CAI method in Cloud services for user workload using Cloudsim framework

AlaguVignesh A, Barath B, R.R Sathiya, P Prakash

Abstract—In sighting the distinct patterns of processing capability in a cloud service is pedantic to enhance the resource management and operable conditions of the servers without compromising the Quality of Service is important. Simulations and models based on practicable parameters are required to understand the impact of the load on new system designs and policies. The proposed scheme and analysis provides a requirement for designing new systems which will be lessaffected by process loads. Classifying, analysis and improving (CAI) is done using real-time data center logs and simulations are done based on user requests and data center configurations. Simulations are created using cloudsim framework. Various simulations are done to provide a comprehensive result to improve the resource allocation for the system.

Index Terms—Cloud service, Resource, Users, Work load, Classification, Neural-network, Cloud-sim, Simulation

I. INTRODUCTION

The growing ultimatum need for scalable and reliable ser- vices for computing has steered to the evolution of Cloudcom- puting. Users needed a salable and cost-effective computing environment that could serve their requirements in an efficient manner. The process of delivering hosted services over the internet is known as Cloud services. These services embrace servers, storage, databases, networking, software, analytics and intelligence. Users can access these services by logging in from any device tha thasaninternetConnection.

Infrastructure as a service, Software as a service and Plat- form as a service are the major services provided by the cloud environment.

Infrastructure as a service(IaaS) fends for processing and storage resources. The advantage of using this model is that it reduces the overall cost, combines the resources, increases the speed of deployment as well as enhances security.

Platform as a service(PaaS) provides an environment for creating cloud ready applications It also provides capabilities like development and deployment.

Software as a service (SaaS) fends for software and applications through a subscription rather than physically installing it on the local machine. Example software as a service are Google apps, Dropbox etc.

Revised Manuscript Received on November 19, 2019

AlaguVignesh A, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham India

Barath B, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham, India

R.R Sathiya, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham, India

P Prakash, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham, India

The common characteristics of a cloud are reduced cost, scalability and Virtualization. These advantages possessed by cloud attract potential customers to migrate to a cloud Environment rather than investing on dedicatedhardware.

There are various ways in which a cloud model is deployed.

Some of the major models are community, hybrid, public and private. A public model is where computing assets are dynamically provisioned on a demand over the internet via applications from a provider who bills on usage basis. A privatecloudenvironmentisamodelwherethecloudsetupis locally maintained within the organization, used by users belonging to that organization. A hybrid cloud consists of a part of computing resource on- site and off-site (public cloud). The advantage of using such a model is that users can opt public cloud for performing specific functions that are too costly to maintain on-site such as disaster recovery and backups. A community cloud is a model where several organizations with similar requirements group together to share a common infrastructure. The usage cost is spread over the organizations in thecommunity.

The service offered by cloud environments is a measured service where resource usage by the user is monitored and disclosed. Users of the cloud environment need to pay as they use. Cloud computing also provides on demand self-service which automatically allocates resources whenever a need for additional resource arises without any human intervention. This eliminates the need for IT administrators to provide and manage computingresources.

II. LITERATURESURVEY

The reduction of total cost for the end user by optimizing the response time which leads to reduced wear and tear of hardware equipment. It also deals with simulation of data-centers in various regions and analysis of network response time. There was a scope for the simulation and modeling by extendingthefunctionalitiesincloudsim[1].

The performance analysis mixed to CPU, HDD, server RAM thus improving the energy utilization of theservers. Then concentrated to increase the number of parameters used for the analysis [2].The cloud computation system uses signal processing techniques to analysis network usage and overall transmission delay between the end user and the data center. Themaindrawbackisitusesacomplexityof $O(n2)$ algorithm. Theconcernwastoimprovethealgorithmcomplexity[3].

The standard parameters such as task and user are used tocreate a model with only 5 percent error. This methodologycan be applied to a data set that contains data for a month,and we can compare the parameters, for this data set with theparameters that we obtained [4]. An analysis of diversity ofcloud computing from a large scale data center are studied.Models are created based on task characteristics of the user,based on different observation over periods. Models are made based on diverse and critical operational parameters [5].Further, design and implementation of a cloud computingservice's resource management system is given. This systemmaps the physical resources to virtual resources accordingly.Here, to group VM's with various characteristic's skewness isused as a metric. The virtual machines are grouped accordinglyto fully utilize the server capacity. For machines with multi resource constraints both green computing and avoidance of overload is achieved by this algorithm [6].

Performanceanalysisofthedatacenterusingtaskparameter submitted by the user and improvement of overall resource usage and the main drawback of this is the analysis is limited to 2-day log which gives inaccurate results in real-time[7]. This work presents the work that advances cloud computing. Clustering approaches are also used to categorize the users[16]. Clustering approaches can be used in any domain or data which can be either structured un-structured or semi- structured [17][18].

Cloud computing can be advanced by the following ways. First, this work can be used to reduce the cost of energy consumption. By reducing energy cost we can build a cloud computing industry that is stronger and more competitive.Sec- ond, the consumers nowadays are more considerate about the environment. The results of this work has depicted that when compared to the traditional method of resource allocation, the method mentioned in this paper is more efficient as it consumes less amount ofenergy[8].

A method for the classification of tasks is developed and this method is then applied to the Google Cloud Backend [9]. The methods for the classification of workloads are: (1) to identify the dimensions of the workload. (2) Construction of classesforthetasksusingasuitablealgorithmlikek-means. (3) Determine the break point within the workload dimension range.(4)Mergingnearbyclassessothattheworkloadnumber will decrease. The efficiency of the Virtual machines and provides a vision of how cloud computing business can be in the future [10]. This provides how virtual machines can be increasedintermsofscalabilityandelasticityfortheuser.

Fuzzy method of classification can identify the possiblearea for each class within the pattern margin using multi-layer feed-forward artificial neural network[11]. Recurrent neural networksaremorepowerfulthanthetraditionalclassifiermod- els. This is because data vector's correlation can be modeled by recurrent neural network. But this may not be practically achievable. The recurrent neural network's performance is better than the traditional model when there is an overlap in the class data. This overlap is overcome by using smoothing techniques[12].

This paper [13] presents the techniques and methods for the construction of workload models. The techniques and methods presented in this paper are only related to system type and the objectives of this study. The problems faced during the workload characterization are already known and have been solved age back. But the same results or conclusions cannot be made for recent architectures. Nowadays it is important to identify the parameter set that will be able to capture the behavior of the workload and it should be able to reproduce the behavior. The hypothesis and how different workloads have a noticeable effect on energy consumption and that a low-cost, scalable, cross-platform software solution for energy measurement is possibly explored[14].

Various numbers of hidden layer determines the performance of the neural network thus producing extremely accurate results [15]. The range of hidden layer units vary from thirty toseventy.Theoutputlayersvaluecanbeusedtofindtheconfidence level which is between 0-1. The confidence value can be used to find the performance of the neural network. If the confidence value is more, then the probability of correct classification increases. For neural networks with less than 30 or more than 70 hidden layer unit's convergence was not achievable.

An efficient resource usage ensures high Quality of Service for the provider as well as the end-user. Analyzing and classifying the users based on tasks helps to develop a model which can be used for economical resource allocation for the virtual machines. Most of the work is based on clustering the users using clustering algorithms for example k-means was used to cluster users based on their similarities in resource usage. To the best of our knowledge, we have not found any existing work that efficiently classifies users using Neural network classification and create server configurations based on the classification output, thus, the motivation for the proposed system.

III. PROPOSEDSYSTEM

A. Classification

Classification is an approach used for supervised learning, the model learns from the input data, uses the insight to classify new observation. This proposed system uses Neural network classification algorithm to classify users based on theirrequestedresources.Theadvantageofusingclassification in this system is that when a new user is added the system can classify him efficiently and can fit him in a data center which has a similar specification as requested by the new user. Fig1showsanexampleofaclassification.

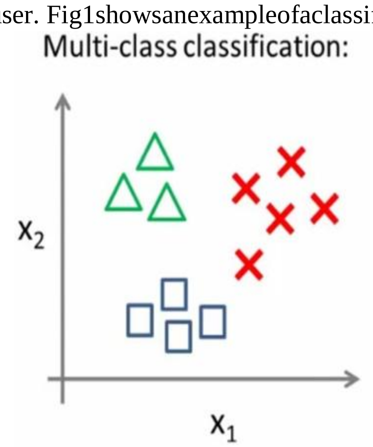


Fig. 1. Classification

B. Modelsimulation

To analyze and model the performance of a cloud data-center under a set of conditions and policies it is essential to simulate a model which would give results similar to real-time usage. The simulation is done using cloud Analyst (Refer Fig 2 and Fig 3**)** which is an extension to cloud sim framework that provides graphical experience. Using this framework data centers can be configured and can be fedwith user requests, then based on the simulation results the configuration can be modified to meet the requirement based onthenumberofuserrequests.

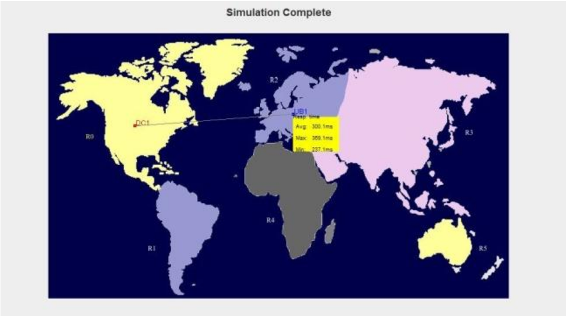


Fig. 2. Cloud Analyst

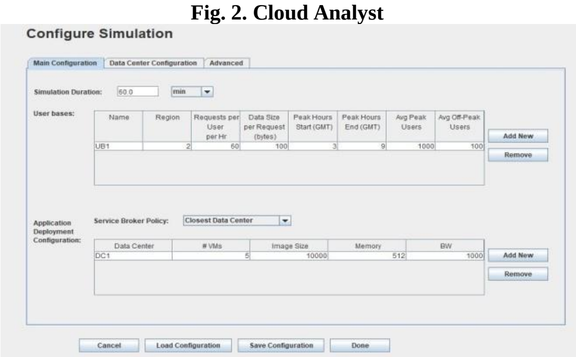


Fig. 3. Configuration settings

C. Verification

A standard validation technique is that part of the data is used to build the model and rest of the data is used to verify whether the system works as learnt. Similar to this based on the simulated model a data center can be designed, operated under the simulated conditions in real time, and verified for correctness of thesimulation.

D. Methodology

The first stepistoclassifytheusersbasedontheirrequested resources as high, medium and low. This is achieved through Neural network classification. The classification works better than clustering in this application because new entries (users) can be classified efficiently than fitting them in a cluster. Classification does not have any overlap thus leading to proper fitting of a user in a data center. Using the classification result the user with the maximum usage is retrieved fromeach classified class and the resource parameters of that user isused to create the configuration model because it allocates the resourceswhichisgoodenoughtomeettheuser'srequirement and other users in the same class. The data set has important parameters User, CPU, Memory, and Disk usage. The data set has to be pre-processed to remove missing values, outliers, etc.togetanaccurateclassification.

The second step is to simulate the workload in a data-center with predefined configurations and see how it behaves to the

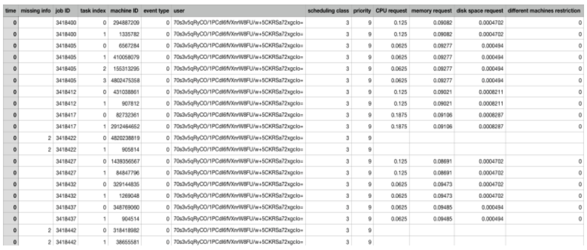


Fig. 4. Original data set

1	user	CPU	Memory	Disk
2	70s3v5qRyCO/1PCdI6fVXnrW8FU/w+5CKRSa72xgclo=	0.094886364	0.07814346	0.000482282
3	y1zb4shvaJurgBq+U8Ba9lm6r6qffGJOHPOT+2c7Zzo=	0.067146667	0.06420167	0.000149255
4	iV68dcBG/5coLZyooOnXWm9MpRg+xiBqLp2MVgNJ8B4=	0.052083333	0.01138967	0.0000382
5	WVtO5qw3sNnP4MeiRUnqr07CekrYMU12Mc7GbsgnjhQ=	0.059868324	0.09155022	0.000469201
6	gUvfmF1WIPzN8t8PAyLthnzpOSxOaYcRSjTVvmShd4o=	0.0546875	0.031335	0.000164475
7	rNyxTd1B3RnDJBIofzoVksjnmaJEN+hW5W+yRHo2xsM=	0.074196874	0.05144176	0.000330306
8	XAk+iT/NLGe4Ow+7ARniw9vei1mUmLS0U2w0PFWXEwk=	0.062885802	0.02894546	0.000306485
9	dTBkGCvrIzbJpD3rFqjUp7wz1+pLdgL7xCt64xcZN0o=	0.030974767	0.03250646	0.000233353
10	+O9sEnpjz/OC05OonUa1cOTSzxVtjMyN6sbMnsTMGzc=	0.061991667	0.01850083	6.39E-05
11	o52i0r1lc0YLJeuQDOy3+gymJ4VbzGMzz+kv+q+HJY4=	0.046875	0.0176775	0.0000844

Fig. 5. Data set after pre-processsing

process load. The simulation is done using cloud analystwhich extends cloud sim framework to give a graphical experience. Using cloud sim we can configure data centers and provide it withnumberofrequestssentbytheuser.Theseparameterscan be modified to see different results and create an impeccable model which can be used to design new systems to handle large amount of data generated by the users. The overall steps areshownasarchitecturediagramsinFig6andFig7

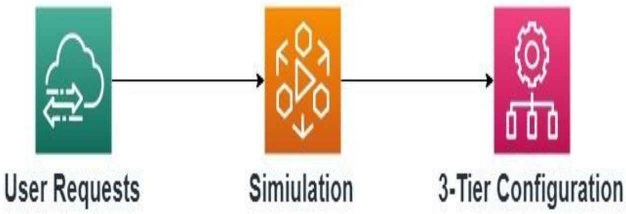


Fig. 6. Classification of users

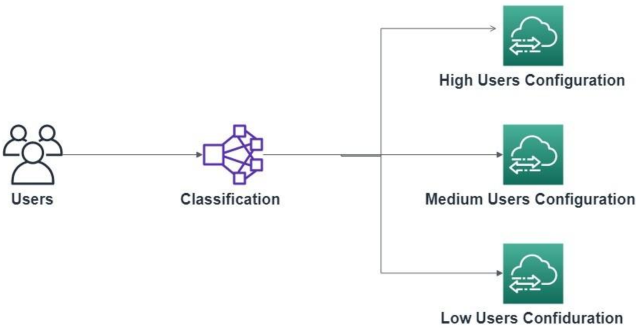


Fig. 7. Model simulation and creating configuration

IV. RESULTANALYSIS

On classifying the users based on their resourcerequirement using neural network the following result (Refer Fig 8) was achieved.Thisisaplotofusersandthecolorcodingrepresents theclasstowhichtheybelongi.e,high,mediumandlow.

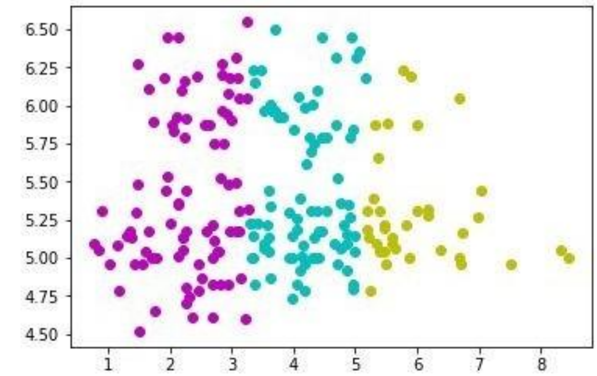


Fig. 8. Classification of Users as (i)High (ii)Medium (iii)Low

TABLE I
MAXIMUM USAGE FROM EACH CLASSIFIED CLASS

Class	CPUusage%	Memoryusage%	Disk usage %
A(Violet)	0.25	0.318366667	0.00686
B(Blue)	0.1875	0.1143	0.001201583
C(Yellow)	0.119549767	0.091550216	0.000801096

On Simulating the Data center with various configuration anduserrequestsperhourthefollowingdatacenterprocessing time (Average, Maximum) was identified and tabulated in (Refer TableII)

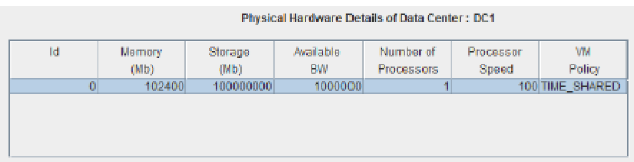
TABLE II
RESPONSE TIME BASED ON DATA-CENTER CONFIGURATION

RAM (MB)	No. CPU cores	Min(ms)	Avg (ms)	Max (ms)
1024	1	5.51	84.09	170.00
2048	2	3.01	34.20	84.01
4096	3	1.26	14.04	28.75
8124	4	1.0110.65	24.00	

Response timing for varied number of requests sent per user to the data center configured with RAM = 1024, No. Cores = 1 is depicted in Table III and configuration settings is shown in Fig9

TABLE III
RESPONSE TIME BASED ON CONFIGURATION 1

No. requests sent per user	Maximum(ms)
60	170.0
120	177.5
240	178.5
480	183.5


Physical Hardware Details of Data Center : DC1

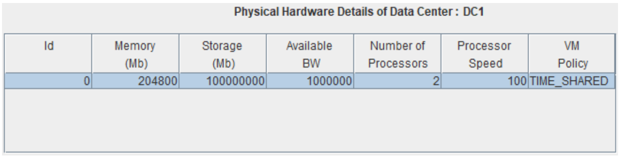
Id	Memory (Mb)	Storage (Mb)	Available BW	Number of Processors	Processor Speed	VM Policy
0	102400	100000000	1000000	1	100	TIME_SHARED

Fig. 9. Data-center configuration 1

Response timing for varied number of requests sent per user to the data center configured with RAM = 2048, No. Cores = 2 is depicted in Table IV and configuration settings are shown in Fig 10.

TABLE IV RESPONSE TIME BASED ON CONFIGURATION 2

No. requests sent per user	Maximum(ms)
60	84.01
120	85.51
240	85.51
480	95.00


Physical Hardware Details of Data Center : DC1

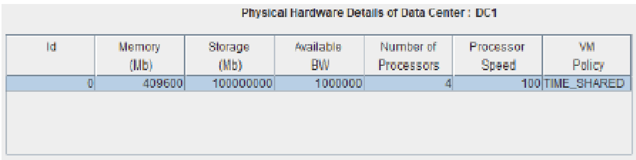
Id	Memory (Mb)	Storage (Mb)	Available BW	Number of Processors	Processor Speed	VM Policy
0	204800	100000000	1000000	2	100	TIME_SHARED

Fig. 10. Data-center configuration2

Response timing for varied number of requests sent per user to the data center configured with RAM = 4096, No. Cores = 4 is depicted in Table V and configuration settings are shown in Fig 10,Fig 9and Fig11.

TABLE V
RESPONSE TIME BASED ON CONFIGURATION 3

No. requests sent peruser	Maximum(ms)
60	28.75
120	38.01
240	38.01
480	38.01
960	38.1


Physical Hardware Details of Data Center : DC1

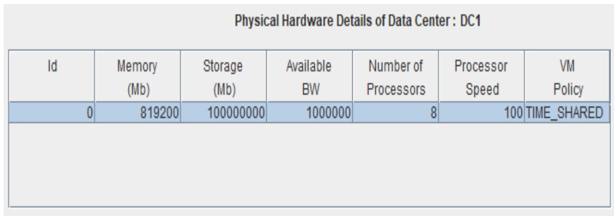
Id	Memory (Mb)	Storage (Mb)	Available BW	Number of Processors	Processor Speed	VM Policy
0	409600	100000000	1000000	4	100	TIME_SHARED

Fig. 11. Data-center configuration 3

Response timing for varied number of requests sent per user to the data center configured with RAM = 8192, No. Cores = 8isdepictedinTableVIandFig12.

TABLE VI
RESPONSE TIME BASED ON CONFIGURATION 4

No. requests sentperuser	Maximum(ms)
60	24.00
120	24.00
240	30.50
480	30.50
960	30.50


Physical Hardware Details of Data Center : DC1

Id	Memory (Mb)	Storage (Mb)	Available BW	Number of Processors	Processor Speed	VM Policy
0	819200	100000000	1000000	8	100	TIME_SHARED

Fig. 12. Data-center configuration 4

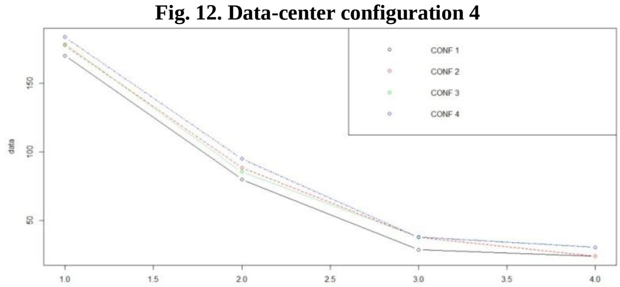


Fig. 13. Configuration VS response time(ms) graph

V. CONCLUSION

In this paper we have presented how the data center would behave on a peak time using simulations. The analysison different configuration directly reflected in the response time, this clearly shows that when a data center is under powered it works good under normal operating conditions and it may subject to overload and throttle during peak time leading to wastage of energy and faster worn out time of hardware. Future data centers can be built efficiently by using optimum resourceswithoutunderfittingoroverfitting. This analysis improves the data-center only by the means of resource allocation, during peak time due to processing load the hardware operates at high load leading to wear out of the hardware and due to high operation of the hardware it consumesmoreenergyleadingtowastageofenergyresources. Future work can be focused on improving the hardware capabilitiestohandletheseprocessingloads.

REFERENCES

1. MnahilKherAlseed Mohammed Alnazir ; Amin Babiker A. Nabi Mustafa ; Hamid Abbas Ali ; Amir Ahmed Omer Yousif, "Performance analysis of Cloud Computing for distributeddatacenterusingcloud-sim",2017.
2. SidakPreet, Amardeep Kaur, "Behavior Pattern Anal- ysisforResourceUtilizationinCloudComputing",2016.
3. Chris Muelder, Biao Zhu, Wei Chen, Senior Member, IEEE, Hongxin Zhang, Member, IEEE,Kwan-Liu Ma, Fellow, IEEE, "Visual Analysis of Cloud Computing Performance Using Behavioral Lines" ,2016.
4. Ismael Solis Moreno, Peter Garraghan, Paul Townend, JieXu,"AnApproachforCharacterizingWorkloadsinGoogle CloudtoDeriveRealistic,ResourceUtilizationModels",2013.
5. Ismael Solis Moreno, Peter Garraghan, Paul Townend, and Jie Xu, Member, IEEE, "Analysis, Modeling and Simu- lation of Workload Patterns in a Large-Scale Utility Cloud", 2013.
6. Zhen Xiao, WeijiaSong, Qi Chen, "Dynamic Resource AllocationUsingVirtualMachinesforCloudComputingEnvi- ronment", Parallel and Distributed Systems IEEETransactions, vol.24,no.6,pp.1107-1117,2013.
7. Iyapparaja M et.al. 2012 Coupling and Cohesion Metrics in Java for Adaptive Reusability Risk Reduction IET Chennai 3rd International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2012),52-57.
8. Anton Jernal, Rajkumar, "Energy-aware resource allo- cation heuristics for efficient management of data centers for Cloud computing", Future Generation ComputerSystems, vol. 28,no.5,pp.755-768,May2012.
9. Mishra, J. L. Hellerstein, W. Cirne and C. R. Das, "Towards characterizing cloud backend workloads: insights from Google compute clusters, ACM SIGMETRICS Perfor- mance Evaluation Review",2010.
10. M. Armbrust et al., "Above the clouds: A berkeleyview of cloud computing," University of California, Berkeley, Tech. Rep., Feb2009.
11. Iyapparaja M, Bhanupriya Sharma, Augmenting SCA project management and automation Framework, IOP Conf. Series: Materials Science and Engineering 263 (2017) 042018 doi:10.1088/1757-899X/263/4/042018,pp-1-8
12. T. L. Burrows and M. Niranjan, "THE USE OF RECURRENT NEURAL NETWORKS FOR CLASSIFICA- TION",1994.M. Calzarossa and G. Serazzi. Workload characteri- zation: A survey. Proceedings of the IEEE, 81(8):1136–1150, Aug1993.
13. M. Iyapparaja, P. Sivakumar. Metrics Based Evaluation for Disease Affection in Distinct Cities. Research J. Pharm. and Tech. 2017; 10(8): 2487-2491.
14. Du-YihTSAI, HiroshiFUJITA, Jiatsuhei HORITA ,Tokiko END0 , Choichiro KID0 and Takeo Ishigaki, "Classification of Breast Tumors in Mammograms using a Neural Network: Utilization of Selected Features", 1993.
15. Sathiya RR. Content Ranking Using Semantic Word Comparison And Structural String Matching. International Journal of Applied Engineering Research. 2015;10(11):28555- 60.
16. Sathiya R.R., Swathi S., Nevedha S., ShanmugaSruthi
17. U. (2016) Building a Knowledge Vault with Effective Data Processing and Storage. In: Suresh L., Panigrahi B. (eds) Proceedings of the International Conference on Soft Computing Systems. Advances in Intelligent Systems and Computing, vol398. Springer, New Delhi.
18. P. Sridhar and Sathiya, R. R., "Crypto-watermarking for secure and robust transmission of multispectral images", in 6th International Conference on Computation of Power, Energy, Information and Communication, ICCPEIC 2017, 2018, vol. 2018-January, pp. 153-163.

AUTHORS PROFILE

AlaguVignesh A. is a B.Tech student at Amrita Vishwa Vidyapeetham, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore India. He is actively involved in research in the areas of Cyber Security, Networks, IoT and Cloud Computing. He is a Certified Ethical Hacker (CEH) and CISCO Certified Network Associate(CCNA R&S).

Barath B. is a B.Tech student at Amrita Vishwa Vidyapeetham, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore India. He is actively involved in research in the areas of Cloud Computing. He is also a Certified Ethical Hacker (CEH).

Sathiya R. R. currently serves as Assistant Professor at Amrita Vishwa Vidyapeetham, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore India. Her research areas include Data mining, Cloud computing and Data analytics in particular with cloud computing.

Dr.PrakashP. currently serves as Assistant Professor at Amrita Vishwa Vidyapeetham, Department of Computer Science and Engineering, Amrita School of Engineering, Coimbatore Campus. His areas of research include Cloud Computing, Automata Theory and Analysis of Algorithms. Dr. Prakash's broad areas of research interest are Cloud computing and Big data analytics.