chandrasekaran[#2]
[#1]*Research Scholar, Department of Computer Science, Annamalai University.*
[#2]*Professor, Department of Computer Science, Annamalai University.*
[1]*art.arthe@gmail.com*
[2]*aurmc@hotmail.com*

Abstract

Data mining is known as analyzing and extracting process the huge datasets to get the optimal data with considering the different kinds of hidden pattern relationship in given data and also consider much other helpful information. The cloud computing gives better frameworks for the cloud provider to employ the application on-demand process and computing the infrastructure with different constrains. Additionally, it also offers the higher flexibility to the cloud user by linking to the different kinds of computing resources and also permitting to the different kinds of IT enabled services. However, it has some risk to provide the privacy policy for the user's data and their security process. However, security among the huge number of cloud users is more significant aspect. Thus, in this paper proposes a novel cloud data security management framework with using soft computing techniques to prevent, contain and detect the unauthorized users in the cloud environment. In this work use a novel heuristic approach such as Genetic Artificial Bee Colony (GABC) algorithm with Secure Multi-party Computation (SMC) protocol in Trusted Cloud Computing Platform (TCCP). The experimental results show that the proposed method considerably minimize the total processing time and also shows that the better security performance during the VM running in cloud environment.

Keywords: **Cloud computing, security, privacy, Genetic Artificial Bee Colony (GABC) algorithm, Multi-party Computation (SMC) protocol, Trusted Cloud Computing Platform (TCCP)**

1. INTRODUCTION

Cloud computing is a one of the well-known technology which is processed based on the sharing of computing resources than having personal devices or local servers to process the applications. Basically in cloud computing, the word "cloud" refers to "the internet", so which means a sort of computing to make use of improving the computing power to process the millions of instruction per second [1]. The cloud computing process uses the networks of a huge group of servers with particularized links to distribute data processing among the different server [2]. In place of installing a software suite for each and every computer, this technology need to install single software in each every computer which permits to use log into web based service and is also hosts all the need programs by the user.

Big data is the word utilized to define massive volumes of unstructured and structured data which are so huge and problematic to manage [3]. In order to examine the cloud complex data and to recognize the patterns it is very significant to securely share, manage and store huge volume of complex cloud data [4]. The cloud comes with a categorical security challenge, for example the data owner might not have process of where the data is in their appropriately located. The reason behind this cloud data managing issue is that if one needs to get the advantage of the cloud computing the data owner must also use the resource allocation and additionally process the scheduling work for controls the given data [5]. Hence it is need to protect the cloud data in the force of unreliable processes. So the cloud includes extensive complexity, here considered that rather than giving a holistic solution to securing the cloud data, it would be ideal to create the considerable improvement in the securing the cloud data that will ultimately give a secure cloud.

Thus, in this paper proposes a novel cloud data security management framework with using soft computing techniques to prevent, contain and detect the unauthorized users in the cloud environment. In this work use a novel heuristic approach such as Genetic Artificial Bee Colony (GABC) algorithm with Secure Multi-party Computation (SMC) protocol in Trusted Cloud Computing Platform (TCCP). The experimental results show that the proposed method considerably minimize the total processing time and also shows that the better security performance during the VM running in cloud environment.

2. RELATED WORK

In [6] author presents approach to mine the data securely utilizing k-means algorithm. This process undertakes that the cloud data is not warehoused in a centralized cloud location but is distributed to different hosts. This proposed method avoids any intermediate data leakage in the procedure of computation while validity of the data mining process and preserving the correctness and the end results.

In [7] author present Access Control (AC) systems are between the maximum critical of network security components. A system’s security and privacy controls are more probable to be conceded because the misconfiguration of access controls policies instead of the failure of cryptographic protocols or primitives. This difficult process becomes gradually severe as software systems become very complex, such as Big Data (BD) processing systems, which are employed to manage a huge amount of sensitive information processing cluster. Fundamentally, access of BD control needs the association among cooperating processing domains to be secure as computing environments that contain of computing units under dispersed AC managements. In this work focused on Velocity, Variety and Volume. Respects for security in shielding BD are typically patch efforts and ad hoc. Even with some presence of security in recent BD systems, AC (Authorization), a critical security component, for protecting BD handling components and their users from the remains elusive, insider attacks. This paper suggests a general purpose AC scheme for distributed BD processing clusters.

In [8] author makes the first attempt to formally address the problem of availability, integrity and authentication. By consuming Tag generation, furthermore one of additional cloud storage service such that Cluster as a Service (CaaS) can make reduced cloud storage space and secure deduplication possible. Without key generation, attribute constructed encryption marks secure data deduplication in the computers cluster.

In [9] author process the cloud computing process with computing resources and services can be professionally utilized and delivered, making the vision of multiplying utility realizable. In different applications, execution of services with more number of tasks has to perform with minimum inter task communication. The approach to create the most of a miscellaneous set of tasks from the available resources in cloud competently is proposed. For this reservation cluster is presented, in which all the unprepared tasks are located and a new representing is done to lessen both the execution time and resource usage. Execution is carried out utilizing CloudSim, a toolkit for simulating and modeling cloud computing environments and assessed the proposed resource provisioning algorithm.

In [10] author proposed a system for secure data mining utilizing well known methods such as homomorphic encryption system, AES algorithm and k means clustering process. In this procedure flow, cloud server is ignorant of data uploaded by the user. And the cloud client only acquires the computational results. Over an experimental evaluation, in this work preserve confidentiality and correctness of final result.

3. TRUSTED CLOUD COMPUTING PLATFORM (TCCP)

Trusted cloud computing platform (TCCP) that gives a closed box execution environment by covering the concept of trusted platform to an entire Infrasructure as a Service (IaaS) backend. The TCCP assurances the integrity and the confidentiality of a user's Virtual Machine (VM), and permit a user to control whether or not the IaaS applies security in cloud platform. The TCCP does the job of leading all trusted nodes on one unit only, in its place the job is distributed among numerous entities, each managing a cluster, such that single unit does become the failure of the whole system, and the system cannot function effortlessly. Distributed trusted Cloud Computing (CC) platform is overcome TCCP issues.

A TCCP for guaranteeing the integrity and confidentiality of computations that are outsourced to IaaS services. The TCCP offers the abstraction of a closed box execution environment for a user's virtual machine, guaranteeing that no cloud provider's restricted administrator can tamper or examine with its Meta data. Furthermore, before inviting the service to introduction a virtual machine, the TCCP permits a customer to remotely and consistently define whether the service backend is running a trusted implementation of TCCP. This competence covers the notion of attestation to the whole service, and therefore it permits a cloud user to authenticate if its computation will run securely.

4. SECURE MULTI-PARTY COMPUTATION (SMC) PROTOCOL

The SMC is one of the mechanism for data privacy preserving in cloud data mining which means the cloud network is joint when the network process is start in cloud environment is shown in figure 1 and 2. It can be definite as, to give computations among different diverse organization in a secure or safe manner. With the SMC, different cloud user can perform jointly some global computation on their private data without any data loss privacy/ security. It gives base for end-to-end development of multiparty protocol.

Let $O_1 \ldots O_n$ where n is defined as organization that wish to perform a joint cloud computation $C1$ on their cloud private data. Since, the process is to performed on the cloud private data, it is significant need which is private data should not be accessible to any other organizations, which means in case $D_1, \ldots, D_n$ be the data associating to a n number of organization and assume D_i be data linked to i^{th} organization, then it is need for computing that D_i should not be process able to any $D_j\ where\ i \neq j\ and\ j = 1, 2 \ldots n$. Therefore, each and every organization only gets the final outputs of joint computation without being aware of given inputs included and the computations made. The privacy is one of the significant concerns for SMC protocols and each and every organization is a means to guarantee them in an easy manner.

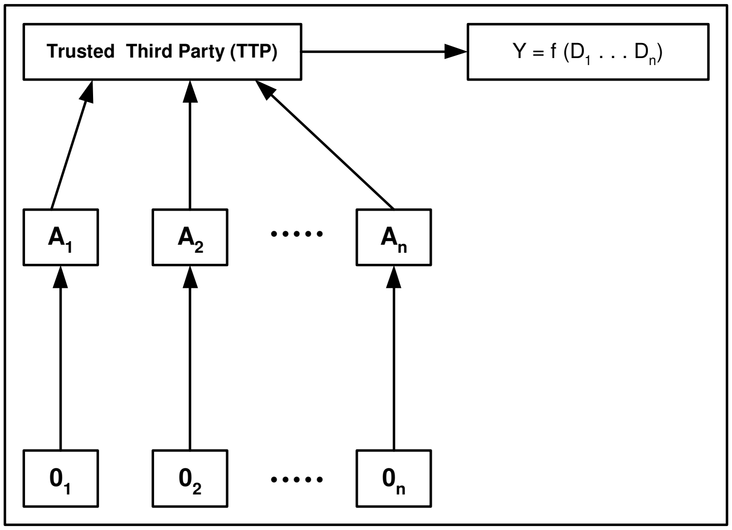

Fig. 1 Secure Multi-party Computation (SMC)

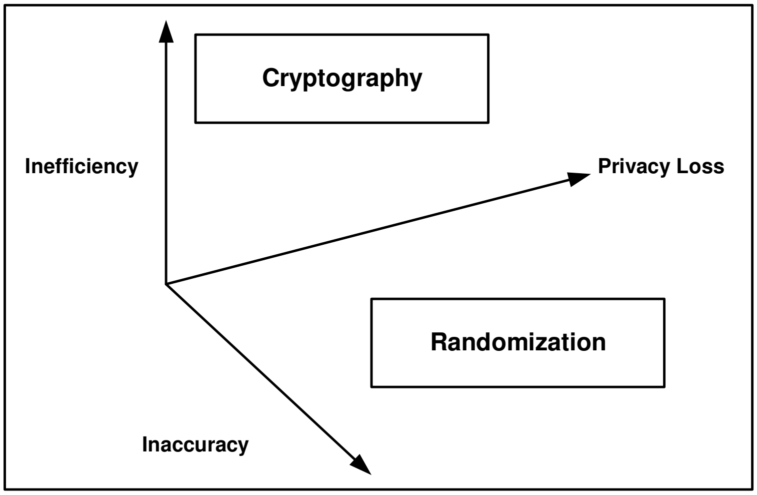

Fig. 2 Secure Multi-party Computation (SMC) process

Secure Multi-party Computation (SMC) Algorithm

```
Define parties $P_1, P_2, P_3, P_4, P_n$ where $n > 1$. Each party involved in computation.
Suppose, each and every party has input data blocks $X_1, X_2, X_3, X_4 \dots X_n$.
Each and every cloud user breaks it data blocks into the $k$ number of segments.
Each and every user utilized distribution function.
Arrange parties in a ring as $P_1, P_2, \dots, P_n$ and selects $P_1$ as the protocol initiator.
Each and every party decides random numbers for each and every segment
$r_{i1}, r_{i2} \dots \dots r_{ik}$.
For $i = 1\ to\ n$
    $SUM = \sum_{i=1}^{n} (D_{il} + r_{il}) * H$
This $SUM$ calculated by trusted party.
Trusted third party sends $SUM$ to party $P_1$.
For i=1 to n
    $SUM = [(SUM - r_{i1}) * H + (D_{i2} + r_{i2}) * H]$
The $n^{th}$ third party sends SUM to first party, since all the parties organized in a ring.
Again, first third party calculates this process
    $SUM = [(SUM - r_{i2}) * H + (D_{i3} + r_{i3}) * H]$
The $n^{th}$ party sends SUM to first party.
For $i = 1\ to\ n$
    $SUM = (SUM - r_{i3})H$
Now, the $n^{th}$ party directs final sum to trusted party.
The trusted party publicizes the broadcasts and result to all the parties
```

5. GENETIC ARTIFICIAL BEE COLONY (GABC) ALGORITHM

The proposed Genetic Artificial Bee Colony (GABC) algorithm with Secure Multi-party Computation (SMC) protocol in Trusted Cloud Computing Platform (TCCP) process is employed in cloud environment. GA is one of the well-known soft computing techniques, which utilize the natural selection mechanism. In the candidate solutions named individuals is computed to better solutions with the genetic operation such as mutation, crossover, and selection. Each and every solution has of properties named chromosomes which can be recombined. In each iterations, solution of the population is computed with considering the every individual fitness value (normally a value associated to the objective function). High fit individuals are surviving and selected for consequent iterations in each and every new iterations (generation), here creating the set of strings utilizing information from the previous ones. This genetic work applied in ABC process for cloud data allocation with secure way.

Let $VM = \{VM_1, VM_2, VM_3, \dots, VM_N\}$ is a set of N VM and $Task = \{task_1, task_2, task_3, \dots, task_k\}$ of K task to be regular and processed in VM. All the machines are unrelated yet are paralleled.

Each services x in R can be defined utilizing coordinates $i\ and\ j$ as:

$$x_{ij} = position(p_i): i = 1,2,\dots,n;\ j = 1,2,\dots.,m \qquad (1)$$

$f(p_i)$ is defined as the variance in computation time of the specific service to total computation time of all the services present in that request of the user and is specified as:

$$f(p_i) = \left[\sum_{j=0}^{\pi} time(p_j)\right] - time(p_j) \qquad (2)$$

fitness of each and every service is calculated by a fitness function

$$fit(p_i) = \begin{cases} \dfrac{1}{1+f(p_i)}, iff(p_i) \geq 0 \\ 1 + abs\,(f(p_i)_i) if\ f(p_i) \leq 0 \end{cases} \qquad (3)$$

The fitness values of each services p_i are calculated using the fitness function.

Genetic Artificial Bee Colony (GABC) Algorithm

```
Cycle1
Initialize the food source positions $x_i$, $1, \ldots SN$
Weigh the nectar amount fitness function $fit_i$ of food sources
Repeat
    Employed Bees‟ Phase
    For each and every employed bee
        Yield new food source positions $v_i$
        Compute the value $fit_i$
        Employ GA selection process
    End For
    Compute the probability values $p_i$ for the solution.
    Onlooker Bees‟ Phase
    For each and every onlooker bee choose a food source depending on $p_i$
        Yield new food source positions $v_i$
        Compute the value $fit_i$
        Employ GA selection mechanism
    End For
    Scout Bee Phase
    If an employed bee becomes scout, after that interchange it with a new random source positions
    Memorize the best solution attained so far & $cycle = cycle + 1$
Until $cycle = MCN$
```

6. PERFORMANCE EVALUATION

In this section, the experimental estimation of the proposed GABC with SMC, TS-GA algorithm, and Round-Robin algorithms is presented.

By utilizing CloudSim toolkit, the proposed GABC with SMC is implemented, and a comparative study has been prepared among three different kinds of algorithms such as Round-Robin (RR), and the TS-GA algorithms. Here using five different kinds of parameters for evaluation process such as efficiency, speedup, resource utilization, and cost and completion time.

Figure 3 and Table 1 shows that the completion time of proposed work GABC with SMC is compared with other two different algorithms such as TS-GA and RR. From the results the proposed work shows that the promising result when compared with other two algorithms which means the proposed work shows that the minimum completion time of data storage process in cloud environment.

Table 1. Completion Time

Number of allocation	GABC with SMC	TS-GA	RR	No of VM
25	132.19	282.51	164.78	8
50	276.12	410.21	568.77	
75	415.97	670.84	572.16	
100	533.7	832.89	944.37	

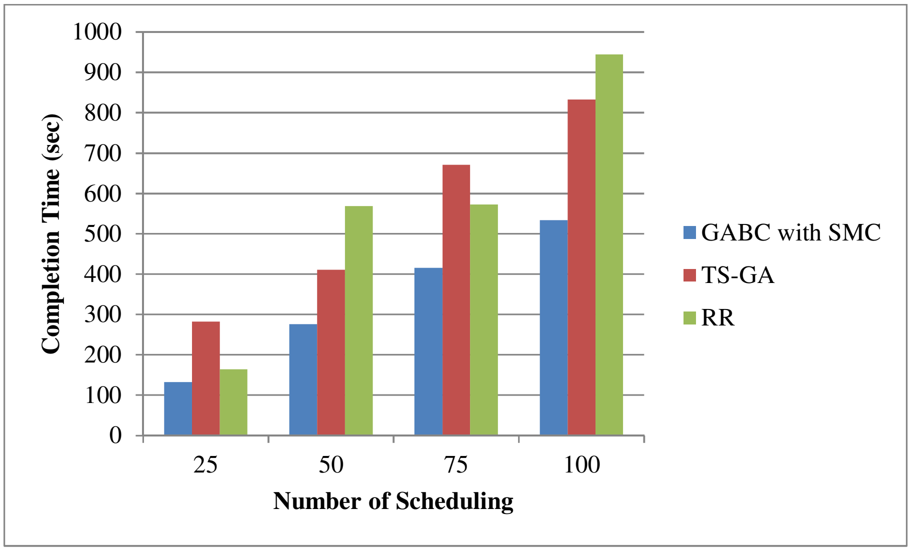

Fig.3 Completion Time

Figure 4 and Table 2 shows that the Execution Cost of proposed work GABC with SMC is compared with other two different algorithms such as TS-GA and RR. From the results the proposed work shows that the promising result when compared with other two algorithms which means the proposed work shows that the minimum Execution Cost of data storage process in cloud environment.

$$Total\ cost = \frac{task\ length * Cost\ per\ seconds}{VM\ mips} + Processing\ cost \qquad (4)$$

Table 2. Execution Cost

Number of allocation	GABC with SMC	TS-GA	RR	No of VM
25	2016.95	3012.76	2814.21	8
50	6432.21	6712.46	6502.88	
75	9416.13	10448.22	9810.48	
100	10414.04	12117.76	12517.22	

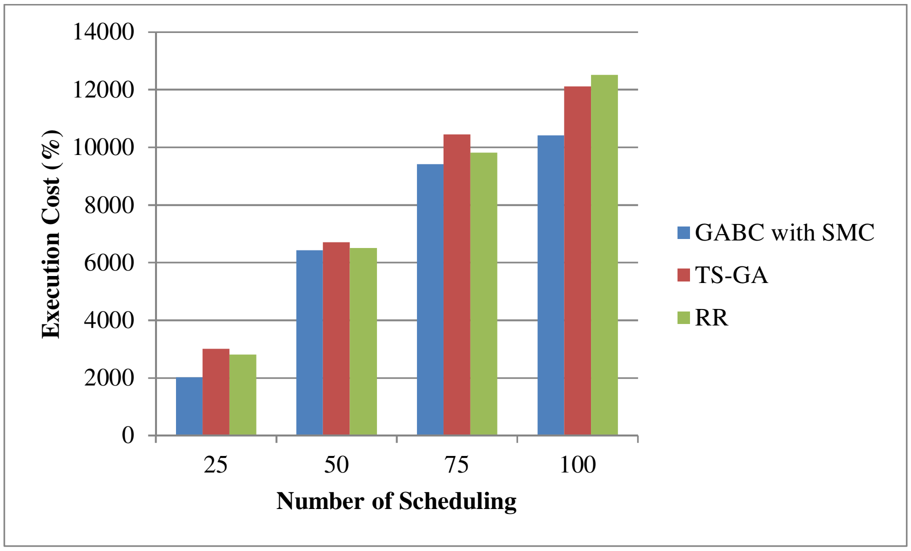

Fig. 4 Execution Cost

Figure 5 and Table 3 shows that the *Resources utilization* of proposed work GABC with SMC is compared with other two different algorithms such as TS-GA and RR. From the results the proposed work shows that the promising result when compared with other two algorithms which means the proposed work shows that the maximum *Resources utilization* of data storage process in cloud environment.

$$Resources\ utilization = \frac{final\ VMs\ avaliable\ time}{Number\ of\ VM * scheduling\ time} * 100 \qquad (5)$$

Table 3 Resource Utilization

Number of allocation	GABC with SMC	TS-GA	RR	No of VM
25	74.89	30.51	54.13	8
50	83.43	46.88	43.36	
75	80.65	37.96	67.4	
100	74.78	47.7	51.04	

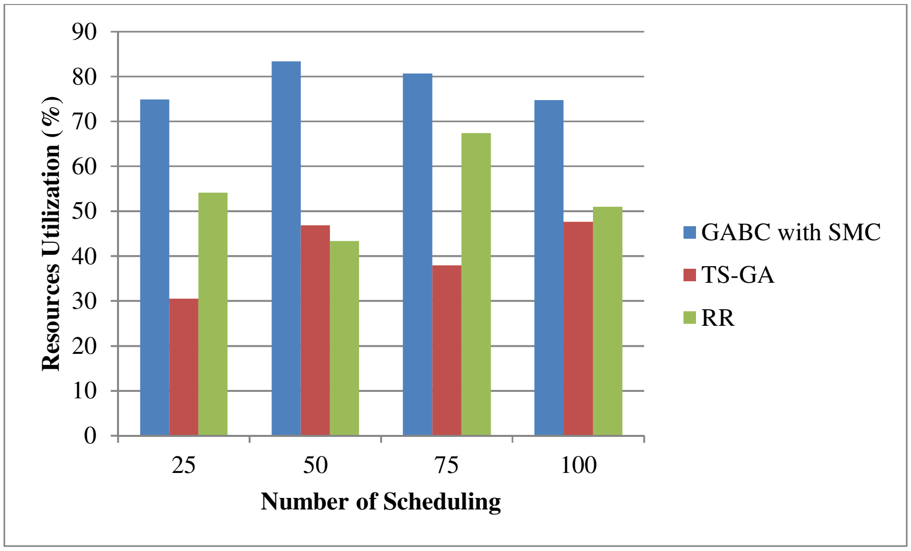

Fig. 5 Resource Utilization

The experimental results show that the proposed method considerably minimize the total processing time and also shows that the better security performance during the VM running in cloud environment.

7. CONCLUSION

This paper proposes a novel cloud data security management framework with using soft computing techniques to prevent, contain and detect the unauthorized users in the cloud environment. In this work use a novel heuristic approach such as GABC algorithm with SMC protocol in TCCP. Here using five different kinds of parameters for evaluation process such as efficiency, speedup, resource utilization, and cost and completion time. The experimental results show that the proposed method considerably minimize the total processing time and also shows that the better security performance during the VM running in cloud environment.

REFERENCES

[1] F. John Krautheim, Dhan anjay S. Phatak, and Alan T. Sherman, "Introducing the Trusted Virtual Environment Module: A New Mechanism for Rooting Trust in Cloud Computing". In TRUST 2010, LNCS 6101, pp. 211–227, 2010, © Springer-Verlag B erlin Heidelberg 2010.

[2] N. Santos, K.P. Gummadi, R. Rodrigues, "Towards Trusted Cloud Computing". In Proc. of the 1s USENIX Workshop on Hot Topics in Cloud Computing, Berkeley, CA, USA, 2009.

[3] Debayan Gupta, Aaron Segal, Aurojit Panda, Gil Segev,Michael Schapira, Joan Feigenbaum, Jenifer Rexford, Scott Shenker, "A New Approach to Interdomain Routing Based on Secure Multi-Party Computation", http://conferences.sigcomm.org/hotnets/2012/papers/hotnets12-final55.pdf.

[4] Assaf Ben-David, Noam Nisan, Benny Pinkas, "FairplayMP – A System for Secure Multi-Party Computation", http://www.cs.stevens.edu/~nicolosi/classes/12fa-cs693/final/f2-1.pdf, 2008.

[5] Z. Zheng, R. Wang, H. Zhong, and X. Zhang, "An approach for cloud resource scheduling based on Parallel Genetic Algorithm," in Computer Research and Development (ICCRD), 2011 3rd International Conference on, 2011, pp. 444-447.

[6] Raunak Joshi, Bharat Gutal, Rajkumar Ghode, Manoj Suryawanshi, Prof U.H. Wanaskar, "Data Mining Using Secure Homomorphic Encryption", International Journal of Advanced Research in Computer and Communication Engineering,Vol. 4, Issue 10, 2015.

[7] Vincent C. Hu, Tim Grance, David F. Ferraiolo, D. Rick Kuhn, "An Access Control Scheme for Big Data Processing", http://csrc.nist.gov/projects/ac-policy-igs/big_data_control_access_7-10-2014.pdf

[8] R.Nalla Kumar, A.Kumari savitha sree, X.Alphonseinbaraj, "Cluster as a Service (CaaS) in Secure Deduplication System", International Journal of Computer Applications Technology and Research, Vol. 4– Issue 1, 18 - 23, 2015.

[9] G. Malathy, Rm.Somasundaram, "Performance Enhancement in Cloud Computing using Reservation Cluster", European Journal of Scientific Research, Vol. 86 No 3, PP.394-401, 2012.

[10] Vikas Maral, Sagar Kale, Ketan Balharpure, Sourabh Bhakkad, Pranav Hendre, "Homomorphic Encryption for Secure Data Mining in Cloud", International Journal of Engineering Science and Computing, PP. 4533-4536, 2016.

AUTHOR PROFILE

Mrs. S. Artheeswari is working as Assistant Professor in Mailam Engineering College, Mailam, Tamilnadu. She has 8 years of experience in academic field. She completed her Bachelor of Technology(IT) in Madras University and Master of Engineering(CSE) in Anna University. Now, doing as a Research Scholar in Annamalai University in the field of Computer Science. Her area of Interest includes Cloud computing, Data

Structures, Security and Database Management System. She also has life member for several association and society. She published 3 papers in in International Journals. She also published many papers in national and international conferences.

Dr. RM. Chandrasekaran is currently working as a Professor, Department of Computer Science & Engineering and also jointly as Controller of Examinations for Annamalai University, Chidambaram. He obtained his Bachelor of Engineering in Computer Science and Master of Engineering from Anna University and Master of Business Administration from Annamalai University. Completed his PhD in Computer Science from Annamalai University, Annamalainagar, India. He has 23 years of teaching experience and 5 years of experience in Research & Development. He also worked as a Registrar in Anna University, Trichy for 3 Years. He worked as software consultant in USA. He was also a Director, Directorate of Distance Education. Annamalai University. Also, he has co-organized two Workshops and two conferences. His area of interest includes Computer Algorithms, Text Data Mining and Software Metrics. He published 10 papers in National Journal and 13 papers in International Journals. He also published many papers in national and international conferences.