RESEARCH PAPER | Computer Science | Volume : 4 | Issue : 4 | Apr 2014 | ISSN - 2249-555X

Multi-Level Security System for Anomaly Detection in Cloud Based Data

KEYWORDS: Data Mining, Cloud, Intrusion Detection System, Anomaly Detection, Multi-Level Security

J.Jabez
Research Scholar, Sathyabama University, Chennai, Tamil Nadu, India.

Dr.G.S.Anandha Mala
Professor & Head, Department of CSE, St. Joseph's College of Engineering, Chennai,Tamil Nadu ,India.

ABSTRACT *In recent years, the Data mining is an essential technique, audit the data by themselves and mine it, and also can be used for intrusion detection. Statistics [Mathematical Analization], ANN- Artificial Neural Network, HMM-Hidden Markov Model, and SVM – Support Vector Machine] are some of the main Data Mining (DM) techniques often used for anomaly and misbehavior detection and these techniques don't have a proper scientific methodology to increase the efficiency of the intrusion detection. Thiswork is a novel approach named MLSS [Multi Level Security System] is proposed for detecting the intruders in the cloud accurately and fast in each level of the Cloud Model. The Cloud Model is defined as Cloud Server, Cloud Engine, Cloud Storage, Cloud Platform, Cloud Software and Cloud Tester. The anomaly can be detected in each model of the cloud by configuring the Software as well as deploying software in the cloud. The experimental result shows the accuracy and efficiency of the MLSS where it is developed in VS2010 Professional Edition integrated with Visual Guard Admin Console.*

Introduction

Information technology has become a key constituent to sustenance critical structure services in various areas of our society. In a determination to share information and rationalize processes administrations are creating multifaceted networked systems that doing as an open network to consumers, contractors, and other professional partners. While most workers of these networks are genuine users an open network disclosures the network to illegitimate entree and use. Augmented network difficulty greater admission and a rising stress on the internet have made network safety a major concern for administrations worldwide.

Safety is an important problem for all networks in businesses and organizations at the current time and all the interruptions are trying in dissimilar ways that can efficaciously access to the data network of those companies, Web services and notwithstanding the progress of manifold ways, to promise that the infiltration of interruption to the society of the grid via the Internet, over the use of firewalls, encryption, etc. But IDS is a comparatively new technique for intrusion discovery methods that have emerged in recent years. The main role of Intrusion Detection System's in a network is to help computer to make and agreement with the network bouts.

Related Works

Rashmi [27] examined an insightful examination of the current status on cloud safety problems based on thorough survey.The author also makes an attempt to describe the security challenges in Software as a Service **(SAAS)** model of cloud computing and also endeavors to provide future security research directions. One of the methods introduced in backup level SAAS [1], is Agentless Method for data Backup andRecovery in cloud model. There are currently a large number of standard bodies with different interests, e.g. IEEE Cloud Computing Standard Study Group [13], ITU Cloud Computing Focus Group [15], Cloud Security Alliance (CSA), Distributed Management Task Force [10], Storage Networking Industry Association [29OpenGrid Forum [25], Open Cloud Consortium [24] and OrganizationStructured InformationStandards [23] and so forth. To promote the wide use of cloud computing, it is necessary toestablish common standards.

In caseof intrusion detection in data mining, there are variousMining methods that have been smeared to intrusion detection because it has theadvantage ofdiscovering useful knowledge that describes a user's or program's behavior from large auditdatasets.Data mining has been used extensible for anomaly detection [18,22]. Statistics [3,9],Artificial Neural Network [5,21] and HMM-Hidden Markov Model [7],Rule Learning [17], Outlier Detection Scheme [28], SupportVector Machines [2], Neuro Fuzzy computing [30],Multivariate Adaptive Regression Splines [33] and Linear Genetic Programming [6] are some of the main data mining techniques widely used for anomaly and misuse detections.None of the above works have proposed a rigorous andscientific approach for increasing and improving the efficiencyof intrusion detection. Our proposed data mining approach results in faster and more accurate intrusion detection.

Very few studies have examined the relative performance of various methods in the context of intrusion detection [19]. The present study addresses this research gap.The primary objective of this proposed paper is to compare the detection accuracy of two data mining methods – MLPand RBF. Comparing data mining methods will provide us with significant insights into selecting the appropriatemodel for detecting intrusions.This research therefore will have a tremendousimpact on electronic commerce and information security embedded models, joined models,meaning of the terms used, are developed to increase the prediction accuracy using variousmachine-supervised learning methods together.Few studies on hybrid or combined models aremade sequentially. For example, Coenenet at. [6] proposed a method used for increasing the rateof direct mailing system. Li an Wang [19] represents a procedure which improves theeffectiveness of the last classification using ANN and set theory. The set theory wasdemonstrated by Pawlak [26]. Wong et al. [34] approved a machine learning function forconditional impedance testing and the information searching phase. The dependency analysisused to reduce the size of the space used for searching in conditional impedance method. A submethod for various main methods is developed by Chen [4], is a hybrid function for text miningby fuzzy theory and the SOM based fuzzy theory is also used by Chen. Versace et al. [32]introduced a novel joining ANN with GA. MC Clean [20] proposed a study about the predictionof probability in the enterprise based mining combined with multivariable statistical with ANNtechnique. Suh et al. [31] recommended and associated numerous ways to chain the classifiersproduced by RFMFrequency and Monetary), logistic reversion, and neural

Volume : 4 | Issue : 4 | Apr 2014 | ISSN - 2249-555X

networks.

Conversano et al. [8] projected a concoction model to increase performance by joining theparameters inferred from multiple arithmetic methods. In supplement to hybrid methods [14] thathave strained to syndicate two entirely dissimilar approaches, hybrid that use one technique innumerous ways have also been deliberate.Hansen and Salaman [11] presented that thesimplification aptitude of a neuralsystem can be considerably enriched through collaborative anamount of neural networks. Indhukhya and Weiss [14] perceived the development of gain valuesof the final nodes in decision trees by manifold re-sampling of decision tree initiation approachesand their grouping using the electivemethod. Kuncheva et al. [16] focused on belongings inwhich prediction correctness was enhanced using hybrid replicas with groupings of networks,and logistic regression models. The presentation of hybrid methods presented development,when the association is low among hybrid replicas is inSuh et al. [31]. Zhang and Zhang [35],describes that a sole data mining method has not been demonstrated suitable for every Instead,numerous techniques may essential to be joined into hybrid schemes that can be used obliginglythrough a specific data mining process,

Symbolization
Some of the symbol used in this paper and its description is given for understanding the paper as good as:

Symbol	Description
ANN	Artificial Neural Network
HMM	Hidden Marko Model
NN	Neural Netwok
DM	Data Mining
CSP	Cloud Service Provider

Existing Methods
The statistical examinations were usedfor data warehousing with minor and intermediate size data.Computing has continuously been tostatistics and it continued so even at times when measured rigorousity was most extreme valuedquality as a data investigation tool. Calculations were originated to be numbers. Statisticalexamination of collective text or only text based data mining presented significantly poor.

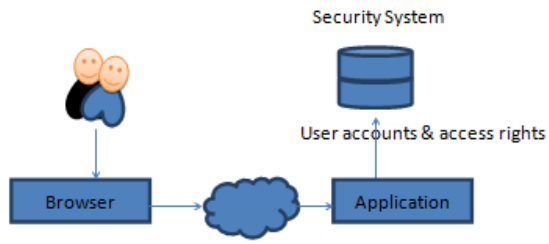
ANN method**,** new network is then exposed to the development of "training". In that phase, neurons apply an iterative procedure to the amount of inputs to regulate the masses of the network in order to enhance the sample data on which the "training" is achieved. After the phase of knowledge from anpresent data set, the new network is ready and it can then be used to makeguesses. We can use the ANN establishedDM forenhancing the training and testing data with manifoldrepetitions. This method is worried only with real-worlddeliberations, that is, with the extrapolative validity of the answer and its functional relevance and not with the nature of the fundamental mechanism or its significance "theory" of the under lying marvels. Also ANN necessitates more time and more examples training phase and it has smallersuppleness.

HMM cannot detect totally unknown attacks. It requires more number of database scan to generate rules. HMM can be used only for detection purpose. It require more graph based, tree based and sub divisions for its detection purpose.

Support vector machine can correctly classify intrusions even if limited sample data are given. It can handle massive number of features. It can classify only discrete features. So, preprocessing of those features is required.

ProblemStatement
This paper tackles a problem of anomaly detection in DM and provides a solution . The solutions can be obtained by applying level by level security in the overall process. The security is divided into four levels where the level by level security when combined and provides a high security for anomaly.

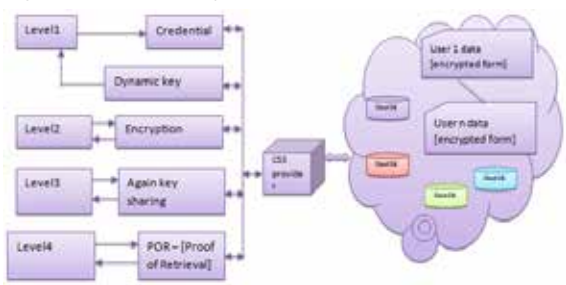
Figure-1: Where to provide Security in Cloud



The security of the complete cloud model is given in the security system which can be a part of the cloud or any third party software deployed on the cloud is depicted in Figure-1.

SystemModel

Figure-2: Proposed System Model



Level1 – user Level Security / **Level2** – Data Level Security / **Level3** – Sharing Level security / **Level4** – Maintenance level security

The system Model says the overall functionality of the proposed system in four stages. This level based security is explained in the section-3.4 in detail.

Proposed Approach
The proposed approach contains FOUR levels of security as:

The first level is the User Level Security [ULS], where secured Authentication and Authorization can be done to find out how the consumer will grip entree into the cloud. Comprising the confirmation of user identifications, strong-minded level of entree and strong-minded place of access is considered.

The second level is the Data Level Security [DLS], where safety construction of the method is deliberated by exhausting AES cipher block binding, which removes the deception that happens today with pinched data. There is no hazard of any data directed within the scheme being interrupted, and replaced.

The third level is Shared Level Security [SLS], where the data of an owner can share documents to their recognized people by distributing key information.

Finally level is the Maintenance Level Security [MLS] is smeared by submitting a Proof of Retrievability Model [POR] and it be contingent on the client proposal system upkeep

Volume : 4 | Issue : 4 | Apr 2014 | ISSN - 2249-555X

and upgrades.

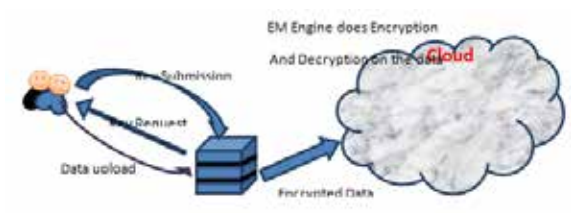
Level-1
When an inter user wants to become a cloud Member, he or she should submit a request and receive a credential application form. Once they answer the credential questions and submit the answer sheet, the cloud engine verify the answers and provide a dynamic key for their sign in activity and furthermore they should do login using that dynamic key as the password. The key may be in any form, of combination of numbers, letters and some of the possible keying in special characters.

Figure-3: Web User Becomes Cloud Member

Once receive the password, the web user can become a cloud user and the person can do all kind of activities like browsing, listening and watching movies etc. If the individual likes to go further to access the IAAS (Infrastructure as a Service), PAAS (Platform as a Service) and SAAS,then he is supposed to submit some more documents and answer the next level credentials and submit a photo for high level security. In the credential the user should make a payment for accessing and using their need in any of the IAAS, PAAS or SAAS. Based on thesize usagetime the payment varies and after the payment in online, they get their name registered with protected infrastructure or protected platform of software according to their request .Once the cloud model is received, the user can upload or download necessary data whenever they need.

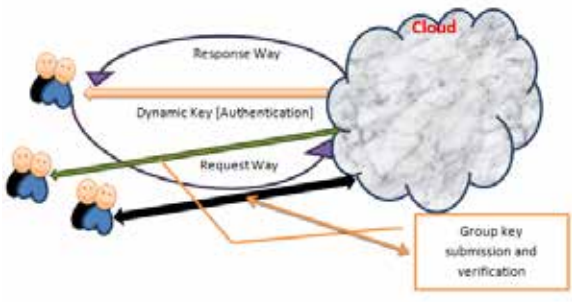
Level-2
In the second level the security is analyzed and provided for the data of the cloud user, and to avoid un-trustiness idea about the cloud should be able to convince and ensure them that their data is highly secured in the cloud than their house bureau. Once the data starts getting uploaded, the data from the user is encrypted by an Encryption Mechanism – [EM] which is deployed in the cloud. The algorithm used in the EM of the cloud model depends on the cloud owner. While the encryption is done by the EM, it gives a change to the user to provide the encryption key. If the user gives the key and do encryption then they can do decryption using the same key by them. If the user is not providing the key, the cloud EM takes the user photo name as the key and encrypts the data, uploads into their occupied place in the cloud (figure-4).

Figure-4: Data Level Security by EM



Level-3
In the third level, if any cloud user likes to share their data with their friends or with their official staff members, the cloud provides a group key for sharing their data, where the group key has its own format, the member ID is prefixed or suffixed with the dynamic key and it counts the number of users sharing the data. Also it maintains the history and Metadata about the user and the data like IP, time, password etc is depicted in Figure-5.

Figure-5: Shared Level Security by EM



Level-4
Maintenance level security is the final security for the location or the space allocated to the particular user is locked by a key. Even though the user passed all the levels, he is supposed to prove them for retrieving the datawhich is known as Proof of Retrieval and is clearly depicted in figure-6

Figure-6: Maintenance Level Security using Poof of Retrieval

Level by Level Algorithm

- *Cloud C is constructed*
- *An User IU is a internet user$_i$ gets pwd$_i$ = [dynamicPwd$_i$]*
- *if loginSuccess == true then*
- *internetUser = cloudUserie. IU = CU* // ULS
- else
- *reLogin*
- end
- *if [CU.payment == success]then check size, period of IAAS*
- *allocateSpace[IAAS].name = CU.name*
- else
- *completePayment*
- end
- *CUser can upload and download data in their allocated IAAS space*
- *cloud.cryptmechanism ← IU.upload(data)* // DLS
- *if cu.download == true then*
- *answer for Q1, Q2, Q3.. Qn ∀ Q_i ∈ set of credential Question for POR*
- End
- *if [Q_1.answer &&Q_2.answer &&Q_3.answer && …… Q_n.answer == correct then*
- *CU.download = ture*
- *CU.data ← POR(CU.data.download)* // POR
- End
- *if userGroup1.Key == groupUseri.pwd then*
- *groupUseri.valid = true*
- end

Volume : 4 | Issue : 4 | Apr 2014 | ISSN - 2249-555X

The ULS, DLS, SLS and the MLS are given in an algorithm form for evaluating the proposed approach using any computer language. This algorithm can be implemented in any language and it can be verified in any number of systems for performance evaluation. Where the IU is the Internet User, after registration, IU receives the pwd, and becomes CU, the Cloud User, and CU should answer for all the Q, the credential questions and make payment for getting cloud services. After received the permission to cloud service, CU can upload and/or download their data into their specific Infrastructure. When the CU wants to download the data they should meet out and get validation from POR is clearly shown in the above algorithm.

Experiments and Results

Figure-7
[a]: New User [b]. Existing User. C]. Dynamic Key Received[d]. Upload / Download options

A Normal web user can become a member of a cloud by register into the cloud-server by filling the credential forms given by the CSP. After the registration they can login to the cloud and apply the basic browsing and searching processes. Once the web user converts into a cloud user they can utilize the **IAAS, SAAS, and PAAS** of the cloud by registering to the second level with the payment. After the user became the cloud user, for further resource utilization they should make the payment level by level as showed in Figure-7.

Figure-7 [a]: Uploading a file [b]. Downloading with POR [c]. Success Download

After the success payment, the CSP offer a dynamic pin number as a password and then the user can upload their files into the IAASbaptized in their [user names] names with associated passwords. If the user stabs to open other user folder then the POR matches the user login info with the file [folder] ID and disavowals (Figure-7).

The performance evaluation can be implemented nearly with 50, 100 number of users and they become cloud user with the clouds space allocated in their names. In that unknowingly there are 12, 5 customersstrained to open the other user's data folder in the existing as well as in proposed respectively. Out of 50 users, there were 12 users who tried to open other's data in existing and 5 users in proposed. They are verified by the POR which documented their IP. If a user attemptsextra than three times their IP address will get congested and he will get a castoffmemo from the POR of the cloud, they are named as vulnerable users. Figure-8shows that the performance comparison among the existing and proposed systems with the total number of users and the number of vulnerable users respectively. It is clearly concluded that the proposed approach reduces the vulnerability from24% into 10% by applying the level by level security for detecting and preventing the Anomaly in Cloud.

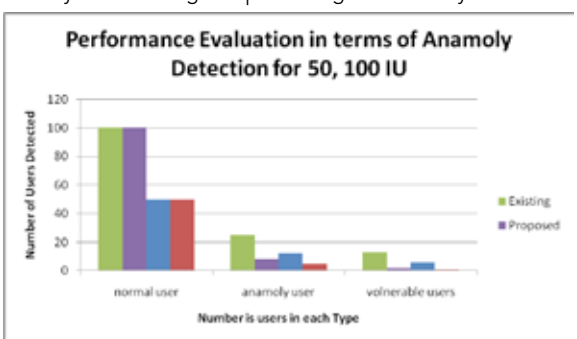


Figure-8: The performance Evaluation of Proposed Approach in term of Anomaly Detection

Conclusion

In this paper the security is completely dynamic, and the key management system for encrypting and decrypting is completelydissimilar than the available systems. After the dynamic validation is over the safety value will getting increased.The study also says that the dynamic authenticationcan be obtained by the ULS, DLS, SLS and MLS, in all the four levels of the CS is associating the user authorizations with IP value and active key. Hence the performance of this approach is high than other approaches like privacy preserving and so on.

REFERENCE

[1]AgentlessRecovery.2013.http://www.ibm.com/developerworks/cloud/library/cl-agentlessrecovery/.[Accessed : January 2013]. | [2] Ajith Abraham. 2001. Neuro-fuzzy systems: state-of-the-art modeling techniques, connectionist models of neurons, learning processes and artificial intelligence, In: Jose Mira, Alberto Prieto(Eds.), Lecture Notes in Computer Science, vol. 2084, Springer-Verlag, Germany, Granada, Spain.269–276. | [3] Anderson, D., Lunt, T.F., Javits, H., Javits, A. Valdes,A. 1995. Detecting Unusual Program Behavior Using the Statistical Components of NIDES. In: NIDES Technical Report, SRI International. | [4] Chen,Y.P. 2003. A hybrid framework using SOM and fuzzy theory for textual classification in data mining;Modeling with Words LNAI.2873: 153–167. | [5] Cho, S.B. and Park, H.J.2003. Efficient anomaly detection by modeling privilege flows with hidden Markov model, Computers and Security. 22 (1):45–55. | [6] Coenen,F., Swinnen,G., Vanhoof, K. and Wets,G. 2000. The improvement of response modeling: combining rule-induction and case-based reasoning, Expert Systems with Applications. 18 (4): 307–313. | [7] Cohen, W.W. 1995. Fast effective rule induction. In: Proceedings of the12th International Conference on Machine Learning. 115–123. | [8] Conversano, C., Roberta, S. and Francesco, M. 2002. Generalized additive multimixture model for data mining , ComputationalStatistics and Data Analysis. 38 (4):487–500. | [9] Debar, H., Becker,M., Siboni,D. 1992. A neural network component for an intrusion detection system, In: Proceedings of 1992 IEEE Computer Society Symposium on Research in Security and Privacy, Oakland,CA.240–250. | [10]DTMF.2013.Distributed Management Task Force. <http://www.dmtf.org/>.[Accessed: January 2013] | [11] Hansen, L.K. and Salaman, P.1990.Neural networks ensembles, IEEE Transactions on Pattern Analysis and Machine Intelligence.12 (10): 993–1001. | [12] Hsu,P.L., Lai,R., Chui,C.C. and Hsu,C.I. 2003. The hybrid of association rule algorithms and genetic algorithm for tree induction: an example of predicting the student course performance, Expert Systems with Application .25 (1): 51–62. | [13] IEEE CCSSG. IEEE.2013. Cloud Computing Standard Study Group.<http :// www. Computer .org / portal/web/sab/cloud>.[Accessed : January 2013] | [14] Indurkhya, N. and Weiss, S.M.1998. Estimating performance gains for voted decision trees, Intelligent Data Analysis .2 (4): 303–310. | [15] ITU, 2013. Cloud Computing Focus Group.<http://www.itu.int/en/ITUT/ focus groups/cloud/Pages/default.aspx>. [Accessed: January 2013] | [16] Kuncheva, L.I., Bezdek, C. and Shutton, M.A.1998.On combining multiple classifiers by fuzzy templates. In: IEEE International Conference on Artificial Neural Networks. 193–197. | [17]Lazarevic,A., Ertoz,L., Kumar,V., Ozgur,A. and Srivastava, J.2003. A comparative study of anomaly detection schemes in network intrusion detection.In: Proceedings of Third SIAM Conference on Data Mining. | [18] Lee, W.and Stolfo, S. 1998. Data mining approaches for intrusion detection.In: Proceedings of the Seventh USENIX Security Symposium, SanAntonio, Texas. | [19] Li,R. and Wang,Z. 2004. Mining classification rules using rough sets and neural networks, European Journal of Operational Research. 157 (2):439–448. | [20] Lin, F.Y. and Mc Clean, S.2001.A data mining approach to the prediction ofcorporate failure, Knowledge-Based Systems.14 (3): 189–195. | [21] Lippmann, R . and Cunningham,S.2000.Improving intrusion detection performance using keyword selection and neural networks,Computer Networks. 34 (4): 594–603. | [22] Lunt, T.F., Tamaru, A., Gilham, F., Jagannathan, R., Jalali, C. and Neuman, P.G. 1992.A Real-time Intrusion Detection Expert System (IDES).In: Technical Report Project 6784, Computer Science Laboratory, SRI International. | [23] OASIS.2013.Organization for the Advancement of Structured Information Standards. <http://www.oasis-open.org/>. [Accessed: January 2013] | [24]OCC.2013.Open Cloud Consortium.<http://www.opencloudconsortium.org/>.[Accessed : January 2013] | [25] OGF.2010.Open Grid Forum, http://www.ogf.org/[Accessed:August 2012] | | [26] Pawlak,Z.1991. Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers. | [27] Rashmi., Sahoo,G. and Mehfuz,S.2013. "Securing Software as a Service Model of Cloud Computing: Issues and Solutions".(IJCCSA).3 (4):4-6. | [28] Sang-Jun Han and Sung-Bae Cho.2003. Detecting intrusion with rule based integration of multiple models, Computers and Security.22 (7): 613–623. | [29]SNIA.2013.Storage Networking Industry Association. <http://www.snia.org/>.[Accessed : January 2013] | [30] SrinivasMukkamala, Andrew Sung and Ajith Abraham.2004. Intrusion detection using ensemble of soft computing paradigms. In: Proceedings of International Conference of Intelligent Systems Design and Applications, Springer-Verlag .239–248. | [31] Suh, E.H., Noh, K.C. and Suh, C.K.1999. Customer list segmentation using the combined response model, Expert Systems with Application. 17 (2): 89–97. | [32] Versace, M., Bhatt, R., Hinds, O. and Shifer, M. 2004. Predicting the exchange traded fund DIA with a combination of genetic algorithm and neural networks, Expert Systems with Application. 27 (3): 417–425. | | [33] Warrender,C., Forrest,S. and. Pearlmutter,B.1999. Detecting intrusion using system calls: alternative data models, in: Proceedings of the IEEE Symposium on Security and Privacy, IEEE Computer Society Press, Oakland, CA. 133–145. | [34] Wong,M.L., Lee,S.Y. and Leung,K.S. 2004. Data mining of Bayesian networksusing cooperative co evolution, Decision Support Systems.38: 451–472. | [35] Zhang, Z. and Zhang, C.2004. Agent- based Hybrid Intelligent Systems, Springer-Verlag, Berlin, Heidelberg. 127 –142. |