Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020)
IEEE Conference Record # 48766; IEEE Xplore ISBN: 978-1-7281-5371-1

Ensuring Anomaly-Aware Security Model for Dynamic Cloud Environment using Transfer Learning

Gavini Sreelatha[1]
Research Scholar, Lincoln University College, Kaula Lampur, Malaysia.
Email id:gsreelatha@stanley.edu.in

Dr.A.Vinaya Babu[2]
Professor, Stanley College of Engineering and Technology for Women, Abids, Hyderabad, India
Email id:avb1222@gmail.com

,Dr.Divya Midhunchakkarvarthy[3]
Associate Professor, Lincoln University College, Kaula Lampur, Malaysia,
Email id:divya@lincoln.edu.in

Abstract— **Cloud concepts such as resource sharing, outsourcing, and multi-tenancy create significant challenges to the security community. Also, trusted third party and web technologies based cloud service provisioning arises new security threats in the cloud environment. Cloud security research still faces the shortcomings in improving the detection accuracy and detecting the new or unknown attacks in the cloud. Machine learning techniques play a significant role in automatically discovering the potential difference between legitimate and malicious data with high accuracy. Hence, it is essential to develop an intelligent security mechanism to learn, adapt, and detect the attacks or anomalies for the distributed and dynamic cloud environment. This work is about security solutions developed by a new mechanism called transfer learning techniques for the cloud environment. The transfer learning model leverages the detection of different types of known and unknown attacks by the utilization of the knowledge from the source domain. Rather learning about the attack from scratch, transfer learning focuses on the transfer of knowledge from source trained attacks to target attacks. This work gives a scope of detecting and solving new attacks on target to be only trained and fix it on the source and maintains qualitative performance.**

Keywords: **Cloud, Cloud Security, Vulnerabilities, Attacks, Machine Learning, Transfer learning.**

I. INTRODUCTION

Nowadays, Cloud computing has become a rapidly growing computational model in the Information Technology (IT) world. It offers on-demand resource access as flexible services to the users [1]. Cloud contains an enormous amount of computing and storage resources in a distributed environment in which there is a higher possibility of coexisting the tasks of a particular user with the tasks of other users due to the execution over a set of Virtual Machines (VMs). In consequence, security has become a significant concern for the cloud users and also, for the cloud service providers. Cloud computing often confronts with several top threats such as the insecure interface, abuse of cloud computing, data loss, service or account hijacking, and malicious insiders [2]. Recently, the increased variations and sophistication has led to serious concern on cyber-attacks such as zero-day attacks and Denial of Service (DoS) tactics, which poses significant threats to the different cloud-based environments especially, industrial, government, and military fields [3]. Cloud computing security must be done on two levels. One is on the provider level and another is on the user level. Cloud computing service provider should make sure that the server is well secured from all the external threats it may come across[4]. Hence, cloud security mechanism needs to focus on protecting the data compromised or misused by both the users in the cloud and the cloud service providers. The existing signature-based attack detection methods [5] lack to support the detection over the increased variety of the cyber-attacks.

With the target of improving the detection accuracy, the conventional security approaches [6, 7] have been applied the machine learning techniques among them transfer learning is used in cloud computing and network environments. Compared to the unsupervised learning methods, the label-based supervised models have accomplished higher accuracy. However, it requires a large number of labeled samples of malicious data [8]. Also, training data based supervised learning models fail to detect the unknown or new anomalies due to the behavior variations in the evolving attacks leads to the feature distributions among the attacks. Consequently, the domain-shift problem occurs that there is a need for retraining the learning model with new training data for the changes in the new attack behaviors [9]. However, acquiring adequate training samples is infeasible for continuously varying attack behaviors in the cloud environment. In essence, traditional machine learning approaches [10, 11] degrade the performance when there is a difference in the data distribution and feature space in the training and testing data.

To address these constraints, several kinds of research have employed the transfer learning methods [12] to improve detection or classification performance. Transfer learning improves the learner performance by transferring the training information from one domain to another related domain.

In recent years, a wide variety of applications and technologies that adopt cloud technology for the application of

processing and storage. The security-enhanced cloud system creates a significant impact on the different field of applications such as financial, medical, military, government, industries and so on. Moreover, time-efficient threat detection and attack classification models in the cloud greatly provoke the countermeasure activities and avoid its dissemination in the massive cloud environment. The security solutions in the cloud provide the potential benefits to both at the individual level and organizational level by avoiding the anonymous email checking and database hacking and crashing

II. RELATED WORK

Recently, the researches on providing security solutions have exclusively increased due to the growing number of insider and outsider attacks in the cloud. This literature briefly reviews the machine learning or data mining based security mechanisms and transfer learning method based security approaches in the cloud and network environments.

A. Machine learning-based security approaches:

A context-based anomaly detection framework [13] enhances the traditional monitoring service in a public or hybrid cloud by addressing the gap between the incidents in workflow layer and infrastructure layer and resolving the stemming constraint in a multi-tenant cloud environment. It estimates the security status of the monitored cloud system by applying the machine learning and complex event processing rules. However, it yields the false-positives at an increased rate in the anomaly detection process.

Supervised lEarning-Based Secure Information Classification (SEB-SIC) model [14] averts the security risk for either customers or financial service providers. It presents the Decision Tree-based Risk Prediction (DTRP) algorithm to classify the information and to predict the risks in the cloud environment. Although, it provokes the additional computation workload when there is an enormous amount of data. DDoS detection system [15] employs the C.4.5 decision tree algorithm along with the signature detection techniques to automatically and effectively detect the DDoS flooding attacks in the cloud environment. Even though the detection model accomplishes accurate and faster detection results by applying the machine learning algorithm in monitoring layer 3 and 4 in the OSI layer model, it requires vast training data.

By analyzing the statistical features of the DDoS attacks such as flooding, spoofing, and brute-force attacks using machine learning techniques, an approach [16] detects the attacks and alleviates the risks of the attacks from the source side of the cloud. It explores the information from both the virtual machines and hypervisors of the cloud server to ensure the scalability in the attack detection. However, labeling the attack types on the entire cloud data is a critical and time-consuming task. Anomaly detection and categorization approach [17] employs two different supervised learning methods such as Linear Regression (LR) and Random Forest (RF) in multi-cloud environments. Even though it accomplishes higher detection accuracy, it degrades the categorization performance due to the similarity among the attack traffic behaviors.

The cloud security approach [18] provides secure data storage, access, and retrieval from the hybrid cloud using the C4.5 decision tree algorithm along with the deduplication algorithm and dynamic access control mechanism. The deduplication and Dynamic Spatial Role-Based Access Control Algorithm ensures the secure storage and retrieval without redundancy and secure access using the decision tree-based user classification respectively. Even though it mitigates the security risks and restricts the data access based on the time and space in the hybrid cloud, it requires a robust cryptographic method to provide secure storage.

Ransomware detection approach [19] performs the volatile memory analysis in virtual machines using volatility framework and generates the descriptive meta-features using a machine learning algorithm. To detect the unknown ransomware in the virtual servers of the private cloud, it analyzes the state of the entire system such as processes, services, threads, privileges, kernel modules, mutexes, handles, DDLs, and callbacks. Despite, the distinct features in the benign and new malware files lead to inaccurate attack detection to the machine learning-based solution. Cloud-based Intelligent Security Technology (CIST) [20] utilizes both the unsupervised and supervised learning methods to improve the performance of network threat detection and classification. To handle both the voluminous data and unknown threat types in the cloud, it applies the hybrid learning method and thus, provides the customized security service to the cloud users. However, generating the labeled data from the analysis of the different normal traffic and attack traffic on the network is difficult due to the privacy and also, time constraints.

Trust-based access control approach [21] employs the machine learning technique to provide secure access to the user in the cloud with the computation of past behavior-based trust values. It predicts the trust values of the user and resources using machine learning techniques along with the consideration of potential parameters such as bogus request, user behavior, forbidden request, unauthorized request, and range specification. It provides faster results even when handling the massive amount of activity logs in the cloud.

B. Transfer learning based security approaches:

Feature-based transfer learning approach [22] learns the invariant features for the attack behaviors to detect the unknown attacks in the network environment. It trains the different attack types and its representation to detect the unseen variants of the new network attacks with the support of HeTL for different feature space. Time series anomaly detection model [23] employs the transfer learning method to transfer the labeled examples from the source domain to the unlabeled target domain based on the measurement of the unexpected and infrequent anomalies. It computes the similarity between the source and target domain using the dynamic time warping method and builds the nearest-neighbor classifier in the target domain to detect the anomalies. Cluster enhanced transfer learning (CeHTL) approach [24] computes the relation

between the known attacks and the new attack to automatically detect the unknown attacks by supporting the different feature space. It improves the detection accuracy for new attacks and ensures the robust detection performance in network environments. Time series anomaly detection approach [25] applies the transfer learning method that incorporates the Convolutional Neural Network (CNN)-based time series segmentation model. It performs the CNN-based anomaly detection to pre-train the massive amount of univariate time series data and fine-tunes the weights on univariate, multivariate, or small-scale data for the detection of unseen anomalies.

From the analysis of the literature review, its pros and cons it is determined that there is considerable attention from the research community in several areas towards cloud security solutions.

Selecting the appropriate security solutions is a difficult task for mitigating the identified vulnerabilities in the dynamic and complex cloud.

This work focus on reducing the training data and supporting the multiple domains is a challenging task of identifying the unknown attacks over the abundance of cloud storage and processing.

III.PROPOSED MODEL

The advent of cloud computing not only exacerbates the processing by both the individuals and organizations but also provides the opportunities to launch cyber crimes in the on-demand cloud environment. Cloud security mechanism has been widely used in many real-time systems due to the increased amount of security threats on most of the popular organizations. For instance, in 2011, the intruder launches Distributed Denial of Service (DDoS) attack on both Amazon.com and eBay. In 2013, a Chinese gang illegally utilized the famous storage service provider Dropbox to distribute the malware by performing the Advance Persistent Threat (APT). Furthermore, the anomaly-aware cloud security mechanisms substantially preserve the cloud environment from the emerging anomalies. Hence, the enhanced transfer learning provides the potential benefits to the cloud environment. The attack in the real time does not seem to be stable at all time. There is always a variation in the probability of the events happening in the network before and after the attacks [26] In essence, it enforces an effective as well as time-efficient detection of new attacks and classification of the existing attacks in the cloud environment

With the increasing utilization of cloud technology in most of the organizations, protecting the data by the organizations over the cloud environment is essential. Machine learning methods play a vital role in providing security solutions to the cloud environment. Several existing machine learning algorithms promote security solutions through its static verification of user behavior in the cloud. Although, the static nature and unpredictability of the new attacks on the conventional machine learning-based security mechanisms become ineffective in the dynamic cloud. Also, ensuring security over the new data with the trained model is a significantly challenging task. The transfer learning has become one of the viable solutions with the small amount of labeled data and higher detection accuracy to deal with this constraint in the emergence of cloud attacks.

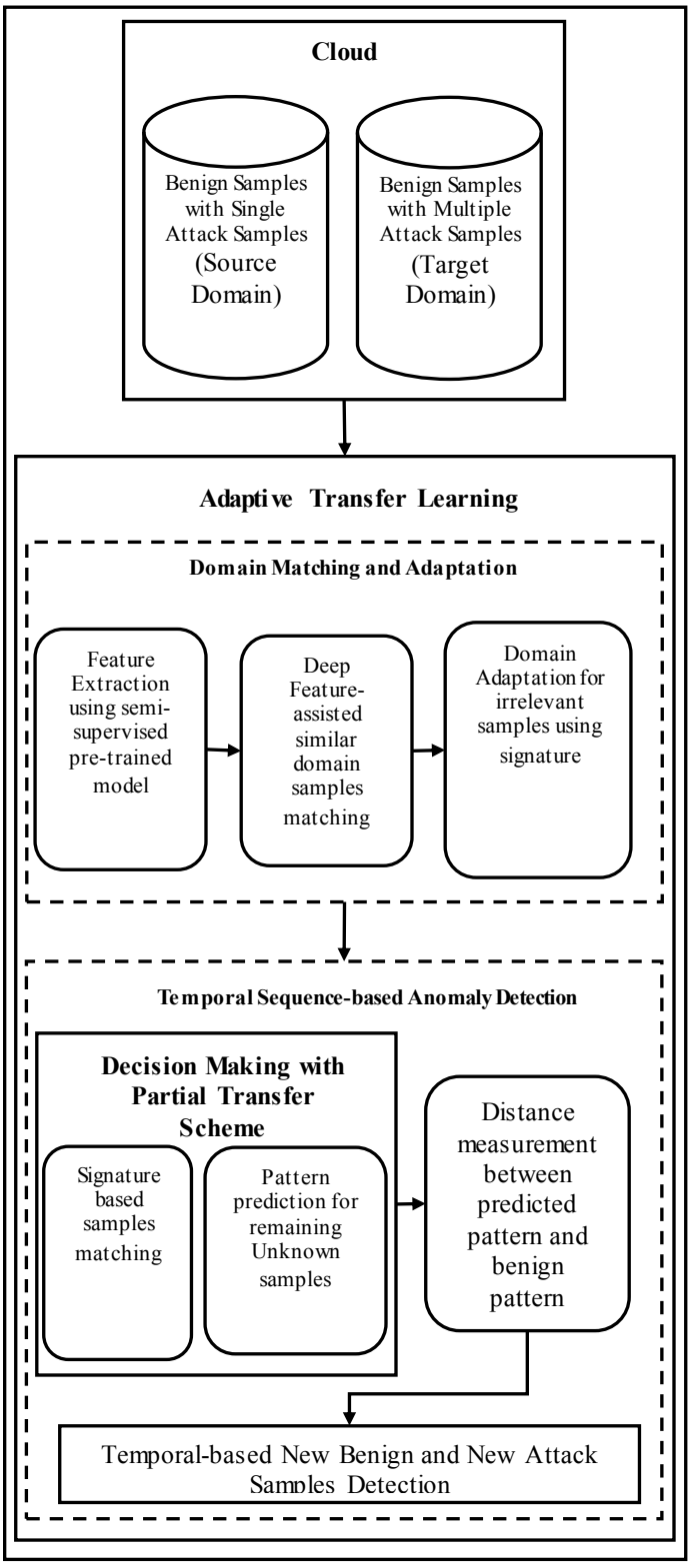

Figure 1: Overall Process of the Research Methodology

Figure 1 illustrates the proposed method involved in the cloud security model. This work to classify the attacks and detect the new attacks in the target domain using transfer learning that performs a classification with the help of minimum amount of training instances in the source domain

The adaptive transfer learning model includes two transfer schemes such as all transfer schemes for the inherently similar feature distributions and partial transfer scheme for inherently distant feature distributions. This work enhances the transfer learning model-based attack detection system to support the multiple types of attack classification and new attack detection. In essence, it models an adaptive transfer learning with the incorporation of semi-supervised learning-based pre-trained model for feature extraction, and domain matching and adapting based target classification. By applying the semi-supervised pre-trained model, the proposed approach preserves the sensitive data as well as handles the large-scale data in the cloud. After extracting the inherent features, it transfers the knowledge of the target domain to match the samples and assign the labels to the matched samples based on the feature distribution. In contrast, the target domain also comprises the samples that are irrelevant to the source domain. For this case, the proposed approach utilizes the signature information of the existing attacks and the deep features extracted from the source domain for the benign samples to decide the target samples.

In essence, this approach matches the signatures and features with the samples in the target domain and matches the attack types for the samples. After determining all the known samples, the proposed approach focuses on the unknown samples that are either new benign samples or new malicious samples. For the remaining unknown samples, it predicts the sequential patterns using the deep learning model and then, computes the distance between the predicted patterns and the existing benign patterns. In consequence, the proposed approach identifies the new benign samples and new malicious samples based on the distance between the distributions of the data through temporal sequence-based decision making. Thus, the transfer learning with help of deep learning model in the cloud security system effectively classifies the known attack types and detects the unknown attacks in the cloud.

Algorithm:

Input: Dataset of attacks given.
Output: Detect the unknown attacks in the cloud.
Create samples attacks from training data.
Match the samples from source domain and assign the labels to the matched samples with target domain.
If matches the signatures and features with the samples in the target domain and matches the attack types for the samples
Then, known samples are determined and handle various attacks.
The unknown samples uses the deep learning model are computed.
All new attacks create new training data in from target domain to source domain and trained to detect the various attacks and unknown attacks in the cloud.
End.

IV.RESULT ANALYSIS

The experimental framework conducts the experiments of the proposed model in any one of the open-source cloud environment. OpenNebula, Eucalyptus, Apache CloudStack, and OpenStack development environment are examples of the open-source clouds. Open source cloud computing software allows the system to create infrastructure, including storage, compute, and network resources. The experimental framework employs Python with Machine learning libraries to implement the proposed algorithm.

Performance metrics

The experimental framework employs four different performance metrics to illustrate the performance of the attack classification and unknown attack detection. Precision and recall show the attack classification performance where as detection accuracy and False alarm rate show the attack detection performance in the cloud.

S.No	Metrics	Performance
1	Precision (P)	It is the ratio between the number of correctly classified malicious samples (CCMS) and the number of classified malicious samples (MS). P=CCMS/MS
2	Recall (R)	It is the ratio between the number of correctly classified malicious samples (CCMS) and the total number of malicious samples (MS). R=CCMS/MS
3	Detection Accuracy (A)	It is the ratio between the number of accurately detected samples (DS) as unknown attacks and the total number of samples(S). A=DS/S
4	False Alarm Rate (F)	It is the ratio between the incorrectly detected (InD) or classified benign samples as attacks and the total number of samples. (S) F=InD/S

Table1: Performance metrics.

Based on the analysis details, this work has effective approach for detecting attacks with might reduce extra workload of repeated attacks and qualitative performance is achieved. This provides the security solution with higher accuracy through an adaptive learning, decision in the cloud. This work address the development of transfer learning techniques on unknown or new attack detection model for dynamic cloud environment

V.CONCLUSION

This work has described a brief introduction about the security threats, and vulnerabilities in the cloud environment and the significance of the attack detection reviewed the

previous research works and discussed the research gaps on the cloud attack classification and new attack detection. From the analysis of the existing works, it is determined that classifying the attacks that are inherently similar and detecting the new attacks is still a challenging task in the massive and heterogeneous nature of the cloud environment. Hence, this work has suggested the cloud security solution by classifying the existing attacks and detecting the new attacks with the assistance of the transfer learning model. It enhances the pre-trained model in the transfer learning and then, adaptively transfers the training information to the target task based on the inherent feature relationship between the source and target domain. Thus, using transfer learning techniques in cloud security model effectively recognizes both the new benign samples and malicious samples in the cloud along with the classification of known attacks. There is more to improve the results in future.

REFERENCES

[1] Varghese Blesson, and Rajkumar Buyya, "Next generation cloud computing: New trends and research directions", Future Generation Computer Systems, Vol.79, pp.849-861, 2018
[2] Singh Saurabh, Young-Sik Jeong, and Jong Hyuk Park, "A survey on cloud computing security: Issues, threats, and solutions", Journal of Network and Computer Applications, Vol.75, pp.200-222, 2016
[3] Juliadotter Nina Viktoria, and Kim-Kwang Raymond Choo, "Cloud attack and risk assessment taxonomy", IEEE Cloud Computing, Vol.2, No.1, pp.14-20, 2015
[4] Sudalai, Sridhar & Smys, S, "A Survey on Cloud Security Issues and Challenges with Possible Measures", International Conference on Inventive Research in Engineering and Technology (ICIRST 2016)
[5] Ahmed Mohiuddin, Abdun Naser Mahmood, and Jiankun Hu, "A survey of network anomaly detection techniques", Journal of Network and Computer Applications, Vol.60, pp.19-31, 2016
[6] Modi Chirag, Dhiren Patel, Bhavesh Borisaniya, Hiren Patel, Avi Patel, and Muttukrishnan Rajarajan, "A survey of intrusion detection techniques in cloud", Journal of network and computer applications, Vol.36, No.1, pp.42-57, 2013
[7] Berman Daniel S., Anna L. Buczak, Jeffrey S. Chavis, and Cherita L. Corbett, "A survey of deep learning methods for cyber security", Information, Vol.10, No.4, p.122, 2019
[8] Kumar Ram Shankar Siva, Andrew Wicker, and Matt Swann, "Practical machine learning for cloud intrusion detection: challenges and the way forward", ACM Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security, pp.81-90, 2017
[9] Garcia-Teodoro Pedro, Jesus Diaz-Verdejo, Gabriel Maciá-Fernández, and Enrique Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges", computers & security, Vol.28, No.1-2, pp.18-28, 2009
[10] Zamani Mahdi, and Mahnush Movahedi, "Machine learning techniques for intrusion detection", arXiv preprint arXiv:1312.2177, 2015
[11] Javadpour Amir, Sanaz Kazemi Abharian, and Guojun Wang, "Feature selection and intrusion detection in cloud environment based on machine learning algorithms", IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC), pp.1417-1421, 2017
[12] Pan Sinno Jialin, and Qiang Yang, "A survey on transfer learning", IEEE Transactions on knowledge and data engineering, Vol.22, No.10, pp.1345-1359, 2009
[13] Gander Matthias, Michael Felderer, Basel Katt, Adrian Tolbaru, Ruth Breu, and Alessandro Moschitti, "Anomaly detection in the cloud: Detecting security incidents via machine learning", In International Workshop on Eternal Systems, Springer, pp.103-116, 2012
[14] Gai Keke, Meikang Qiu, and Sam Adam Elnagdy, "Security-aware information classifications using supervised learning for cloud-based cyber risk management in financial big data", IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), pp.197-202, 2016
[15] Zekri Marwane, Said El Kafhali, Noureddine Aboutabit, and Youssef Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments", IEEE 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech), pp.1-7, 2017
[16] He Zecheng, Tianwei Zhang, and Ruby B. Lee, "Machine learning based DDoS attack detection from source side in cloud", IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), pp.114-120, 2017
[17] Salman Tara, Deval Bhamare, Aiman Erbad, Raj Jain, and Mohammed Samaka, "Machine learning for anomaly detection and categorization in multi-cloud environments", IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), pp.97-103, 2017
[18] Praveena D., and P. Rangarajan, "A machine learning application for reducing the security risks in hybrid cloud networks", Multimedia Tools and Applications, pp.1-13, 2018
[19] Cohen Aviad, and Nir Nissim, "Trusted detection of ransomware in a private cloud using machine learning methods leveraging meta-features from volatile memory", Expert Systems with Applications, Vol.102, pp.158-178, 2018
[20] Kim Hyunjoo, Jonghyun Kim, Youngsoo Kim, Ikkyun Kim, and Kuinam J. Kim, "Design of network threat detection and classification based on machine learning on cloud computing", Cluster Computing, Vol.22, No.1, pp.2341-2350, 2019
[21] Khilar Pabitr Mohan, Vijay Chaudhari, and Rakesh Ranjan Swain, "Trust-Based Access Control in Cloud Computing Using Machine Learning", Springer, In Cloud Computing for Geospatial Big Data Analytics, pp.55-79, 2019
[22] Zhao Juan, Sachin Shetty, and Jan Wei Pan, "Feature-based transfer learning for network security", In MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM), pp.17-22, 2017
[23] Vercruyssen Vincent, Wannes Meert, and Jesse Davis, "Transfer learning for time series anomaly detection", In CEUR Workshop Proceedings, Vol.1924, pp.27-37, 2017
[24] Zhao Juan, Sachin Shetty, Jan Wei Pan, Charles Kamhoua, and Kevin Kwiat, "Transfer learning for detecting unknown network attacks", EURASIP Journal on Information Security, Vol.2019, No.1, 2019
[25] Wen Tailai, and Roy Keyes, "Time Series Anomaly Detection Using Convolutional Neural Networks and Transfer Learning", arXiv preprint arXiv:1905.13628, 2019
[26] S. R. Mugunthan, "Soft Computing Based Autonomous Low Rate Ddos Attack Detection And Security For Cloud Computing", Journal of Soft Computing Paradigm (JSCP) (2019) Vol.01/ No. 02