EAI Endorsed Transactions
on Energy Web

Research Article EAI.EU

Contemporary PCA and NBA based Hybrid Cloud Intrusion Detection System

Sivakami Raja[1*], S. Gokul Pran[2], N. Pandeeswari[1], P. Kiruthiga[3], D. Nithya[3], G. MuthuPandi[4]

[1]PSNA College of Engineering and Technology, Silvarpatti, Tamilnadu.
[2]Veerammal Engineering College, Dindigul, Tamilnadu.
[3]Velalar College of Engineering and Technology, Erode, Tamilnadu.
[4]School of Engineering Presidency University, Bangalore.

Abstract

INTRODUCTION: Cloud computing offers on-demand services, from which consumers can avail networked storage and computer resources. Due to the fact that cloud is accessed through internet, its data are prone to internal and external intrusions. Cloud Intrusion Detection System will now be able to classify each pattern of testing dataset as either normal or intrusive and in case of intrusion, it will identify the type of intrusion. By comparing each of these actual results with the expected results of testing dataset. It is strongly observing the inside-activities of a network. Hence, it is suitable for detecting intrusions in cloud environment.
OBJECTIVES: Hybrid Cloud Intrusion Detection System can function well for a very huge dataset and it can also detect unknown attacks. To achieve the better performance in the cloud setting by utilizing this Cloud Intrusion Detection System. METHODS: To overcome performance issues, Principal Component Analysis and Network Behaviour Analysis are proposed.
RESULTS: The experimental and performance assessment show that the proposed model is well planned, efficient and effective in finding cloud environment intrusions. An Intrusion Detection System (IDS) monitors all incoming and outgoing network activity to identifies any signs of intrusion in your system that could compromise your systems.
CONCLUSION: Experiments are performed using a standard benchmark KDD-cup dataset and the findings are addressed. IDS helps the Network Administrator to track down bad guys on the Internet whose very purpose is to bring your network to a breach point and make it vulnerable to attacks.

Keywords: Cloud computing, intrusion detection, principal component analysis, network behaviour analysis, genetic algorithm.

Received on 27 January 2021, accepted on 07 February 2021, published on 19 February 2021

doi: 10.4108/eai.19-2-2021.168727

Corresponding author. Email: rsivakami@psnacet.edu.in

1. Introduction

In an era of internet-based technology, cloud computing has become a central point of attraction. Due to the convenience and efficiency offered by cloud, individuals and organizations are interested in using the storage and software services of clouds. However, the security threats of remote cloud data centres have not yet been clearly identified. Therefore, it is often regarded as a difficult challenge to establish an effective technique to achieve cloud protection with optimum accuracy. This challenge contributes to the need to invent an awareness that can retain different cloud provisions free from security threats. For the need of satisfying security requirements of cloud computing intrusion detection system (CIDS) has been developed by adopting a combination of soft computing approaches. The cloud intrusion detection dataset is separated into dual sets called training dataset and testing dataset in the technique of cloud intrusion detection. The training dataset is used to teach the CIDS about the patterns of intrusive and normal behaviours of users. Once the training is complete, the Cloud Intrusion Detection System will be able to identify any unknown

activity as either normal or intrusive. This ability of Cloud Intrusion Detection System is measured by allowing the testing dataset to flow through the trained Cloud Intrusion Detection System. By using the knowledge which is acquired at the training stage, Cloud Intrusion Detection System will now be able to classify each pattern of testing dataset as either normal or intrusive and in case of intrusion, it will identify the type of intrusion. By comparing each of these actual results with the expected results of testing dataset, the mean squared error (MSE) is measured by which the accuracy of Cloud Intrusion Detection System is assessed.

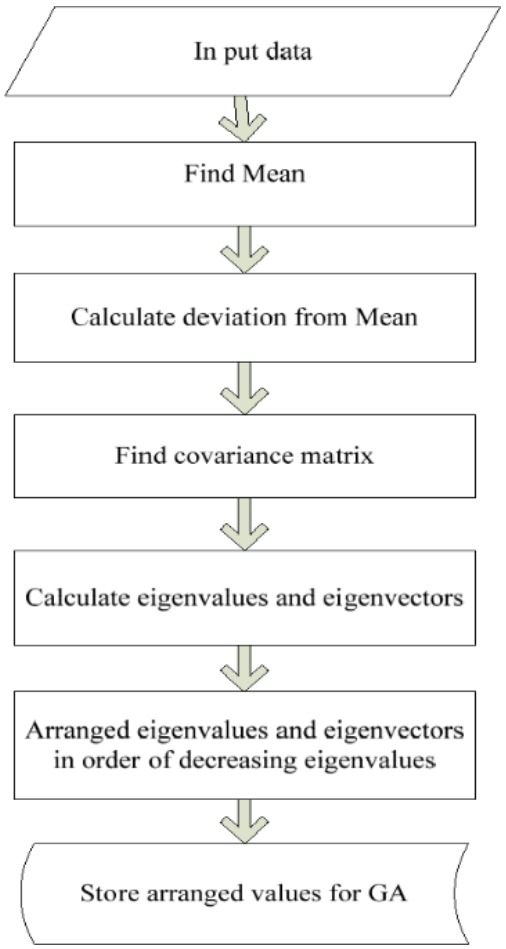


Figure 1. PCA based approach for feature selection

Several existing intrusion detection systems, namely User-to-Root (U2R) and Remote-to-Local (R2L) attacks, are weak in detecting low frequency attacks. Studies prove that they can be detected by employing a hybrid approach. Hence, a hybrid system using Principal Component Analysis (PCA), Network Behaviour Analysis (NBA) and Genetic Algorithm (GA) has been adopted for achieving high efficiency in detecting both high-frequent and low-frequent attacks.

Techniques for intrusion detection are split into techniques that are based on signatures and anomalies. Modern intrusion detection techniques, common datasets used for IDS performance assessments and the evasion methodologies adopted by intruders are presented in [1]. Cloud architecture, intrusions in the cloud, challenges, cloud intrusion detection techniques, types are analysed in [2]. Further, a comparison between numbers of Cloud Intrusion Detection System is illustrated in terms of their type, position, success rate, execution time and dataset. A hybrid intrusion detection technique is suggested in [3], which employs both signature-based and anomaly-based techniques. This framework uses distributed approach for embedding intrusion detectors within a cloud environment and centralized approach for generating detection alarms.

More advanced intrusion detection techniques are portrayed in [4]. Soft computation methods using neuro-fuzzy structures of type-1, type-2 and interval type-2 are employed in [5] to detect intrusions in a cloud environment. Extra Trees feature classifier-based intrusion detection is recommended in [6]. This process is formulated to work in three stages for detecting several attacks. A Cloud Intrusion Detection System with reference to smart mobile systems and mobile clouds is put forward in [7]. This system further addresses the security requirements in connection to the communication between cloud services and smart devices. To integrate intrusion detection techniques within cloud, fuzzy neural network (FNN) built genetic algorithm approach is conferred in [8]. This system is able to build up knowledge of fuzzy rules as of dataset to classify invasions in a cloud atmosphere. Depsky, Secret Sharing and Key Aggregate Cryptosystems algorithms have been proposed in [9] for cloud security. In [11] new hybris intrusion detection system has been introduced with one class support vector machine.

2. Proposed Intrusion Detection System

Each pattern is of 41 variables which are used for defining 24 types of attacks and normal accesses. Besides this, the size of dataset is 744MB, which is also large. Hence the usage of all 41 variables with the full dataset will result in a complicated IDS model which can experience overheads during online intrusion detection.

Though an enormous amount of intrusion detection approaches are available, performance is still an issue. It becomes vital to achieve satisfactory success rate in the field of intrusion detection. As feature selection plays a key role in reaching this goal, Principal Component Analysis is found to provide a straightforward, universal, and prevailing framework for selecting good subsets of features, which can lead to improved detection rates.

Towards this goal, the Principal Component Analysis space is searched using Genetic Algorithm for selecting a subset of principal components. This methodology is presented in figures 1 and 2. Traditionally, a subset of top principal components is selected to symbolize the whole set. Since this method is independent of classification technique, errors occur. Hence, the proposed framework is tested on the cloud

intrusion detection framework to demonstrate the significant improvement in performance analysis. PCA results in 22 parameters.

The standards of these 22 Knowledge Discovery in Databases (KDD) dataset variables are initially normalized within the range [0, 1]. As genetic algorithm is found to be suitable for this purpose, it has been adopted to eliminate the redundant and least significant variables. The application of GA into PCA optimizes feature set selection identifies 12 features, via dropping the number of important structures, thereby eliminating the least significant features. The identified parameters are Service, src_bytes, dst_bytes, logged_in, Count, srv_count, serror_rate, srv_rerror_rate, srv_diff_host_rate, dst_host_count, dst_host_srv_count and dst_host_diff_ srv_ rate This approach boosts the detection rate.

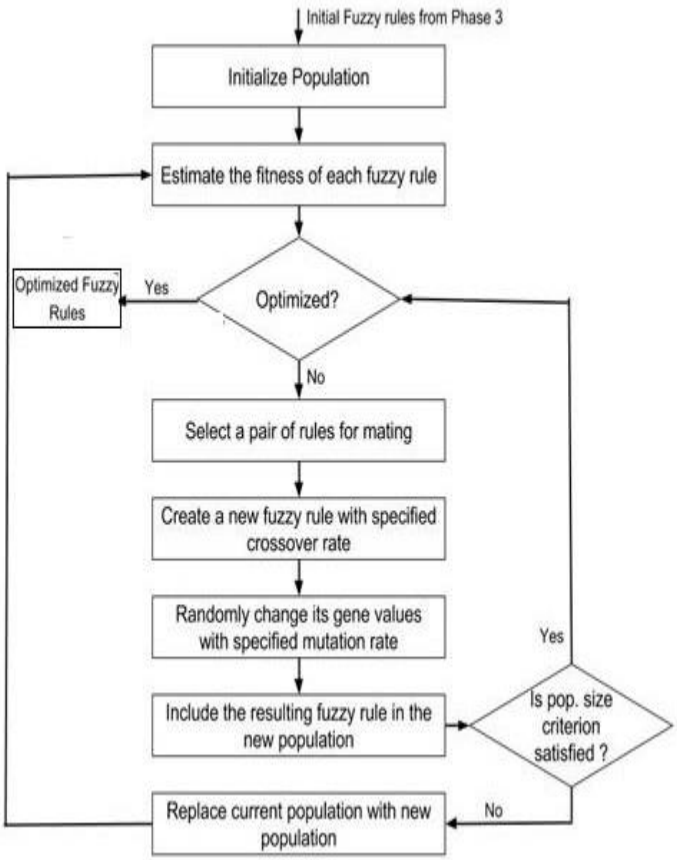


Figure 2. GA based approach for feature set optimization

Towards this goal, the Principal Component Analysis space is searched using Genetic Algorithm for selecting a subset of principal components. This methodology is presented in figures 1 and 2. Traditionally, a subset of top principal components is selected to symbolize the whole set. Since this method is independent of classification technique, errors occur. Hence, the proposed framework is tested on the cloud intrusion detection framework to demonstrate the significant improvement in performance analysis. Principal Component Analysis results in 22 parameters.

The standards of these 22 KDD dataset variables are initially normalized within the range [0, 1]. As genetic algorithm is found to be suitable for this purpose, it has been adopted to eliminate the redundant and least significant variables. The application of Genetic Algorithm into Principal Component Analysis optimizes feature set selection identifies 12 features, via dropping the number of important structures, thereby eliminating the least significant features. The identified parameters are Service, src_bytes, dst_bytes, logged_in, Count, srv_count, serror_rate, srv_rerror_rate, srv_diff_host_rate, dst_host_count, dst_host_srv_count and dst_host_diff_ srv_ rate This approach boosts the detection rate.

Network behaviour analysis (NBA) strengthens the security of an underlying system by closely tracking system traffic and identifying any abnormal action or deviation from usual process. As Network Behaviour Analysis accumulates data from every node of an underlying environment, it is strongly observing the inside-activities of a network. Hence, it is suitable for detecting intrusions in cloud environment.

3. Experimental Results

Initially, the set of cloud services are created using network behaviour analysis as shown in figure 3. Additionally, consumers are also established out of which some are designated as intruders as depicted in figure 4. A cloud intrusion detection system is embedded with the proposed framework to monitor all activities happening within the cloud services. It classifies each activity as either intrusive or non-intrusive. As illustrated in figure 5, the Cloud Intrusion Detection System successfully classifies intruding consumers and they are eliminated from service grants.

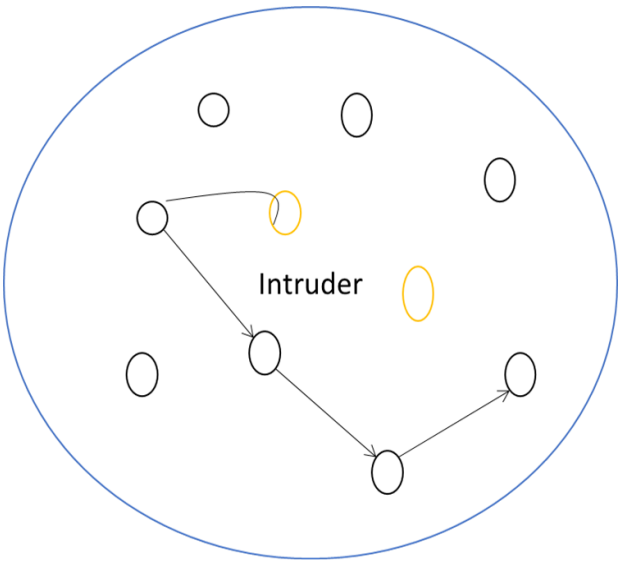


Figure 3. Cloud service providers and consumers

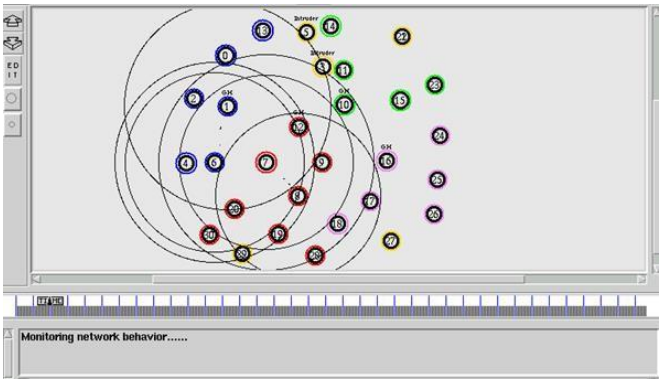


Figure 4. Cloud service providers and consumers during runtime

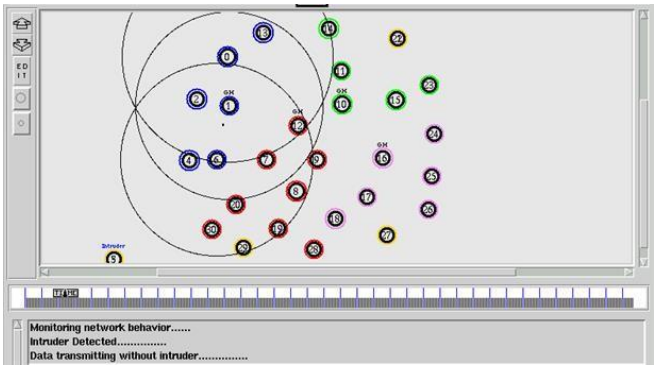


Figure 5. Intrusion identification

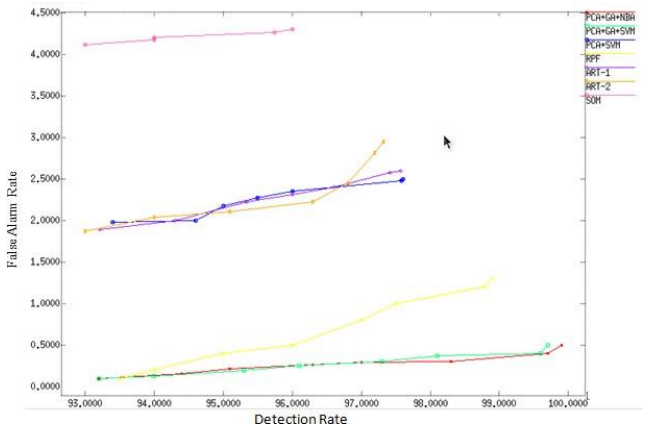
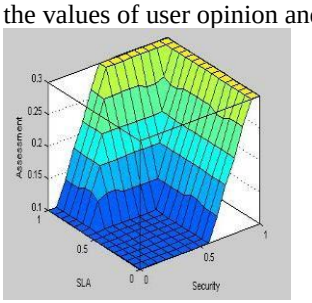


Figure 6. Performance analysis

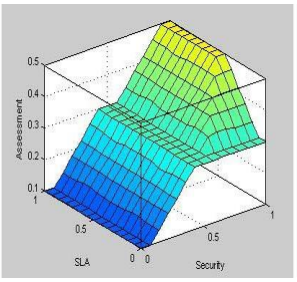
Figure 6 demonstrates the comparison between the planned framework and other approaches in terms of false alarm rate and intrusion detection proportion. The proposed approach achieves a reduced false alarm rate.

A framework for evaluating the confidence of cloud facility benefactors is suggested in. This framework manipulates four parameters namely, SLA, security, performance and opinion of the user. The 3D depictions of our outcomes show the development of trust index in relation to the input limitations Service Level Agreements (SLA), security, performance and opinion of the user are secured.

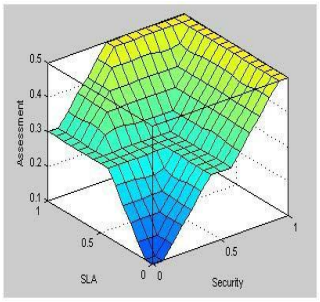
Figure 7 determines the evolution of trust assessment for different performance values with reference to negative user opinion. From Figure 7, we infer that, the maximum trust what we gain is 0.5 irrespective of the value of user opinion. For poor performance, the trust index (0.3) is in distrust region, as shown in Figure 7a. Figure 7c shows that, even for good performance value, the highest trust index is only 0.5. But it exceeds 0.3 for medium level values of SLA and security. In addition, it is steadily increasing from 0 with respect to the increase in SLA and security parameters. On the other side, for medium performance, in spite of getting 0.5 as a maximum trust index, the increase is found to be uneven. Figures 8 and 9 demonstrate the progress of trust, for neutral and positive user opinions which illustrates the relationship between the values of user opinion and trust.



a. Performance = poor

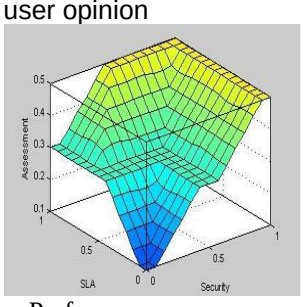


b. Performance = medium

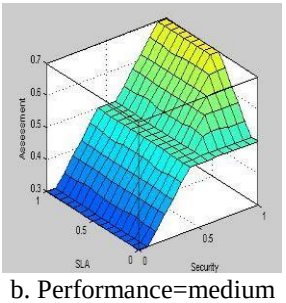


c. Performance = good

Figure 7. Evolution of trust with respect to negative user opinion



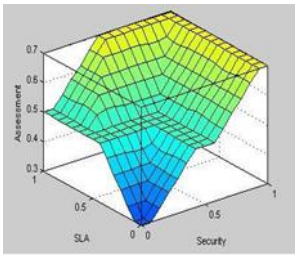
a. Performance = poor



b. Performance=medium

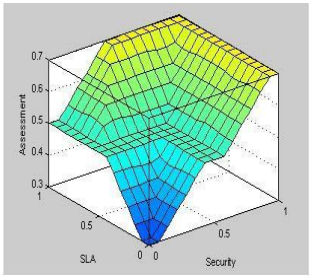
Fig. 8 and Fig. 9 show the progression of trust evaluation with respect to SLA and Performance, where customers have given neutral and positive opinions, respectively. Even for zero membership value of security, their trust value is uniformly increased by 0.2 in Fig. 8 for neutral opinion and by 0.4 in Fig. 9 for positive opinion. This shows the significance of user opinion in the process of trust evaluation. This provides

justification for the reason why we give major importance in measuring the believability of customers who provide opinion or feedback about the efficient service provision of cloud service providers.

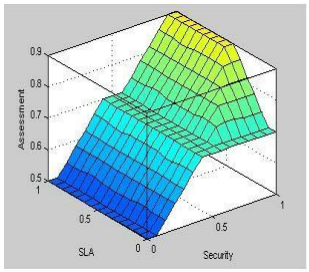


c. Performance = good

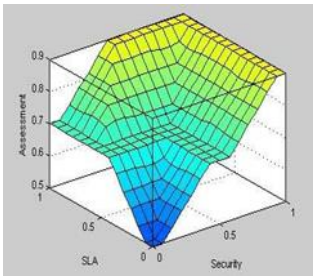
Figure 8. Evolution of trust with respect to neutral user opinion



a. Performance = poor



b. Performance = medium



c. Performance = good

Figure 9. Evolution of trust with respect to positive user opinion

4. Conclusion

An intrusion detection scheme is proposed for secure cloud setting. It is designed to detect malicious access on the cloud virtual network with PCA and GA. This approach employs Principal Component Analysis for feature set extraction and genetic algorithm for its optimization. Network Behaviour Analysis is employed for establishing cloud service providers and cloud service consumers. The results of the experimental and performance assessment show that the proposed model is well planned, efficient and effective in finding cloud environment intrusions. From the results, it is understood that this hybrid CIDS can function well for a very huge dataset and it can also detect unknown attacks. It is also possible to achieve the better performance in the cloud setting by utilizing this CIDS.

References

[1] Alkadi, O., Moustafa, N. and Turnbull, B., 2020. A Review of Intrusion Detection and Block chain Applications in the Cloud: Approaches, Challenges and Solutions. *IEEE Access*, *8*, pp.104893-104917.

[2] Prabavathy, S., Sundarakantham, K. and Shalinie, S.M., 2018. Design of cognitive fog computing for intrusion detection in Internet of Things. *Journal of Communications and Networks*, *20*(3), pp.291-298.

[3] Loukas, G., Vuong, T., Heartfield, R., Sakellari, G., Yoon, Y. and Gan, D., 2017. Cloud-based cyber- physical intrusion detection for vehicles using deep learning. *IEEE Access*, *6*, pp.3491-3508.

[4]Sadaf, K. and Sultana, J., 2020. Intrusion detection based on auto encoder and isolation Forest in fog computing. *IEEE Access*, *8*, pp.167059-167068.

[5] Mishra, P., Varadharajan, V., Pilli, E.S. and Tupakula, U., 2018. Vmguard: A vmi-based security architecture for intrusion detection in cloud environment. *IEEE Transactions on Cloud Computing*, *8*(3), pp.957-971.

[6] Zhang, Z., Wen, J., Zhang, J., Cai, X. and Xie, L., 2020. A many objective-based feature selection model for anomaly detection in cloud environment. *IEEE Access*, *8*, pp.60218-60231.

[7] Gao, Y., Liu, Y., Jin, Y., Chen, J. and Wu, H., 2018. A novel semi-supervised learning approach for network intrusion detection on cloud-based robotic system. *IEEE Access*, *6*, pp.50927-50938.

[8] Chkirbene, Z., Erbad, A., Hamila, R., Mohamed, A., Guizani, M. and Hamdi, M., 2020. TIDCS: A dynamic intrusion detection and classification system based feature selection. *IEEE Access*, *8*, pp.95864-95877.

[9] Alkadi, O., Moustafa, N. and Turnbull, B., 2020. A Review of Intrusion Detection and Blockchain Applications in the Cloud: Approaches, Challenges and Solutions. *IEEE Access*, *8*, pp.104893-104917.

[10] Varadharajan, V. and Tupakula, U., 2016. On the design and implementation of an integrated security architecture for cloud with improved resilience. *IEEE Transactions on Cloud Computing*, *5*(3), pp.375-389.

[11] Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J.; Alazab, A. Hybrid Intrusion Detection System Based on the Stacking Ensemble of C5 Decision Tree Classifier and One Class Support Vector Machine. *Electronics* **2020**, *9*, 173.